

GENES

Benjamin Lewin
Editor, Cell

John Wiley and Sons
New York Chichester Brisbane
Toronto Singapore

Б. ЛЬЮИН Гены

Перевод с английского
канд. биол. наук А. Л. Гинцбурга,
д-ра биол. наук Т. С. Ильиной,
канд. биол. наук Э. С. Каляевой
и
канд. биол. наук Т. Ю. Перслени

под редакцией
чл.-корр. АН СССР Г. П. Георгиева



Москва «Мир»
1987

ББК 28.04 + 28.070

Л91

УДК 575 + 577.2

Льюин Б.

Л91 Гены: Пер. с англ.—М.: Мир, 1987.—544 с., ил.

Книга американского автора представляет собой фундаментальное руководство по молекулярной биологии и генетике, содержащее четкое, полное и ясное изложение современных представлений о структуре генов, организации генома про- и эукариот, механизме синтеза белка и регуляции экспрессии генов, а также о мобильных элементах генома и перестройках ДНК.

Перевод книги сделан по первому изданию, но дополнен новым материалом из второго издания, вышедшего в США в 1985 г.

Предназначена для генетиков, биохимиков, специалистов по биотехнологии, студентов и преподавателей биологических и медицинских вузов.

Л 2001010000-303 128-88, ч. 1
041 (01)-87

ББК 28.04 + 28.070

Редакция литературы по биологии

© 1983 by John Wiley and Sons, Inc. All Rights Reserved. Authorized translation from English language edition published by John Wiley and Sons, Inc.

© перевод на русский язык с дополнениями, «Мир», 1987

ПРЕДИСЛОВИЕ РЕДАКТОРА ПЕРЕВОДА

Предлагаемая вниманию советского читателя книга Бенджамина Льюина широко используется в качестве руководства для подготовки студентов и аспирантов на биологических факультетах многих университетов в США и других странах мира.

Эта книга представляет собой глубокое и детальное описание структуры и функции генетического аппарата как прокариотических, так и эукариотических организмов. Б. Льюин сам в настоящее время не ведет экспериментальной научной работы; он является главным редактором одного из наиболее престижных молекулярно-биологических журналов — «Cell». Благодаря этому через руки Льюина проходит огромное число первоклассных работ, выполняемых в лабораториях всего мира. В результате создается благоприятная возможность для их глубокого осмысления и обобщений. Помимо этого журнал регулярно помещает мини-обзоры, где в очень сжатой форме рассказывается о новейших достижениях науки. Льюин уже опубликовал ряд замечательных книг, в которых обобщил достижения в области молекулярной биологии гена. Книги эти были с удовлетворением встречены мировой научной общественностью. Настоящая книга является, несомненно, наиболее глубоким произведением автора. Обсуждаемые в книге вопросы затрагиваются также и в публикуемом одновременно издательством «Мир» пятитомном издании «Молекулярная биология клетки» (Албертс Б. и др.). Однако ни одна из этих книг не может заменить другую. Так, в частности, вопросы излагаемые в данной книге, рассматриваются гораздо подробнее, чем в «Молекулярной биологии клетки».

Построена книга «Гены» по тому же принципу, что и «Молекулярная биология клетки» — короткие главы, каждая из которых содержит определенное четкое утверждение и его экспериментальное доказательство. Такой стиль изложения весьма эффективен для усвоения материала читателем и для понимания того, как делается наука.

Читатель найдет в этой книге подробные сведения о механизмах трансляции, транскрипции, репликации, амплификации, рестрикции-модификации и рекомбинации генов, о сплайсинге про-мРНК, о процессинге белков, о структуре и функционировании обычных генов, множественных генов и мобильных генетических элементов, о регуляции экспрессии генов, прежде всего регуляции транскрипции, о структурной организации хромосом и, наконец, о механизмах иммунного ответа.

В целом книга написана прекрасным языком, вдумчиво иллюстрирована. Автор нигде не жертвует строгостью изложения результатов в угоду излишнему упрощению. Особенно привлекает умение автора сочетать все необходимые «базисные» сведения с самыми злободневными вопросами современной молекулярной биологии. Такой стиль является залогом того, что книга еще не один год будет служить важнейшим учебником по молекулярной биологии и генетике и в ней найдут для себя много важного не только студенты и аспиранты, но и зрелые научные работники. Для них «Гены» Льюина тоже станут, как я думаю, настольной книгой.

Книгу переводили: А. Л. Гинцбург (гл. 6–13, 39, словарь терминов); Т. С. Ильина (гл. 14–16, 31–38); Э. С. Каляева (гл. 1–5, 28–30) и Т. Ю. Переслени (гл. 17–27).

В 1985 г. в США вышло в свет 2-е издание данной книги, которое отличается от первого как наличием нового материала, так и самой структурой книги. Значительная часть нового материала, представляющая наибольший интерес, переведена и включена в предлагаемое издание.

Г. П. Георгиев

ПРЕДИСЛОВИЕ

Название этой книги само по себе делает предисловие практически излишним. «Гены» — книга про сто о генах, о том, какие достижения имеются в новой научной области, бурное развитие которой едва ли не оттеснило на второй план традиционную генетику. Моя задача — разобраться в огромной массе накопленной к настоящему времени информации, выделить общие принципы и охарактеризовать состояние дел в этой захватывающей области исследований. В книге рассматриваются следующие вопросы: что такое ген, как он воспроизводится, как экспрессируется, каким образом регулируется его экспрессия.

Основная идея книги заключается в том, что существует множество способов сохранения определенного гена в геноме, причем в различных организмах реализуются разные способы. С учетом дальнейшего развития современных исследований в книге уделено равное внимание молекулярной биологии прокариот и эукариот. И те и другие теперь — действующие лица одной и той же истории. Для начала мы рассмотрим вопрос о том, каким образом информация, заключенная в гене, реализуется в виде белка, а затем, уже на молекулярном уровне, вернемся назад, непосредственно к самой ДНК.

Представляя собой законченное введение в молекулярную биологию гена, эта книга не требует наличия специальных знаний и в то же время находится на уровне современных представлений. Я надеюсь, что такой обзор всего имеющегося материала облегчит проникновение в эту быстро развивающуюся область исследований и читатель сможет без труда перейти к изучению более сложных работ. Принимая во внимание объем специальной литературы по данным вопросам, мы отказались от включения в текст ссылок на оригинальные работы, по-

скольку это привело бы только к загромождению книги. Поэтому каждая глава включает библиографический указатель наиболее полезных обзоров и некоторых оригинальных статей, прочтение которых позволит глубже проникнуть в предмет исследований.

Все наиболее важные места изложения я попытался проиллюстрировать схемами и, где это возможно, дать с помощью иллюстраций некоторое представление о соотношениях и взаимосвязи рассматриваемых элементов. Это предисловие будет, разумеется, неполным, если я не выражу признательности John Valbalis за художественное воплощение этого замысла.

Одним из удовольствий, полученных мною от написания этой книги, было последующее обсуждение ее с моими друзьями и коллегами, сделавшими критические замечания. Немало поправок было внесено благодаря не жалеющим своих усилий Sankar Adhya, Sidney Altman, French Anderson, David Clayton, Nicholas Cozzarelli, Bernard Davis, Igor Dawid, Arg Efstratiadis, Nina Federoff, Alice Fulton, Joe Gall, Nicholas Gillham, Philip Hanawalt, Ira Herskowitz, Lee Hood, Joel Huberman, George Khoury, Nancy Kleckner, Marilyn Kozak, Charles Kurland, Art Landy, Jeffrey Miller, Masayasy Nomura, Charles Radding, Jeff Roberts, Rich Roberts, Gerry Rubin, Robert Schimke, David Schlessinger, David Shafritz, Phil Sharp, Allen Smith, Phang-C. Tai, Susumu Tonegawa, Harold Varmus, Alex Varshavsky и Hal Weintraub.

И наконец, вряд ли есть необходимость говорить о том, что писать эту книгу было бы отнюдь не таким удовольствием, если бы не исполненное энтузиазма участие в этом процессе членов моей семьи.

Б. Льюин

Часть I

ПРИРОДА ГЕНЕТИЧЕСКОЙ ИНФОРМАЦИИ

«Называя структуру хромосомных нитей «кодовым письмом», мы имеем в виду, что всепроникающий разум может по этой структуре предсказать, разовьется ли данное яйцо в подходящих условиях в черного петуха или пеструю курицу, в муху или в растение кукурузы, в жука, мышь или в женщину... Но этот термин «кодовое письмо», конечно, слишком узок. Хромосомные структуры служат, кроме того, и инструментом, осуществляющим то развитие, которое они же предопределяют. Они и кодекс законов, и исполнительная сила, или (используя другую аналогию) они и архитектурный проект, и строительная бригада в одно и то же время».

Эрвин Шрёдингер, 1945

Глава 1

ЧТО ТАКОЕ ГЕН?

ГЕНЕТИЧЕСКАЯ ТОЧКА ЗРЕНИЯ

В концепции гена сфокусированы результаты примерно ста лет работы, потребовавшейся для того, чтобы разгадать основу наследственности. Проведенные исследования были столь многоплановы, что нельзя однозначно и удовлетворительно ответить на вопрос: что такое ген?

Ген — это последовательность ДНК, несущая информацию об определенном белке. Еще совсем недавно это казалось вполне достаточным (если не полным) биохимическим определением. Последовательность ДНК можно было бы определить как непрерывную линейную последовательность нуклеотидов, коллинеарную аминокислотной последовательности соответствующих белков. Однако, как теперь уже установлено, последовательность, в которой закодирован белок, не всегда непрерывна. Она может прерываться вкрапленными в нее некодирующими участками. Таким образом, гены могут состоять из отдельных кусков, соединяющихся воедино в процессе генной экспрессии.

Ген можно идентифицировать по кластеру (группе) мутаций, каждая из которых предотвращает образование соответствующего белка. Хотя данное положение остается справедливым и в отношении прерывистых генов, свойства их кластеров в этом случае оказываются значительно сложнее. Такое утверждение достаточно спорно, поскольку большинство прерывистых генов обнаружено в таких ситуациях, в которых детальный генетический анализ невозможен. Однако очевидно, что генетический взгляд на ген необходимо пересмотреть.

У всех видов большое число генов, составляющих геном, организовано в сравнительно небольшое число хромосом. Генетический материал каждой хромосомы представляет собой чрезвычайно длинную последовательность ДНК, содержащую множество линейно расположенных генов. Довольно долго оставалось загадкой, сколько всего генов существует в геноме. Представление о том, что каждый ген может существовать сам по себе, как уникальное целое, пришлось пересмотреть, когда было обнаружено, что во многих случаях кластеры родственных генов организованы в небольшие семейства.

Обычно считалось, что геном довольно стабилен и изменения в его общей структуре и организации происходят крайне редко — лишь в растянутой эволюционно-временной шкале. Однако это противоречит недавно полученным данным о том, что в некоторых случаях в ДНК периодически происходят перестройки и что существуют относительно мобильные компоненты генома.

Сама концепция гена, таким образом, подверглась эволюции. И хотя многие из традиционных положений данной концепции остались без изменений, были обнаружены исключения, свидетельствующие о том, что ни одно из положений нельзя считать абсолютным. Со времени открытия гена как неизменной единицы наследственности его свойства описывали, исходя из того, что положение гена в хромосоме фиксировано. Из этого в свою очередь было развито представление о том, что генетический материал хромосомы представлен непрерывной нитью ДНК, соответствующей множеству генов. Следуя за раз-

витием этих идей, мы сможем дать операциональное описание гена, но это вряд ли поможет нам сформулировать краткое определение.

Элементарный фактор наследственности

«Постоянные признаки, появляющиеся в данной группе растений, могут быть получены во всех комбинациях в соответствии с математическими законами сочетаний. Те признаки, которые при передаче потомству проявляются и, следовательно, обнаруживаются у гибридов, называют *доминантными*, а те, которые оказываются в скрытом состоянии, называют *рецессивными*. Выражение «рецессивный» было выбрано потому, что такие признаки как бы отстают или полностью исчезают у гибридов, но вновь появляются в их потомстве без каких-либо существенных изменений и в предсказуемом соотношении. Промежуточные формы ни в одном из экспериментов обнаружены не были.»

Эта (сокращенная) цитата из замечательной статьи Менделя, опубликованной в 1865 г., вводит основную концепцию генетики: существует единица наследственности в виде некоторого фактора, который передается от родителей потомству. Этот дискретный фактор не что иное, как хорошо нам теперь известный ген. Работа Менделя пролила свет на общее поведение гена при наследовании в ряду поколений. Из этой работы следовало, что отдельный живой организм — это способ, которым ген экспрессируется и увековечивает себя. Суть этого представления изображена на **рис. 1.1** в виде чередования поколений.

В отличие от запутанных ранее представлений о наследственности работа Менделя утвердила генетику как экспериментальную науку, введя в нее несколько новых подходов. Во-первых, исходный набор растений был «чистым по признаку»: каждое растение давало в потомстве только растения с данным признаком и никогда — с противоположным. Сначала было изучено наследование только одного признака. После этого можно было одновременно исследовать наследование двух признаков или более. Свои эксперименты Мендель поставил на количественную основу, проводя учет потомков каждого типа в каждом скрещивании.

Свойства генов, выявленные в работе Менделя, сформулированы в его первом и втором законах. Первый закон рассматривает особенности индивидуального гена. Организм имеет две копии каждого гена, т. е., выражаясь современным языком, он *диплоиден*. Только одна из двух копий гена попадает от родителя к потомку через *гаметы* (половые клетки). Гаметы, сливаясь, образуют *зиготу* (оплодотворенное яйцо), несущую по одной копии от каждого родителя. В результате каждый организм получит одну отцовскую копию гена и одну материнскую.

Ген может существовать в альтернативных формах, что проявляется в различии признаков — например, красная окраска цветка в противовес белой. Эти формы гена

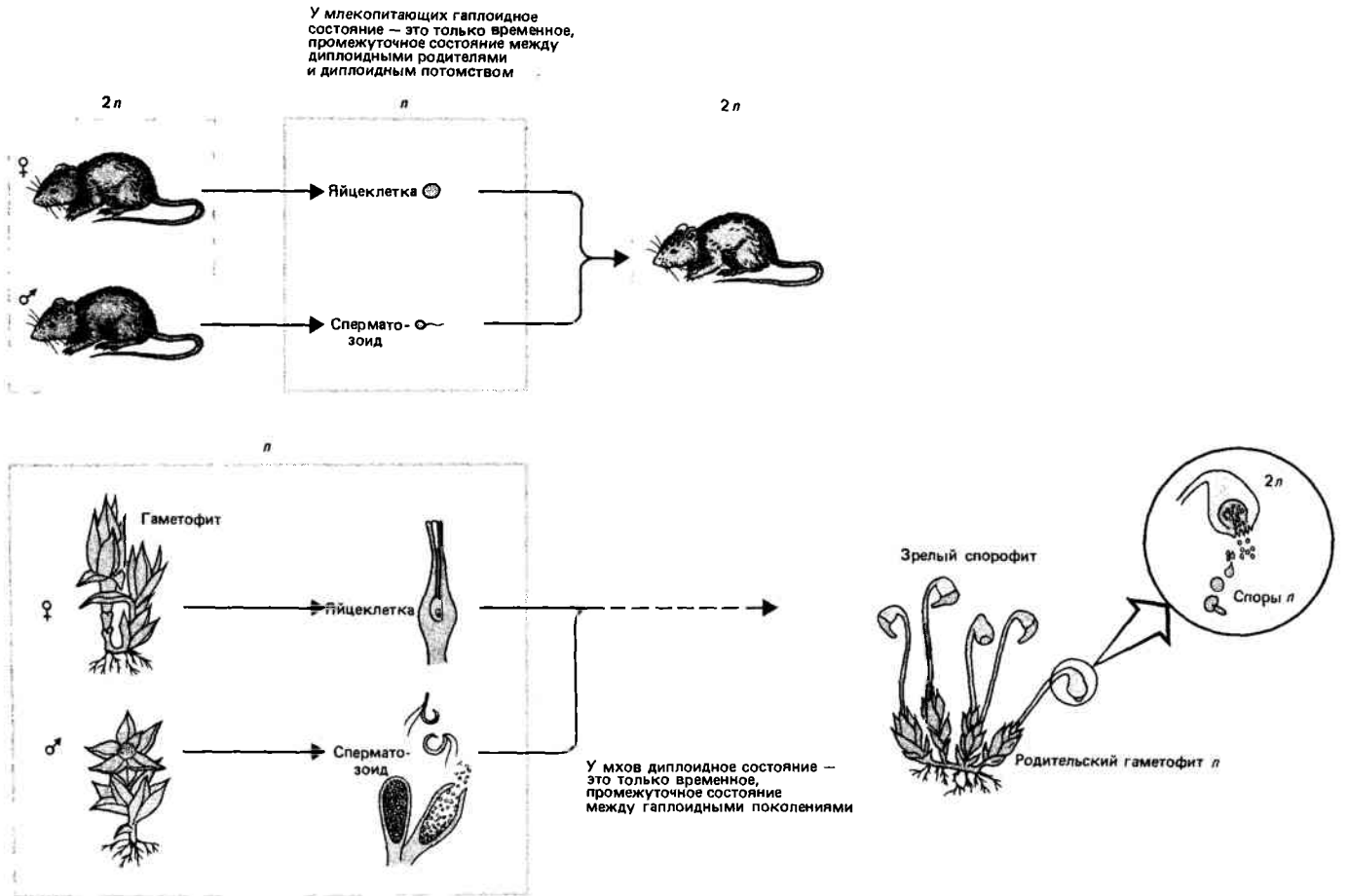


Рис. 1.1. Когда эукариоты «увекочивают» свои гены путем чередования диплоидного и гаплоидного состояний, одно из состояний преобладает у «взрослого» организма, тогда как другое связано с образованием гамет или зигот. Млекопитающие

и мхи — это крайние примеры. У первых взрослое поколение диплоидно, у вторых — гаплоидно.

Красным цветом указаны диплоидные (2n) ткани, серым — гаплоидные (n).

называют **аллелями**. Закон *независимого расщепления* утверждает, что присутствуя одновременно в одном растении, эти аллели не изменяют друг друга, но расщепляются, переходя в различные гаметы, при формировании следующей генерации.

У организма, чистого по признаку, т.е. у **гомозиготы** оба аллеля **одинаковы**. Но скрещивание двух родителей, каждый из которых гомозиготен по *разным* аллелям, приведет к образованию гибрида, или **гетерозиготы**. Если один аллель **доминантен**, а другой **рецессивен**, то организм по своему внешнему виду, или **фенотипу**, будет только доминантного типа (так что гетерозиготы не отличаются от чистого по признаку доминантного родителя).

Но, согласно первому закону Менделя, генетическая конституция, или **генотип**, гибрида включает в себя оба аллеля. Как показано на **рис. 1.2**, это можно установить, если скрестить между собой два гибрида и получить второе поколение. Критическое значение при этом имеет тот факт, что аллели не смешиваются друг с другом («промежуточных форм не обнаружено»), а сохраняют свою физическую целостность, взаимодействуя только на уровне экспрессии.

☞ Хотя в опытах Менделя исследованные признаки проявляли (к счастью) полное доминирование, так бывает не всегда. **Доминирование** аллелей может быть **неполным** (ча-

стичным) или совсем отсутствовать (что иногда называют **кодоминантностью**). В последнем случае гетерозиготы отличаются от гомозигот промежуточными свойствами. Например, скрещивание растений львиного зева с красными и белыми цветами дает розовые гибриды. Хотя в этом случае нет доминирования, наблюдается все та же закономерность — гибриды первого поколения (F1) одинаковы. В дальнейших скрещиваниях гибридов образуются те же численные соотношения, как и при полном доминировании, хотя вместо двух фенотипов можно различить три.

Независимость различных генов

Во втором законе Менделя утверждается *независимость комбинации различных генов*. При скрещивании растения, доминантного по двум различным признакам, с растением, у которого эти признаки рецессивны, F1 по-прежнему состоит из одинаковых растений доминантного типа. Но при последующем скрещивании гибридов F1 появляются два класса растений. Один класс состоит из растений исходных **родительских типов**. Второй класс включает растения новых фенотипов, с доминантными признаками одного родителя и рецессивными признаками другого. Эти фенотипы называют **рекомбинантными**

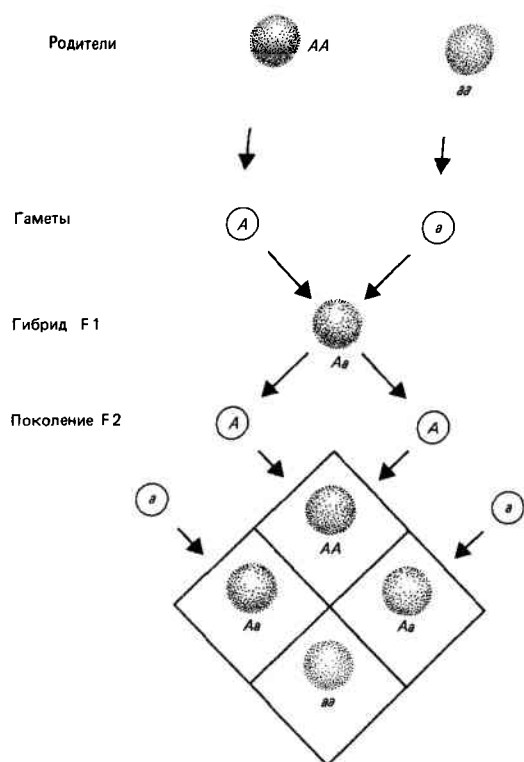


Рис. 1.2. Аллели расщепляются независимо.

Два родителя гомозиготны: AA имеет две копии доминантного аллеля; aa имеет две копии рецессивного аллеля. Каждый образует только один тип гамет, так что F1 (первое гибридное поколение) состоит из одинаковых гибридных особей Aa. Поскольку A доминирует над a, фенотип потомков Aa такой же, как у родителя AA (показано цветом). Фенотип рецессивной гомозиготы aa показан отсутствием цвета.

Каждый гибрид F1 образует гаметы A и a в одинаковом количестве. При скрещивании объединение этих гамет приведет к образованию F2 (второе гибридное поколение), которое будет состоять из 1AA : 2Aa : 1aa. Так как AA и Aa имеют одинаковый фенотип, получается классическое соотношение доминантного и рецессивного фенотипов—3:1.

В тех случаях, когда фенотип гетерозиготы занимает промежуточное положение между родительскими AA и aa, F1 будет другим, а в F2 отношение фенотипов будет: 1 доминантный : 2 промежуточных : 1 рецессивный.

типами. Они встречаются в обеих возможных (реципрокных) комбинациях.

На рис. 1.3 показано, что соотношение потомков различного типа можно объяснить, предположив, что в процессе образования гамет происходит совершенно случайное объединение одного из аллелей первого признака с одним из аллелей второго. Четыре возможных типа гамет образуются в равных соотношениях; при формировании зиготы следующего поколения они объединяются случайно.

Типичное соотношение фенотипов скрывает за собой большее разнообразие генотипов, что может быть установлено путем **возвратного** (анализирующего) скрещивания с рецессивным родителем. По существу, возвратное скрещивание позволяет прямо установить генотип исследуемого организма (рис. 1.4).

Закон независимого распределения генов, как следует из сказанного выше, говорит, что поведение любой пары (или большего числа) генов можно полностью предсказать с помощью правил комбинаторики. Распределение между потомками (наследование) одного гена не влияет на распределение другого. Эта концепция подразумевает, что распределение генов происходит со *статистической вероятностью*, а не предопределено точно. Чем больше проведено скрещиваний, тем точнее будет совпадать с предсказанным соотношением типов в потомстве.

Роль хромосом в наследственности

Открытие Менделя не сразу оценили по достоинству, поскольку не было никаких представлений о физической основе постулируемых факторов. Хромосомная теория наследственности, предложенная одновременно Саттоном и Бовери (Sutton, Boveri), разрешила долгие споры о возможной роли хромосом. Уже существовали некоторые, хотя довольно туманные, предположения о том, что хромосомы как-то участвуют в механизмах наследования. В 1903 г. наконец стало ясно, что по своим свойствам они точно совпадают с дискретными единицами наследственности, описанными Менделем.

Клеточная теория, утвердившаяся в середине девятнадцатого века, предполагала, что все организмы состоят из клеток и что эти клетки могут возникать только из предсуществующих клеток. Ранние цитологические исследования показали, что «типичная» клетка состоит из плотного ядра, отделенного мембраной, от менее плотной окружающей цитоплазмы. Внутри ядра с помощью определенных красителей можно было различить зернистые участки—**хроматин**. Вскоре после работы Менделя было установлено, что хроматин представлен определенным числом нитевидных образований, или **хромосом**.

В большинстве клеток хромосомы видны только в процессе клеточного деления. Два типа деления, присущие организмам, размножающимся половым путем, объясняют и постоянство генетического материала, и процесс наследования в том виде, как это было предсказано законами Менделя.

Клеточный цикл у растущего организма состоит из двух этапов. Более длинный период—**интерфаза**, когда клетка синтетически активна и занята воспроизведением своих компонентов. Затем следует короткий период—**митоз**, интерлюдия, во время которой фактически завершается процесс разделения на две дочерние клетки. Клетки, возникающие в результате ряда митотических делений и образующие целый организм, называют **соматическими клетками**.

Каждая дочерняя клетка, начинающая свою жизнь после митоза, содержит по две копии каждой хромосомы. Их называют **гомологичными**. Общее число хромосом в клетке, известное как **диплоидный набор**, обозначают $2n$. Типичная соматическая клетка существует в диплоидном состоянии (кроме того периода, когда она готовится к делению или уже делится).

В интерфазе растущая клетка удваивает свой хромосомный материал. Однако это становится очевидным только в последующем митозе. В митозе каждая хромосома разделяется вдоль по длине, образуя две копии—**сестринские хроматиды**. В этот момент клетка содержит $4n$ хромосом, организованных в $2n$ пар сестринских хроматид. Иными словами, в клетке имеется по две (гомологичные) копии каждой пары сестринских хроматид. На рис. 1.5 показана последовательность процессов, обеспечивающих митотическое деление. Суть заключается в том, что сестринские хроматиды растаскиваются к противоположным полюсам клетки, так что каждая дочерняя клетка получает по одной копии каждой сестринской хроматиды. Теперь это самостоятельные хромосомы. $4n$ хромосомы, существовавшие в начале деления, разделились на два набора по $2n$ хромосом. Этот процесс повторяется в следующем клеточном цикле. Таким образом, митотическое деление гарантирует постоянство набора хромосом в соматических клетках.

1. Что такое ген? Генетическая точка зрения

Рис. 1.3. Различные гены комбинируются независимо.

Один родитель гомозиготен по двум доминантным генам — по гену *A*, определяющему цвет, и гену *B*, определяющему форму (круглая форма). Другой родитель гомозиготен по рецессивным аллелям *a* и *b* (отсутствие цвета, морщинистая форма). Потомство F1 одинаково — с доминантным фенотипом.

Родители F1 образуют гаметы, в которых происходит независимое расщепление аллелей и независимая комбинация генов, так что образуется равное количество гамет каждого из четырех возможных типов. Они соединяются случайно и образуют 9 генотипических классов, которые из-за отношений доминантности проявляются в виде четырех фенотипических классов: 9 окрашенных гладких : 3 окрашенных морщинистых : 3 бесцветных гладких : 1 бесцветный морщинистый.

Заметим, что каждый реципрокный генотип представлен одинаковым числом особей, например два родительского типа (по одному *AABB* и *aabb*) и два рекомбинантных (по одному каждого рекомбинантного класса *AAbb* и *aaBB*). Соотношение 3 : 1 сохраняется для каждого независимо расщепляющегося признака.

Число фенотипических классов будет больше, если один или оба признака не проявляют полного доминирования. (В этом случае гетерозиготы отличаются от любой гомозиготы.) Их будет меньше, если два гена действуют на один признак. Так, если бы оба гена *A* и *B* были нужны для развития окраски, то у гибридов F2 наблюдалось бы соотношение 9 окрашенных : 7 бесцветных.

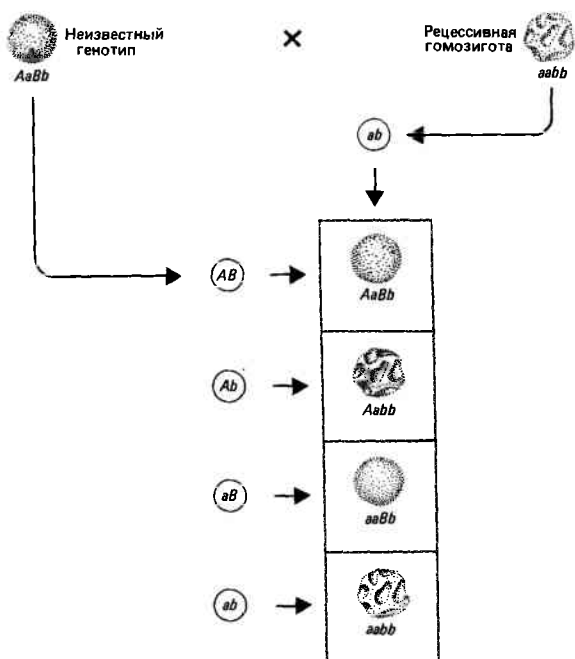
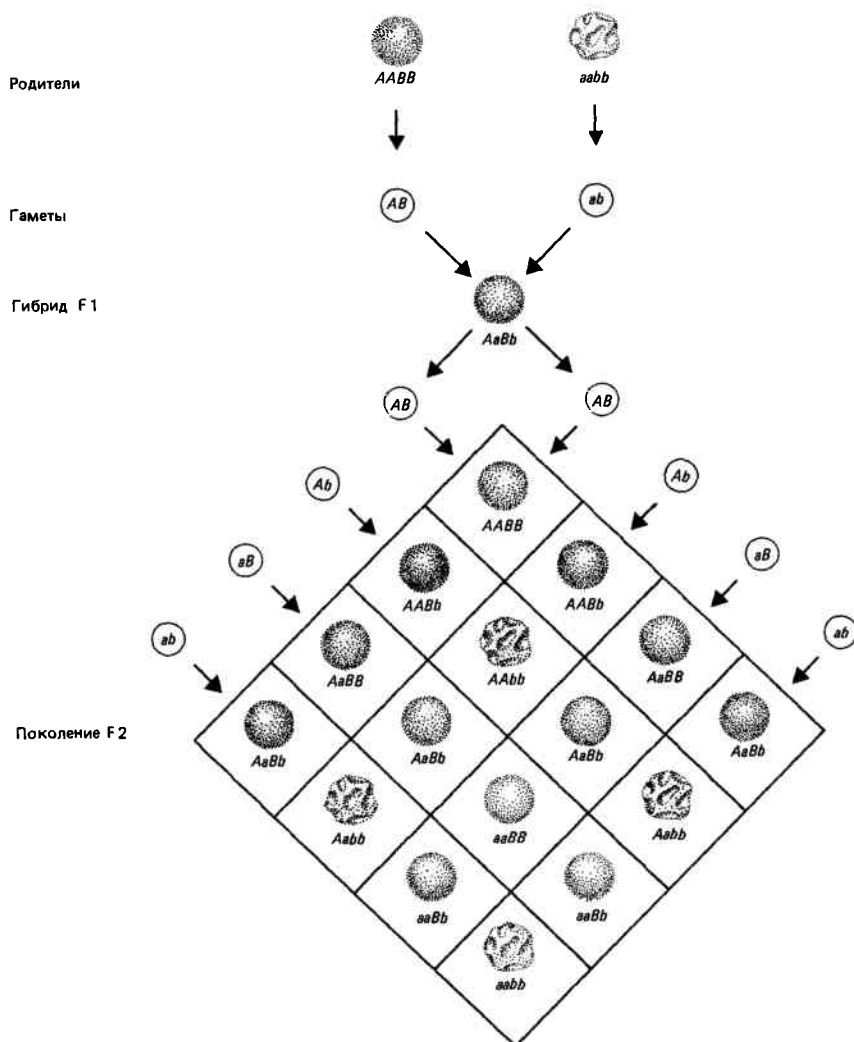


Рис. 1.4. Возвратное скрещивание с рецессивной гомозиготой используется для определения генотипа.

Соотношение гамет, образованных растением с неизвестным генотипом, отражает обычно независимое расщепление аллелей и независимую комбинацию генов. Генотип каждой гаметы выясняется непосредственно в скрещивании с гаметой, несущей только рецессивные аллели.

Иной цели служит процесс образования половых клеток (гамет, т. е. яйцеклеток или сперматозоидов), который происходит при другом типе деления. Этот путь деления изображен на рис. 1.6. В процессе мейоза образуются клетки, содержащие гаплоидное число (*n*) хромосом. Этот процесс включает в себя два последовательных деления. И в этом случае хромосомы сначала удваиваются, так что клетка приступает к делению в состоянии $4n$.

В начале первого деления гомологичные пары сестринских хроматид конъюгируют (соединяются попарно), образуя биваленты. Каждый бивалент содержит все четыре копии одного гомолога, находящиеся в клетке. В первом делении каждый бивалент расщепляется на две пары сестринских хроматид, разделенных в свою очередь на полухроматиды. В результате образуются два набора по $2n$ хромосом, из *n* пар сестринских хроматид каждый.

Далее следует второе мейотическое деление, в котором оба набора $2n$ делятся еще раз. Это деление формально подобно митотическому делению, поскольку в различные дочерние клетки попадает по одной из копий каждой пары сестринских хромосом.

Общий результат мейоза заключается в том, что исходное число хромосом $4n$ делится на четыре гаплоидные клетки, дающие начало зрелым яйцеклеткам или сперматозоидам.

При формировании гамет гомологи отцовского и материнского происхождения разделяются так, что каждая гамета получает только один из двух гомологов своих ро-

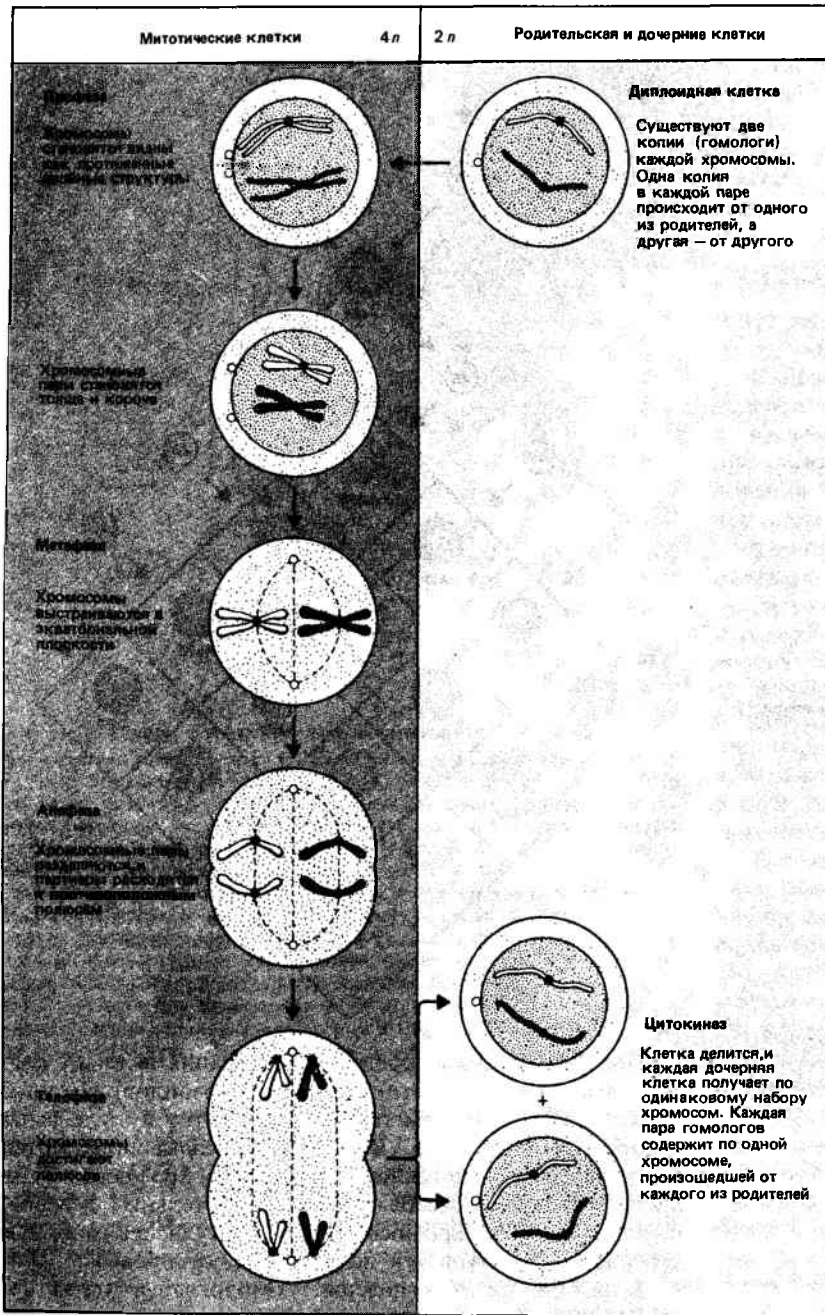


Рис. 1.5. Митоз обеспечивает постоянство хромосомного состава клетки.

Одна пара гомологичных хромосом (обозначена белым) происходит от одного из родителей, другая (окрашена) — от второго родителя. Диплоидная клетка имеет по две копии каждой хромосомы (общее число $2n$). Перед митозом хромосомы удваиваются, и профаза, следовательно, начинается с состояния $4n$. Два члена каждой пары разделяются на две отдельные хромосомы, которые затем расходятся в разные дочерние клетки. При этом восстанавливается диплоидное число $2n$, так что дочерние клетки имеют тот же набор хромосом, что и родительская клетка.

дителей. Существенным моментом в этом процессе является тот факт, что в полном соответствии с законами Менделя наблюдается *независимость* сегрегации *негомологичных* хромосом. Каждый член одной гомологичной пары входит в гамету в случайном сочетании с любым членом другой гомологичной пары.

Подытожим параллели между хромосомами и менделевскими единицами наследования. Гены встречаются в аллельных парах — по одному аллелю от каждого родителя в каждой паре; диплоидный набор хромосом образуется из двух гаплоидных родительских наборов. Распределение неаллельных генов в гаметах происходит независимо; негомологичные хромосомы подвергаются независимой сегрегации. Критическое условие, заключающееся в том, чтобы каждая гамета получала полный гаплоидный набор, выполняется независимо от того, рас-

считываем ли мы этот процесс как сегрегацию хромосом или как распределение элементарных факторов наследственности.

Гены располагаются в хромосомах

Только установив, что определенный ген всегда присутствует в определенной хромосоме, можно доказать, что гены находятся в хромосомах. Такое доказательство было получено благодаря свойствам одного варианта плодовой мушки *Drosophila melanogaster*, обнаруженного Морганом в 1910 г. Этот белоглазый самец появился спонтанно в линии мух с обычными красными глазами.

Поскольку обычно встречаются мухи с красными глазами, их назвали **диким типом**; белые же глаза — это **мутантный** фенотип. Событие, приводящее к возникновению

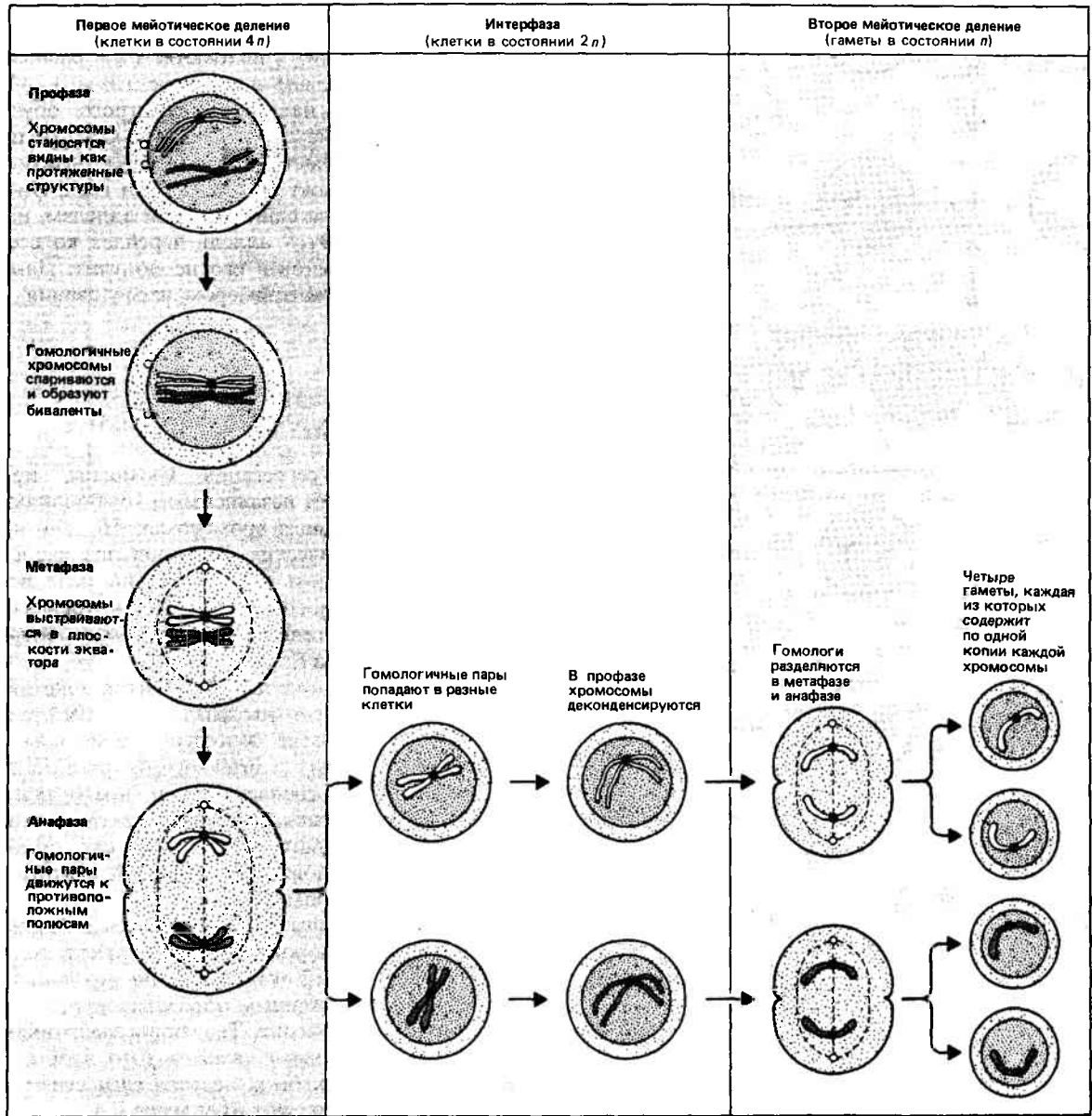


Рис. 1.6. В мейозе число хромосом уменьшается вдвое. На рисунке показано поведение одной пары гомологичных хромосом. Оба члена пары были удвоены перед началом профазы. Во время первого деления происходит спаривание гомологичных хромосом, а затем каждая пара расходит в разные клетки. Во время второго деления дублированные члены каждой пары расходятся в разные клетки, так что

каждая гамета получает по одной копии. В клетке обычно находится несколько пар гомологичных хромосом, и члены каждой пары комбинируются между собой независимо друг от друга. Таким образом, подбор гомологичных хромосом (материнской и отцовской) в каждой гамете происходит случайно.

мутантного фенотипа, представляет собой генетическое изменение, называемое **мутацией**. В результате большинства мутаций (но не всех) функционирование соответствующего гена нарушается частично или полностью. Поэтому изучение мутаций, как правило, сводится к изучению неактивных аллелей. Иногда происходят обратные мутации (реверсии), т.е. такие изменения в генетическом материале, которые восстанавливают исходное состояние, — явление, называемое **реверсией** к дикому типу.

Мутацию *white* удалось локализовать в определенной хромосоме потому, что она оказалась связанной с полом. Обычно гомологичные пары хромосом при расхождении в мейозе образуют идентичные гаплоидные наборы. Однако у многих организмов, размножающихся половым

путем, существует одно исключение из этого общего правила: самцы и самки могут иметь видимые различия в хромосомных наборах. Чаще всего у *одного* из полов одна хромосомная пара состоит из разных хромосом.

Такую пару называют **половыми хромосомами**, а остальные гомологичные пары — **аутосомами**. Хромосомные наборы двух полов можно записать следующим образом: $2A + XX$ и $2A + XY$, где гаплоидный набор аутосом обозначают как *A*, а две половые хромосомы — как *X* и *Y*. Пол с хромосомным набором $2A + XX$ называют **гомогаметным**; у него образуются только гаметы $A + X$. Пол с набором $2A + XY$ называют **гетерогаметным**, так как у него образуются в одинаковом количестве гаметы двух типов: $A + X$ и $A + Y$. В результате случайного со-

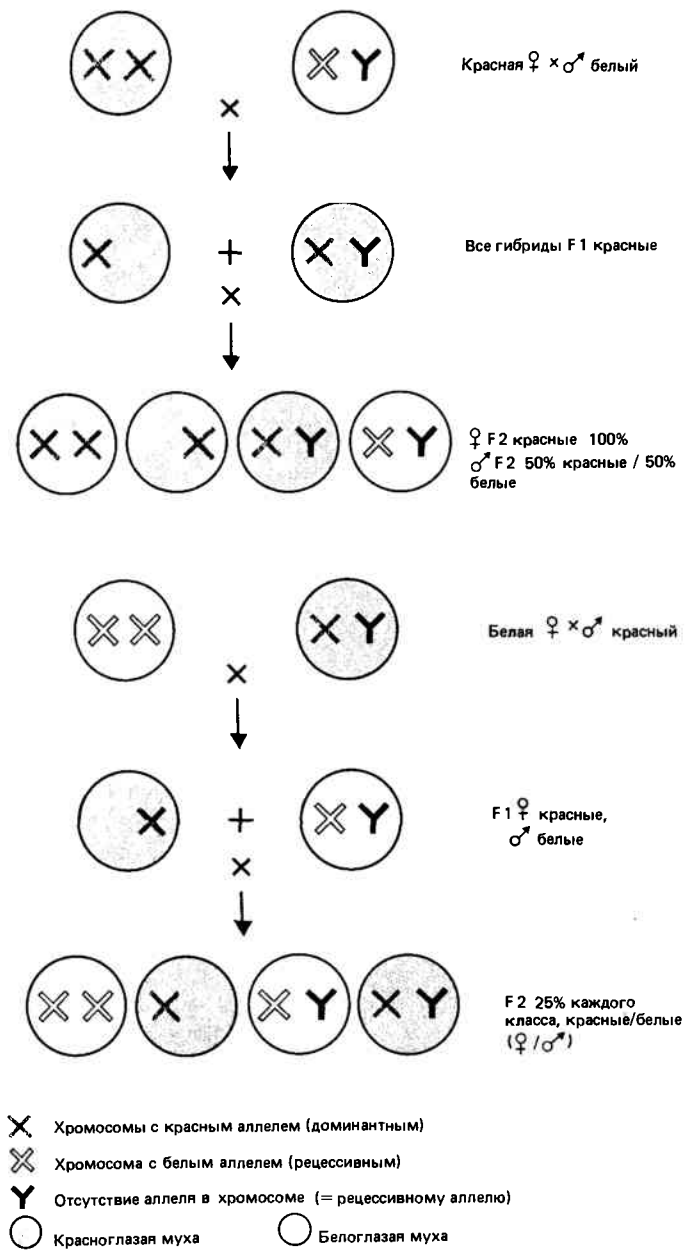


Рис. 1.7. Гены X-хромосомы наследуются сцепленно с полом. Красная или белая окраска глаз у самца зависит только от X-хромосомы, полученной им от матери. Фенотип самки зависит от того, получит ли она доминантный аллель от одного из родителей. При этом наблюдается характерный тип перекрестного наследования, сцепленного с полом, — от отца к дочерям, от матери к сыновьям (крисс-кросс).

единения гамет одного пола с гаметами другого пола постоянно сохраняется равное соотношение полов при образовании зигот. У *Drosophila* гомогаметным полом являются самки.

Из законов Менделя следует, что результаты генетического скрещивания должны быть одинаковы независимо от того, кому из родителей принадлежит данный аллель. Но реципрокные скрещивания с признаком «белые глаза» у *Drosophila* дают различные результаты (рис. 1.7).

Скрещивание белоглазый самец × красноглазая самка дает лишь красноглазое потомство F1, что и можно было предсказать, если красный цвет доминантен, а белый рецессивен. Но в потомстве F2 все появившиеся белоглазые мухи оказались самцами. В реципрокном скрещи-

вании красноглазые самцы × белоглазые самки все самцы F1 были с белыми глазами, а все самки — с красными. При их скрещивании в потомстве F2 в равном соотношении появились белые и красные глаза у обоих полов.

Этот пример наследования строго обусловлен половыми хромосомами. Если аллели красных и белых глаз находятся в X-хромосоме, и при этом Y-хромосома вообще не содержит локуса для окраски глаз, фенотип самца будет определяться единственным аллелем, находящимся в X-хромосоме. Этот аллель перейдет ко всем дочерям, но ни один из сыновей его не получит. Данный случай является типичным примером наследования, **сцепленного с полом**.

Гены линейно выстроены вдоль хромосом

Независимая сегрегация хромосом, происходящая в мейозе, объясняет независимую комбинацию генов, расположенных в разных хромосомах. Но, как известно, общее число генетических факторов значительно больше, чем число хромосом. Следовательно, если все гены расположены в хромосомах, то каждая хромосома должна содержать много генов. Каковы же взаимоотношения между этими генами?

О стремительном темпе развития генетики в начале века свидетельствует тот факт, что в 1913 г., всего лишь через три года после сообщения о первом белоглазом мутанте, Стертевант (Sturtevant) сообщил об изучении наследования шести сцепленных с полом мутаций. В 1911 г. Морган (Morgan) показал, что каждый из этих факторов ведет себя в скрещиваниях так же, как красная и белая окраска глаз. Следовательно, каждый фактор должен располагаться в X-хромосоме.

Некоторые из факторов оказались связанными между собой. Вопреки второму закону Менделя, доля родительских генотипов в F2 оказалась выше предполагаемой, что объяснялось пониженным образованием рекомбинантных типов в этом поколении. Тенденция некоторых признаков оставаться связанными, вместо того чтобы комбинироваться независимо, была названа **сцеплением**. На рис. 1.8 показано, как измеряют сцепление.

Морган предположил, что причина генетического сцепления факторов — это «просто механический результат их локализации в одной хромосоме». Он предположил также, что образование генетических рекомбинантов можно отождествить с процессом **кроссинговера**, наблюдаемого в мейозе. В раннем мейозе, на стадии, когда четыре копии каждой хромосомы представлены в виде бивалента, между близко расположенными (конъюгированными) гомологичными парами происходит попарный перекрест генетического материала, названный **хиазмой**. Этот процесс схематически изображен на рис. 1.9.

Забегим на несколько лет вперед. Только в 1931 г. было формально доказано, что рекомбинация обусловлена кроссинговером. У кукурузы и у дрозофилы были получены подходящие **транслокации**, при которых оторванная часть одной хромосомы прикрепилась к другой. Это позволяет отличить по внешнему виду хромосому с транслокацией от нормальной хромосомы. Как изображено на рис. 1.10, в подобных скрещиваниях можно показать, что образование генетических рекомбинантов происходит только при физическом обмене между соответствующими областями хромосом.

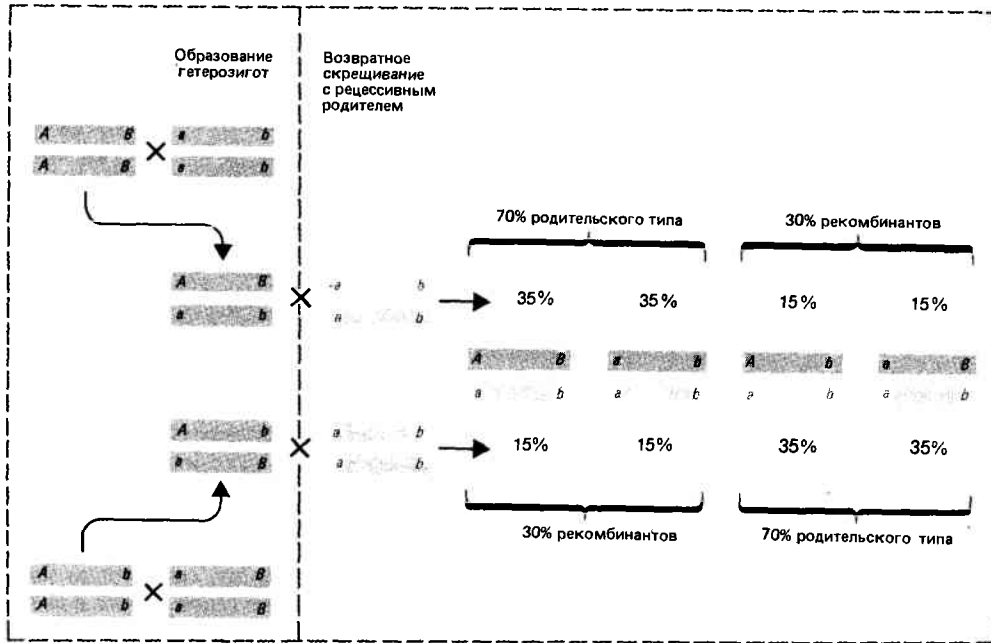


Рис. 1.8. Сцепление можно измерить при помощи возвратного скрещивания (бэккросс) с двойной гомозиготой.

Хромосомы изображены в виде окрашенных полосок. Гены гетерозиготного родителя обозначены черными буквами, гены родителя, рецессивного по двум признакам, — цветными буквами. Это скрещивание относится к тому же типу, что и скрещивание, показанное на рис. 1.4.

В скрещивании, изображенном сверху, одна хромосома несет оба доминантных аллеля, другая — оба рецессивных (скрещивание с аллелями «в фазе притяжения»). Таким образом, родительские типы представляют собой AB и ab , рекомбинантные типы — Ab и aB .

Если вероятность образования хиазмы между двумя точками в хромосоме зависит от расстояния между ними, гены, расположенные ближе друг к другу, будут наследоваться вместе. По мере увеличения расстояния между двумя генами возрастает и вероятность кроссинговера между ними. Таким образом, если рекомбинация обусловлена кроссинговером, гены, находящиеся близко друг к другу, должны быть тесно сцеплены, причем генетическое сцепление будет уменьшаться по мере физического удаления. И наоборот, генетическое сцепление можно использовать как меру физического расстояния.

Исходя из представления о том, что гены одной хромосомы сцеплены между собой, Стертевант (Sturtevant) предложил использовать частоту рекомбинации в качестве единицы **расстояния на карте** для измерения относительной локализации генов. Эта единица расстояния выражается как процент рекомбинации:

$$\text{Расстояние на карте} = \frac{\text{Число рекомбинантов} \times 100}{\text{Общее число потомков}}$$

Расстояние измеряется в условных **единицах карты**; одна единица карты (сантиморганида) соответствует 1% рекомбинантных потомков. Основные закономерности рекомбинации были установлены при изучении всех шести сцепленных с полом признаков вместе. Индивидуальные расстояния на карте **аддитивны**. Иными словами, если два гена A и B находятся на расстоянии 10 ед. друг от друга, а расстояние от B до C равно 5 ед., то расстояние между A и C , полученное при прямом измерении,

будет близко к 15. Поэтому гены можно расположить в **линейном порядке**. Критическое значение для построения генных карт имеет тот факт, что расстояние между генами не зависит от использованных **аллелей**, а зависит только от положения генных **локусов**. Лocus — это **место** в хромосоме, в котором находится ген данного признака. Все различные альтернативные формы гена, т.е. аллели, используемые при картировании, находятся в **одном и том же** локусе. Таким образом, генетическое картирование позволяет установить взаимную локализацию генных локусов, расположенных в хромосоме в определенном месте и в линейном порядке. В экспериментах по картированию один и тот же результат получается независимо от конкретной комбинации аллелей (рис. 1.8). Стертевант заключил, что его результаты «служат новым доводом в пользу хромосомной теории наследственности, поскольку они убедительно показывают, по крайней мере математически, что исследуемые факторы расположены в виде линейных рядов».

Этот вывод представляет большой интерес. Можно построить генетическую карту, в которой хромосома будет изображена в виде линейного ряда генов. Однако это еще не доказывает, что генетическое содержание хромосомы представлено **физически** непрерывным рядом генов.

Рассматривалось множество других моделей хромосомы, пока не стало очевидным, что хромосома состоит только из одной-единственной нити генетического материала.

Не все пары генов, локализованных в одной хромосоме, сцеплены. На основе второго закона Менделя макси-

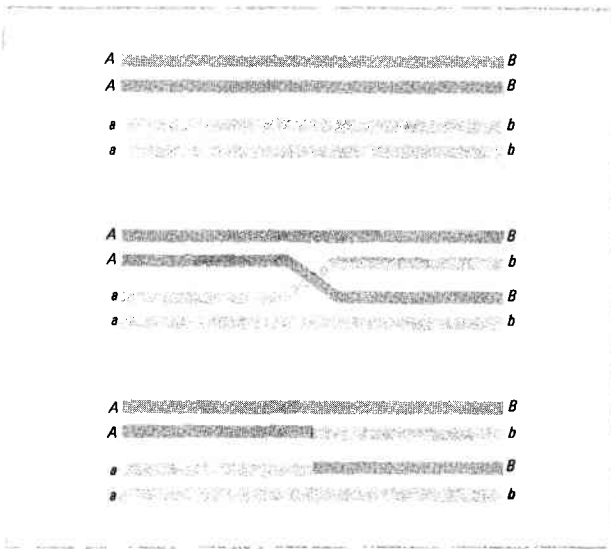


Рис. 1.9. Образование хиазмы приводит к образованию рекомбинантов.

Каждая полоска изображает хромосому, и каждому родителю соответствуют две гомологичные хромосомы. В профазе мейоза они соединяются, образуя **бивалент**. Изображение в верхней части схемы представляет собой увеличенный вариант картины, показанной во втором положении рис. 1.6. В средней части диаграммы показан **кроссинговер** между двумя хромосомами. В результате образуются две рекомбинантные хромосомы, несущие гены Ab и aB (нижняя часть схемы). Другие две хромосомы сохраняют исходный родительский тип (AB и ab).

В некоторых отношениях образование хиазмы аналогично процессу генетической рекомбинации. Образование одной хиазмы, например, уменьшает вероятность возникновения близости другой хиазмы. Это явление, называемое **положительной интерференцией**, проявляется в уменьшении числа двойных рекомбинантов в трехфакторном скрещивании. (В скрещивании $ABC \times abc$ образуется класс двойных рекомбинантов AbC и aBC , при этом в одной и той же хромосоме происходит сразу два перекреста: один — между A и B и другой — между B и C .) Меньшее число двойных рекомбинантов соответствует меньшей частоте двойных перекрестов по сравнению с произведением одиночных перекрестов (которые должны преобладать, если оба перекреста происходят независимо). Положительная интерференция проявляется на таких протяженных участках генетической карты, за пределами которых уже нельзя пользоваться величиной прямого сцепления, так как она приближается к максимальному значению — 50% рекомбинации. При образовании перекрестов между соседними маркерами в один перекрест вовлечена одна пара из четырех хромосом бивалента, а в другой перекрест — вторая, причем оба перекреста происходят независимо.

мальное число рекомбинантов не может превышать 50%. (Хотя и существует высокая вероятность рекомбинации между далеко отстоящими генами одной хромосомы, каждое рекомбинационное событие затрагивает только две из четырех конъюгировавших хромосом. Поэтому число кроссинговеров не может превысить 50%.) Итак, факторы, отстоящие далеко друг от друга, могут не проявлять сцепления при скрещиваниях, хотя и находятся в одной хромосоме. Однако их сцепленность можно обнаружить благодаря промежуточным факторам. В результате генетическая карта может быть растянута за пределы 50 ед. карты, которые получились бы при прямом измерении расстояния между двумя крайними генами. На практике из-за искажения результатов, получаемых при прямом определении сцепления двух удаленных друг от друга маркеров, для построения карт обычно используют близко расположенные факторы, а затем производят дополнительную коррекцию непосредственно полученной величины.

Группа сцепления включает в себя все гены, которые прямо или опосредованно сцеплены между собой. Гены, расположенные рядом, сцеплены между собой прямо; гены же, отстоящие друг от друга более чем на 50 ед. карты, распределяются практически независимо. Поскольку сцепление действует на протяженном участке, каждый новый идентифицированный ген в организме можно поместить рядом с группой генов, уже картированных ранее. Таким образом, гены распадаются на дискретное число групп сцепления. Гены одной из групп сцепления всегда независимо комбинируются с генами других групп.

Число групп сцепления совпадает с числом хромосом. Относительная длина групп сцепления аналогична относительным размерам хромосом. На рис. 1.11 в качестве примера рассматриваются хромосомы дрозофилы, у которых оказалось легко измерить длину. Сформулированная Менделем концепция гена как дискретного элементарного фактора наследственности может быть расширена в концепцию, рассматривающую хромосому как протяженную единицу, состоящую из многих генов. Физическое положение генов лежит в основе их генетического поведения.

Генетические карты непрерывны

На тех генетических картах высших организмов, которые были установлены в первой половине этого века, гены были расположены как бусы на нитке. Они были закреплены в определенном порядке, и генетическую рекомбинацию представляли как обмен соответствующими участками нитки между гомологичными хромосомами. При этом ген со всеми своими свойствами выступает как таинственный объект (бусинка), чье взаимоотношение с ее окружением (ниткой) совершенно не ясно.

Из-за небольшого числа потомков, получаемых в результате скрещивания, при построении рекомбинационных карт достигается лишь низкое разрешение. Рекомбинация между близко расположенными точками происходит с настолько низкой частотой, что ее редко можно наблюдать между мутациями, затрагивающими один и тот же признак и расположенными, возможно,

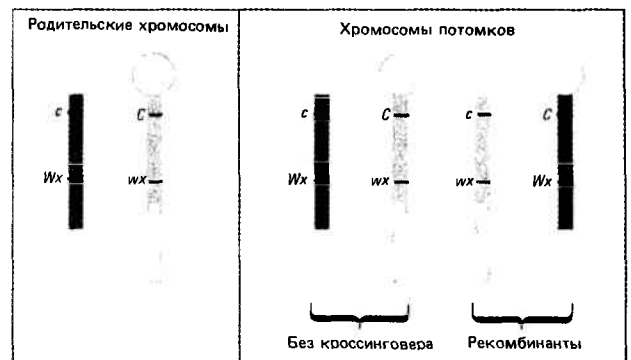


Рис. 1.10. Генетическая рекомбинация происходит в результате кроссинговера, т.е. физического перекрестного обмена.

У одного из мутантов кукурузы хромосома, несущая аллели C и wx , получила «вздутие» на одном конце и протяженную транслокацию участка другой хромосомы — на другом конце. В гетерозиготе с нормальной хромосомой, несущей аллели c и Wx , рекомбинация между генетическими маркерами всегда связана с образованием нового типа хромосом. Именно это должно происходить при физическом обмене участками хромосом.

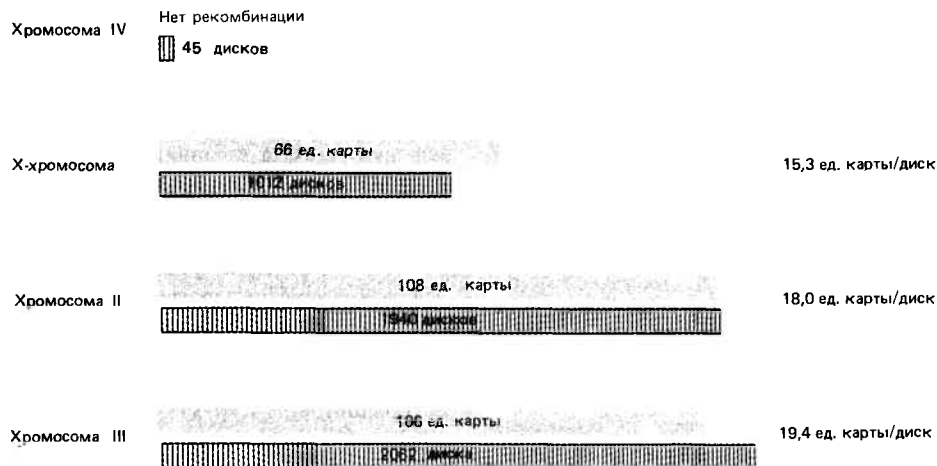


Рис. 1.11. Карта сцепления отражает физический размер хромосомы.

Каждая хромосома *Drosophila melanogaster* видна в виде ряда дисков (это

особое состояние хромосомы обсуждается в гл. 28). Относительное число дисков пропорционально физической длине каждой хромосомы, так же как относительное число единиц карты.

в одном гене. И тут возникает принципиальный вопрос: имеют ли гены и межгенные промежутки одинаковую структуру? В более конкретном, практическом смысле этот вопрос можно поставить так: происходит ли рекомбинация внутри гена и можно ли ее использовать для определения линейного порядка мутаций в одном гене?

Чтобы ответить на этот вопрос, пришлось перейти к микроорганизмам, у которых в каждом скрещивании можно получить большое число потомков. Бактериофаг T4—это вирус, инфицирующий и затем убивающий бактерию *Escherichia coli*. Заражение одной бактерии ведет к образованию примерно 100 потомков фага менее чем за 30 мин. Фаги высвобождаются из клетки при лизисе бактерий (выход фага).

Строение индивидуального локуса было исследовано в серии опытов с генами *rII* фага T4, участвующими в лизисе бактериальной клетки. Мутации в этих генах вызывают **быстрый лизис**. Были подобраны такие условия опыта, когда при совместном инфицировании бактерий двумя *rII*-мутантами фага потомство образуется *лишь* при условии, что в результате рекомбинации мутантных фагов появится фаг дикого типа. Частота рекомбинации при этом также зависела от расстояния между мутациями, как и в случае эукариотической хромосомы.

Высокая разрешающая способность этого подхода, благодаря селективной методике, позволяет количественно оценить даже самые редкие рекомбинационные события, а это дает возможность измерять расстояния по карте для любой пары мутаций. (Единственным ограничением этого метода является частота обратных мутаций, или «реверсий», к дикому типу. Эта частота составляет примерно 0,0001%. Таким образом, чтобы достичь минимального уровня измерений, превышающего в 10 раз фон спонтанных мутаций, можно измерять только рекомбинационные частоты выше 0,001%.)

Было проанализировано примерно 2400 мутаций, что дало возможность идентифицировать 304 различных мутантных сайта. (Если две мутации не рекомбинировали между собой, их принимали за две спонтанные мутации *одного и того же* генетического сайта.) Исследованные мутации можно было расположить в линейном порядке. Из этого следовало, что сам ген имеет тот же линейный

принцип устройства, что и гены в хромосоме. Из этих данных вытекает, что генетическая карта линейна как внутри генов, так и между ними, т.е. она представляет собой непрерывную последовательность, включающую в себя все гены.

Один ген—один белок

До 1945 г. было лишь известно, что гены—это основные единицы наследственности. Но каким путем они выполняют свою функцию, оставалось неясным. Гены можно было идентифицировать, только исходя из мутаций, вызывающих некоторые отклонения от нормы в фенотипе, причем степень отклонения варьировала от изменения одного признака (такого, как цвет глаз) до крайне резких морфологических перестроек, затрагивающих ряд тканей организма. Почти во всех случаях эффект мутации анализировали только на описательном уровне. Однако предполагали, что существует некоторая общая основа мутационных воздействий.

Идея о том, что действие мутационных факторов опосредовано ферментами, возникла почти одновременно с возрождением генетики. В работах, проведенных с 1902 по 1908 г., Геррод (Garrod) высказал мнение, что болезнь человека—алкаптонурия—обусловлена нарушением какой-то метаболической реакции, катализируемой ферментом. Его фраза—«врожденные ошибки метаболизма»—включает в себе концепцию, согласно которой генетический дефект может привести к нарушению определенного метаболического процесса, обуславливая тем самым наблюдаемый фенотип. В последующие три десятилетия накопились примеры влияния специфических мутаций на определенные биохимические реакции. Основная трудность исследований этого периода состояла в том, что приходилось довольствоваться случайно отобранными мутациями, не всегда пригодными для биохимического изучения.

Систематические попытки связать гены с ферментами предприняли Бидл и Эфрусси (Beadle, Ephrussi) в тридцатых годах, когда они установили, что нормальная красная окраска глаз у дрозофилы развивается в результате серии дискретных стадий и что блокирование этого про-

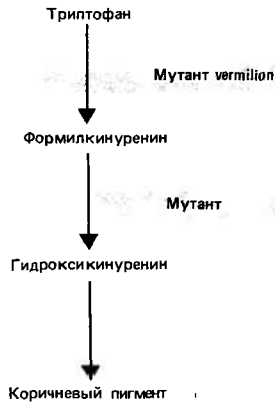


Рис. 1.12. Гены контролируют метаболические реакции. Действие двух мутаций, затрагивающий окраску глаз у *D. melanogaster*, нарушает различные стадии превращения триптофана в коричневый пигмент. Каждая мутация инактивирует определенный фермент, выполняющий соответствующую реакцию. При этом происходит накопление промежуточного продукта, который влияет на окраску глаз.

цесса на разных стадиях приводит к образованию различных мутантных окрасок. Но еще довольно долго не удавалось выяснить весь путь биосинтеза пигмента. Бидл и Татум (Beadle, Tatum) попытались подойти к проблеме с другой стороны. Бидл вспоминает: «Неожиданно мне пришло в голову, что можно перевернуть ход исследований и, вместо того чтобы пытаться разобраться в химии известных генетических повреждений, нам следует отобрать мутантов, у которых блокированы известные химические реакции. *Neurospora* была подходящим объектом для такого подхода, так как ее можно было вырастить в культуральной среде известного химического состава».

Таким образом, с помощью рентгеновского облучения были индуцированы мутанты *Neurospora*, не способные расти на минимальной среде. Биохимическую природу дефекта можно было установить. Для этого было достаточно определить, добавление какого именно вещества в среду позволяет расти мутантному штамму. Было установлено, что у каждого мутанта блокирована определенная метаболическая стадия и каждой такой стадии у штамма дикого типа соответствовал один определенный фермент. Если какой-то субстрат утилизируется в результате целой серии метаболических реакций, мутация, нарушающая любую из этих реакций, может блокировать всю серию реакций. Об этом свидетельствует накопление метаболического предшественника того соединения, которое образуется на заблокированной ступени синтеза. На рис. 1.12 изображен пример пути биосинтеза глазного пигмента дрозофилы, последовательные этапы которого хорошо согласуются с рассмотренной выше схемой.

К 1945 г. результаты этого анализа легли в основу гипотезы: «Один ген — один фермент». Согласно этой гипотезе, каждую метаболическую ступень катализирует отдельный фермент, за образование которого отвечает один ген. Мутация в гене может привести к потере активности соответствующего белка. Поскольку мутация — событие случайное и, следовательно, может повредить любой участок гена, то наиболее вероятно, что она нарушит функцию гена. Поэтому в результате большинства мутаций образуются нефункциональные гены. Однако это отнюдь не единственное последствие мутации: в результате мутации иногда происходит не потеря функции, а ее изменение.

С этих позиций легко объяснить природу рецессивных мутаций: у них нарушена функция гена, потому что мутация препятствует образованию нужного фермента. Однако, как показано на рис. 1.13, в гетерозиготе, содержащей один ген дикого типа и один мутантный ген, ген дикого типа может обеспечить необходимую функцию. Поэтому ген дикого типа доминантен. (Иными словами, для образования необходимого количества белка достаточно наличия одного аллеля дикого типа. Если же это не так, т.е. когда один аллель обеспечивает меньшее количество белка, чем то, которое имеет место при наличии двух аллелей, наблюдается промежуточный фенотип. В таких случаях ген дикого типа только частично доминантен или кодминантен.)

Прямого доказательства того, что ген действительно детерминирует структуру белка, пришлось ждать до 1957 г., когда Инграм (Ingram) показал, что серповидноклеточная анемия, наследуемая как моногенный признак, обусловлена изменением аминокислотного состава белка гемоглобина.

Некоторые белки состоят более чем из одной субъединицы. Их называют **мультимерными**. Если субъединицы белка одинаковы, то белок — **гомомультимер**, детерминируемый одним геном. Если же субъединицы белка различны, то белок называют **гетеромультимером**. Гемоглобин служит примером белка, состоящего более чем из одного типа полипептидных субъединиц. Группа гема связана с двумя α -субъединицами и двумя β -субъединицами. Каждый тип субъединиц — иная полипептидная цепь и продукт другого гена. Таким образом, функция гемоглобина может быть подавлена мутацией в любом из генов, кодирующих либо α -, либо β -субъединицу. В связи с этими данными гипотеза «один ген — один фермент» была сформулирована в более общем виде, применительно к любому гетеромультимерному белку. Теперь она получила более точное выражение: «**один ген — одна полипептидная цепь**».

Новое определение: цистрон

Если рецессивная мутация — это любое изменение в гене, препятствующее образованию активного белка, то в каждом гене может произойти множество таких мутаций, так как многие аминокислотные замены могут суще-

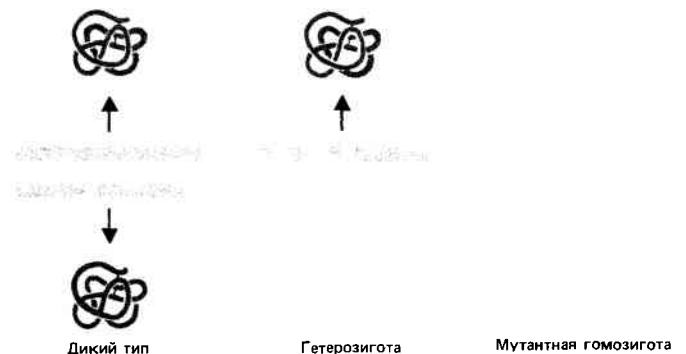


Рис. 1.13. Гены кодируют белки.

У особи дикого типа оба гена активны (окрашенные полоски) и продуцируют белок. В гетерозиготе доминантный аллель активен и тоже продуцирует белок, но рецессивный аллель (светлая полоска) не дает активного белка. В гомозиготе, несущей два рецессивных аллеля, совсем нет соответствующего белка и данная функция полностью отсутствует.

ственно изменить структуру белка и нарушить его функцию. Если серия мутаций нарушает один и тот же признак, и при этом располагается рядом на генетической карте, эти мутации можно считать множественными аллелями одного гена. В этом случае можно получить гетерозиготу по двум мутантным аллелям.

Но каким образом отличить множественные аллели от мутаций в генах, тесно сцепленных и затрагивающих один и тот же фенотипический признак? Основной тест, показывающий принадлежность мутаций к разным генам, — это тест на комплементацию в гетерозиготе. На практике это сводится к скрещиванию, в котором оба родителя гомозиготны по разным мутациям.

Если мутации находятся в одном и том же гене, от каждого родителя в зиготу попадают разные мутантные аллели, так что у гетерозиготы будет генотип

$$\frac{m_1}{m_2},$$

в котором нет аллеля дикого типа.

Но если мутации лежат в разных генах, родительские генотипы можно обозначить как $(m_1 +)$ и $(+ m_2)$, где каждый имеет аллель дикого типа одного гена и мутантный аллель — другого гена. (Знак плюс используют для обозначения дикого типа.) Тогда у гетерозиготы будет генотип

$$\frac{m_1 +}{+ m_2},$$

в котором оба родителя вместе обеспечивают по «дикой» копии каждого гена. Такая гетерозигота имеет дикий фенотип, а о генах говорят, что они комплементируют друг друга.

Рассмотрим более детально комплементационный тест с помощью рис. 1.14, на котором представлены исследуемые генотипы в двух возможных конфигурациях. При *цис*-конфигурации обе мутации находятся в *одной и той же* хромосоме. При *транс*-конфигурации они расположены в *противоположных* хромосомах.

Сначала рассмотрим пример, когда мутации лежат в одном и том же гене. *Транс*-конфигурация соответствует тесту, который мы только что описали. Обе копии гена в этом случае мутантные. При *цис*-конфигурации, однако, один геном производит дважды мутантный белок, а другой — белок дикого типа. Таким образом, если исследуемые мутации лежат в одном гене, фенотип гетерозиготы определяется их взаимным расположением: при *транс*-конфигурации фенотип — мутантный, а при *цис*-конфигурации — дикого типа. И наоборот, если мутации лежат в разных генах, их взаимное расположение значения не имеет. Для каждого гена в обеих конфигурациях есть по одной мутантной копии и по копии дикого типа.

Из этого следует, что при *цис*-конфигурации результат не зависит от того, находятся ли мутации в одном или в разных генах. *Цис*-вариант используют в основном как формальный контроль для выявления функции дикого типа, по отношению к которому судят о наличии или отсутствии комплементации при *транс*-конфигурации).

Процедура проверки комплементации двух мутаций состоит в том, чтобы выяснить, будет ли у *транс*-гетерозиготы дикий фенотип (мутации в разных генах) или же мутантный (мутации в одном гене). Чтобы выполнить этот эксперимент с эукариотами, нужно сконструировать соответствующую двойную гетерозиготу; в случае виру-

сов достаточно одновременно инфицировать клетку-хозяина мутантами двух типов. Если мутантные гены комплементируют друг друга при *транс*-положении, то они, очевидно, детерминируют различные белки и, следовательно, дополняют недостающие друг у друга функции, обеспечивая появление дикого фенотипа. Продукты, дополняющие друг друга при *транс*-конфигурации, называются *транс*-активными. Они представляют собой, очевидно, диффундирующие белки, способные действовать совместно, независимо от того, происходят ли они из одного генома или из разных.

Если две мутации не комплементируют при *транс*-конфигурации, то делается вывод, что обе они нарушают одну и ту же функцию. Такие мутации относят к одной группе комплементации. При анализе *rII*-области фага T4 все мутации распределились на две группы комплементации; в каждой из этих групп мутантные сайты расположились в линейном порядке. Внутри группы *rIIA* и внутри группы *rIIB* никакие две мутации не комплементировали друг друга, но любая мутация *rIIA* была комплементарна любой мутации *rIIB*. Эти данные говорят о том, что *rII*-мутации соответствуют двум соседним группам комплементации, затрагивающим одну и ту же фенотипическую функцию.

Описанный выше метод носит название *цис/транс*-тест или тест на комплементацию [функциональный тест на аллелизм.—Пер.]. С его помощью определяют группу комплементации как протяженную единицу, все части которой должны быть дикого типа и находиться в одной хромосоме. Единичная мутация в любой части может нарушить функцию. Эта генетическая единица получила название цистрон. Согласно этой терминологии, две мутации в *одном и том же* цистроне не могут комплементировать; факт комплементации двух мутаций указывает на их принадлежность к двум разным цистронам.

Причиной появления нового термина послужило стремление ввести определения для генетических единиц: единица мутации, единица рекомбинации, единица функции. Исследование *rII*-системы четко показало, что единицы мутации и рекомбинации меньше единицы функции, детерминированной данным цистроном. Этот термин теперь обычно считают равнозначным старому термину «ген», и в настоящее время он практически вышел из употребления.

Однако иногда наблюдается исключение из правила, что комплементировать могут только различные гены, — это в том случае, когда ген кодирует полипептид, представляющий собой субъединицу гомомультимерного белка. В клетке дикого типа активный белок состоит из нескольких идентичных субъединиц. В клетке, содержащей два различных мутантных аллеля, их продукты могут смешиваться, образуя мультимерные белки из субъединиц обоих типов. Иногда происходит взаимная компенсация мутаций, и в таком случае белок со смешанными субъединицами может быть активен, тогда как белки, содержащие только по одному типу мутантных субъединиц, неактивны. Такое явление называют межаллельной комплементацией.

Подводя итог, можно сказать, что описанный генетический анализ привел к представлению о гене как о серии последовательных сайтов в геноме. Предполагалось, что эта серия сайтов участвует в образовании единичного *транс*-активного продукта, т. е. полипептидной цепи. Но вскоре мы увидим, что это определение стало уже недостаточно точным.

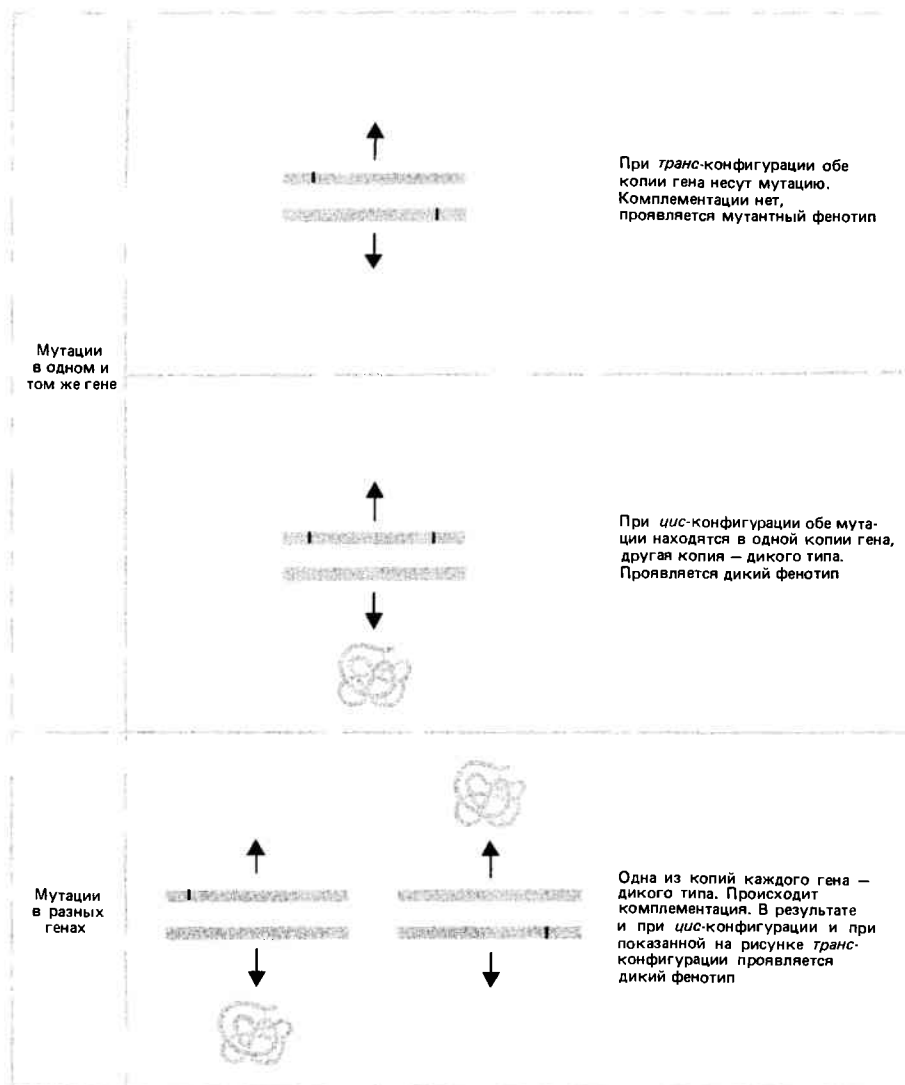


Рис. 1.14. Функциональная единица цистрон определяется с помощью *цис/транс*-теста.

Гены дикого типа изображены серыми сплошными полосками; мутации показаны цветными вертикальными линиями. Белки дикого типа также показаны серым цветом, мутантные белки показаны коричневым цветом. При *транс*-конфигурации обе копии гена несут мутацию. Комплементации нет, проявляется мутантный фенотип.

При *цис*-конфигурации обе мутации находятся в одной копии гена, другая копия – дикого типа. Проявляется дикий фенотип. Одна из копий каждого гена – дикого типа. Происходит комплементация. В результате и при *цис*-конфигурации и при показанной на рисунке *транс*-конфигурации проявляется мутантный фенотип.

Замечания, касающиеся терминологии

Прежде чем мы перейдем к молекулярному анализу гена, рассмотрим терминологию, используемую для описания генетических локусов и детерминированных ими признаков.

Под термином **генетический маркер** часто подразумевается тот или иной ген, используемый для картирования или для идентификации конкретного локуса. Можно сказать, например, что клетка несет определенный набор маркеров, т.е. аллелей.

Буквенные символы генотипа *всегда набирают курсивом*; фенотип обозначают с помощью тех же символов, но набранных прямым шрифтом. Под термином «дикий тип» подразумевают нормальную активную форму гена, или нормальный фенотип. Для обозначения «дикого типа» можно использовать также знак плюс над названием локуса. Иногда для обозначения аллеля дикого типа ис-

пользуют знак «+» сам по себе (без символа соответствующего локуса), а локус указывают только для мутантного аллеля. Полностью дефектную форму гена (или отсутствие конкретного фенотипического признака) иногда обозначают с помощью знака минус в верхнем индексе. При описании множества мутантных аллелей, чтобы различать их друг от друга, вводят дополнительные индексы. Например, локус *w* (белые глаза) у *D. melanogaster* можно обозначить w^+ (красный цвет глаз) в случае аллеля дикого типа, w^1 – в случае мутантного аллеля (глаза цвета слоновой кости – ivory), w^2 – в случае другого мутантного аллеля (глаза абрикосового цвета – apricot) и т.д.

Для обозначения генов пользуются сокращениями, причем для обозначения различных генетических систем используются разные обозначения. Для большинства эукариотических систем нет ограничений в используемой форме сокращения. Согласно общему правилу, сокраще-

ние, принятое для обозначения доминантного аллеля, пишется с заглавной буквы (для дрожжей символ целиком пишется заглавными буквами), обозначения рецессивных аллелей – строчными.

При описании бактерий используют стандартную трехбуквенную номенклатуру. Одно и то же сокращение из трех строчных букв служит символом для всех генов, нарушающих один и тот же признак. Каждый индивидуальный ген обозначают дополнительной заглавной буквой. Существует, например, серия генов *lac*, известных как *lacZ*, *lacY* и *lacA*. Аллель дикого типа гена *lacZ* будет обозначен как *lacZ⁺*, а мутант – как *lacZ⁻*. Те же самые сокращения используются для описания фенотипа, но в этом случае они пишутся не курсивом, а прямо и начинаются с заглавной буквы. Так, дикий и мутантный *lac*-аллели определяют соответствующие фенотипы *Lac⁺* и *Lac⁻*. Продукт гена – это фенотипический признак данно-

го гена, и изображается он соответственно; например, белок *LacZ* – это фермент β-галактозидаза.

Рекомендуемая литература

Современная точка зрения о соотношении понятий ген/цистрон стала формироваться начиная со статьи Бензера (*Bensler S.*, Proc. Natl. Acad. Sci. USA, **41**, 344, 1955). Эта работа часто обсуждается в различных обзорах. Достаточно полно она освещена в книге *Hayes W.*, Genetics of bacteria and their viruses, Wiley, New York, 1968. [Имеется перевод: Хэйс У., Генетика бактерий и бактериофагов.–М.: Мир, 1965.] Данные о бактериофагах рассматриваются в книге, вышедшей под редакцией Кэрнса, Стента, Уотсона (*Cairns, Stent, Watson*, Phage and the origins of Molecular Biology, Cold Spring Harbor Laboratory, New York, 1966).

Глава 2

ЧТО ТАКОЕ ГЕН? БИОХИМИЧЕСКАЯ ТОЧКА ЗРЕНИЯ

Методами формальной генетики было установлено, что ген – это дискретный фактор наследственности, часть хромосомы, и что он переходит от родителя к потомку. Решающую роль в выяснении природы гена сыграло представление о том, что ген отличается от признака, который он определяет. Различное влияние мутаций на фенотип объясняют, исходя из того, что каждая мутация препятствует синтезу определенного белка. Физическая же природа генетического материала, ответственного за образование данного белка, служила предметом многочисленных споров. Когда же наконец физическая природа была установлена, она показалась неправдоподобной.

Значительным толчком к изучению природы гена послужил обзор Шрёдингера (*Schrödinger*), опубликованный в 1945 г.; в этом обзоре рассматриваются свойства генетического материала с точки зрения физика: «Невероятно маленькая группа атомов, слишком маленькая, чтобы к ней можно было применить законы статистики, играет доминирующую роль в очень упорядоченных и регламентированных событиях, происходящих в живом организме... Ген слишком мал... чтобы передаваемая им способность к упорядоченному и регламентированному поведению происходила на основании законов физики». Далее Шрёдингер развил ранее высказанные соображения Дельбрюка (*Delbrück*) о том, что, исходя из одних только законов физики, вряд ли можно объяснить свойства генетического материала, особенно его стабильность в ряду бесчисленных поколений. Никто не сомневался, что ген подчиняется тем законам физики, которые уже известны, но при этом думали, что изучение его свойств может привести к открытию новых законов физики. Именно эта перспектива и привела многих физиков в биологию.

Шрёдингер описал свойства генов таким образом, что с очень небольшими изменениями обрисованная им картина вполне соответствует современным представлениям: «Мы можем представить структуру гена в виде гигант-

ской молекулы, способной только к локальному изменению, которое сводится к перестройке атомов и приводит к образованию изомерной молекулы. Перестройка [мутация] может повреждать только небольшую часть гена, но при этом возможно множество различных перестроек». И хотя Шрёдингер представлял себе ген как аperiодический кристалл, поскольку тогда казалось, что это единственная физическая структура, способная удовлетворять требованиям, предъявляемым к генетическому материалу, ранее представление о гене как о гигантской молекуле объясняло, каким образом ген может быть протяженной частью хромосомы.

Теперь, конечно, мы знаем, что гены состоят из ДНК и что их структура сохраняется и воспроизводится не изолированно от остальной клетки, а зависит от ферментов, влияющих на точность воспроизведения и выполняющих необходимые каталитические функции. Все эти процессы подчиняются уже известным законам физики и химии, и не было никакой необходимости искать какие-то новые законы.

Генетический материал – это ДНК

Несмотря на то что ранние генетические исследования были сконцентрированы на высших организмах, первый случай переноса генетической информации обнаружили у бактерий. Позднее тот же самый подход использовали и применительно к высшим организмам. Оказалось, что у двух основных групп живых организмов – прокариот и эукариот – генетический материал находится в одной и той же форме.

Прокариоты (бактерии) характеризуются отсутствием ядра (или каких-либо других компартментов, таких, как митохондрии, хлоропласты и т. д.). Бактериальная клетка

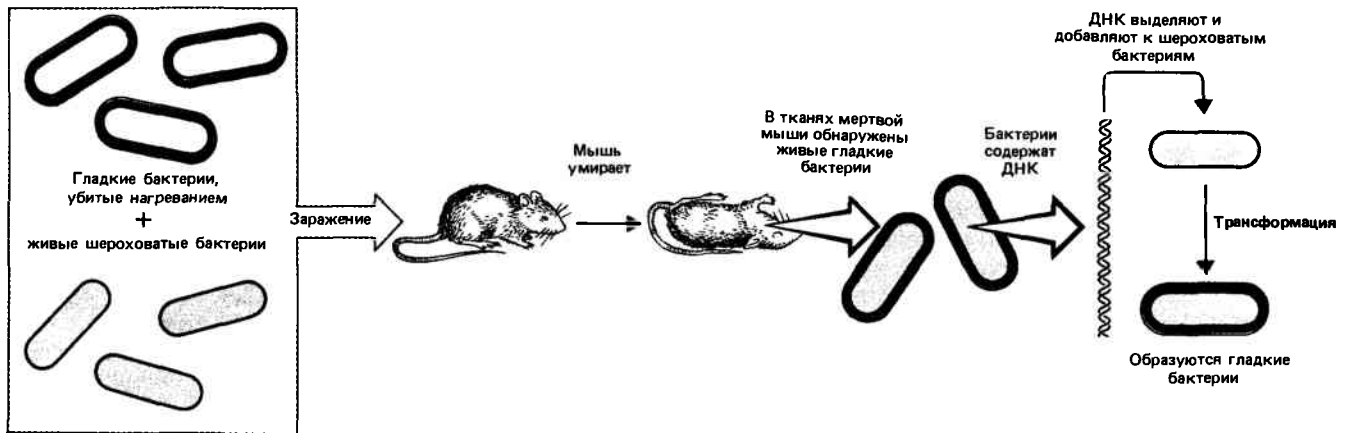


Рис. 2.1. Трансформирующий фактор—это ДНК.

Ни инактивированные гладкие бактерии, ни живые шероховатые (мутанты) не патогенны для мыши. Однако, если мышью заразить их смесью, она погибнет и из нее могут быть выделены живые гладкие бактерии.

Трансформацию убитых шероховатых бактерий в вирулентные, гладкие бактерии можно произвести *in vitro*, если добавить ДНК, выделенную из гладких бактерий.

представляет собой единый компартмент, в котором нельзя различить хромосомы, хотя можно увидеть нуклеотид, содержащий генетический материал. Эукариотические клетки имеют ядерную мембрану, которая ограничивает от цитоплазмы генетический материал. В процессе деления клетки мембрана исчезает и появляются хромосомы. Кроме ядра в эукариотической клетке есть и другие дискретные компартменты.

К эукариотам относят и так называемых микробных эукариот, представляющих собой одноклеточные организмы, а также многоклеточных растений и животных. Считалось, что основа наследственности у всех эукариот должна быть одинаковой; что же касалось бактерий, то тут возникали некоторые сомнения.

Явление **трансформации** впервые обнаружил в 1928 г. Гриффит (Griffith) при изучении пневмококковой инфекции мышей. Вирулентность (инфекционность) этой бактерии определяется капсульным полисахаридом, расположенным на поверхности клеточной стенки. Вирулентные бактерии образуют гладкие колонии, обозначаемые буквой S (от англ. smooth—«гладкий»). Несколько «типов» пневмококков несут различные капсульные полисахариды. Любой из этих типов может дать мутантов, не способных к образованию капсульных полисахаридов. Такие бактерии растут в виде шероховатых колоний, обозначаемых буквой R (от англ. rough—шероховатый). Эти пневмококки авирулентны, т. е. не убивают мыши при инфицировании (поскольку из-за отсутствия капсульного полисахарида они фагоцитируются в организме животного). Но если мышью инфицируют бактериями R вместе с инактивированными нагреванием бактериями S (которые сами по себе не патогенны для мыши), животное погибает от пневмококковой инфекции. Погибшие животные содержат вирулентные бактерии S (рис. 2.1).

Мертвые бактерии S принадлежали к типу I. Живые бактерии R были производными бактерий типа II. Вирулентные бактерии, обнаруженные после смешанной инфекции, были типа IS. Это означало, что какой-то компонент мертвых бактерий S типа I может **трансформировать** живых бактерий R так, что они начинают синтезировать капсульный полисахарид типа I.

Как только этот результат удалось воспроизвести в бесклеточной системе, где вместо живых бактерий S были использованы их экстракты, стало возможным

очистить и определить активный компонент, названный **трансформирующим фактором**. В 1944 г. Эйвери (Avery) и его коллеги в своих классических исследованиях показали, что трансформирующий фактор—это **дезоксирибонуклеиновая кислота (ДНК)**. Новое доказательство трансформирующей активности ДНК было получено несколько позже, когда удалось очистить фермент дезоксирибонуклеазу, разрушающую ДНК. Было показано, что добавление этого фермента необратимо инактивирует трансформирующий фактор. Возможная роль примесей белка в препаратах ДНК была исключена, поскольку добавление фермента трипсина (разрушающего все белки) не влияло на трансформацию.

О том, насколько неожиданным было это открытие, можно судить хотя бы по тому, что тогда еще не знали даже, что ДНК входит в состав пневмококков, хотя уже в течение нескольких десятилетий было известно, что ДНК является основным компонентом эукариотической хромосомы. Описывая методику выделения ДНК, Эйвери отмечает: «Когда концентрация спирта достигает примерно 9/10 объема, отделяется волокнистое вещество, которое при помешивании наматывается на стеклянную палочку, как нитка на катушку, тогда как другие примеси остаются в растворе в виде гранулированного осадка. Волокнистое вещество затем снова растворяют и процедуру повторяют несколько раз. Оказалось, что это вещество обладает высокой реакционной способностью и при элементарном анализе совпадает по свойствам с очищенной ДНК. (Кто бы мог догадаться!) Насколько мне известно, этот тип нуклеиновой кислоты пока не был обнаружен у пневмококков, хотя она была найдена у других бактерий».

Значение полученного результата точно сформулировано в оригинальной статье, посвященной этому исследованию. «Индукующее вещество по своим химическим и физическим свойствам является высокополимерной и вязкой формой ДНК. С другой стороны, капсульное вещество типа III, синтез которого вызывается этим трансформирующим агентом, состоит главным образом из полисахарида, не содержащего азота... Таким образом, ясно, что индуцирующее вещество и вещество образующееся различны по химическим свойствам и специфичны по биологическому действию. Оба вещества необходимы для определения специфичности клетки, частью которой

они являются». Различие между генетическим материалом и продуктами его экспрессии было установлено в результате последующих исследований.

Нежелание признать за ДНК роль генетического материала привело к слишком натянутым доводам против очевидного вывода, сделанного Эйвери и потом подтвержденного дальнейшими исследованиями. Что касается самой трансформации, общий характер вывода подтверждался тем, что с помощью ДНК удалось передать способность к образованию и других типов капсульных полисахаридов, характерных для пневмококков. Кроме того, была обнаружена трансформация и у других бактерий, причем по многим отличающимся признакам. Однако эти данные истолковывали как довод в пользу участия ДНК в образовании полисахаридов у бактерий, а не как доказательство ее генетической роли вообще. И на самом деле, как мы теперь знаем, гены, переносимые в первых опытах по трансформации, отвечали за образование фермента, катализирующего одну из стадий синтеза капсулярного полисахарида (UDPG-дегидрогеназы). Следующий признак, переданный с помощью трансформации, был совершенно отличен от первого — это была устойчивость к пенициллину.

Дальнейшие доказательства роли ДНК

После того как трансформация у бактерий была доказана, необходимо было перейти к следующему шагу — доказать генетическую роль ДНК в совершенно иной системе. Фаг T2 — это вирус, инфицирующий бактерию *E. coli*. Когда фаги добавляют к бактериальной культуре, они адсорбируются на наружной поверхности, вводят определенное вещество в бактерию и затем, примерно через 20 мин, бактерия разрывается (лизует), высвобождая большое число фаговых потомков.

В 1952 г. Херши и Чейз (Hershey, Chase) инфицировали бактерий *E. coli* фагом T2, у которого радиоактивным изотопом метили либо ДНК-компоненты (^{32}P), либо белковые компоненты (^{35}S). Фаги смешивали с бактериями и неадсорбированные частицы удаляли методом центрифугирования. Затем инфицированные бактерии подвергали сильному встряхиванию с помощью блендера и разделяли полученный препарат путем центрифугирования на две фракции. Одна фракция содержала пустые фаговые оболочки, отделившиеся от клеточной стенки бактерий, другая — сами бактерии.

На рис. 2.2 показано, что метка ^{35}S была связана с оболочками фага. Большая же часть метки ^{32}P оказалась внутри инфицированных бактерий. В потомстве фага, образованном после инфицирования, было найдено примерно 30% исходной метки ^{32}P . А от исходного белка, меченного ^{35}S , в фаговом потомстве было обнаружено всего лишь менее 1%. Этот эксперимент прямо показывает, что родительская фаговая ДНК проникает в бактерию и затем становится частью фагового потомства. Именно так должно происходить наследование генетического материала.

Вопрос о том, можно ли вирусы отнести к «живым» организмам, довольно долго служил предметом споров. Классический пример с фагом T2 (напомним, что фаги — это бактериальные вирусы), показывает, что, инфицировав бактерию, вирус способен использовать «аппараты» клетки-хозяина для воспроизведения самого себя во многих копиях. Генетический материал фагов ведет се-

бя аналогично клеточному геному: его признаки точно воспроизводятся и подчиняются тем же правилам, что делает возможным линейное картирование мутаций. Таким образом, на примере фага T2 еще раз подтвержден общий вывод о том, что генетическим материалом служит ДНК.

Необходимо обратить внимание на следующее обстоятельство. Хотя эти эксперименты и доказательны в отношении генетической роли ДНК, в них не так полно, как в опытах по трансформации, исключается возможность загрязнения использованного препарата ДНК белковыми примесями. В опытах по трансформации, улучшая очистку ДНК, можно свести содержание примеси до ничтожно малой величины. При инфицировании фагом в бактериальную клетку вместе с ДНК всегда вносится небольшое количество белка, хотя в потомстве обнаруживается только меченая ДНК. Однако со времени работ Эйвери точка зрения биологов уже изменилась и результаты Херши и Чейз были сразу восприняты как доказательство генетической роли ДНК.

Стало несомненным, что у бактерий и бактериофагов генетическим материалом служит ДНК. Но как обстоит дело у высших организмов? В течение долгого времени все доказательства носили чисто умозрительный характер. ДНК присутствует в хромосомах; ее количество постоянно во всех соматических клетках, и только половина этого количества обнаруживается в половых клетках. Все это точно соответствует ожидаемому поведению генетического материала, но не может служить доказательством.

В обсуждении своих результатов Эйвери сделал замечание более широкого плана, касающееся не только его непосредственных результатов. «Если мы правы», — писал он, — «это означает, что нуклеиновые кислоты являются не только структурно важными, но и функционально активными веществами, определяющими биохимическое поведение клеток. Поэтому, используя известные химические соединения, можно индуцировать предсказуемые наследственные изменения в клетках». Он имел в виду систему бактериальной трансформации, но сходные результаты теперь получены и в случае клеток высших организмов.

Первые доказательства были получены, когда удалось показать, что хромосомы можно выделить из одной клетки и добавить к другим, просто смешав их с клетками: с очень низкой частотой в реципиентных клетках начинали проявляться признаки, внесенные хромосомами донора. Затем появилась возможность выполнить эти эксперименты с очищенной ДНК. А теперь, увеличивая чистоту трансформирующего препарата, можно фактически добавлять к клеткам чистые гены, т. е. ДНК, содержащую только тот ген, который был утрачен этими клетками. Затем можно получить трансформантов, содержащих и экспрессирующих данный ген. Такую процедуру можно использовать применительно к любому гену, при условии, что можно тестировать активность соответствующего фермента. Пример одной из стандартных систем изображен на рис. 2.3.

Хотя по причинам исторического характера описываемые нами опыты на эукариотических клетках были названы **трансфекцией**, они в точности соответствуют трансформации у бактерий. В этих экспериментах было не только установлено, что ДНК является у эукариот генетическим материалом, но и была доказана возможность переноса генов между разными видами с сохранением их

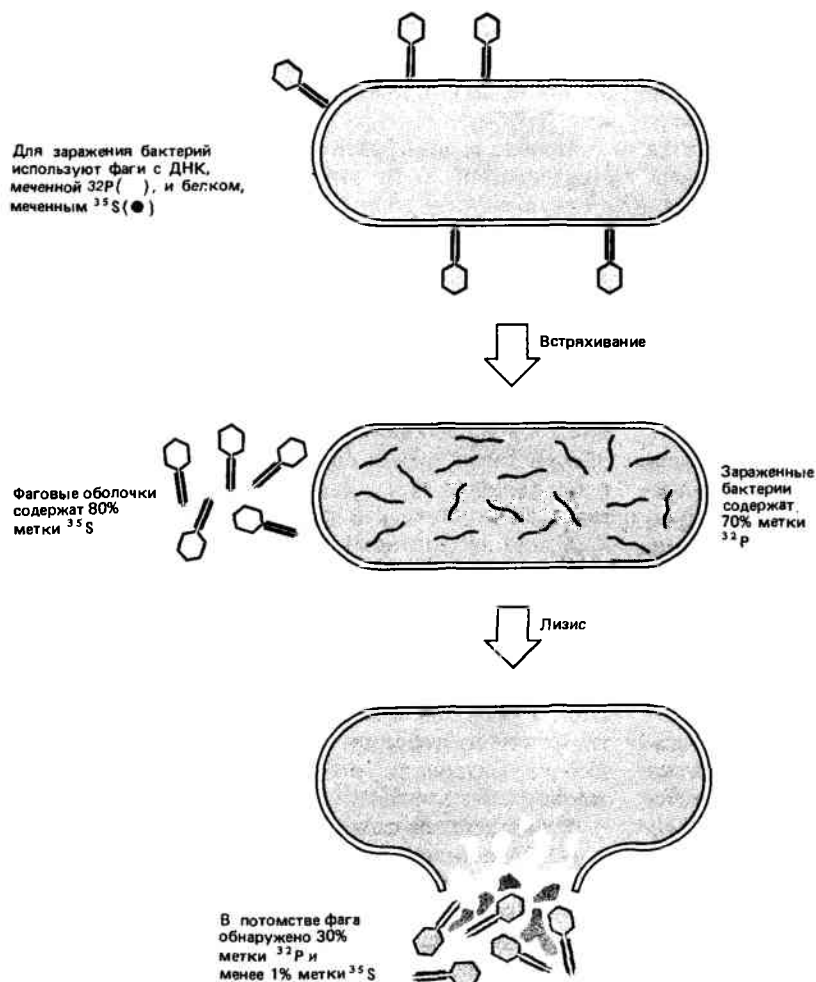


Рис. 2.2. Генетическим материалом фага T2 является ДНК.

Когда ДНК фага поместили изотопом ^{32}P , а белок – изотопом ^{35}S и такими фагами заразили бактерий, оказалось, что метка ^{32}P включилась в бактерии и обнаруживалась в потомстве фага, тогда как ^{35}S у потомков фага не обнаружили.

функциональных свойств. Сначала такие опыты можно было проводить только с индивидуальными клетками, адаптированными к росту в культуре. Однако теперь гены вводят в яйцеклетки мыши методом микроинъекции и потом их обнаруживают в виде стабильного компонента генома взрослой мыши. (Трансфекция обсуждается в гл. 38.)

Существуют некоторые вирусы, у которых в качестве генетического материала используется **рибонуклеиновая кислота (РНК)**. Хотя по своему химическому строению РНК несколько отличается от ДНК, она тем не менее выполняет у этих вирусов ту же роль. В качестве основного генетического материала клетки всегда используют ДНК; что же касается вирусов, то одни из них используют ДНК, другие же – РНК. Во всех случаях генетическим материалом являются нуклеиновые кислоты.

Компоненты ДНК

Одной из причин скептицизма по отношению к возможной генетической функции ДНК было ошибочное представление о ее структуре. Было известно, что ДНК содержит четыре типа нуклеотидов, однако полагали, что она построена из регулярно повторяющихся тетра-нуклеотидных единиц. К тридцатым годам Касперсон (Casperson) показал, что ДНК состоит из очень больших молекул, намного больших, чем молекулы белков. Однако монотонность ее структуры казалась препятствием, не

позволяющим выполнять какую-либо генетическую функцию, для которой предполагается, конечно, разнообразие форм.

Но затем в работах Чаргаффа (Chargaff) было показано, что четыре основания, найденные в ДНК, находятся в варьирующих количествах. Эти количества различны у разных организмов и характерны для каждого вида. На основе этих наблюдений была выдвинута концепция, согласно которой **генетическая информация** заложена в последовательности оснований ДНК, и эта последовательность каким-то образом определяет, или **кодирует**, последовательность аминокислот в белке. К пятидесятым годам эта концепция генетической информации получила всеобщее признание. Основная задача теперь заключалась в том, чтобы построить модель ДНК, объясняющую, каким образом в последовательности ДНК может быть записана последовательность белка.

Нуклеиновые кислоты состоят из последовательности химически связанных **нуклеотидов**. Каждый нуклеотид содержит гетероциклическое кольцо из атомов углерода и азота (**азотистое основание**), пятиуглеродное сахарное кольцо (**пентозу**) и **фосфатную** группу.

В нуклеиновых кислотах обнаружено два типа пентоз. Именно характером пентозы, входящей в состав молекулы, и различаются ДНК и РНК. От характера пентозы зависит и название двух типов нуклеиновых кислот. В ДНК пентозой является 2-дезоксирибоза, а в РНК – это

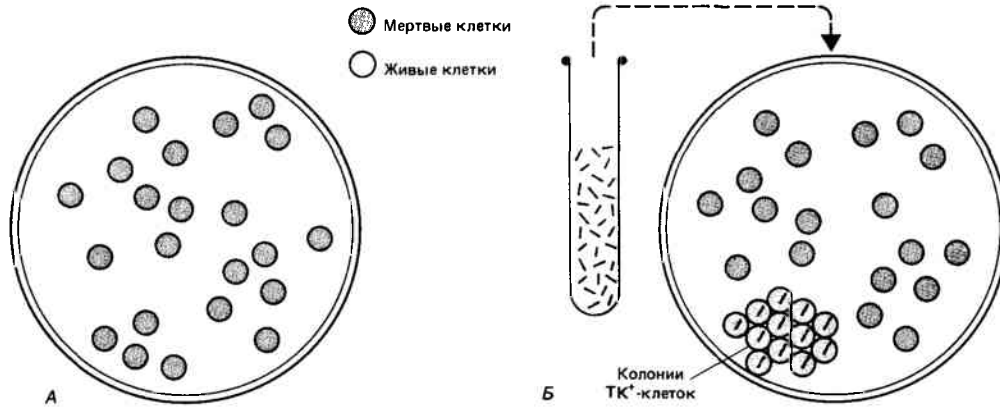


Рис. 2.3. С помощью ДНК можно произвести трансфекцию эукариотических клеток.

Метод напоминает бактериальную трансформацию. К эукариотическим клеткам, растущим в культуре, добавляют ДНК. У клеток, включивших эту ДНК, проявляются новые фенотипические признаки. А. Эукариотиче-

ские клетки, утратившие ТК-ген, не образуют фермента тимидинкиназы и погибают в отсутствие тимина. Б. При добавлении к клеткам препарата ДНК, содержащего ТК-гены, некоторые клетки получают эти гены и начинают расти, образуя в результате колонии.

рибоза. Различие заключается в отсутствии или наличии гидроксильной группы во втором положении сахарного кольца (рис. 2.4).

Как показано на рисунке, азотистые основания делятся на два типа: пиримидиновые и пуриновые основания, называемые для краткости **пиримидины** и **пурины**. Пиримидины состоят из шестичленного кольца, а у пуринов по два конденсированных кольца: одно—пятичленное и второе—шестичленное. Каждая нуклеиновая кислота синтезируется из оснований только четырех типов. Одни и те же самые пурины (аденин и гуанин) входят в состав и ДНК и РНК. Два пиримидина, входящие в состав ДНК,—это цитозин и тимин, а в РНК вместо тимина находится урацил. Тимин отличается от урацила только наличием метильной группы в пятом положении пиримидинового кольца. Основания обычно обозначают их начальными буквами. Так, ДНК содержит А, G, C, T, а РНК содержит А, G, C, U. (В некоторых случаях отдельные основания могут быть модифицированы после того, как включатся в нуклеиновую кислоту. Это явление описано для ДНК в гл. 30, а для РНК—в гл. 7.)

Азотистые основания соединены с пентозным кольцом гликозидной связью между С-1 пентозного кольца и N-3 пиримидина или N-9 пурина. Чтобы избежать путаницы в нумерации атомов азотистых оснований и сахаров, положения атомов углерода пентозы пишут со штрихом (').

Нуклеотиды соединены в **полинуклеотидную цепь**, остов которой состоит из перемежающихся остатков сахара и фосфата. Как показано на рис. 2.5, атом в 5'-положении одного пентозного кольца соединен с атомом в 3'-положении следующего пентозного кольца через фосфатную группу. Принято говорить, что сахарофосфатный остов состоит из (5'-3')-связей. Концевой нуклеотид на одном конце цепи имеет свободную 5'-группу, на другом конце—свободную 3'-группу. Последовательности нуклеиновых кислот принято писать именно в таком направлении: от 5'-конца к 3'-концу.

На рис. 2.5 показано, что азотистые основания «торчат» из сахарофосфатного остова. Заметим, что фосфатные группы заряжены отрицательно и нуждаются в нейтрализации—ионами металлов и(или) положительно заряженными белками.

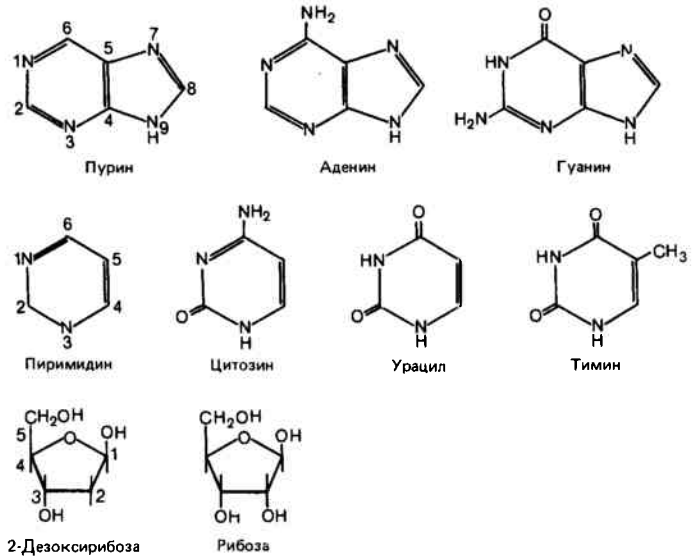


Рис. 2.4. В состав нуклеиновых кислот входят азотистые основания (пурины и пиримидины), сахар (дезоксирибоза или рибоза) и фосфатный остаток.

Показана нумерация положений атомов в молекулах оснований и сахаров. Молекулу, состоящую из азотистого основания и сахара, называют нуклеозидом, а молекулу, состоящую из основания, сахара и фосфатного остатка,—нуклеотидом. Буква «d» обозначает дезоксирибозу, а если ее нет, значит сахарный остаток представлен рибозой. Отсутствие буквы «d» соответствует наличию 2'-ОН-группы. Для обозначения нуклеиновых кислот приняты следующие обозначения:

Основание	Нуклеозид	Нуклеотиды (и принятые сокращения)
Аденин	Аденозин	Адениловая кислота = AMP (или dAMP)
Гуанин	Гуанозин	Гуаниловая кислота = GMP (или dGMP)
Цитозин	Цитидин	Цитидиловая кислота = CMP (или dCMP)
Тимин	Тимидин	Тимидиловая кислота = dTMT
Урацил	Уридин	Уридиловая кислота = UMP

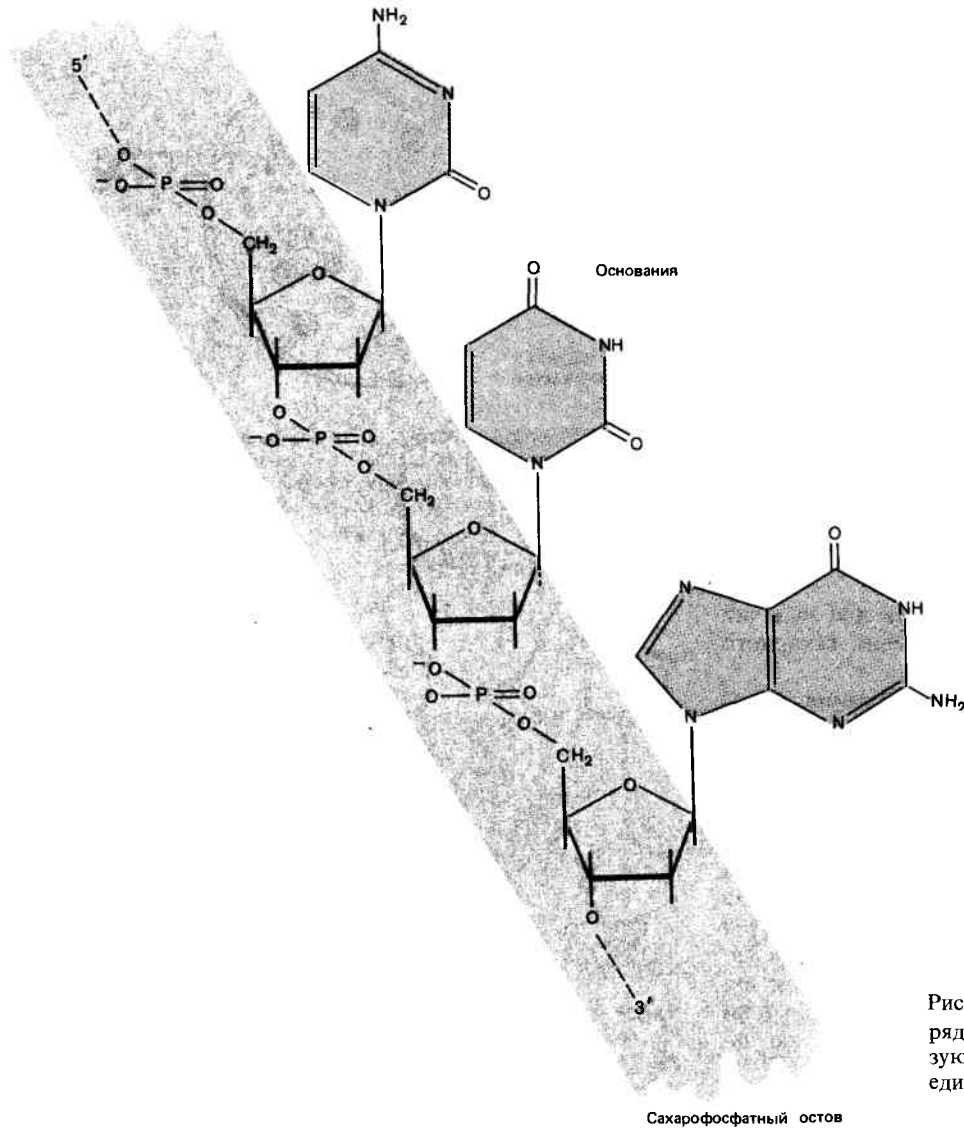


Рис. 2.5. Полинуклеотидная цепь состоит из ряда (5'-3')-сахарофосфатных связей, образующих остов молекулы, к которому присоединяются азотистые основания.

Если разорвать ДНК или РНК на отдельные нуклеотиды, разрыв может произойти с любого конца фосфодиэфирной связи. В зависимости от обстоятельств фосфатная группа может остаться присоединенной либо с 5'-, либо с 3'-конца пентозы (рис. 2.6). Фрагмент нуклеиновой кислоты, состоящий из сахара и основания, называют **нуклеозидом**. Соответственно от нуклеиновой кислоты можно отделить два типа нуклеозидов: нуклеозид-3'-монофосфат и нуклеозид-5'-монофосфат.

Все нуклеозиды могут быть связаны более чем с одной фосфатной группой в 5'-положении. Пример показан на рис. 2.6. Связи между первой (α) и второй (β), а также между второй (β) и третьей (γ) фосфатными группами высокоэнергетические (богаты энергией) и служат источником энергии для различных клеточных процессов. Нуклеозидтрифосфаты—это те соединения, из которых синтезируются нуклеиновые кислоты.

ДНК—двойная спираль

Открытие двойной спирали было одним из наиболее волнующих событий в молекулярной биологии. Две группы фактов легли в основу модели ДНК, построенной Уотсоном и Криком (Watson, Crick) в 1953 г.

Во-первых, данные рентгеноструктурного анализа показали, что ДНК имеет форму регулярной спирали, делающей полный оборот каждые 34 Å, с диаметром около 20 Å. Поскольку расстояние между соседними нуклеотидами составляло 3,4 Å, то в одном витке должно быть 10 нуклеотидов.

Плотность ДНК свидетельствовала о том, что спираль должна состоять из двух полинуклеотидных цепей. Постоянный диаметр спирали можно было бы объяснить тем, что в каждой цепи основания направлены внутрь спирали, причем расположены так, что пурин всегда находится против пиримидина, обеспечивая ситуацию, при которой не может возникнуть контакта пурин—пурин (слишком широко) или пиримидин—пиримидин (слишком узко).

Во-вторых, критическую роль сыграло предшествовавшее наблюдение Чаргаффа о том, что независимо от действительного количества каждого основания относительное содержание G всегда было равно отношению содержания C, а содержание A—содержанию T. (Таким образом, любая ДНК может быть охарактеризована по ее составу как **отношение** $(G + C)/(A + T)$, которое, варьируя от 26 до 74%, остается характерным для каждого вида.)

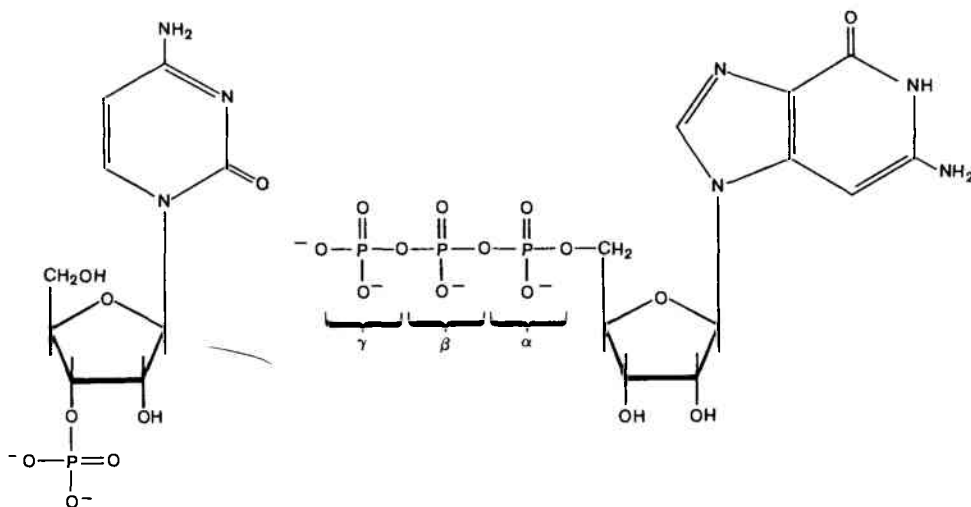


Рис. 2.6. Фосфатная группа в нуклеотидах может находиться в 5'- или в 3'-положении.

Слева изображен цитидин-3'-монофосфат, а справа – гуанозин-5'-трифосфат. Дифосфаты сокращенно обозначают NDP, а трифосфаты – NTP.

Уотсон и Крик предположили, что две полинуклеотидные цепи в ДНК не связаны ковалентно, а соединяются **водородными связями**, возникающими между азотистыми основаниями. На **рис. 2.7** показано, что в своей обычной форме G может образовывать водородную связь специфически только с C, тогда как A специфически соединяется только с T. Эти реакции называют **спариванием оснований**, а об основаниях, способных спариваться (G с C и A с T), говорят, что они **комплементарны**. Для осуществления специфического спаривания основания должны находиться в соответствующей форме. Перемещения водородного атома позволяют каждому основанию существовать в различных **таутомерных формах**. Основания, входящие в состав двойной спирали, имеют аминогруппы (NH_2) и оксогруппы ($\text{C}=\text{O}$) в отличие от таутомеров, имеющих иминогруппы (NH) и енольные группы ($\text{C}=\text{OH}$).

Согласно модели двойной спирали, две полинуклеотидные цепи **антипараллельны** (т.е. идут в противоположных направлениях; **рис. 2.8**). Поэтому, рассматривая спираль вдоль оси, можно увидеть, что одна цепь идет в направлении 5'-3', а другая – в направлении 3'-5'.

Основания имеют плоскую форму и располагаются парами перпендикулярно оси спирали. Свободная энергия водородных связей между спаренными основаниями вносит свой вклад в термодинамическую стабильность двойной спирали. Вдоль спирали основания **уложены стопкой друг на друга**; энергия, необходимая для поддержания спиральной структуры частично обеспечивается гидрофобными **межплоскостными взаимодействиями оснований (стэкинг-взаимодействия)**.

Каждая пара оснований повернута на 36° вокруг оси спирали относительно следующей пары оснований. Таким образом, 10 пар оснований составляют полный оборот в 360° . Две цепи, закручиваясь друг относительно друга, образуют двойную спираль, в которой имеется две бороздки – **малая бороздка** (около 12 \AA шириной) и **большая бороздка** (около 22 \AA шириной), – которые хорошо видны на приведенной модели (**рис. 2.9**). Двойная спираль **правосторонняя**: если смотреть вдоль оси спирали, повороты следуют по часовой стрелке. Это описание соответствует модели ДНК, известной как **В-форма**.

Одно из обязательных требований к генетическому материалу состоит в том, что он должен быть способен к точному самовоспроизведению при каждом клеточном делении. По этому поводу в конце своей статьи Уотсон и Крик сделали классический вывод: «Нельзя не заметить, что постулированное нами специфическое взаимодействие непосредственно предполагает возможность копирования генетического материала». Эта мысль была подхвачена в последующей статье, где они подчеркнули, что соединение двух полинуклеотидных цепей только водородными связями позволяет им разделяться. Из специ-

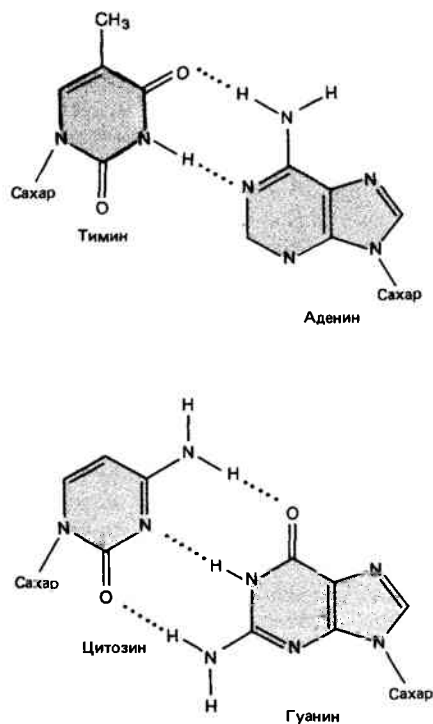


Рис. 2.7. При специфическом спаривании оснований образуются две водородные связи между А и Т и три водородные связи между G и C.

Других пар оснований в ДНК не бывает.

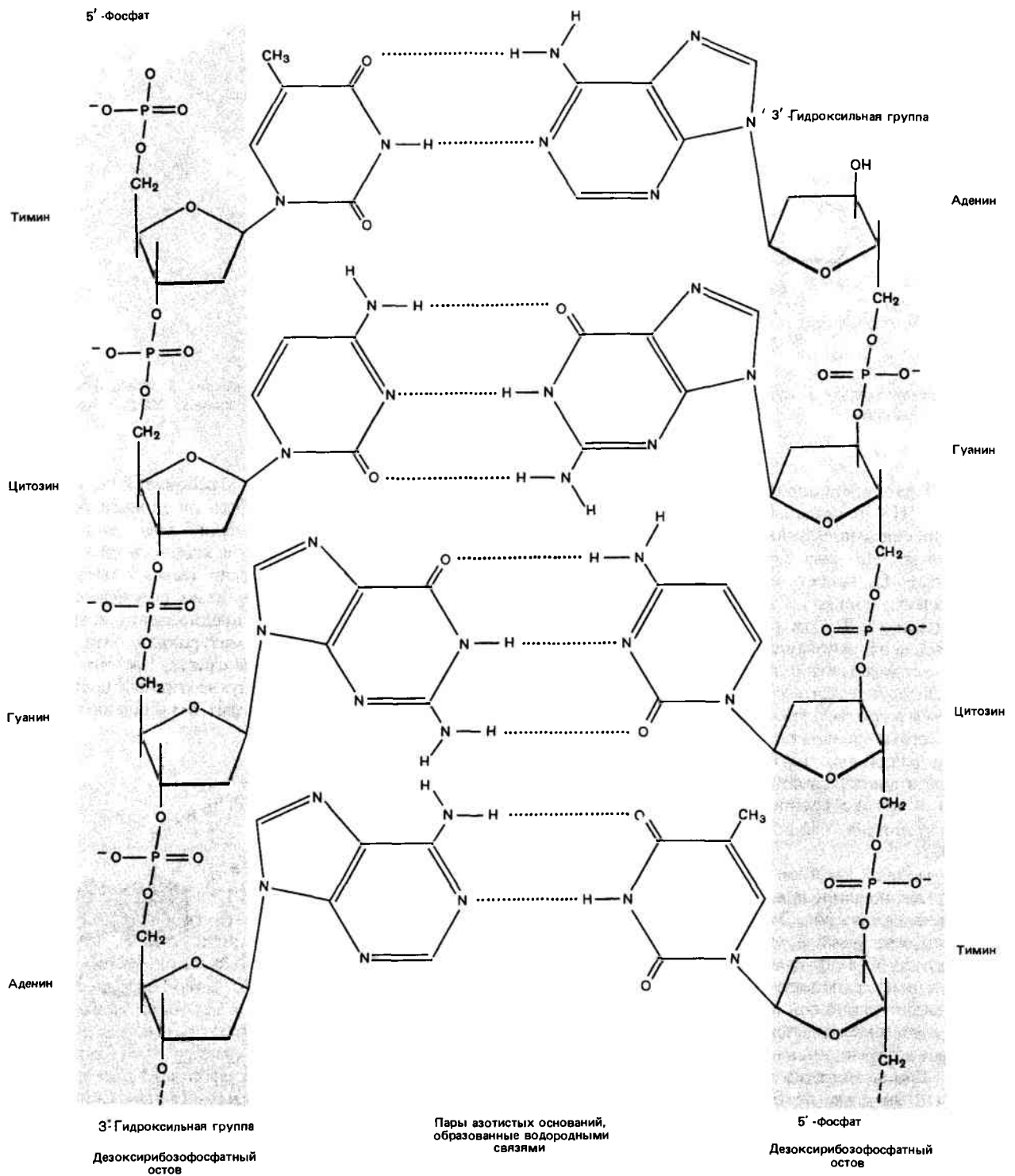


Рис. 2.8. Двойная спираль имеет постоянную ширину, поскольку пурины всегда расположены против пиримидинов, образуя специфические пары оснований G—C и A—T. Сахарофосфатный остов расположен снаружи и заряжен отрицательно.

Основания повернуты внутрь спирали и располагаются стопкой, друг над другом. Левая цепь сверху вниз имеет направление 5'-3', а правая — 3'-5'. Рисунок схематичен, и закручивание двух цепей в спираль не показано.

фичности спаривания вытекает, что каждая цепь может служить **матрицей** для синтеза комплементарной цепи (рис. 2.10). Такой механизм был назван **полуконсервативным синтезом** (см. гл. 31). Таким образом, информация, необходимая для воспроизведения последовательности ДНК, заложена в ее структуре.

Об альтернативных двуспиральных структурах

Рассмотренную структуру ДНК обычно считают одним из несомненных фактов молекулярной биологии. Но недавно стало очевидным, что некоторые параметры классической В-формы нужно пересмотреть и даже что ДНК может образовывать другие типы двуспиральных структур.

Проблема определения точных значений параметров двойной спирали связана с тем, что все первоначальные значения были получены при изучении дифракции рентгеновских лучей на волокнах ДНК. По этим данным установлены основные характеристики, такие, как число пар оснований на виток, расстояние между соседними парами вдоль оси спирали. Однако этим методом нельзя определить положения отдельных атомов, как это делают, используя рентгеноструктурный анализ. Поэтому модели уточняют, согласуя расчетную картину дифракции рентгеновских лучей с экспериментальными данными. Теоретически на основе одних и тех же данных всегда можно построить разные модели. Дело в том, что модель удовлетворяет *усредненным* данным и, следовательно, в определенных областях структуры возможны вариации. Например, до сих пор мы говорили о ДНК как о длинной, жесткой двойной спирали, но мы знаем, что в действительности она должна быть свернута и плотно уложена, для того чтобы уместиться в клетке. При этом детали ее структуры могут изменяться.

Сейчас по крайней мере кажется, что двухцепочечная ДНК – всегда двойная спираль. Однако была предложена и другая модель, хорошо согласующаяся с данными дифракции рентгеновских лучей. Согласно этой модели, две антипараллельные цепи ДНК также соединяются путем комплементарного спаривания оснований, но лежат «бок о бок», вместо того чтобы закручиваться в непрерывную двойную спираль. Существование ДНК в виде двойной спирали было подтверждено экспериментами, в которых прямо измеряли число пар оснований на виток. Оказалось, что их 10,4 вместо 10,0, предсказанных классической В-моделью. Это различие вызвало необходимость несколько изменить угол вращения между соседними парами оснований вдоль спирали до $34,6^\circ$, так что отрезок спирали, в пределах которого совершается полный виток на 360° , стал несколько длиннее.

Особенно важно, что значение 10,4 является средним для ДНК как целой молекулы при определенных условиях. Изменение условий или даже последовательности отдельных оснований может привести к большему или меньшему закручиванию спиральной структуры в соответствующих участках. Действительно, методом рентгеноструктурного анализа было показано, что молекула, состоящая из 12 пар оснований, содержит 10,1 пары оснований на виток, что обеспечивается слабым сдвигом каждой пары оснований, при котором улучшаются межплоскостные (стекинг) взаимодействия между основаниями, по сравнению с первоначальной моделью.

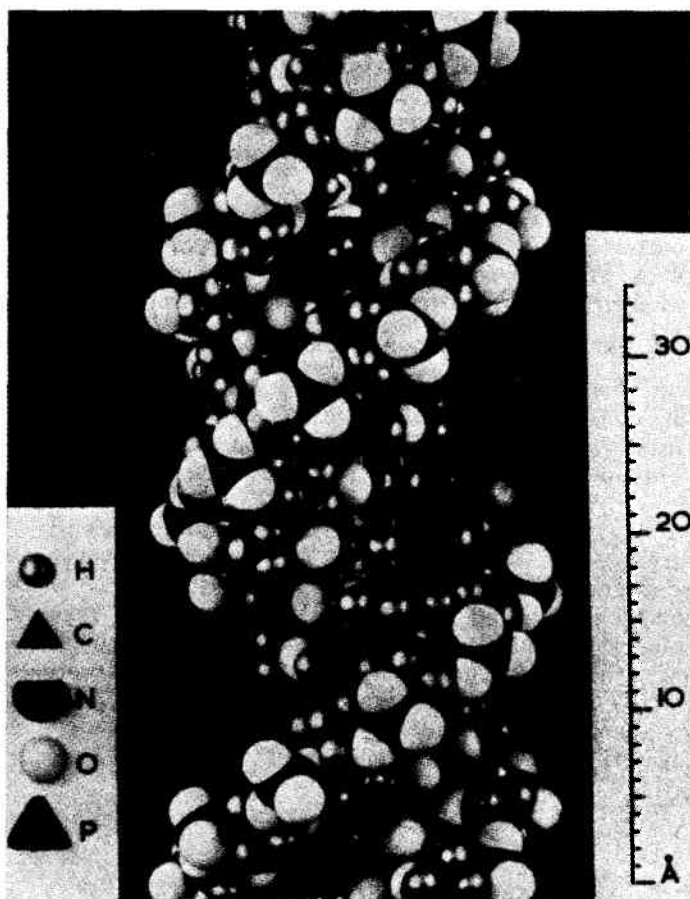


Рис. 2.9. Две цепи ДНК образуют двойную спираль. (Фотография любезно представлена Maurice Wilkins.)

Ввиду таких вариаций идея о существовании единственной структуры двуспиральной ДНК сменилась представлением, допускающим наличие семейства структур, каждая из которых имеет характерный тип, но проявляет различия по главным параметрам n (число нуклеотидов на виток) и h (расстояние между соседними повторяющимися элементами). Вариации обусловлены изменением вращения отдельных групп вокруг связей, обладающих свободой вращения. В пределах каждого семейства структур параметры могут слабо варьировать. Например, для В-ДНК значение n может составлять от 10,0 до 10,6.

В течение долгого времени были известны три структурные формы ДНК, способные к взаимопревращениям при изменении соответствующих условий. Общие свойства этих форм суммированы в табл. 2.1.

В-форма спирали, для которой Уотсон и Крик построили свою модель, характерна для волокон ДНК при очень высокой относительной влажности (92%) и в растворах низкой ионной силы. Считают, что именно в такой форме ДНК обычно находится в живой клетке.

А-форма обнаружена в волокнах ДНК при 75% влажности и нуждается в присутствии ионов натрия, калия или цезия, несущих противоположный заряд. Основания, располагавшиеся строго перпендикулярно оси спирали в В-форме, в А-форме наклонены по отношению к оси спирали, и их число на виток больше. А-форма интересна с биологической точки зрения, так как ее конформация, очевидно, близка к структуре гибридов ДНК – РНК и дву-

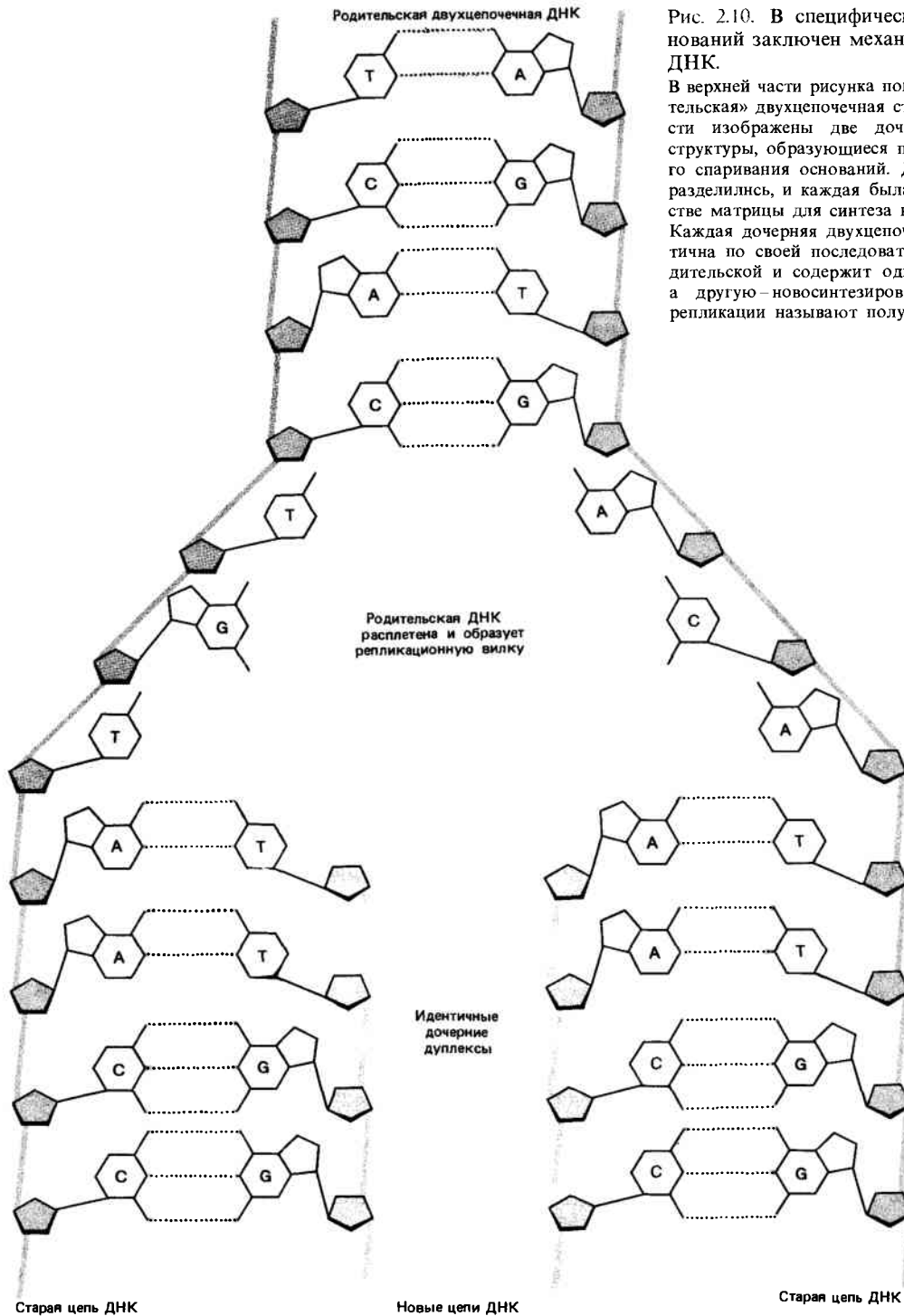


Рис. 2.10. В специфическом спаривании оснований заключен механизм для репликации ДНК.

В верхней части рисунка показана исходная «родительская» двуцепочечная структура. В нижней части изображены две дочерние двуцепочечные структуры, образующиеся путем комплементарного спаривания оснований. Две родительские цепи разделились, и каждая была использована в качестве матрицы для синтеза комплементарной цепи. Каждая дочерняя двуцепочечная структура идентична по своей последовательности исходной родительской и содержит одну родительскую цепь, а другую — новосинтезированную. Такой способ репликации называют полуконсервативным.

спиральных участков РНК (обсуждается далее). Причина этого заключается в том, что 2'-гидроксильная группа мешает РНК принять В-форму.

С-форма образуется, когда ДНК находится при 66% влажности в присутствии ионов лития. У нее меньше пар оснований на виток, чем у В-ДНК.

В этих трех формах могут находиться все ДНК независимо от их нуклеотидной последовательности. Следующие формы характерны только для молекул ДНК с определенными особенностями в составе пар оснований.

D-форма и *E-форма* (возможно, крайние варианты

одной и той же формы) имеют наименьшее число пар на виток (8 и 7,5) и обнаружены только в определенных молекулах ДНК, не содержащих гуанина.

Z-форма представляет собой наиболее резкий контраст с классическим структурным семейством. Эта форма *левоспиральная*, тогда как все остальные — *правоспиральные*. Она имеет наибольшее число пар на виток, т.е. менее скрученная и более тонкая. Свое название форма получила из-за зигзагообразной (*zigzag*) линии, которую образует сахарофосфатный остов вдоль спирали. В отличие от этого остов В-формы ДНК образует плавно изги-

Таблица 2.1
ДНК может существовать в виде структурных семейств нескольких типов

Тип спирали	Число пар оснований на виток	Угол вращения ¹⁾ одной пары, градусы	Расстояние между парами оснований, Å	Диаметр спирали, Å
A	11	+ 32,7	2,56	23
B	10	+ 36,0	3,38	19
C	9 ¹ / ₃	+ 38,6	3,32	19
Z	12	- 30,0	3,71	18

¹⁾ Угол вращения показан знаком (+) в случае правосторонней спирали и знаком (-) в случае левосторонней спирали.

бающуюся линию, что хорошо видно на рис. 2.11, где сравниваются В- и Z-формы ДНК.

Z-форма двойной спирали найдена в полимерах, образованных чередующимися пурин-пиримидиновыми последовательностями. Два исследованных полимера состояли из многократно повторенных динуклеотидных пар: poly-d $\begin{pmatrix} GC \\ CG \end{pmatrix}_n$ и poly-d $\begin{pmatrix} AC \\ TG \end{pmatrix}_n$. (Буква «d» указывает, что это дезоксиформы, т.е. что это ДНК, а не РНК. Верхний и нижний ряды букв соответствуют двум цепям ДНК.) Z-форма существует только при очень высоких концентрациях соли (что объясняется необходимостью противостоять повышенному электростатическому отталкиванию

между нуклеотидами, сжатыми из-за уменьшения диаметра двойной спирали Z-ДНК).

Вполне возможно, что Z-форма ДНК имеет определенное биологическое значение. Об этом свидетельствуют следующие данные: замена остатка С в полимере

poly-d $\begin{pmatrix} GC \\ CG \end{pmatrix}_n$ на 5-метилцитозин значительно повышает стабильность Z-ДНК при низкой концентрации соли.

Единственное различие между 5-Ме-С и С заключается в наличии метильной группы при С-5. Эта модификация цитозина происходит *in vivo* в результате метилирования динуклеотидной последовательности $\begin{pmatrix} CG \\ GC \end{pmatrix}$

в некоторых участках ДНК. Отсюда вытекает, что Z-ДНК может существовать *in vivo* при подходящих условиях, т.е. когда последовательность соответствующего состава окажется в условиях, способствующих образованию Z-формы.

До сих пор мы рассматривали ДНК как структуру совершенно изолированную. В действительности она связана с белками, которые могут оказывать значительное влияние на возможность перехода из В-формы в Z-форму. Например, ДНК, связанная с гистонами (основные хромосомные белки эукариотического ядра), не переходит из одной формы в другую в тех условиях, когда это наблюдают у свободной ДНК. Таким образом, одним из условий, необходимых для образования Z-ДНК *in vivo*,

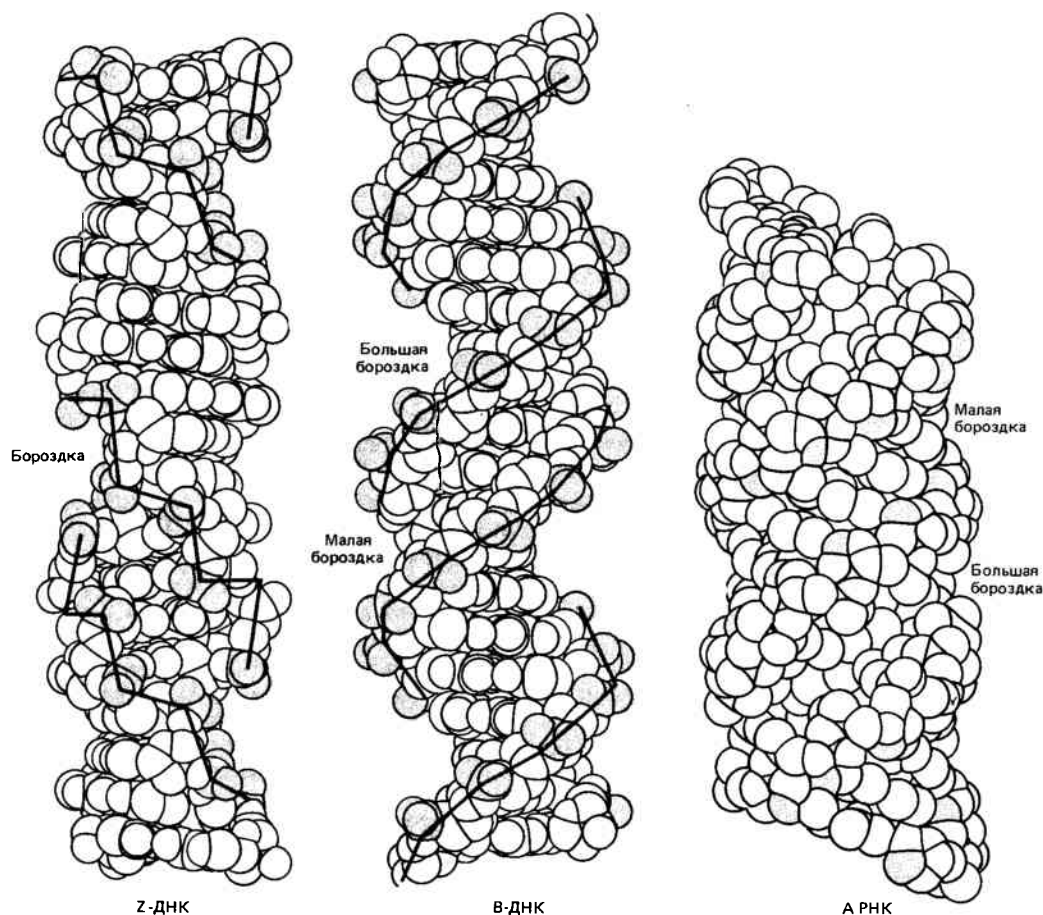


Рис. 2.11. Нуклеиновые кислоты могут образовывать несколько типов двойной спирали.

по-видимому, является присутствие особых белков, стабилизирующих ее структуру.

Является ли Z-ДНК единственной левосторонней формой двойной спирали? Модели для левосторонних вариантов традиционных А-, В- и С-форм построены так, что, возможно, они и существуют. Следует помнить, что все эти формы сохраняют самую существенную в биологическом смысле особенность двойной спирали: комплементарность спаривания оснований, определяющую специфичность соединения цепей.

При построении моделей каждую форму ДНК рассматривают независимо как конструкцию, в которой находится вся молекула в определенных условиях. Но вряд ли это действительно имеет место *in vivo*. Большинство клеточной ДНК находится, по-видимому, в В-форме с определенными модификациями параметров спирали, варьирующими локально. И только отдельные короткие участки спирали переходят в другие формы.

Возможно ли, чтобы переход из одной формы в другую происходил внутри одной и той же молекулы ДНК? При определенных условиях участки в 26–32 дуэлетов

GC/CG может переходить в Z-форму, тогда как участки ДНК с двух сторон остаются в классической В-форме. Эти данные подтверждают идею о том, что конформационные переходы могут происходить *in vivo* в специфических участках.

Получены антитела, способные отличать Z-форму от В-формы ДНК. Эти антитела связываются с определенными хромосомами *Dr. melanogaster*. Реакцию легко наблюдать из-за необычного строения этих хромосом, у которых более плотные участки (диски) контрастируют с менее плотными (междисками) (см. гл. 28). Участки Z-ДНК находятся в междисках. Из этого наблюдения следует, что Z-ДНК существуют в естественных условиях, хотя размер индивидуального отрезка Z-ДНК в хромосоме пока не известен.

Двойная спираль может подвергаться суперспирализации

До сих пор мы говорили о ДНК как о линейной двухцепочечной молекуле. Однако ДНК может находиться в виде **кольцевой молекулы** или чего-то подобного. У мелких вирусов геном действительно представляет собой кольцевую ДНК, в которой обе цепи двойной спирали замкнуты в непрерывное кольцо. В бактериальном и эукариотическом геномах ДНК может находиться в виде больших петель. При этом важно следующее: каждая петля удерживается у основания таким образом, что

опять не образуется свободных концов (см. гл. 28). Значение этой формы двойной спирали состоит в том, что она накладывает дополнительные ограничения на структуру.

Если двухцепочечную ДНК скрутить вокруг оси двойной спирали, то можно получить супервитки. В качестве аналогии обычно рассматривают полоску резины, скрученную вокруг своей оси и образующую напряженную структуру, в которой резиновая полоска (двойная спираль) местами образует крестообразные структуры. При этом возникают конформации, изображенные на **рис. 2.12**. В таком скрученном состоянии структура остается только в том случае, если нет свободных концов и раскручивание невозможно.

В гл. 32 мы рассмотрим, каким образом измеряют супервитки и как они образуются в ДНК. Здесь нужно только напомнить, что всякий раз, когда образуется суперспираль, нить скручивается относительно своей оси и чем больше супервитков, тем сильнее скручивание.

Для резиновой полоски безразлично, в каком направлении мы ее закрутим с образованием супервитков. (Оба края резиновой полоски одинаковы.) Однако двойная спираль — и так уже скрученная структура (что видно из переплетения двух цепей), поэтому ее реакция на скручивание зависит от того, в какую сторону оно происходит.

Отрицательные супервитки закручивают ДНК вокруг ее оси *против* часовой стрелки — в направлении, обратном правосторонней двойной спирали. Это в принципе означает, что напряжение скручивания можно частично ослабить, регулируя саму структуру двойной спирали. При этом может получиться форма с меньшим углом вращения на пару оснований, а местами может даже нарушиться спаривание оснований. ДНК с отрицательными супервитками называют **недокрученной** (*underwound*). Крайним случаем служит локальный переход правосторонней спирали в левостороннюю.

Не следует забывать, что такие события не происходят спонтанно в ДНК, а могут случаться только под влиянием внешних условий. Для простоты можно считать, что в результате суперспирализации молекула приобретает избыток энергии. Таким образом, если нужно произвести некоторые изменения в структуре ДНК, например удержать какой-то участок в одноцепочечном состоянии в результате связывания его с белками или перевести его в какую-либо иную форму, например в Z-форму, на это потребуются *меньше энергии*, если ДНК находится в состоянии отрицательной суперспирали. Так, степень спирализации может влиять на равновесие между разными структурами. Действительно, было обнаружено, что отрицательно суперспирализованная ДНК проявляет склонность к изменениям такого рода.

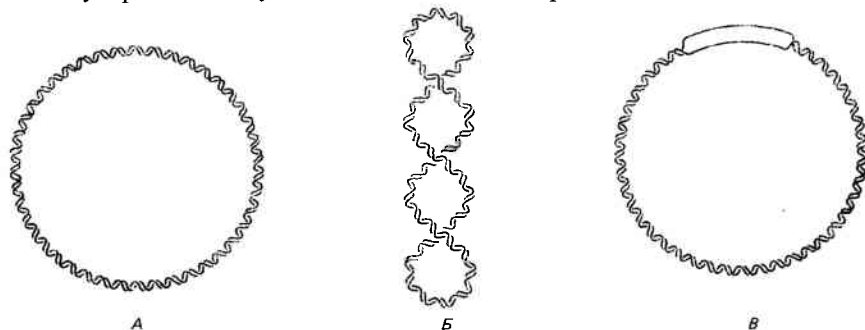


Рис. 2.12. При суперспирализации двуспиральная молекула ДНК закручивается сама на себя. Разделение оснований снимает отрицательную суперспирализацию (сверхскручивание).

А. Кольцевая ДНК без суперспирализации. Б. Отрицательно суперспирализованная ДНК. В. Отрицательная суперспирализация может привести к разделению цепей.

Суперспирализация влияет на структуру двойной спирали

Изменения структуры ДНК *in vivo* происходят не спонтанно — они осуществляются в таких условиях, при которых на структуру ДНК налагаются различные ограничения. Некоторые из этих воздействий приводят к отрицательной суперспирализации.

Введение отрицательных супервитков требует затраты энергии. Суперспирализованная молекула обладает большей энергией благодаря наличию супервитков. Следовательно, супервитки можно рассматривать как форму запаса энергии. Таким образом, если в структуре ДНК нужно произвести какое-то изменение, *требуемое* энергии, в случае отрицательно суперскрученной ДНК понадобится затратить *меньше* энергии. Практически это означает, что отрицательная суперспирализация может влиять на стабильность структуры. Суперспирализованная ДНК может подвергаться структурным изменениям в таких условиях, при которых релаксированная ДНК остается без изменений.

Одним из таких изменений является разделение цепей, облегчаемое недокрученностью молекул с отрицательной суперспирализацией. Излишек энергии, которым обладают отрицательно суперспирализованные молекулы, может быть использован для разделения цепей ДНК.

Можно ли количественно измерить эффект отрицательной суперспирализации? Все исследованные геномы проявляют некоторую отрицательную суперспирализацию. В типичном случае *in vivo* один отрицательный супервиток приходится на каждые 200 пар оснований, что соответствует плотности суперспирализации $-0,05$. Этому значению равно количество энергии около -9 ккал/моль.

Энергия, необходимая для расплетения ДНК, зависит от последовательности пар оснований. Это, по существу, величина, обратная количеству энергии, высвобождаемой при образовании двойной спирали. Таким образом, чтобы разделить 10 пар оснований, нужно затратить 12–50 ккал/моль. Значит, реально существующий уровень суперспирализации соответствует количеству энергии, которое может обеспечить расплетение только совсем немногих пар оснований.

Крайним случаем раскручивания правосторонней двойной спирали является ее превращение в левостороннюю спираль. Именно это происходит при переходе В-формы ДНК в Z-форму. Действительно, отрицательно суперспирализованная ДНК более предрасположена к переходу в Z-форму, чем релаксированная ДНК. В кольцевой молекуле ДНК, содержащей блоки повторов CG, их переход в Z-форму при физиологических концентрациях соли происходит лишь в том случае, если плотность отрицательной суперспирализации достигает критического значения. Когда C-остатки метилированы, для такого перехода требуется меньшее число отрицательных супервитков.

Чем длиннее отрезок, содержащий дуплеты CG , тем легче он превращается в Z-форму. Например, для конверсии отрезка из 32 дуплетов нужна плотность суперспирализации $-0,05$, тогда как конверсия 14 дуплетов происходит при плотности $-0,07$. Из этого следует, что относительно больше энергии требуется для *инициации* превращения В-формы ДНК в Z-форму, тогда как отно-

сительно мало энергии нужно для перевода в Z-форму последующих дуплетов CG внутри данного отрезка ДНК.

В соответствии с этими результатами свободная энергия, необходимая для перехода В-формы ДНК в Z-форму, складывается из двух значений: ΔG в точке перехода, равной 7,7 ккал/моль, и ΔG на каждую дополнительную пару оснований, равной 0,45 ккал/моль. Количество необходимой энергии представляет собой величину такого же порядка, что и энергия, обеспечиваемая отрицательной суперспирализацией. Таким образом мы видим, что суперспирализация потенциально представляет собой значительную силу, способную оказывать влияние на структуру двойной спирали.

Что происходит в месте перехода отрезка Z-ДНК в В-ДНК? Очевидно, в этом месте находится по меньшей мере одна пара оснований (а вероятно, их несколько), члены которой разделены и не соответствуют форме двойной спирали. Необходимость поддерживать неспаренное состояние в таких местах перехода частично объясняется большой затратой энергии на образование каждого участка перехода.

Другой возможный результат суперскручивания — это образование крестообразных структур в участках ДНК, содержащих палиндромы. Для поддержания крестообразной структуры необходимо значительное количество энергии — порядка $\Delta G = +18$ ккал/моль. Однако еще большая энергия активации расходуется на разрушение двойной спирали, необходимое для образования структуры креста: примерно $\Delta G = +50$ ккал/моль. При физиологических условиях эта реакция протекает также очень медленно. В действительности кажется, что степень суперспирализации, нужная для образования структуры креста, может оказаться намного больше той, которая реально встречается в живой клетке.

Противоположный эффект достигается, если ДНК закручена в *той же самом* направлении, что и сама двойная спираль. Введение **положительных супервитков** усиливает напряженность структуры, повышая напряжение скручивания, что еще сильнее закручивает двойную спираль. Положительно суперспирализованную ДНК называют **перекрученной** (overwound). Такого состояния молекулы можно добиться различными воздействиями *in vitro*, но оно не встречается в естественных условиях.

РНК тоже имеет вторичную структуру

Какое структурное разнообразие доступно для РНК? Ее первичная структура такая же, как и у ДНК: полинуклеотидная цепь с 5'-3'-сахарофосфатными связями. Но обычно она существует в виде *одноцепочечной* полинуклеотидной цепи, а не в виде двойной спирали из антипараллельных цепей. Таким образом, по составу оснований она не подчиняется правилу $[G] = [C]$ и $[A] = [U]$. Что касается биофизических свойств РНК, то наличие вторичной структуры у нее выражено гораздо слабее. Однако спаривание оснований может происходить внутри и между молекулами РНК. Мы уже упоминали, что такие двухцепочечные участки РНК, вероятно, существуют в А-форме.

Если за какой-то последовательностью оснований сле-

дует комплементарная ей последовательность, полинуклеотидная цепь может сложиться и образовать антипараллельную двухцепочечную структуру. Такую структуру называют **шпилькой**. Она состоит из спаренных оснований, образующих двуспиральный участок — **стебель**, часто с **петлей** из неспаренных оснований на одном конце. Конкретный пример рассматривается ниже (рис. 2.13).

Если комплементарны два более удаленных участка цепи РНК, они могут соединиться и также образовать двуспиральную область. При этом образуется стебель с очень длинной одноцепочечной петлей. Кроме того, две различные молекулы РНК могут иметь комплементарные участки, которые спариваются при подходящих условиях. Обычно это достаточно короткие участки.

Такие взаимодействия имеют важное биологическое значение, и примеры каждого из них рассматриваются в этой книге по ходу изложения. Из этих примеров следует, что спаривание оснований занимает центральное место во всех процессах, протекающих с участием нуклеиновых кислот.

Если вторичная структура РНК нерегулярна, то можно ли ее предсказать, исходя из последовательности оснований, или установить экспериментально? Иногда в молекуле РНК может быть несколько участков, потенциально способных к спариванию друг с другом во взаимоключающих комбинациях. В этом случае необходимо установить, какая комбинация возникает в действительности.

Общую протяженность двухцепочечных участков можно установить, исходя из физических свойств молекулы, так, как это описано в следующем разделе. Однако на основе таких измерений нельзя узнать, какие именно индивидуальные отрезки участвуют в спаривании. Одноцепочечные и двухцепочечные участки РНК различаются по чувствительности к некоторым нуклеазам. Это дает возможность установить, какие конкретные участки молекулы вовлечены в спаривание. Такие данные, однако, имеют ограниченное применение, и в основном анализ вторичной структуры РНК зависит больше от развития теоретических представлений, чем от информации, получаемой экспериментальным путем.

Вероятность спаривания оснований на определенном участке может быть предсказана на основе **правил**, описывающих взаимодействие между основаниями. При существовании альтернативных структур можно оценить их относительную стабильность. Конечно, при этом РНК рассматривают как изолированную и стабильную структуру, игнорируя любые другие факторы, которые могут взаимодействовать с РНК *in vivo*, влияя на ее конформацию (например, связывание с белками). Однако применение этих правил позволяет оценить принципиальную возможность существования определенной структуры.

В основе этих правил лежит расчет **свободной энергии** образования каждой структуры. Вкратце свободная энергия — это термодинамическая величина, соответствующая такому количеству энергии, которое может высвободиться в результате реакции. Формально это параметр ΔG , имеющий **отрицательное** значение и соответствующий энергии связи. Например, если имеются две альтернативные структуры со значениями свободной энергии $\Delta G = -21$ и $\Delta G = -35$, то при прочих равных условиях более вероятно образование последней.

Общая свободная энергия структуры подсчитывается как сумма свободных энергий спаривания индиви-

дуальных оснований, участвующих в контакте при образовании двухцепочечной структуры. Для такого подсчета каждой паре оснований приписывают определенное значение ΔG . Поскольку пары G—C имеют три водородные связи, и при образовании каждой из них высвобождается определенное количество энергии, они более стабильны. Значение ΔG для G—C-взаимодействия равно $-2,4$ ккал/моль. Для пар A—U $\Delta G = -1,2$ ккал/моль, так как они удерживаются только двумя водородными связями. Пары, образующие в РНК такие неканонические сочетания, как G—U, имеют $\Delta G = 0$. Другими словами, G может находиться против U в двуспиральной РНК, не нарушая спирали и не влияя на ее образование.

В действительности оказывается, что основным источником свободной энергии при двуспиральном спаривании служит гидрофобное взаимодействие оснований, уложенных стопкой вдоль по спирали (межплоскостное, или стэкинг-взаимодействие). Каждая комбинация пар оснований имеет различную свободную энергию, на которую влияет не только состав пар, но и порядок их расположения. Например, двойные последовательности $\begin{matrix} GG \\ CC \end{matrix}$ и $\begin{matrix} GC \\ CG \end{matrix}$ освобождают больше свободной энергии ($\Delta G = -5,0$), чем $\begin{matrix} CG \\ GC \end{matrix}$ ($\Delta G = -3,2$), хотя все они состоят из двух GC-пар. Аналогично последовательность $\begin{matrix} AA \\ UU \end{matrix}$ освобождает меньше энергии ($\Delta G = -1,2$), чем альтернативные последовательности того же состава $\begin{matrix} AU \\ UA \end{matrix}$ и $\begin{matrix} UA \\ AU \end{matrix}$ ($\Delta G = -1,6$). (Принято, что при написании последовательных двойных спиралей верхний ряд соответствует направлению 5'—3', а нижний—3'—5'.)

По этой причине свободную энергию рассчитывают для каждого дуплета соседних пар оснований в потенциальной двуспиральной структуре. При подсчете каждую пару включают в два дуплета — по одному с каждой стороны. (Если пара оснований лежит на конце спирали, она участвует в подсчете только один раз, так как состоит только в одном дуплете.) Сумма свободной энергии всех дуплетов составляет общую энергию образования структуры. И в этом случае считается, что комбинации, содержащие пары G—C, освобождают больше энергии, чем те, которые содержат пары A—U. Таким образом, стабильность двуспиральных участков увеличивается пропорционально GC-содержанию. Однако точное значение зависит от последовательности оснований.

Пример расчета свободной энергии показан на **рис. 2.13**. В отличие от ДНК, которая имеет совершенную двуспиральную структуру, двуспиральные участки РНК образуются обычно из двух отдельных цепей, не имеющих полной комплементарности. Это означает, что по ходу двойной спирали встречаются помехи, нарушающие спаривание. Неспаренный участок образует **петлю шпильки**, если он расположен между соседними комплементарными последовательностями, находящимися в противоположных направлениях. **Внутренние петли** образуются, если в потенциально комплементарных последовательностях встречаются некомплементирующие вставки. Даже одно лишнее неподходящее основание в потенциально комплементарных цепях вызывает **дефект** в двойной спирали. Поскольку неспаренные области мешают образованию двойной спирали, все такие *помехи* учитывают при подсчете и включают в общую сумму,

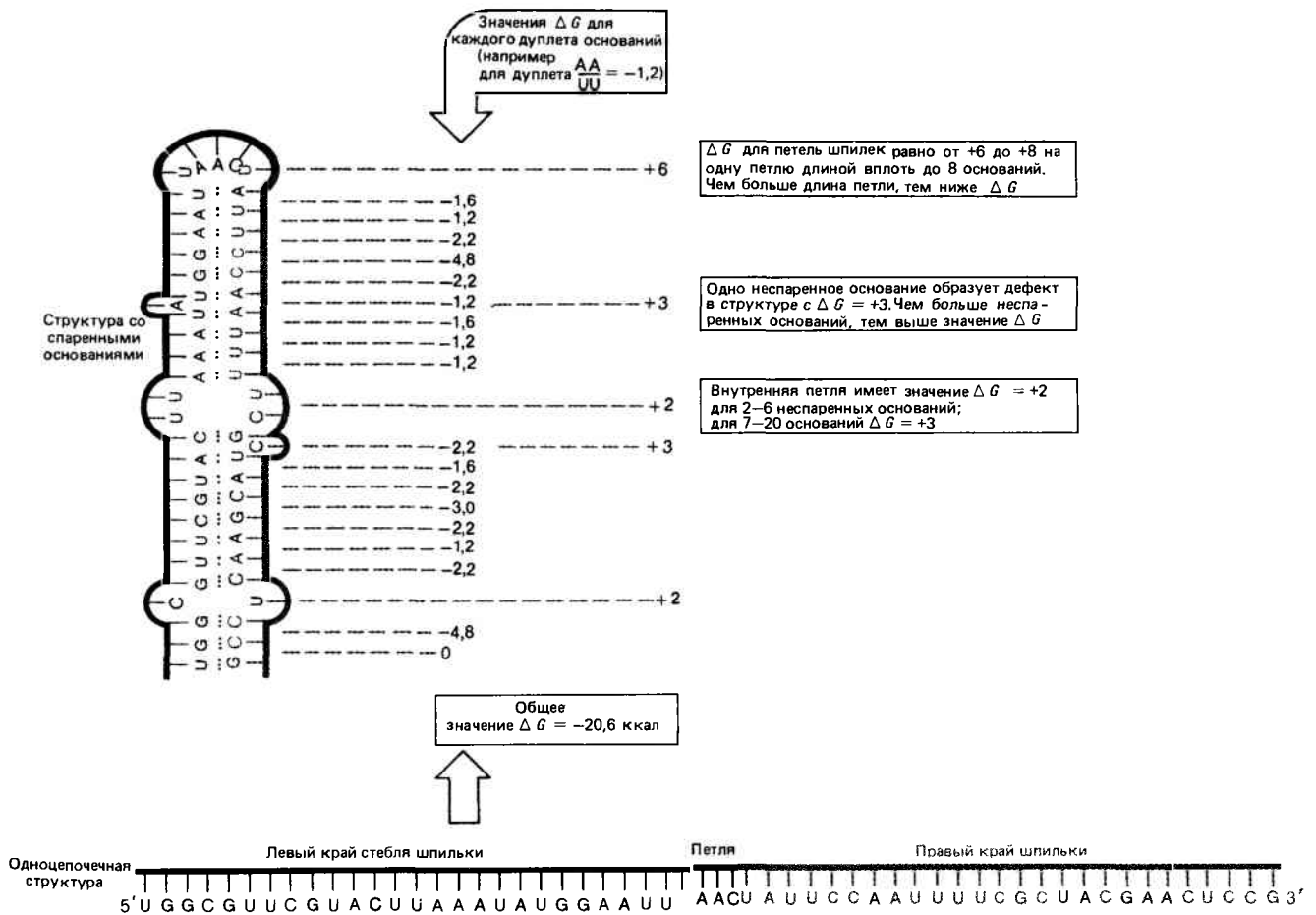


Рис. 2.13. Свободную энергию потенциально двуспирального участка высчитывают как сумму значений ΔG . Эти значения могут быть отрицательными (энергия высвобождения в ре-

зультате спаривания оснований) или положительными (нужна дополнительная энергия, для того чтобы удержать неспаренные участки в пределах соответствующей двуспиральной конформации).

причем каждому типу помех в отдельности придается *положительное* значение свободной энергии. Иными словами, требуется дополнительная *затрата* энергии, чтобы удержать эти участки в пределах данной структуры. Эта энергия – обратного знака по сравнению со свободной энергией спаривания оснований и стэкинг-взаимодействий. Все конкретные величины приведены на рисунке. Общая сумма свободной энергии должна обязательно быть отрицательной величиной, иначе вторичная структура не сможет образоваться.

ДНК можно денатурировать и ренатурировать

Структуру двойной спирали ДНК (или РНК), скрепленную водородными связями, можно разрушить нагреванием. Поскольку спаренные цепи не связаны между собой ковалентно, после разрыва всех водородных связей две полинуклеотидные цепи ДНК полностью разделяются. Процесс разделения цепей называют **денатурацией** или **плавлением**.

Денатурация происходит в узком интервале температур и отражается в кардинальном изменении многих физических свойств ДНК. Особенно полезным оказалось изменение оптической плотности. Гетероциклические коль-

ца нуклеотидов поглощают свет в ультрафиолетовой области (с максимумом, близким к 260 мкм, характерным для каждого основания). Но поглощение, характерное для самой ДНК, почти на 40% меньше, чем поглощение смеси свободных нуклеотидов того же состава. Это явление, называемое **гипохромным эффектом**, обусловлено взаимодействием электронных систем оснований в результате их стэкинг-взаимодействий при параллельном расположении в двойной спирали. Любое отклонение от двуспирального состояния немедленно сказывается в изменении величины этого эффекта. Иными словами, происходит сдвиг оптической плотности в сторону значения, характерного для свободных оснований. Таким образом, за денатурацией ДНК можно наблюдать, исследуя ее гиперхромность.

Среднюю точку температурного диапазона, при которой происходит разделение цепей ДНК, называют точкой плавления и обозначают $T_{пл}$. Пример кривой плавления, полученной по измерению оптического поглощения, показан на **рис. 2.14**. Кривая всегда имеет одну и ту же форму, но ее положение на температурной шкале зависит от состава оснований ДНК и условий, использованных для денатурации.

Когда ДНК находится в растворе примерно в физиологических условиях, $T_{пл}$ лежит в диапазоне 85–95°C. Точ-

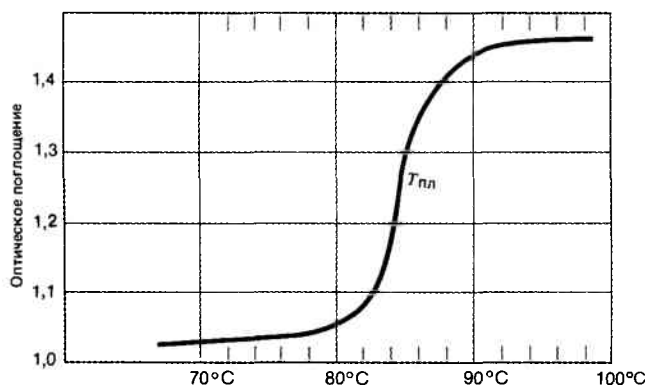


Рис. 2.14. Денатурация ДНК характеризуется значением $T_{пл}$. Экспериментально за денатурацией ДНК можно следить по изменению ее оптической плотности. Оптическое поглощение — в условных единицах.

ное значение зависит от состава оснований. Пары G—C, связанные тремя водородными связями и имеющие большую свободную энергию образования, оказываются более тугоплавкими, чем пары A—T, связанные двумя водородными связями. Зависимость $T_{пл}$ от состава оснований линейна. При каждом увеличении содержания G—C пар значение $T_{пл}$ увеличивается на $0,4^{\circ}\text{C}$. Так, ДНК, на 40% состоящая из G—C (что типично, например, для генома млекопитающих), будет денатурировать при $T_{пл}$ около 87°C в обычных условиях, тогда как ДНК, содержащая 60% G—C, при тех же условиях будет иметь $T_{пл}$ около 95°C .

Решающее влияние на $T_{пл}$ оказывает ионная сила раствора. $T_{пл}$ возрастает на $16,6^{\circ}\text{C}$ при каждом десятикратном увеличении концентрации моновалентных катионов. Чаще всего реакцию проводят в $0,12\text{ M}$ фосфатном буфере, что обеспечивает концентрацию моновалентных ионов Na^+ , равную $0,18\text{ M}$.

Значение $T_{пл}$ сильно меняется при добавлении в реакционную смесь таких веществ, как формамид, которые дестабилизируют водородные связи. Их присутствие позволяет снизить $T_{пл}$ до 40°C и тем самым избежать дополнительных повреждений ДНК (таких, как разрыв цепей), вызываемых повышением температуры.

Чрезвычайно полезное свойство денатурации ДНК — это обратимость процесса. При определенных условиях две разделенные комплементарные цепи могут восстано-

вить двойную спираль. Это явление называют **ренатурацией** (рис. 2.15). Ренатурация зависит от специфичности спаривания оснований между комплементарными цепями. Реакция происходит в две стадии. Сначала короткие комплементарные последовательности двух цепей случайно соединяются друг с другом и образуют двуспиральный участок. Затем область спаривания, подобно застежке-молнии, распространяется вдоль молекулы и образуется длинная двухцепочечная структура. Реконструкция двойной спирали завершается восстановлением первоначальных свойств, утраченных при денатурации ДНК.

Нуклеиновые кислоты гибридуруются путем спаривания оснований

В ренатурации участвуют две комплементарные последовательности, которые были разделены при денатурации. Однако метод ренатурации можно использовать применительно к любым комплементарным последовательностям, способным образовать двухцепочечную структуру при совместном **отжиге**. Если исследуемые кислоты взяты из различных источников, это явление обычно называют **гибридизацией** — например, в том случае, когда подвергаются отжигу ДНК и РНК. В основе гибридации лежит тот же принцип спаривания комплементарных оснований, который обеспечивает репликацию ДНК или ренатурацию молекул. Действительно, способность к гибридации двух препаратов нуклеиновых кислот служит строгим тестом на комплементарность их последовательностей.

В общих чертах метод гибридации заключается в совместной инкубации двух одноцепочечных препаратов нуклеиновых кислот и последующем измерении количества образованных дуплексов. Существуют два основных способа проведения этой реакции: **гибридизация в растворе** и **гибридизация на фильтре**.

Метод гибридации в растворе ясен из названия: препараты одноцепочечных молекул смешивают в растворе. При большом количестве материала за реакцией можно следить по изменению оптической плотности. Если в реакции участвует небольшое количество материала, один из компонентов метят радиоактивным изотопом и по количеству метки в двухцепочечной ДНК судят об эффективности гибридации. Для такого измерения

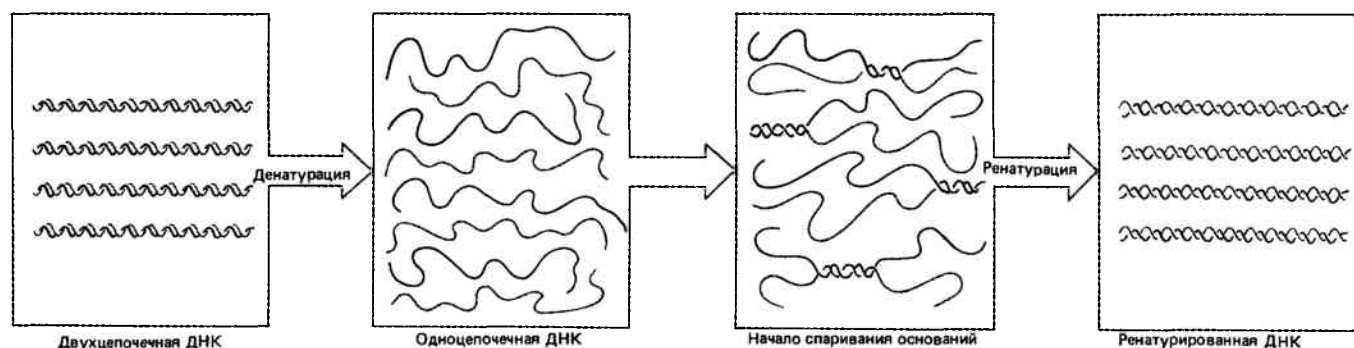


Рис. 2.15. Отдельные цепи денатурированной ДНК могут ренатурировать с образованием двуспиральной молекулы.

A — двухцепочечная ДНК, Б — денатурация, В — начало спаривания оснований, Г — ренатурированная ДНК.

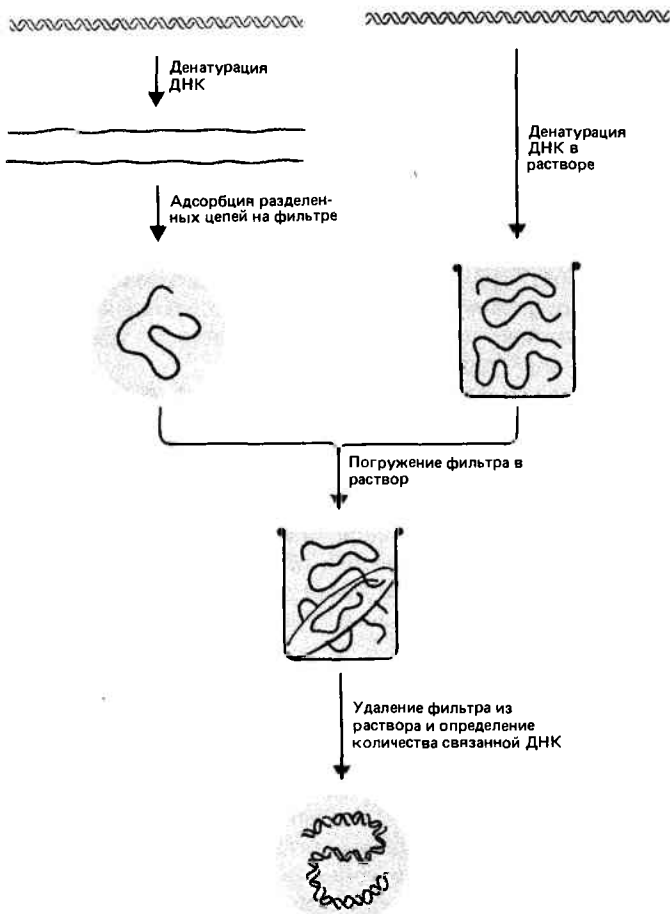


Рис. 2.16. С помощью гибридизации на фильтре можно определить, содержит ли данный препарат ДНК (или РНК) последовательности, комплементарные нуклеиновой кислоте, иммобилизованной на фильтре.

оставшиеся одноцепочечные молекулы подвергают специфической деградации или разделяют одно- и двухцепочечные молекулы методом хроматографии.

Определенные затруднения в использовании этого метода возникают в том случае, когда один или оба исследуемых препарата содержат двухцепочечную ДНК. Например, если две двухцепочечные ДНК денатурировать и затем смешать вместе, то могут одновременно происходить реакции двух типов. Единичные цепи могут ренатурировать с образованием *исходных* молекул, но, кроме того, каждая цепь может гибридизоваться с комплементарной последовательностью *другой* ДНК. Поскольку обе реакции конкурируют между собой, трудно оценить степень гибридизации.

Это затруднение можно преодолеть, если иммобилизовать один из препаратов ДНК таким образом, чтобы он не мог ренатурировать. Для этой цели используют нитроцеллюлозные фильтры, которые адсорбируют одноцепочечную ДНК и не адсорбируют РНК. Фильтры с адсорбированной одноцепочечной ДНК обрабатывают специальным образом, для того чтобы предотвратить дальнейшую адсорбцию одноцепочечных молекул. На рис. 2.16 показана схема эксперимента, в котором препарат ДНК денатурировали, полученные одноцепочечные молекулы адсорбировали на фильтре и затем добавляли второй препарат денатурированной ДНК (или препарат РНК). Связывание второй нуклеиновой кислоты происхо-

дит только в том случае, если она может комплементарно спариваться с ДНК, первоначально адсорбированной на фильтре.

Этот метод был использован во многих из тех опытов, которые будут рассматриваться в последующих разделах. Обычно к фильтрам добавляют радиоактивно меченные препараты РНК. Эффективность гибридизации измеряют по метке, оставшейся на фильтре. Этот метод очень полезен, так как, за исключением некоторых вирусных геномов, состоящих из РНК, все РНК в клетке образуются в результате копирования ДНК. Техника гибридизации дает в руки исследователя эффективный инструмент, позволяющий выяснить, с какого участка синтезируется данная РНК.

Молекулярная основа мутаций

Выделение мутантов имело огромное значение для изучения гена. До недавнего времени гены можно было идентифицировать только по мутациям, которые оказывались летальными или же блокировали развитие некоторого видимого или измеримого фенотипического признака. В результате картирования мутаций возникло предположение, что группа мутаций, нарушающих один и тот же признак, может быть тесно сцеплена, а использование теста на комплементарность показало, что каждая такая группа составляет функциональную генетическую единицу. Какова же природа этих мутаций?

Открытие генетической роли ДНК естественным образом породило представление о том, что мутации возникают в результате изменений в последовательности нуклеотидов. Мы скоро увидим, что последовательность нуклеотидов в гене отвечает за образование определенной последовательности аминокислот соответствующего белка. Изменение нуклеотидной последовательности ведет к изменению последовательности аминокислот и далее к изменению или нарушению активности белка.

Некоторые мутации возникают в результате нормальных процессов в клетке или при взаимодействии с окружающей средой. Такие мутации называют **спонтанными мутациями**; они с определенной частотой встречаются у любого организма. Частоту мутаций можно увеличить, воздействуя некоторыми соединениями. Наследственные изменения, возникшие под действием таких соединений, называют **индуцированными мутациями**, а сами соединения — **мутагенами**. Большинство мутагенов действует, прямо реагируя с определенными азотистыми основаниями или включаясь в нуклеиновую кислоту. Об эффективности мутагена судят по его способности увеличивать спонтанную («фоновую») скорость мутирования.

Любая пара оснований в ДНК может мутировать. Изменение, которое затрагивает только одну пару оснований, называют **точковой мутацией**. Точковые мутации могут быть двух типов. Замены одного пурина на другой пурин или одного пиримидина на другой пиримидин называют **транзисиями**. Например, пара G—C может быть заменена на A—T или наоборот. Это наиболее часто встречающийся класс точковых мутаций. К **трансверсиям** относят замены пурина пиримидином и наоборот, т. е. пара A—T превращается в T—A или C—G.

Одной из причин транзисий служит химическое превращение одного основания в другое. К обычному классу мутагенов относят соединения, взаимодействующие с одним или более основаниями и изменяющие их тип спаривания. Например, основной эффект азотистой кислоты за-

ключается в окислительном дезаминировании цитозина, превращающегося в урацил. Как показано на **рис. 2.17**, в следующем цикле репликации U спаривается с A вместо G, с которым соединялся бы исходный C. Таким образом, пара C—G замещается парой A—T (поскольку в следующем цикле репликации в пару с A становится T). Азотистая кислота может также дезаминировать аденин, вызывая обратную транзицию A—T в G—C.

Другой путь возникновения транзиций—это случаи **ошибочного спаривания**, приводящие к возникновению неканонических пар и, следовательно, к дефектам в уотсон-криковской спирали. В нормальном цикле репликации такая ошибка может случайно произойти вследствие включения неправильного основания. Спонтанная частота ошибок определяется прежде всего точностью фермента ДНК-полимеразы, отвечающей за репликацию (см. гл. 32). Существует также более ограниченный «репаративный» синтез ДНК, который активируется в результате генетической рекомбинации или повреждения ДНК (см. гл. 34). Различные системы репарации характеризуются разной частотой ошибок. Например, одна из репаративных систем *E. coli* особенно часто делает ошибки, и, следовательно, ее активация может стимулировать образование мутаций. Мы не располагаем достаточной информацией о частоте возникновения мутаций такого рода.

Некоторые мутагены являются аналогами обычных оснований, но способны к неканоническому спариванию.

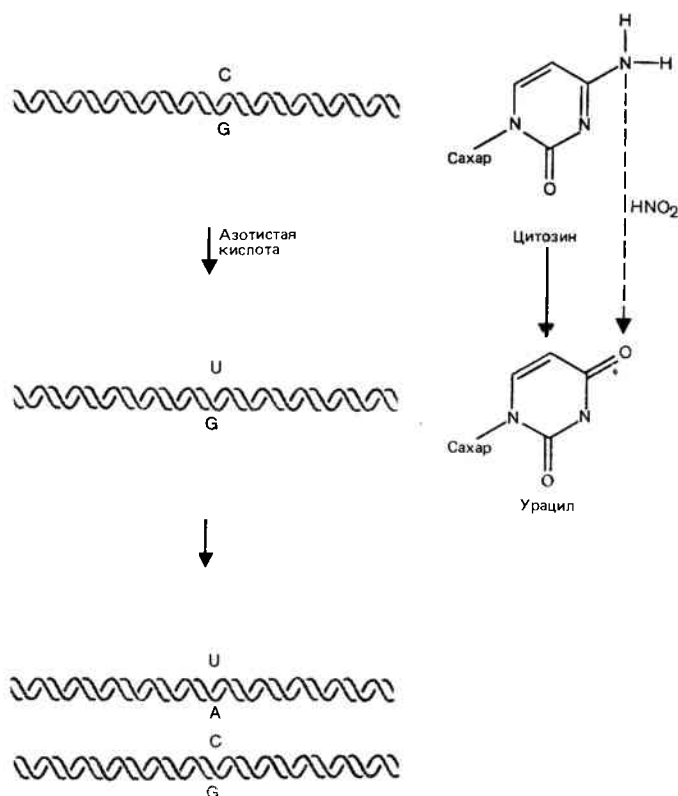


Рис. 2.17. Мутации можно индуцировать с помощью химической модификации основания.

Азотистая кислота окислительно дезаминирует цитозин и превращает его в урацил. При репликации ДНК один дочерний дуплекс получает исходную пару C—G, а у другого образуется пара U—A, которая в последующих циклах репликации превратится в A—T-пары.

Они могут включаться в ДНК при репликации. Из-за ошибочного спаривания при их включении или при последующей репликации они могут вызвать целую серию транзиций. Схема такого мутагенного действия рассматривается на примере бромурцила (BrdU на **рис. 2.18**). Это соединение включается в ДНК вместо тимина. Но из-за того что BrdU достаточно эффективно спаривается с гуанином (вместо аденина), его присутствие ведет к образованию замен пар A—T на G—C. В каждом цикле репликации эти замены возникают с определенной вероятностью, поэтому BrdU, однажды включенный в ДНК вместо тимина, продолжает индуцировать транзиции в последующих циклах репликации.

К другому классу мутагенов относятся акридиновые соединения. Их связывание с ДНК деформирует структуру двойной спирали так, что это приводит к вставке дополнительных оснований или к выпадению оснований при репликации. В результате каждого мутагенного события, индуцированного акридином, исчезает или прибавляется одна пара оснований. Мутации такого типа называют **мутациями сдвига рамки** в отличие от точковых мутаций, т.е. замены одного основания на другое (см. гл. 4).

Довольно долго считали, что основным типом мутаций в индивидуальных генах являются точковые мутации. Однако теперь мы знаем, что по крайней мере у бактерий достаточно часто происходят **вставки (включения, инсерции)** дополнительного материала. Источником этих вставок служат **мобильные элементы**—последовательности ДНК, способные перемещаться из одного места в другое (см. гл. 36). Аналогичные события происходят и у эукариот. Так, оказалось, что самая первая обнаруженная мутация (*белые глаза* у дрозофилы) возникла в результате вставки участка чужеродной ДНК в локус, контролирующей образование глазного пигмента. В местах включения таких вставок часто наблюдается **выпадение (делеция)** части или целиком всей вставки, а иногда и соседних участков ДНК.

Точковые мутации и мутации типа вставок (или делеций) различаются по их реакции на обработку мутагенами. Под действием мутагенов увеличивается частота точковых мутаций, но эти соединения обычно не влияют на изменения, вызванные мобильными элементами. Однако вставки и делеции могут возникать и по другим причинам—например, в результате ошибок во время репликации и рекомбинации, хотя и с меньшей вероятностью. Самые маленькие вставки и делеции индуцируют акридины. (Для описания изменений, затрагивающих только одну пару оснований, пользуются той же терминологией, что и при описании вставок (и делеций) длинных отрезков ДНК.)

Мутации концентрируются в горячих точках

До сих пор мы говорили о мутациях как об отдельных изменениях в последовательности ДНК, влияющих на активность той генетической единицы, в которой они возникают. Если рассматривать мутации как способ инактивации генов, то у большинства генов и спонтанная, и индуцированная частота мутирования примерно одинакова. Точнее, частота мутирования есть функция размера гена, т.е. мишени для мутаций. Но когда мы рассматриваем мутационные сайты внутри определенной последовательности ДНК, следует задать вопрос: все ли пары

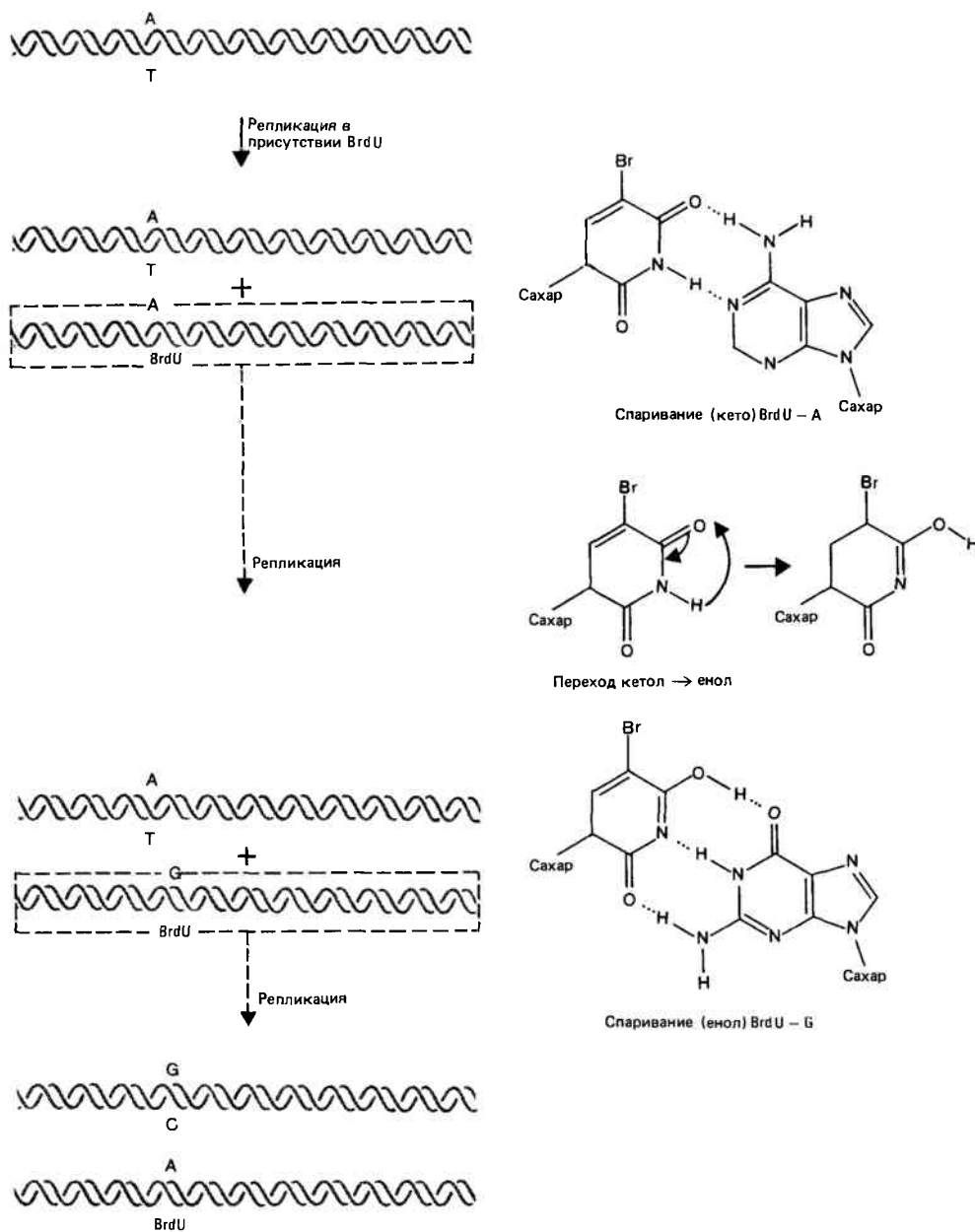


Рис. 2.18. Мутации могут возникать в результате включения в ДНК аналогов оснований.

В молекуле бромурацила метильная группа тимина замещена атомом брома. В результате такой замены увеличивается частота кето-енольных

переходов и BrdU спаривается то с гуанином, то с аденином. Таким образом, если BrdU включен в ДНК вместо тимина, он может вызывать замены А на G в каждой генерации с определенной частотой. Это приводит к замене пар А—Т на G—С.

оснований чувствительны в одинаковой степени или некоторые из них мутируют с большей вероятностью, чем другие?

Ответ на этот вопрос можно получить, выделив большое число независимых мутаций в одном гене. Это означает, что каждый полученный мутант возник в результате отдельного мутационного события. Затем определяют сайт каждой мутации (обычно методом генетического картирования, но теперь часто и прямым анализом последовательности ДНК). Большинство мутаций распределяется по разным сайтам, но некоторые попадают в один и тот же сайт. Две независимо отобранные мутации могут возникнуть в результате одинаковых или различных изменений. В первом случае одно и то же мутационное

событие происходит более одного раза, во втором—три разные точечные мутации могут произойти в результате изменения одной пары оснований.

Статистическая вероятность возникновения в отдельном сайте более чем одной мутации (здесь «сайт»—одна пара оснований) описывается одноударной кривой согласно распределению Пуассона. При этом в некоторых сайтах могут возникнуть одна, две или три мутации, а в других—ни одной. Все зависит от общего числа проанализированных мутаций. Но в некоторых сайтах оказывается намного больше мутаций, чем ожидается при случайном распределении. Их количество может быть в 10 и даже в 100 раз выше случайного. Такие сайты называют горячими точками.

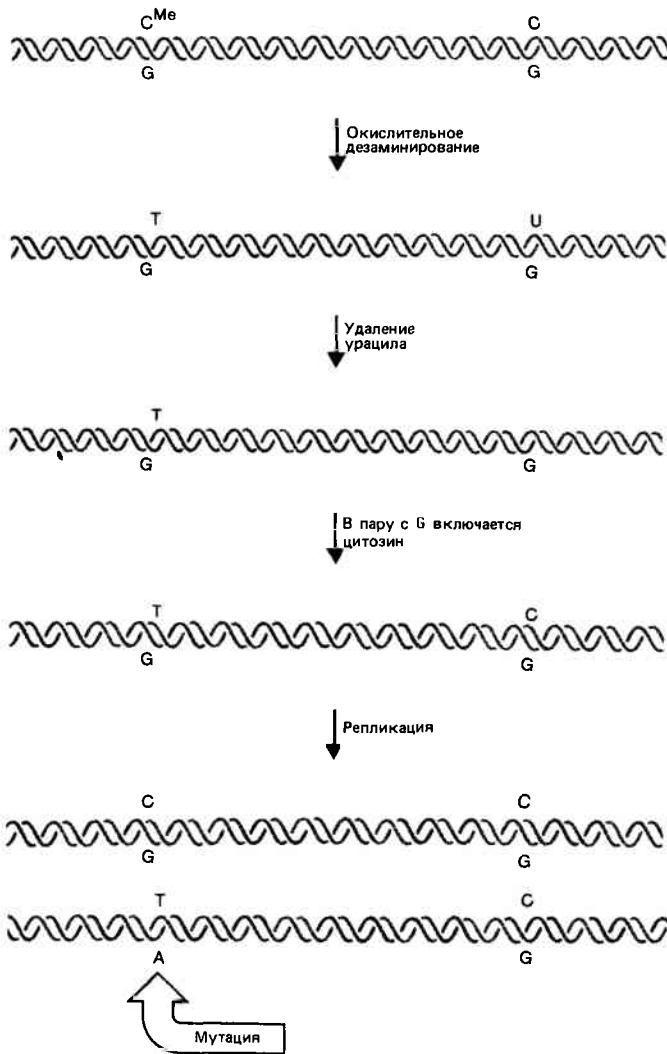


Рис. 2.19. В результате дезаминирования 5-метилцитозина образуется тимин, что приводит к транзиции С—G в Т—А. В то же время дезаминирование цитозина не вызывает замены оснований, поскольку урацил, образованный из цитозина, удаляется из ДНК.

Долгое время природа горячих точек оставалась загадкой. Считали, что горячие точки должны соответствовать каким-то последовательностям, особенно чувствительным к спонтанным и индуцированным мутациям. Однако причина этой чувствительности была не ясна. Действительно, сайты одних типов оказывались горячими точками для спонтанных мутаций и мутаций, индуцированных некоторыми мутагенами, тогда как другие горячие точки реагировали на другие мутагены.

Главной причиной спонтанных точковых мутаций у *E. coli* оказались необычные основания в ДНК. Кроме известных четырех оснований, включающихся в ДНК под действием ферментов репликации и репарации, иногда находят **модифицированные основания**. Название отражает их природу: они образуются в результате химической модификации одного из четырех оснований, *присутствующих в ДНК*. В модифицированной форме они не включаются в ДНК. Наиболее распространенное модифицированное основание — это 5-метилцитозин, который находят у про- и эукариот и о котором мы уже упоминали в связи с Z-формой ДНК. Модификацию производит фермент

метилаза, который добавляет метильную группу к цитозину только в определенных участках ДНК, модифицируя незначительную долю всех остатков С.

В одном детально исследованном гене (ген *lacI*) все горячие точки спонтанных точковых мутаций попали в сайты, которые у дикого типа содержат 5-метилцитозин. Во всех случаях мутации возникли в результате транзиции G—C в A—T. В штаммах *E. coli* с нарушенным метилированием горячие точки не были обнаружены.

Причина таких замен заключается в том, что 5-метилцитозин подвергается спонтанному дезаминированию с заметной частотой. Как показано на рис. 2.4, замена аминогруппы на оксогруппу превращает его в тимин. При этом возникает ошибочное спаривание G с T, что при последующей репликации приводит к образованию одной пары нуклеотидов дикого типа G—C и одной мутантной пары A—T.

Та же самая реакция дезаминирования превращает цитозин в урацил. Однако у *E. coli* есть фермент урацил-ДНК-гликозидаза, удаляющий остаток урацила из ДНК (см. гл. 32). В результате остается неспаренный остаток G, и система репарации по правилам обычного спаривания включает цитозин. Вероятно, эта система достаточна для защиты ДНК от последствий спонтанного дезаминирования (хотя и недостаточна для защиты от мутагенного действия азотистой кислоты; см. рис. 2.17). Но в случае дезаминирования 5-метилцитозина получается тимин, который является обычным компонентом ДНК. В этих условиях система репарации бездействует, и в результате возникает мутация.

На рис. 2.19 сопоставляются результаты дезаминирования цитозина и 5-метилцитозина. Эти данные по-новому освещают тот факт, что в ДНК используется T, тогда как в РНК—U. Возможно, это связано с необходимостью сохранять стабильность последовательности ДНК. Использование тимина имеет то преимущество, что любое дезаминирование С немедленно узнается, так как при этом образуется основание U, не характерное для ДНК.

Другая важная горячая точка в гене *lacI* находится в таком сайте, в котором последовательность CTGG повторяется подряд три раза. Мутации возникают в результате делеции или добавления одной четырехбуквенной единицы.

Мутант	CTGGCTGG
	↑ Делеция
Дикий тип	CTGGCTGGCTGG
	↑ Вставка
Мутант	CTGGCTGGCTGGCTGG

Такие мутации могут возникать в результате «проскальзывания» при репликации, если одна цепь не точно совпадает с другой. Например, первый четырехчленник одной цепи может спариться со вторым четырехчленником другой цепи. В этом случае может произойти добавление или потеря четырехчленных единиц.

Некоторые мобильные элементы имеют предпочтительные участки включения, играющие роль горячих точек мутаций. Однако такие участки различны для различных элементов, а большинство мобильных элементов включается в ДНК случайным образом (гл. 36).

Частота мутирования

Спонтанные мутации, инактивирующие функцию гена, возникают у бактерий с частотой от 10^{-6} до 10^{-7} раз за время одной генерации в расчете на один локус. У нас нет точных данных о скорости мутирования в случае эукариот. Однако обычно считают, что она примерно такая же, как и у бактерий.

Мы не знаем, какая доля спонтанных мутаций у бактерий обусловлена точковыми мутациями и какая — включениями мобильных элементов. Ясно только, что и те и другие играют в этом важную роль. То же самое касается и эукариот, у которых до настоящего времени еще слишком мало генов изучено так детально, чтобы можно было судить о природе мутаций.

В случае бактериальных генов, содержащих 1000 пар оснований каждый, частота мутирования соответствует изменению одного нуклеотида из 10^9 – 10^{10} в расчете на поколение. Если говорить только о точковых мутациях, такой подсчет является сильным упрощением, поскольку не все мутации приводят к заметным изменениям фенотипа. Мутации, не приводящие к заметным изменениям, называют **молчащими мутациями**. Различают два типа молчащих мутаций. В случае мутаций одного типа замена основания не приводит к аминокислотной замене в соответствующем белке. При мутациях второго типа замена основания сопровождается заменой соответствующей аминокислоты, однако эта замена не влияет на функциональные свойства белка. Мутации этого типа называют **нейтральными** заменами.

Мутации, инактивирующие ген, — это **прямые мутации**. Вызванное ими повреждение может быть устранено в результате **обратных мутаций**, которые подразделяются на два типа. Точное восстановление исходной структуры называют **истинной обратной мутацией (истинной реверсией)**. Например, если пара А—Т заменена парой G—C, другая мутация, восстановившая пару А—Т, восстановит и последовательность дикого типа. Однако эффект первой мутации может быть компенсирован мутацией в другой части гена. Такие мутации называют **вторичными реверсиями**. Например, замена одной аминокислоты может нарушить функцию гена; вторичное же изменение может компенсировать первое и восстановить активность белка.

Прямые мутации вызывают различные изменения, инактивирующие ген, тогда как обратные мутации должны восстановить функцию белка, поврежденного данной прямой мутацией. Таким образом, требования, предъявляемые к обратной мутации, гораздо более специфические, чем требования к прямой мутации. Частота обратных мутаций соответственно ниже, чем прямых, обычно в 10 раз. Иногда возникают также мутации в других генах, которые помогают каким-то образом обойти эффект первой мутации. Их называют **супрессорными мутациями** или более формально — **внегенными супрессорами**.

(Здесь термин *супрессия* употребляется в обычном смысле, в отличие от его использования при описании поведения мутаций, индуцированных акридиновым оранжевым).

Получение ревертантов служит важной характеристикой, позволяющей отличать точковые мутации и вставки от делеций. Так, точковые мутации могут ревертировать в результате восстановления исходной пары оснований или какого-то изменения во вторичном сайте. Реверсия вставки может произойти путем удаления вставки. Делеция же части генетического материала не может быть

точно восстановлена. Такие делеции играют важную роль в генетическом картировании. Генетическую протяженность делеции определяют по ее неспособности **комплементировать серию ряда расположенных мутаций (из-за нехватки соответствующего отрезка ДНК)**. Такая делеция комплементирует точковые мутации, расположенные с обеих сторон. На **рис. 2.20** показана делеция, размер которой определен в генетическом эксперименте по способности комплементировать серию точковых мутаций.

Получив серию **частично перекрывающихся делеций**, можно картировать любую точковую мутацию, используя тест на комплементацию. Схема такого опыта показана на **рис. 2.21**. Если две делеции не комплементируют точковую мутацию, это означает, что мутационный сайт должен находиться в общем для обеих делеций участке. Если же одна делеция комплементирует, а другая нет, то значит мутационный сайт находится в том участке гена, где делеции не перекрываются.

Говоря о природе мутаций, следует упомянуть еще одну деталь. В идеале хотелось бы уметь направленно получать мутации в любом гене организма. Но тут мы сталкиваемся с ограничениями двух типов.

Во-первых, необходим критерий, по которому можно отличить мутанта от дикого типа. Это нужно для того, чтобы не подбирать случайно попавшихся мутантов, а создавать методы их систематического **отбора**. Простейшим критерием служат видимые изменения фенотипа, например изменение цвета глаз. Другие критерии могут быть установлены путем использования различных методов отбора, например путем подбора таких условий роста, при которых необходимо присутствие или, наоборот, отсутствие какого-либо фермента. Возникающая тут проблема связана главным образом с возможностью наличия таких функций, о существовании которых мы не подозревали и поэтому не можем ни увидеть их, ни подобрать к ним селективных тестов. В предельном случае можно предположить, что существуют гены, чьи продукты не нужны и чье отсутствие не имеет никаких последствий. Как же такие гены обнаружить?

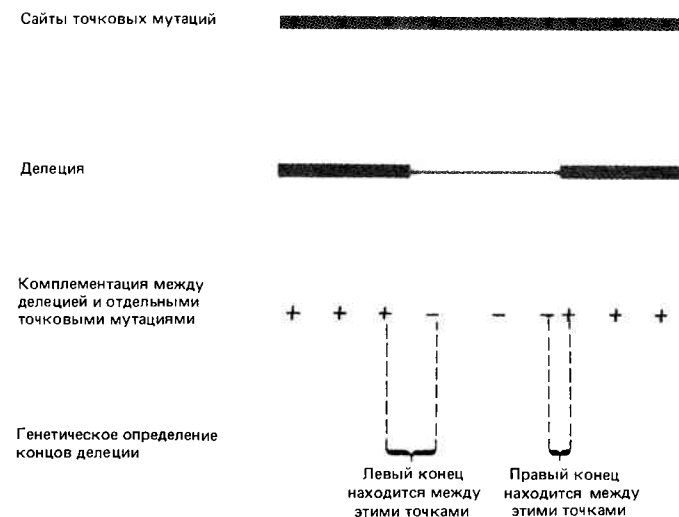


Рис. 2.20. Делеция может комплементировать серию точковых мутаций с обеих сторон от делетированной области, но она не комплементирует точковые мутации, находящиеся в пределах делетированной области. Границы делеции определяются по смене положительного результата в тесте на комплементацию на отрицательный.



Рис. 2.21. Метод делеционного картирования можно использовать для локализации точковых мутаций между концами перекрывающихся делеций.

Если точковая мутация не комплементируется делецией, мутационный сайт должен находиться в пределах делетированной области. Если точковая мутация комплементирует делецию, мутационный сайт должен лежать вне делетированной области. Сравнивая серию делеций по их способности комплементировать точковую мутацию, можно локализовать точковую мутацию.

Вторая трудность касается *жизненно важных* функций. Мутация в таком гене, вероятно, просто убьет организм. Для отбора мутантов по их жизненно важным функциям нужно получить **условно-летальные мутации**. Эти мутации, будучи летальными в одних условиях, либо совсем не влияют на жизнеспособность, либо оказывают

более слабое действие в других условиях, обеспечивающих продолжение жизни клетки. Другими словами, одна и та же мутантная клетка может быть выращена в двух условиях: **пермиссивных**, в которых мутантный фенотип не проявляется, и в **непермиссивных**, в которых она погибает или становится совершенно больной. Тем не менее, если гибель наступает не сразу, такие клетки можно исследовать. (В связи с этим следует отметить, что условные мутации не ограничиваются жизненно важными функциями и могут быть обнаружены в любом гене.)

Рекомендуемая литература

Впервые идея молекулярно-биологического подхода к проблемам генетики была сформулирована в книге Шрёдингера (*Schrödinger E. What is life? Cambridge University Press, Cambridge, England, 1945*). [Имеется перевод: Шрёдингер Е. Что такое жизнь?—М.: ИЛ, 1947.]

Воспоминания, посвященные последующему развитию событий, включая и открытие ДНК, содержатся в книге, вышедшей под редакцией Кейрнса, Стента и Уотсона (*Cairns, Stent, Watson, Phage and the Origins of Molecular Biology, Cold Spring Harbor Laboratory, New York, 1966*). Историческая оценка этого периода дана в книгах Олби (*Olby, The Path to the Double Helix, MacMillan, London, 1974*) и Джэдсона (*Judson, The Eighth Day of Creation (Knopf, New York, 1978)*). Интересное обсуждение последних работ по структуре ДНК можно найти в работе Кантора (*Cantor, Cell, 26, 293–295, 1981*). О перспективах первых работ с мутациями можно прочесть в специальном обзоре Дрейка и Бальза (*Drake, Balz, Ann. Rev. Biochem., 45, 11–37, 1976*). Исследование горячих точек началось с работы Бензера (*Benzer, Proc. Nat. Acad. Sci. USA, 47, 403–416, 1961*), а об их соответствии модифицированным основаниям сообщили Кулондр и др. (*Coulondre et al., Nature, 274, 775–780, 1978*). Подвижные элементы обсуждаются в главах 36–38 данной книги.

Глава 3

ЧТО ТАКОЕ ГЕН?

МОЛЕКУЛЯРНАЯ СТРУКТУРА

Хромосома содержит длинную, непрерывную цепь ДНК, вдоль которой располагается множество генов. Мутации, вносящие изменения в последовательность пар оснований, позволяют распознавать различные участки генетического материала. На карте сцепления хромосом расстояния между мутациями представлены частотой их рекомбинации.

Насколько точно генетическая карта соответствует действительной длине генетического материала? В гигантском масштабе хромосомы генетическая карта достаточно адекватно отражает соответствующую молекулярную структуру. Однако единица карты зависит от относительной частоты рекомбинации, которая может отличаться у каждого вида. Следовательно, генетические

карты разных организмов нельзя сравнивать между собой.

Сопоставление общей длины всех групп сцепления в геноме с содержанием геномной ДНК дает некоторое представление о частоте генетических обменов относительно длины ДНК. Для сравнения можно принять, что 50% ед. карты (50% рекомбинации) соответствуют среднему расстоянию между независимо рекомбинирующими единицами. Значения для различных геномов сведены в **табл. 3.1**.

У бактериофагов генетические обмены совершаются *чрезвычайно* часто (хотя следует оговориться, что количественные значения частоты весьма приблизительны). Кроме того, среди огромного числа потомков можно об-

Таблица 3.1

Частота генетической рекомбинации у эукариот намного меньше, чем у прокариот

Виды	Среднее расстояние между независимо рекомбинирующими единицами (в парах оснований ДНК)	Число пар оснований на 1 ед. карты
Бактериофаг (T4)	$1,0 \cdot 10^4$	200
Бактерия (<i>E. coli</i>)	$1,2 \cdot 10^5$	2 400
Дрожжи (<i>S. cerevisiae</i>)	$2,5 \cdot 10^5$	5 000
Плодовая мушка (<i>D. melanogaster</i>)	$2,5 \cdot 10^7$	500 000
Домовая мышь (<i>M. musculus</i>)	$9,0 \cdot 10^7$	1 800 000

наружить рекомбинантов, возникающих с очень низкой частотой. Эти особенности позволяют проводить детальное внутригенное картирование. В гл. 1 мы говорили о том, что можно получить рекомбинацию между мутациями в соседних парах оснований. У бактерий частота генетических обменов также велика, и также можно обнаружить рекомбинантов среди большого числа потомков, что позволяет проводить детальное внутригенное картирование.

У низших эукариот, таких по крайней мере, как дрожжи, частота рекомбинации остается высокой, но у высших эукариот она резко снижается (в сто раз). Пока еще мы не располагаем достаточной информацией, чтобы оценить диапазон значений частоты рекомбинации для высших эукариот, но на двух примерах (плодовая мушка и домовая мышь) можно проиллюстрировать трудности, возникающие при детальном генетическом картировании. Предположим, что ген плодовой мушки имеет длину в 5000 пар оснований. Это соответствует 0,01% рекомбинации (принимая, что $2,5 \cdot 10^7$ пар оснований соответствует 50% рекомбинации). Таким образом, чтобы обнаружить рекомбинацию между мутациями, локализованными на противоположных концах гена, нужно найти 1 рекомбинантную муху среди 10 000 потомков. При таких условиях внутригенное картирование у плодовой мушки едва ли возможно. В отношении мыши это совсем не практикуется, поскольку частота рекомбинации еще ниже, а количество потомков от одного скрещивания меньше.

Следовательно, картирование на уровне тонкой структуры возможно только у прокариот и, вероятно, у низших эукариот. Таким методом нельзя получить информацию об организации гена у высших эукариот.

Прямые исследования структуры гена

Даже у прокариот информация, получаемая от внутригенного картирования, ограничена природой рекомбинационного акта. На уровне внутригенного картирования частота рекомбинации частично зависит от природы мутаций, использованных в скрещивании, и может в большой степени определяться последовательностью ДНК в данном участке. Другими словами, здесь вместо идеального свойства независимости аллелей, которое мы обсуждали в гл. 1, проявляется эффект специфичности аллелей. Поэтому наши представления о гене с позиций генетической карты искажаются особенностями рекомбинационных систем.

Следовательно, расстояния между мутациями по кар-

те не обязательно соответствуют разделяющему их расстоянию на ДНК. Использование рассмотренного подхода ограничено, кроме того, доступностью мутаций. Пустые промежутки на генетических картах могут на деле соответствовать участкам генома, в которых просто не обнаружено мутаций; кроме того, они могут быть следствием локального увеличения частоты рекомбинаций. Нет основания думать, что у эукариот положение было бы иным, если бы удалось провести внутригенное картирование. Таким образом, и по теоретическим соображениям, и по практическим причинам необходим другой источник информации. Каким образом можно установить молекулярную структуру гена и соотнести ее со структурой белкового продукта? Насколько далеко друг от друга расположены соседние гены? И как мы можем выявить участки, расположенные между ними?

Генетическое картирование ставит своей конечной целью определение нуклеотидной последовательности гена и прилегающих к нему участков вплоть до соседних генов. Для этого нужно прежде всего выделить ДНК, соответствующую данному гену. Сначала это сделали на вирусах, у которых легко выделить всю геномную ДНК. С тех пор, однако, выделение любого типа клеточной ДНК стало почти общедоступной процедурой в том случае, если есть соответствующий белковый продукт (хотя при малом количестве белка выделение специфической ДНК может занять довольно много времени).

Затем следует построить карту ДНК. Для этого ДНК разрезают в определенных точках, расстояния между которыми можно точно измерить. Такие точные разрывы возможны благодаря **рестриктирующим ферментам (ферменты рестрикции, рестриктазы)**, мишенью которых служат короткие специфические последовательности ДНК. Карта ДНК, полученная в результате локализации точек разрыва, называется **рестрикционной картой**. Она представляет собой линейную последовательность сайтов, в которых определенные рестриктирующие ферменты специфически узнают свои мишени. Расстояние между сайтами рестрикции измеряют прямо в нуклеотидных парах ДНК (сокращенно п.н. для коротких отрезков и т. (п.) н.—тысяча (пар) нуклеотидов для длинных отрезков).

После построения физической карты можно перейти к определению последовательности ДНК между близко расположенными точками (обычно около 300 п. н.). Выбирая подходящие точки, отдельные расшифрованные последовательности можно соединить в последовательность целого гена. Затем установленную последовательность ДНК сравнивают с последовательностью соответствующего белка. Это позволяет выявить участки ДНК, отвечающие за структуру белка. Продолжив определение последовательности в одном из направлений, определяют расстояние до следующего гена.

Сравнивая между собой последовательность ДНК дикого типа и его мутантного аллеля, можно точно локализовать мутационный сайт и установить природу мутации. Таким методом определяется соответствие между генетической картой (целиком основанной на мутационных сайтах) и физической картой (составленной на основе нуклеотидной последовательности ДНК). В конечном итоге карту какого-либо участка генома можно выразить в нуклеотидах ДНК, а не в относительных единицах карты, принятых в формальной генетике. Действительно, теперь можно идентифицировать гены по их белковому продукту и даже иногда только по нуклеотидной последовательности. Таким образом, при построении карты генома мы

больше не зависим целиком от мутаций как от необходимого сырья. Однако мутации по-прежнему играют важную роль при идентификации функций генных продуктов.

Рестриктирующие ферменты расщепляют ДНК на специфические фрагменты

Ключом к рестрикционному картированию служат свойства одного класса ферментов, обнаруженных у бактерий (более детально эти ферменты рассматриваются в гл. 34 в связи с их естественным распространением). Мы уже говорили, что рестриктирующие ферменты режут двухцепочечную ДНК в специфических участках. Каждый рестриктирующий фермент имеет свою особую мишень. Обычно это специфическая последовательность нуклеотидов длиной от 4 до 6 пар оснований. Фермент режет ДНК в каждой точке, где встречается такая последовательность. Разные рестриктирующие ферменты имеют различные последовательности-мишени. Сейчас в распоряжении исследователей находится большой набор рестриктаз. (Их применение детально описано в гл. 19.)

Если взять определенную молекулу ДНК и обработать ее подходящим рестриктирующим ферментом, то в строго определенных местах произойдут разрывы, которые разделят ДНК на ряд отдельных фрагментов. Эти фрагменты можно фракционировать по размеру методом **гель-электрофореза**. Для этого препарат разрезанной ДНК наносят сверху на агарозный гель. При наложении электрического поля фрагменты начнут перемещаться вниз по гелю со скоростью, зависящей от их длины. Чем короче фрагмент, тем быстрее он движется. (Пройденное расстояние обратно пропорционально логарифму длины фрагмента.)

В результате электрофореза в геле образуется ряд **полос**. Те полосы, которые располагаются сверху геля, соответствуют большим фрагментам, а те, которые снизу — маленьким. Если гель **прокалибровать**, то можно определить длину любого индивидуального фрагмента ДНК. Для этого на одну из дорожек геля наносят контрольную пробу. Контрольная проба содержит смесь стандартных фрагментов известного размера (их часто называют **маркерами**). Соотношение между размером исследуемого фрагмента и пройденным им расстоянием устанавливают по скорости миграции маркеров.

Пример использования этого метода показан на **рис. 3.1**. Индивидуальную молекулу ДНК длиной 5000 п.н. обрабатывают отдельно двумя рестриктазами — А и В. Разрезанные фрагменты ДНК разделяют с помощью электрофореза. Можно видеть, что фермент А разрезал ДНК-субстрат на четыре фрагмента размером 2100, 1400, 1000 и 500 п.н., тогда как фермент В образовал три фрагмента размером 2500, 1300 и 1200 п.н. Можно ли на основании этих данных расположить сайты рестрикции в определенном порядке и построить рестрикционную карту?

Существует несколько методов сравнения фрагментов, полученных при расщеплении двумя различными рестриктазами. Общая схема одного метода, использующего процедуру двойного расщепления, показана на **рис. 3.2**. Суть этого метода заключается в том, что исследуемая ДНК обрабатывается не одной, а двумя рестриктазами. Наиболее полный вариант такой процедуры со-

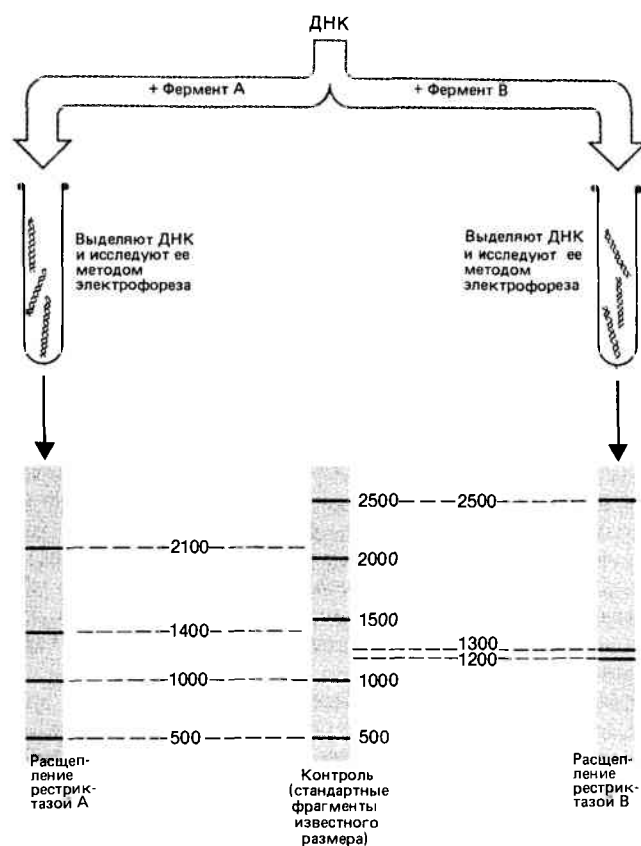


Рис. 3.1. Рестриктирующие ферменты расщепляют ДНК на отдельные фрагменты, которые можно разделить методом электрофореза в агарозном геле.

Размер индивидуальных фрагментов, образованных ферментом А (слева) или ферментом В (справа), определяют сравнением с положением фрагментов известного размера, показанных на контрольной дорожке электрофореза (в центре).

стоит в том, чтобы экстрагировать каждый фрагмент, образующийся в результате расщепления одним ферментом (А или В), а затем обработать его вторым ферментом. Смесь фрагментов, полученную после расщепления, анализируют с помощью электрофореза.

Каждый гель на рис. 3.2 обозначен в соответствии с использованным ферментом и размером фрагментов, которые были экстрагированы из гелей, изображенных на рис. 3.1. Так, «А-2100» означает фрагмент в 2100 п.н., полученный в результате обработки исходной молекулы ДНК ферментом А. После того как этот фрагмент извлекли из геля и обработали рестриктазой В, образовались два фрагмента размером 1900 п.н. и 200 п.н. Другими словами, разрез, сделанный ферментом В, находится на расстоянии 200 п.н. от ближайшего разреза ферментом А с одной стороны и на расстоянии 1900 п.н. — от ближайшего разреза ферментом А — с другой. Относительное положение разрезов видно при анализе чувствительности фрагмента В-2500 к ферменту А. Фрагмент В-2500 разрезается на два фрагмента размером 1900 п.н. и 600 п.н. Таким образом, фрагмент 1900 п.н. образуется в результате двух разрезов — в А-сайте с одного конца и В-сайте с другого. Поскольку этот фрагмент образуется в результате расщепления двух разных фрагментов (А-2100 и В-2500), можно думать, что эти фрагменты **перекрываются** и область размером 1900 п.н. является для них общей.

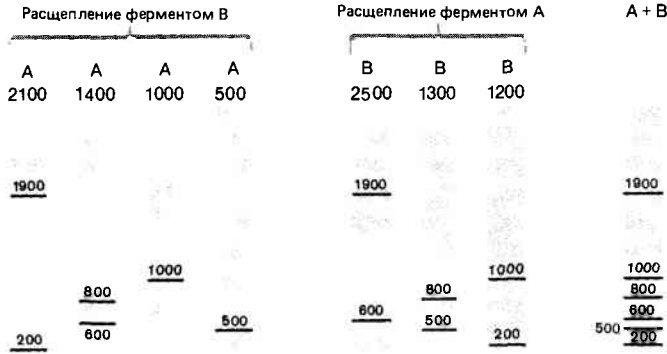


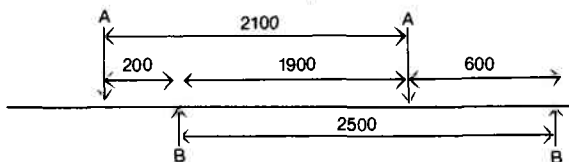
Рис. 3.2. Сравнение результатов обработки ДНК двумя рестриктазами дает возможность определить положение сайтов расщепления одним ферментом по отношению к сайтам расщепления другим ферментом.

Четыре фрагмента, полученные в результате действия рестриктазы А (см. рис. 3.1), элюировали из геля и обрабатывали рестриктазой В (четыре геля слева). Также три полоски, полученные после обработки рестриктазой В, выделили и обработали рестриктазой А (три геля в центре). В контрольной пробе интактную ДНК обрабатывали двумя рестриктазами одновременно (гель справа).

После того как проведен подобный анализ всех фрагментов, становится ясно, что каждый фрагмент, полученный из исходных А-фрагментов под действием фермента В, можно также обнаружить в образцах, содержащих фрагменты, полученные в результате расщепления исходных В-фрагментов ферментом А. В геле, содержащем все продукты двойного расщепления (правый гель на рис. 3.2), каждый фрагмент встречается один раз. Эти данные позволяют однозначно расположить на карте сайты рестрикции.

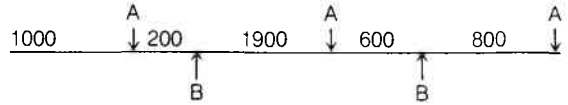
Построение рестрикционной карты

Использование перекрывающихся фрагментов — это ключ к рестрикционному картированию. Рассмотрим для примера уже известный нам фрагмент в 1900 п.н. Он образуется либо в результате обработки фрагмента А-2100 ферментом В, либо после обработки фрагмента В-2500 ферментом А. Из данных о расщеплении фрагментов А-2100 и В-2500 мы знаем, что с одной стороны на расстоянии 200 п.н. от фрагмента 1900 находится следующий А-сайт, а с другого конца, на расстоянии 600 п.н., — следующий В-сайт. На основе этих данных мы можем нарисовать карту, указав на ней вертикальными линиями сайты рестрикции ферментами А и В и размеры фрагментов между ними.



Следует иметь в виду, что при реконструкции больших фрагментов из маленьких всегда получается хорошее совпадение размеров (в пределах разрешающей способности метода). Характерной чертой рестрикционных карт является их полная *аддитивность по длине*.

Теперь мы можем продолжить картирование влево и вправо, определив, из каких фрагментов образуются участки 200 п.н. и 600 п.н. с левой и правой сторон. Фрагмент 200 п.н. получается при расщеплении фрагмента В-1200 рестриктазой А. Отсюда вытекает, что следующий В-сайт должен находиться на расстоянии 1000 п.н. влево. Фрагмент 600 п.н. образуется при расщеплении фрагмента А-1400 рестриктазой В. Тогда следующий А-сайт должен находиться через 800 п.н. вправо. Теперь карта будет выглядеть так:



Чтобы завершить построение карты, нужно установить происхождение двух концевых фрагментов. Слева фрагмент 1000 п.н. происходит либо из фрагмента В-1200, либо из фрагмента А-1000. Поскольку В-фермент не разрезает фрагмента А-1000, этот фрагмент должен находиться с краю карты. Другими словами, начиная с левого конца первый А-сайт расположен через 1000 п.н., а первый В-сайт — через 1200 п.н. По этой причине на карте, построенной нами ранее, мы не обозначили В-сайт слева от фрагмента А-1000 (хотя ранее мы формально рассматривали этот конец как В-сайт).

На правом конце карты фрагмент в 800 п.н. получен после обработки фрагмента В-1300 рестриктазой А. Следовательно, правее его располагается фрагмент в 5000 п.н. Поскольку рестриктаза В не расщепляет фрагмента А-500, он служит правым концом и завершает построение карты.

Все данные, полученные при сравнении специфического действия рестриктаз А и В, сведены в рестрикционную карту на рис. 3.3. Таким образом, молекула ДНК размером в 5000 п.н. поделена на серии участков определенной длины, расположенные между участками, специфически узнаваемыми двумя рестриктирующими ферментами.

Некоторые тонкости рестрикционного картирования

При построении рестрикционных карт обычно используют несколько рестриктаз, поэтому приходится анализировать сложные соотношения между фрагментами, полученными под действием различных ферментов. Процедуру картирования можно значительно упростить с помощью специальных приемов.

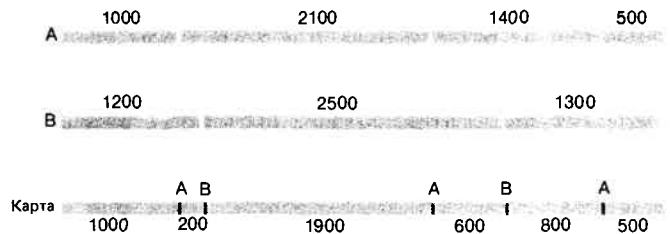


Рис. 3.3. Рестрикционная карта ДНК представляет собой линейную последовательность сайтов расщепления, находящихся на определенном расстоянии друг от друга. На карте показано расположение сайтов расщепления ферментами А и В, которые были определены в результате анализа.

Один из таких приемов заключается в **неполном** расщеплении. В определенных условиях эксперимента рестриктаза узнает и расщепляет не все свои мишени в каждой молекуле ДНК. Например (рис. 3.3), при частичном расщеплении ДНК ферментом A могут образоваться фрагменты 3100 п.н., 1400 п.н. и 500 п.н. Сопоставив их с тем, что получается после полного расщепления, можно сразу расположить фрагменты 2100 п.н. и 1100 п.н. рядом, как составляющие фрагмента 3100 п.н. Таким путем можно последовательно расположить сайты рестрикции одного фермента. При неполной рестрикции можно также получить фрагмент 3500 п.н.; тогда можно расположить рядом фрагменты 2100 п.н. и 1400 п.н.

Другой удобный прием — это введение **концевой метки** с помощью радиоактивного фосфора. Так называют ферментативное присоединение радиоактивной метки (^{32}P) к 5'- или 3'-концу молекулы ДНК. В этом случае концевые фрагменты определяются по включенной концевой метке. Таким образом, фрагменты A-1000 и A-500, а также B-1200 и B-1300 можно сразу расположить по краям карты.

Сопоставление фрагментов часто производят с помощью метода гибридизации нуклеиновых кислот. Например, фрагменты A-2100 и B-2500 содержат перекрывающиеся участки ДНК. Следовательно, они будут гибридизоваться друг с другом.

Для применения всех этих методов нужно иметь в распоряжении полный набор фрагментов, из которых должна быть построена карта. Однако иногда это условие невозможно выполнить. Если несколько специфических сайтов одного фермента находятся слишком близко друг к другу (например, на расстоянии 50 п.н.), образуются мелкие фрагменты, которые нельзя обнаружить при электрофорезе в агарозном геле. Это, конечно, приведет к расхождению при суммировании полученных значений молекулярных масс всех фрагментов. Однако расхождение в молекулярных массах само по себе не обязательно должно настораживать, так как всегда существует некоторое расхождение при сравнении размеров фрагментов.

Комбинируя методы концевой метки и неполного расщепления, можно определить расположение всех фрагментов, полученных под действием одной рестриктазы. Сочетая этот метод с методом перекрывающихся фрагментов, а также используя несколько рестриктаз, можно быть уверенным, что построенная рестрикционная карта включает всю ДНК и ни один из мелких фрагментов не потерян.

Дополнительные трудности возникают из-за того, что не все индивидуальные фрагменты можно увидеть в геле в виде отдельной полосы. Так, иногда образуются фрагменты с одинаковой молекулярной массой (или гораздо чаще с настолько близкими размерами, что при электрофорезе они не разделяются). Для того чтобы убедиться в уникальности каждой полосы, следует обратить внимание на интенсивность ее окраски. Полоски, состоящие из нескольких фрагментов, должны быть более интенсивно окрашены. Фрагменты с одинаковой молекулярной массой можно различить, используя дополнительные рестриктазы.

Подробную рестрикционную карту можно связать с генетической картой. Крупные генетические изменения, такие, как делеции или вставки, можно обнаружить методом рестрикционного картирования. Они приведут к уменьшению или увеличению соответствующих рестрикционных фрагментов. Кроме того, иногда в результате генетиче-

ских перестроек исчезают некоторые сайты рестрикции или же возникают новые.

В заключение следует сказать о номенклатуре рестриктирующих ферментов. Рестриктазы обозначают сокращениями из трех, а иногда из четырех букв, соответствующих их происхождению. К сокращению добавляют римские цифры, чтобы различить ферменты одинакового происхождения. Например, HpaI — рестриктаза из *Haemophilus parainfluenzae*, а EcoRI — фермент из *E. coli*.

Сайты рестрикции можно использовать в качестве генетических маркеров

Генетическая и рестрикционная карты характеризуют один и тот же материал с разных точек зрения. Однако, сопоставив карты обоих типов, можно установить физическую основу генетической карты. Кроме того, результаты рестрикционного расщепления можно объединить с генетической картой.

Рестрикционная карта представляет собой серию сайтов линейно расположенных в ДНК, расстояние между которыми соответствует реальному расстоянию вдоль нуклеиновой кислоты. *Рестрикционную карту можно получить для любой последовательности ДНК, независимо от того, идентифицированы ли на ней мутации и известны ли что-нибудь о ее функции.*

Генетическая карта представляет собой серию сайтов, в которых встречаются мутации. *Выявление этих сайтов основано на том, что изменение в последовательности оснований вызывает соответствующее изменение в фенотипе. Расстояние между этими сайтами определяется по частоте рекомбинации, которая зависит от ряда параметров в каждом конкретном участке.*

Чтобы сопоставить рестрикционную и генетическую карты, нужно сравнить рестрикционные карты ДНК дикого типа и соответствующих мутантов. Связь между генетической и физической картами можно установить на основе одного рестрикционного картирования, если имеются мутанты, сильно изменяющие генетическую карту. Например, маленькая делеция или вставка, нарушающая генетическую карту, вызовет соответственно уменьшение или увеличение размера того рестрикционного фрагмента, в котором она находится. Более протяженная делеция (или вставка) может удалять (или вводить) серии рестрикционных сайтов. Такой пример разобран на рис. 3.4.

Точковые мутации труднее локализовать на рестрикционной карте. Иногда они могут изменить сайты-мишени для рестриктирующего фермента. В противном же случае их не удастся определить, так как размеры рестрикционных фрагментов ДНК дикого типа и мутантов оказываются одинаковыми. Поэтому для локализации таких замен оснований часто необходимо определить последовательность оснований в ДНК.

Построение генетической карты основано на генетической неоднородности популяции. Сосуществование более, чем одного варианта называют **генетическим полиморфизмом**. Любой сайт, в котором существуют множественные аллели в качестве стабильных компонентов популяции, считается полиморфным. Например, плодовая мушка *D. melanogaster* полиморфна по серии аллелей локуса *white*, w^+ , w^1 , w^a и т.п.

Полиморфизм на уровне фенотипа объясняется одно-временным существованием в одной популяции мух алле-

ля дикого типа и серии мутантных аллелей. Рассмотрим основу такого полиморфизма среди мутантных аллелей. Эти мутации изменяют продукт гена таким образом, что функция данного белка оказывается неполноценной, что приводит к изменению фенотипа. Если мы сравним рестрикционные карты или последовательности ДНК соответствующих аллелей, то они тоже окажутся полиморфными в том смысле, что каждая карта или последовательность будет отличаться от другой.

Хотя это никак не проявляется в фенотипе, но дикий тип сам по себе может быть полиморфным. Может иметь место существование измененных вариантов аллеля дикого типа, различающихся по таким изменениям в последовательности ДНК, которые нарушают функции белка и поэтому не обнаруживаются в виде фенотипических вариантов. С точки зрения генотипа популяция мух может иметь значительный полиморфизм. Может существовать много вариантов с различной последовательностью ДНК в и-локусе. Некоторые из них заметны, поскольку они приводят к изменению фенотипа; другие скрыты, так как они не имеют видимого проявления.

Существование полиморфизма в геноме можно установить, сравнивая рестрикционные карты разных индивидуумов. Критерием полиморфизма в этом случае служит изменение в характере расположения фрагментов, образованных в результате расщепления рестриктирующим ферментом. На рис. 3.5 показано, что если в геноме одного индивидуума имеется сайт-мишень для рестриктирующего фермента, отсутствующий у другого, то в результате дополнительного расщепления в первом геноме образуется два фрагмента, соответствующих в совокупности одному целому фрагменту во втором геноме.

Поскольку рестрикционная карта не зависит от функции гена, полиморфизм на этом уровне можно обнаружить независимо от того, приводит ли изменение нуклеотидной последовательности к изменению фенотипа или нет. В действительности только незначительная часть случаев рестрикционного полиморфизма в геноме, по-видимому, проявляется в виде фенотипических изменений. Возможно, что в большинстве случаев имеет место такое изменение последовательности нуклеотидов, которое не

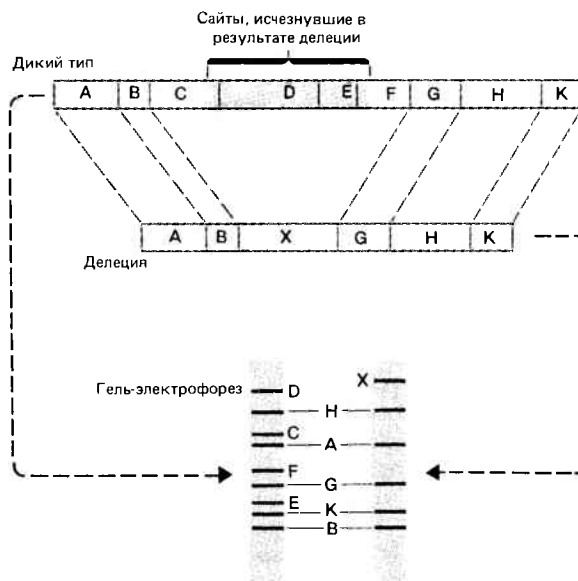


Рис. 3.4. В результате делеции исчезает серия рестрикционных сайтов, а рестрикционные фрагменты, расположенные по ее концам, сближаются.

Делеция целиком удаляет фрагменты D и E, которые имеются среди рестрикционных фрагментов дикого типа, но не у мутанта, несущего делецию. Фрагменты C и F дикого типа также нет у мутанта с делецией. Однако вместо них у мутанта есть новый фрагмент X, который содержит часть последовательности этих двух фрагментов. Фрагменты, расположенные по обоим концам делеции, A и B слева и G, H и K справа, одинаково представлены как в диком, так и в мутантном рестрикционных переварах.

В случае вставки эффект будет прямо противоположным; у мутанта будут присутствовать дополнительные фрагменты, которых нет у дикого типа.

влияет на образование белков (например, из-за того, что такие сайты расположены между генами).

Различие в рестрикционных картах двух индивидуумов можно использовать в качестве генетических маркеров точно так же, как и любые другие маркеры. Единственное отличие состоит в том, что вместо того, чтобы изучать некий фенотипический признак, исследуют непо-

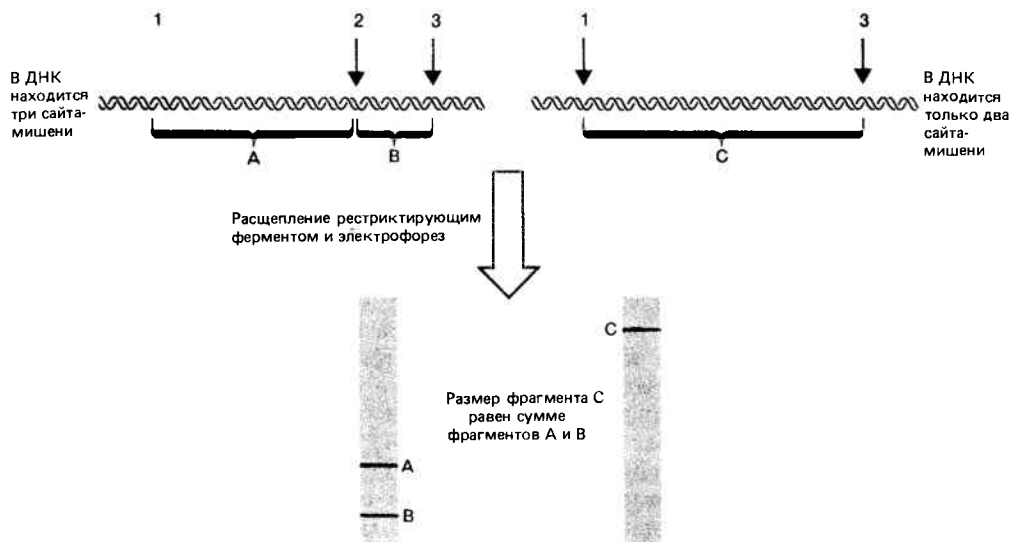


Рис. 3.5. Изменения в ДНК, нарушающие сайты рестрикции, определяют по образованию других фрагментов рестрикции.

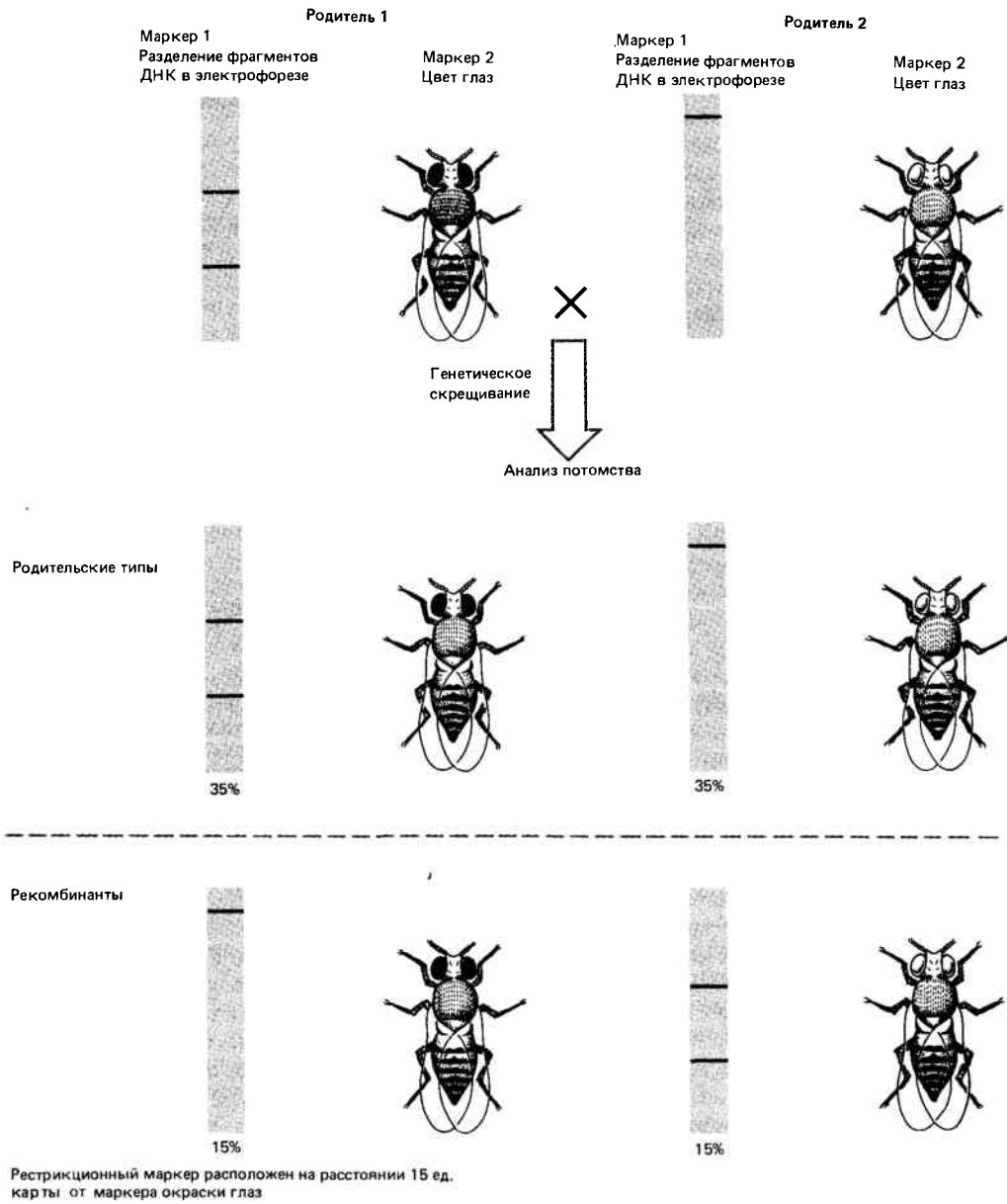


Рис. 3.6. Рестрикционный полиморфизм можно использовать в качестве генетического маркера для измерения рекомбина-

ционного расстояния по отношению к маркеру, выражающемуся в фенотипе (такому как окраска глаз).

средственно генотип по его рестрикционной карте. На рис. 3.6 показано измерение частоты рекомбинации между рестрикционным и фенотипическим маркерами.

Тот факт, что рестрикционные маркеры сохраняются при изменениях генома, затрагивающих фенотип, лежит в основе чрезвычайно эффективного метода идентификации генетических локусов на молекулярном уровне. Типичным примером могут служить мутации с известным фенотипическим эффектом, которые локализованы на генетической карте, хотя функция соответствующего гена или белка не известна. К этой категории относятся некоторые тяжелые заболевания человека. Это, например, кистозный фиброз, хорea Гентингтона и многие другие серьезные и даже смертельные болезни, которые наследуются по законам Менделя. Во всех этих случаях молекулярная природа мутантной функции неизвестна и, вероятно, она сможет быть выяснена только после того,

как удастся установить и охарактеризовать соответствующие гены.

Если рестрикционные полиморфные сайты распределены в геноме случайно, некоторые из них окажутся около интересующего нас гена-мишени. Такие рестрикционные маркеры можно идентифицировать по их тесному сцеплению с мутантным фенотипом. Если мы сравним рестрикционные карты ДНК у больных и здоровых людей, то может оказаться, что у больных определенный рестрикционный сайт всегда присутствует (или всегда отсутствует). Гипотетический пример изображен на рис. 3.7. В этом случае сцепление рестрикционного маркера с фенотипом составляет 100%. Это означает, что рестрикционный маркер находится так близко от мутантного гена, что никогда не отделяется от него при рекомбинации.

Идентификация такого маркера может иметь два важных последствия. Это может привести к выделению

гена. И это может привести к разработке диагностического метода определения данной болезни.

В том случае, если два локуса рекомбинируют редко или не рекомбинируют совсем, рестрикционный маркер должен находиться на генетической карте вблизи от данного генома. Хотя понятие «вблизи» в генетическом смысле может оказаться достаточно большим расстоянием в парах нуклеотидов ДНК (табл. 3.1), тем не менее это может служить отправной точкой, от которой можно проследить последовательность ДНК до самого гена (гл. 17).

Некоторые болезни человека, которые хорошо изучены генетически, но слабо охарактеризованы с молекулярной точки зрения, трудно поддаются диагнозу. Если же рестрикционный маркер тесно сцеплен с фенотипом, то его наличие можно использовать для диагностики болезни в пренатальной стадии или позднее.

Естественно, что большой размер генома человека значительно осложняет определение индивидуального полиморфизма. На практике оказывается даже трудно установить, какую часть генома следует изучать. Однако методы манипуляции с ДНК получили такое мощное развитие (гл. 17), что удалось достичь значительного прогресса в идентификации генов для нескольких болезней человека.

Определение нуклеотидной последовательности ДНК

Мы уже говорили о том, что благодаря специфичности спаривания оснований (G—C и A—T) двойная спираль ДНК сохраняет постоянную структуру независимо от последовательности нуклеотидных пар. Это означает, что индивидуальные молекулы ДНК различаются последовательностью нуклеотидов, а не значительными изменениями в структуре. (В противоположность белкам, у которых общая структура белка зависит от последовательности аминокислот, но именно эта общая структура в целом определяет биологическую функцию.) Совсем небольшие различия в нуклеотидной последовательности ДНК могут иметь очень важное значение, поскольку замена одной пары оснований вызывает мутацию.

Можно ли точно узнать нуклеотидную последовательность ДНК? Первые попытки определения нуклеотидной последовательности были повторением метода определения аминокислотной последовательности белков: расщепление молекулы на мелкие фрагменты, выяснение нуклеотидного состава во фрагментах и (на основе данных о перекрывающихся фрагментах) восстановление полной последовательности. Однако в случае белков дело обстояло значительно проще. Наличие в белковой молекуле 20 аминокислот обуславливает большое разнообразие фрагментов, тогда как в состав молекулы ДНК входит всего четыре основания. Поэтому описанный метод можно было использовать только для очень коротких фрагментов нуклеиновых кислот.

Однако сейчас стало возможным прямо определять последовательность ДНК. Более того, эта процедура проще, чем определение белковой последовательности.

Все используемые методы определения нуклеотидной последовательности сводятся к тому, чтобы получить серию одноцепочечных молекул ДНК, различающихся по размеру на одно основание. Такие молекулы можно разделить с помощью электрофореза в акриламидном геле. Полосы, соответствующие молекулам различной длины, образуют «лестницу» длиной до 300 нуклеотидов.

Для получения таких наборов различающихся фрагментов используют в основном два подхода. В одном случае ДНК химически расщепляют по одному основанию. Другой подход заключается в том, чтобы синтезировать нужную последовательность ДНК *in vitro*, специфически останавливая синтез на заданном основании.

Для определения нуклеотидной последовательности молекулы ДНК ее обрабатывают четырьмя разными способами, каждый из которых специфически действует только на один тип оснований. Продукты расщепления подвергают электрофорезу, помещая их на четыре соседние дорожки одного геля. Полоски, соответствующие молекулам определенного размера, появляются только на одной из четырех дорожек. Это позволяет определить нуклеотид, находящийся в данном положении на ДНК. Переходя от одной полоски к полоске большего размера, можно «прочитать» всю последовательность нуклеотидов.

При использовании химической обработки ДНК сна-

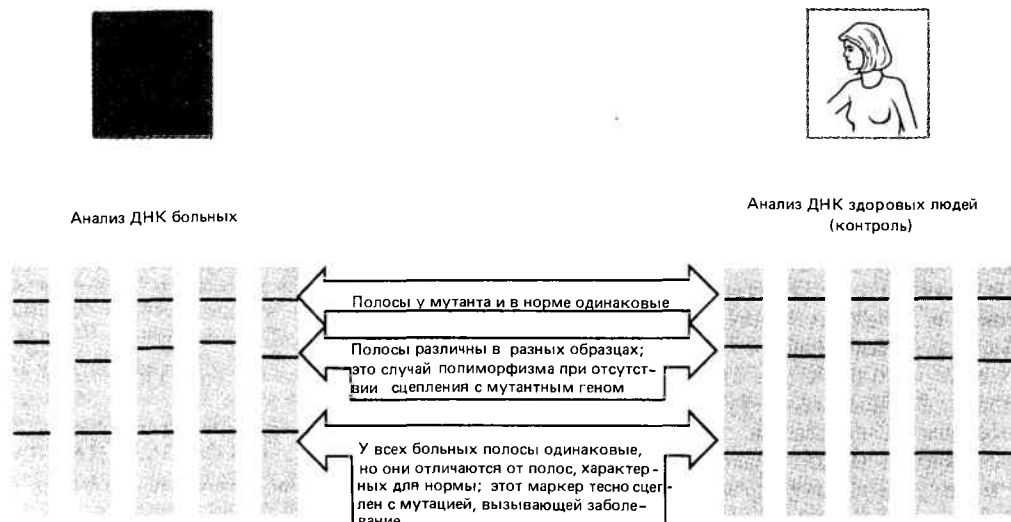


Рис. 3.7. Если рестрикционный маркер всегда (или обычно) связан с фенотипическим признаком, то этот сайт рестрикции дол-

жен находиться рядом с геном, отвечающим за данный фенотип.

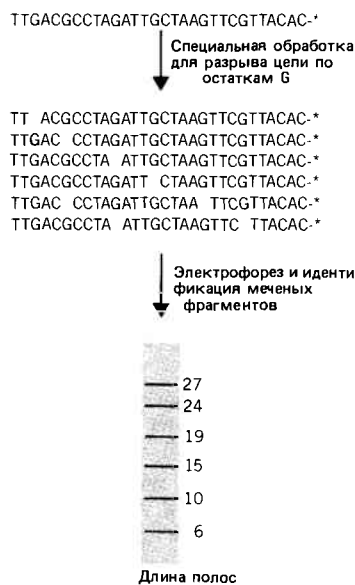


Рис. 3.8. Положение отдельного основания в одноцепочечных фрагментах ДНК можно определить с помощью электрофореза, пометив концы фрагментов радиоактивной меткой и обработав ДНК соответствующим реагентом.

В качестве примера показан опыт, где полинуклеотид был разорван по остаткам G. Некоторые молекулы имели только один разрыв, другие — несколько. Однако в каждом случае один конец анализируемых молекул был помечен (показано звездочкой). Набор меченых фрагментов в условиях этого опыта состоял из молекул, размер которых соответствовал расстоянию от одного конца до каждого положения G в исходной ДНК. При разделении в электрофорезе фрагменты разной длины образуют полосы, положение которых в геле зависит от их длины.

чала исследуемый фрагмент метят с одного конца радиоактивным фосфором и затем специфически расщепляют полинуклеотид по одному из четырех оснований. Суть этого метода заключается в том, чтобы расщепление происходило случайно в каждом одном из чувствительных положений с вероятностью примерно 1–2%.

Таким образом, сначала препарат ДНК содержит большое число идентичных молекул, меченных с одного конца. После химической обработки каждая молекула будет разрезана только в небольшом числе участков, содержащих данное основание. Но весь препарат в целом будет состоять из набора молекул, у которых разрыв случайно произошел практически в каждом положении данного основания (рис. 3.8).

Полученные молекулы подвергают электрофорезу и радиоактивно меченные фрагменты идентифицируют радиоавтоматически. Каждый фрагмент образует полосу, размер которой зависит от расстояния места разрыва до меченого конца. В результате каждого типа химической обработки образуется целая серия полосок, указывающих на положение всех соответствующих оснований по отношению к меченому концу молекулы.

В исходном методе специфическое расщепление по пуринам производили, используя диметилсульфат. Это соединение метилирует N7-положение гуанина примерно в 5 раз эффективнее, чем N3-положение аденина. При нагревании метилированное основание удаляется из ДНК, образуя разрыв в полинуклеотидной цепи. Поскольку реакция происходит более эффективно с гуанином, остатки G образуют темные полоски, тогда как остатки A — светлые полоски (что соответствует более редкому рас-

щеплению). Противоположная картина будет наблюдаться в том случае, если для удаления метилированных оснований вместо нагревания использовать обработку кислотой. Это дает возможность определить местоположение A и G по отдельности.

Для анализа пиримидинов используют гидразин, который эффективно реагирует с цитозином и тиминном. Однако при высокой концентрации соли реакция идет только с цитозином. В итоге можно получить две серии полос, одна из которых будет соответствовать только C, а другая — C + T.

Как показано на рис. 3.9, четыре реакционные смеси подвергают дальнейшему анализу. Сначала проводят электрофорез на параллельных дорожках. На каждой из четырех дорожек образуются полосы, соответствующие молекулам, специфически разорванным в местах расположения одного из четырех оснований. Последовательность ДНК читают от полоски к полоске по мере увеличения размера на одно основание. Поскольку внизу геля находятся самые короткие фрагменты, последовательность читают снизу вверх.

Насколько точно можно таким способом определить последовательность ДНК? Как мы уже говорили, ряд химических реакций позволяет определить последователь-

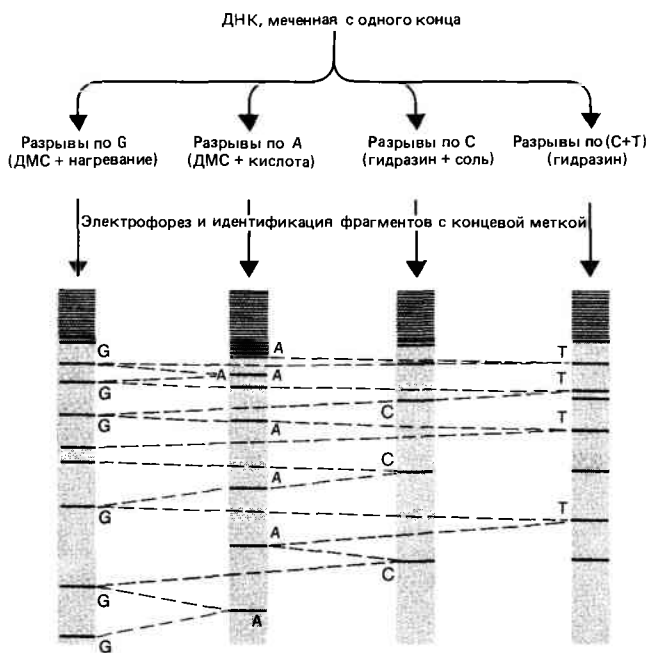


Рис. 3.9. Используя набор реакций, специфичных для каждого отдельного основания, можно определить последовательность ДНК.

Одну и ту же ДНК использовали в четырех независимых реакциях и затем реакционную смесь разделили в электрофорезе. Последовательность читают сразу по всем четырем дорожкам геля, переходя снизу вверх от полоски к полоске, как показано зигзагом. Положение G, A и C определяют по трем левым дорожкам. На правой дорожке находится сумма C и T, поэтому разница третьей (C) и четвертой (C + T) дорожек позволяет определить положение T. Таким образом, последовательность ДНК, изображенную на рисунке, можно прочитать:

конец — N₂₀ — GAGCATGACGGTAGCTAGAGTA...

В самой нижней части геля обычно находятся фрагменты размером около 20 п.н., поэтому чтение последовательности начинается с этого расстояния от меченого конца фрагмента. Ближе к верхней части геля полосы располагаются теснее, и последовательность уже нельзя прочитать.

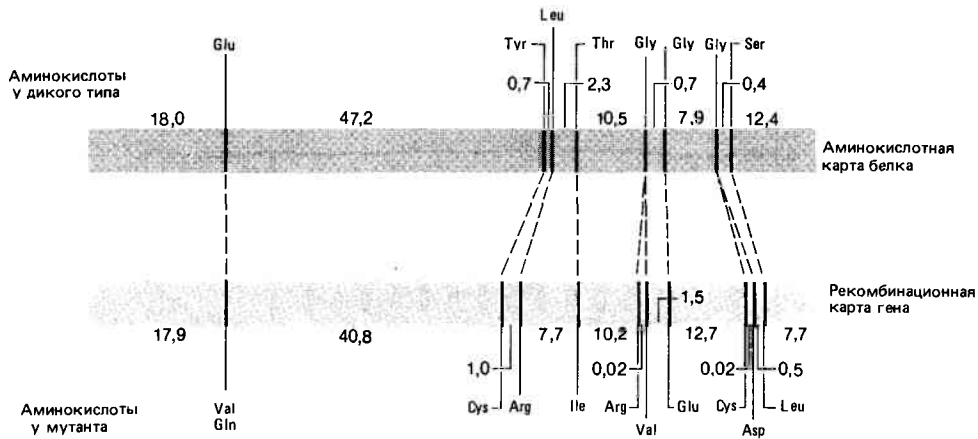
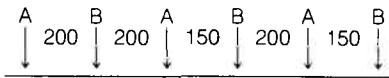


Рис. 3.10. Рекомбинационная карта гена триптофан-синтазы соответствует аминокислотной последовательности этого белка. Вертикальными черточками показаны мутационные сайты, определенные по аминокислотным заменам в белке и по картированию мутаций в гене. Расстояние между черточками указывает на расстояние между мутациями на карте (выраженное в процентах рекомбинации). Общий размер гена составляет 3,9 ед. карты, что соответствует 268 аминокислотам данного белка. Следовательно, одна аминокислота в сред-

нем соответствует 0,015 ед. карты. (При анализе индивидуальных пар мутаций наблюдаются отклонения от этой величины.) На нижней схеме можно видеть, что в двух случаях получены мутации, разделяемые при рекомбинации, которые вызывают замену одной и той же аминокислоты на верхней схеме (показано расходящимися пунктирными линиями). В одном случае глицин заменен на аргинин или валин, в другом — на цистеин и аспарагиновую кислоту. Формально это указывает на то, что единица рекомбинации (одна пара оснований) меньше единицы, кодирующей данную аминокислоту (три пары оснований).

ность одной цепи ДНК. Конечно, всегда существует вероятность ошибки. Однако для повышения точности анализа можно независимо проанализировать обе цепи. Выявляют все участки, в которых не наблюдается комплементации и в которых, следовательно, возможны ошибки, и их уточняют. При таком методе определения ошибки очень редки.

До сих пор мы рассматривали определение последовательностей в случае определенных, но коротких фрагментов ДНК. А как перейти к определению нуклеотидной последовательности длинных молекул? Так же как и при рестрикционном картировании, метод состоит в использовании перекрывающихся фрагментов. Предположим, что у нас есть серия последовательных фрагментов, расположенных следующим образом:



Мы можем прямо определить последовательность каждого из этих фрагментов, но, для того чтобы быть уверенным, что мы не потеряли мелких фрагментов, образованных в результате расщепления ферментом двух близко расположенных участков (проблема, обсуждавшаяся в предыдущем разделе), важно продолжить определение последовательности *через места соединения фрагментов*. Для этого нужен еще один набор фрагментов, в котором оставались бы интактными места соединения А—В. И так же, как можно построить рестрикционную карту перекрывающихся фрагментов, можно построить и общую последовательность ДНК, анализируя перекрывание отдельных последовательностей ДНК.

Колинеарны ли гены и белки?

Представление о том, что каждая цепь ДНК служит матрицей при синтезе комплементарной цепи (при репликации), существенно и для определения природы гена, и для установления его связи с процессом белкового синтеза. Каждый ген экспрессируется путем образования

РНК-посредника, которая служит матрицей при синтезе белка (см. гл. 5). Здесь важно запомнить, что последовательность нуклеотидов в матричной РНК точно соответствует последовательности аминокислот в белке. У бактерий последовательность нуклеотидов в матричной РНК (мРНК) точно соответствует последовательности одной из цепей ДНК и комплементарна другой цепи — той, которая направляет ее синтез с помощью того же механизма, который используется при репликации, — комплементарного спаривания оснований.

В двух экспериментах, выполненных на бактериях, было формально показано, что ген и его продукт **колинеарны**; иными словами, последовательность нуклеотидов в гене точно соответствует последовательности аминокислот в белке. Такой ход мысли аналогичен тому, что на основании однотипности частот рекомбинации внутри одного гена и между генами делается вывод о том, что структура хромосомы в генах и межгенных промежутках однотипна. Все это дает основание рассматривать бактериальный геном как длинную молекулу ДНК, последовательные участки которой детерминируют различные белки.

В первом эксперименте сравнивали расстояние между мутационными сайтами с расстоянием между аминокислотными заменами в белке. Расстояния на генетической карте измеряли в единицах рекомбинации, а расстояния на белке — по числу аминокислот, разделяющих аминокислотные сайты. На **рис. 3.10** показано, что порядок семи мутационных сайтов соответствует порядку семи аминокислотных замен. Расстояние на рекомбинационной карте между мутациями относительно совпадало (как это не удивительно) с действительным расстоянием между аминокислотными заменами в молекуле белка.

В другом эксперименте использовали мутации, ведущие к преждевременной остановке синтеза белка. Если ген и белок колинеарны, длина фрагмента белка, синтезированного у мутанта, будет соответствовать расстоянию от начала гена до данной мутации. В каждом исследованном случае обнаруживали начало белка, но конец

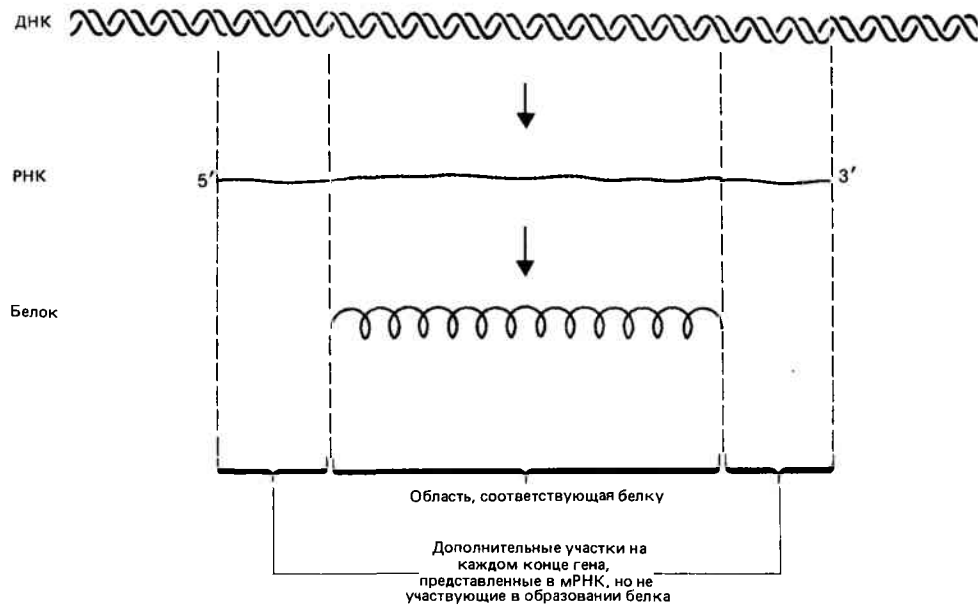


Рис. 3.11. Размер гена больше последовательности, кодирующей белок.

белковой молекулы отсутствовал. Фрагменты белков, синтезированные в каждом мутанте, расположили в порядке увеличения их длины. Оказалось, что по мере удаления точки мутации от начала гена увеличивается размер синтезированного фрагмента белка.

Сейчас эксперименты такого типа кажутся несовершенными по сравнению с возможностью прямо определять нуклеотидную последовательность ДНК. Если сравнить последовательность ДНК у бактерий с последовательностью соответствующего белка, то обнаруживается полное соответствие. Каждый ген состоит из непрерывного отрезка ДНК, размер которого точно соотносится с размером определяемого им белка.

В действительности участок ДНК обычно бывает длиннее, чем это нужно для образования белка, так как мРНК может содержать на обоих концах некоторые дополнительные последовательности, не используемые для синтеза белка. Обычно под «геном» подразумевают всю последовательность, копируемую в мРНК. Ген, который экспрессируется как целая единица, изображен на рис. 3.11. (Более сложные случаи, когда несколько генов экспрессируются как части единого целого, рассмотрены в гл. 9). Несмотря на все сделанные оговорки, мы можем дать следующее определение гена: ген — это такая последовательность ДНК, участвующая в образовании белка, в которой участок, точно соответствующий белку, является непрерывным.

Гены эукариот могут быть прерывистыми

Мы уже говорили о том, что группа комплементации всегда соответствует набору мутаций, лежащих в непрерывном отрезке ДНК, прямо соответствующем белковому продукту. Однако правильность этого простого представления была поставлена под сомнение после того,

как в 1977 г. были обнаружены **прерывистые гены**. Первые данные о наличии таких прерывистых генов были получены в результате сравнительного анализа структуры ДНК и соответствующей информационной РНК. Оказалось, что внутри гена (в определении, сделанном выше) имеются последовательности, которых нет в мРНК. Это явление обнаружено только у эукариот.

Ген состоит из последовательностей ДНК двух типов (рис. 3.12). Одни — **экзоны** — это участки, которые транскрибируются в мРНК, используемую для синтеза соответствующего белка. Другие — **интроны** — это участки, транскрипты которых не обнаруживаются в зрелой мРНК. Таким образом, в процессе экспрессии гена появилась новая ступень, которой нет у бактерий. С ДНК считывается копия РНК, точно соответствующая последовательности генома. Но это только РНК-предшественник (про-мРНК); она непосредственно не используется для синтеза белка. Только после удаления из этой РНК интронов образуется зрелая мРНК, состоящая из одних только экзонов. Этот процесс называют **сплайсингом РНК**.

Как это меняет наши представления о гене? Экзоны всегда соединены вместе в том же порядке, в котором они находятся в ДНК. Это означает, что между индивидуальными экзонами и соответствующими участками полипептидной цепи сохраняется коллинеарность гена и белка.

Порядок, в котором расположены мутации в гене, оказывается таким же, как и порядок аминокислотных замен в соответствующем белке. Однако *расстояния* между мутациями в гене и аминокислотными заменами в белке не совпадают.

Последствия такой организации генов для построения групп комплементации могут быть серьезными. Но это зависит от функции интронов. Мутации, нарушающие синтез белка, должны располагаться в гене в виде кластеров, каждый из которых соответствует отдельному экзо-

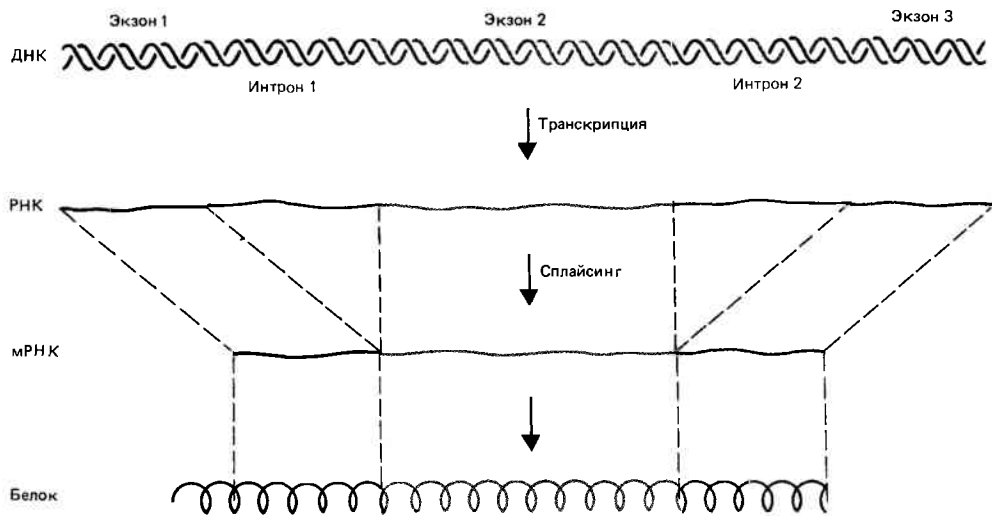


Рис. 3.12. Прерывистые гены экспрессируются в виде РНК-предшественника, из которой затем удаляются интроны; эк-

зоны, соединяясь вместе, образуют мРНК.

ну. (Раньше этого не замечали по той причине, что у высших организмов вообще трудно картировать большое число мутаций.) Ни одна из этих мутаций не может комплементировать другую, так как все они локализируются в одной молекуле РНК, и соединение экзонов вместе при сплайсинге является *внутримолекулярной* реакцией. Экзоны *различных* молекул РНК не соединяются вместе. Это исключает возможность сплайсинга последовательностей, соответствующих различным аллелям. Таким образом, существование сплайсинга не изменяет определения группы комплементации.

Какова природа мутаций в интронах? Мутации в интронах не могут прямо нарушать структуру белка, так как интроны не входят в состав мРНК; они могут препятствовать образованию мРНК — например подавляя сплайсинг. Такой эффект интронных мутаций ограничивается только определенными аллелями, которые не комплементируют другие интронные мутации и должны относиться к той же группе комплементации, что и экзоны.

Но если окажется, что интрон кодирует какой-то белок, тогда мутация, нарушающая функцию этого белка, будет относиться к независимой группе комплементации. В этом случае мутации в экзонах образуют серию кластеров, относящихся к одной группе комплементации. А между двумя такими кластерами будет находиться ряд мутаций, относящихся к другой группе комплементации.

В прерывистых генах эукариотического генома большинство интронов, по-видимому, не способно выполнять какую-либо независимую функцию в синтезе белка. Таким образом, случаи смешивания различных групп комплементации не могут быть обычным явлением. Тем не менее такое имеет место в митохондриях, и, следовательно, рассматриваемую возможность нельзя считать всего лишь гипотезой.

Не все гены эукариот прерывисты. Некоторые гены, подобно бактериальным генам, точно соответствуют своему белковому продукту. Пока еще нет четких представлений о соотношении прерывистых и непрерывных генов в геноме эукариот, хотя, по мере того как накапли-

ваются данные, создается впечатление, что генов первого типа больше.

Перекрывающиеся и альтернативные гены

Известно несколько случаев, когда отдельная последовательность ДНК участвует в кодировании более чем одного белка. При дальнейшем изложении материала мы рассмотрим такие случаи более детально. Пока же нам следует познакомиться с тем, какого рода взаимоотношения возможны между белком и ДНК.

Перекрывающиеся гены встречаются в относительно простых ситуациях, когда один ген является одновременно частью другого гена. Например, первая (или вторая) половина гена может независимо кодировать белок, представляющий собой часть (первую или вторую половину) белка, кодируемого целым геном. Здесь нет особой *генетической* проблемы (хотя требуется некоторая корректировка в процессе синтеза белка; см. гл. 7). Аналогичный результат можно получить при частичном расщеплении уже синтезированного белка. Такие ситуации встречаются в вирусных и бактериальных геномах.

В некоторых случаях перекрывание бывает более сложного типа. У вирусов прокариот и эукариот обнаружены гены, у которых одна и та же последовательность ДНК принимает участие в синтезе двух *негомологичных* белков. Как мы увидим в гл. 4, это происходит в результате считывания одной и той же последовательности в двух различных рамках. Мутация в перекрывающейся области в зависимости от ее типа может нарушать функцию одного или обоих генов и входить сразу в две группы комплементации.

В обычном прерывистом гене каждый экзон кодирует отдельную аминокислотную последовательность, представляющую часть белка, и ни один интрон не участвует в образовании белка. Роль этих участков гена различна. Однако в некоторых случаях какая-то последователь-

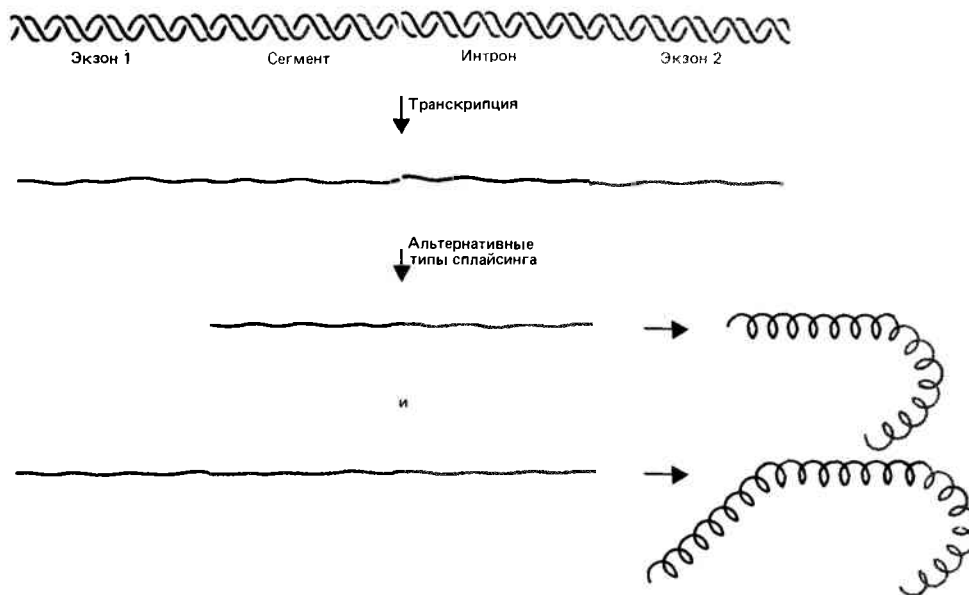


Рис. 3.13. Гены, имеющие альтернативные пути экспрессии, бывает трудно разграничить.

В первом случае экзон 1 и экзон 2 соединяются вместе, а участки, обозначенные «сегмент» и «интрон», удаляются из РНК. Во втором случае соединяются экзон 1, сегмент и экзон 2, а удаляется только интрон. Та-

ким образом, область, обозначенная «сегмент», служит интроном в первом случае и экзоном – во втором. В результате этих двух типов сплайсинга образуются два белка, которые имеют одинаковые концы, но различаются по средней части. Так одна последовательность ДНК может кодировать более одного белка.

ность ДНК может участвовать в альтернативных процессах, и ее нельзя просто охарактеризовать как экзон или интрон. У вирусов высших организмов, а иногда и в самих эукариотических клетках, обнаружены *альтернативные* типы экспрессии генов. В этих случаях возможно переключение в процессе сплайсинга. При образовании мРНК один конкретный экзон может быть соединен с любым из нескольких других экзонов. В результате образуются различные белки, у которых одна часть общая, а другая варьирует. Такой пример рассматривается на рис. 3.13, где показано, что некоторые участки про-мРНК в одном случае ведут себя как экзоны, а при другом типе сплайсинга оказываются интронами. У мутаций в таких участках – очень сложные комплементарные свойства, и не всегда возможно отнести их к какой-либо независимой группе комплементации.

Что такое ген?

Вопрос о том, что подразумевается под данным названием, особенно уместен в отношении гена. Очевидно, что больше нельзя говорить о гене как о непрерывной последовательности ДНК, однозначно кодирующей определенный белок. В тех случаях, когда говорят, что отрезок ДНК отвечает за образование одного определенного белка, под полной последовательностью ДНК теперь подразумевают совокупность «гена» – от первой точки, представленной в мРНК, до самого конца: экзоны, интроны и т.п.

В тех же случаях, когда речь идет о перекрывающихся или альтернативных формах экспрессии, следует переиначить замечание, которым мы закончили гл. 1. Вместо того чтобы говорить «один ген – один полипептид», следует сказать «один полипептид – один ген». Тогда мы можем считать «геном» последовательность (включая не только

экзоны, но и интроны), действительно отвечающую за образование полипептида, даже если известно, что часть этой последовательности участвует в синтезе другого белка, т.е. входит составной частью в *его* ген. В таких случаях можно использовать термин «перекрывающиеся гены».

Используя термин «ген» для обозначения последовательности ДНК, ответственной за синтез данного белка, мы теперь не можем вкладывать в это понятие представления о структурной организации и комплементарных свойствах, которые мы подразумевали совсем недавно.

Используя термин «ген» для обозначения последовательности ДНК, ответственной за синтез данного белка, мы теперь не можем вкладывать в это понятие представления о структурной организации и комплементарных свойствах, которые мы подразумевали совсем недавно.

Рекомендуемая литература

Принципы рестрикционного картирования были впервые изложены в работе Данна, Сака и Натанса (*Danna, Sack, Nathans. J. Mol Biol.*, **78**, 363–376, 1973) и в обзоре Натанса и Смита (*Nathans, Smith, Ann. Rev. Biochem.*, **46**, 273–293, 1975). Методы анализа последовательности ДНК рассматриваются в обзоре Ву (*Wu, Ann. Rev. Biochem.*, **47**, 607–734, 1978). Наибольшее распространение получили два метода определения нуклеотидной последовательности. Первый из них, описанный в этой книге, предложен Максамом и Гилбертом (*Maxam, Gilbert, Proc. Nat. Acad. Sci. USA*, **74**, 560–564, 1974), второй – Сэнгером и др. (*Sanger, Nicklen, Coulson, Proc. Nat. Acad. Sci. USA*, **74**, 5463–5467, 1977). Современные представления о количественности гена и белка, лежащие в основе традиционных методов молекулярной генетики, были развиты на основе работы Яновского и др. (*Yanofsky et al., Proc. Nat. Acad. Sci. USA*, **57**, 296–298, 1967) и Сарабаи и др. (*Sarabhai et al., Nature*, **201**, 13–17, 1964). Генетический смысл прерывистой структуры гена обсуждается в гл. 7 книги Льюина (*Lewin, Gene Expression*, **2**, Eucaryotic Chromosomes, Wiley, New York, 1980).

Глава 4

РАСШИФРОВКА ГЕНЕТИЧЕСКОГО КОДА

Еще 2500 лет назад Аристотель высказал предположение о том, что гаметы – это отнюдь не миниатюрные варианты будущего организма, а структуры, содержащие *информацию* о развитии эмбриона (хотя он признавал только исключительную важность яйцеклетки в ущерб сперматозоиду). Однако развитие этой идеи в современных исследованиях стало возможным лишь после опытов Менделя, показавших, что гены представляют собой элементарные факторы наследственности. Но как эти факторы обуславливают развитие различных фенотипических признаков?

С открытием генетической роли ДНК была заложена основа концепции о том, что наследуется **генетическая информация**. Критическое значение при этом имеет тот факт, что структура ДНК не зависит от последовательности пар оснований. Таким образом, последовательность оснований в полинуклеотидной цепи имеет важное значение не для самой структуры ДНК, а постольку, поскольку она **кодирует** последовательность аминокислот в белке. Представление о том, что каждый белок состоит из постоянного числа определенных аминокислот, возникло в 50-х годах на основе работы Сэнгера (Sanger) по расшифровке структуры инсулина.

Сумма каталитических и структурных активностей различных белков клетки определяет ее фенотип. Конечно, кроме последовательностей, кодирующих белки, в ДНК находятся специфические области, которые узнаются регуляторными молекулами, обычно белковыми. Функция этих областей зависит непосредственно от их последовательности, и не опосредована кодом. Генетическая информация организма состоит из последовательностей обоих типов, из генов, экспрессирующихся в виде белков, и из участков, функционирующих непосредственно как регуляторные зоны.

Согласно общепринятому в молекулярной биологии представлению, каждая соматическая клетка организма содержит одинаковый набор хромосом и, следовательно, обладает одинаковой генетической информацией. Однако не вся эта информация экспрессируется в каждой клетке. Значит, генетический материал можно рассматривать как хранилище генетической информации. В каждой клетке экспрессируется только часть этого материала, остальная же информация хранит молчание. Недавно, однако, было показано, что в определенных пределах возможно изменение в содержании и экспрессии генетической информации (см. часть 10).

**ДНК нужна только для того,
чтобы кодировать
последовательность аминокислот**

Взаимоотношения между последовательностью гена и последовательностью белка устанавливаются с помощью **генетического кода**. Каким образом в последова-

тельности нуклеотидов может быть представлена последовательность аминокислот?

Двадцать аминокислот, обычно обнаруживаемых в белках, показаны на **рис. 4.1**. Аминокислоты соединяются вместе **пептидными** связями, образующимися в результате конденсации аминогруппы (NH_2) одной аминокислоты с карбоксильной группой



другой (**рис. 4.2**). Белковую последовательность условно записывают от аминокислоты со свободной NH_2 -группой (N-конец) до аминокислоты со свободной COOH -группой (C-конец).

Любые другие аминокислоты, например относительно редко встречающийся гидроксипролин, образуются в результате специфической модификации одной из 20 обычных протеиногенных аминокислот. Различные виды модификаций аминокислот, такие, как фосфорилирование, аденилирование и метилирование, выполняются специфическими ферментами уже в составе белковой молекулы.

Последовательность аминокислот, входящих в полипептидную цепь, составляет ее **первичную структуру**. Но критическое значение для каждого белка имеет его способность складываться в определенную **конформацию**, что сопровождается образованием активных центров или других особенностей структуры, необходимых для выполнения специфической функции в клетке. Формирование **вторичной структуры** происходит в результате свободного вращения относительно химических связей в первичной структуре полипептидной цепи. Термины «конформация» и «вторичная структура» отражают пространственную организацию полипептидного остова. Под **третичной структурой** понимают законченную трехмерную организацию всех атомов полипептидной цепи, включая боковые группы и полипептидный остов. Следующий уровень организации белков – это **мультимерные белки**, состоящие из агрегатов нескольких полипептидных цепей. Конформацию мультимерных белков принято называть **четвертичной структурой**.

Общий принцип организации белковой молекулы заключается в том, что структуры высшего порядка определяются непосредственно структурой низшего порядка. Это означает, что в первичной последовательности аминокислот заложена информация, необходимая для образования правильной конформации белка. Сюда относится образование ковалентных и нековалентных связей. Например, существенным моментом в формировании вторичной структуры является образование дисульфидных (S—S) мостиков между остатками цистеина. Каждый цистеин находится в полипептидной цепи отдельно, но в результате свертывания в специфическую конформацию две аминокислоты оказываются рядом и происходит конденсация двух SH-групп.

Какие условия необходимы для того, чтобы первичная структура образовала правильную структуру высшего по-

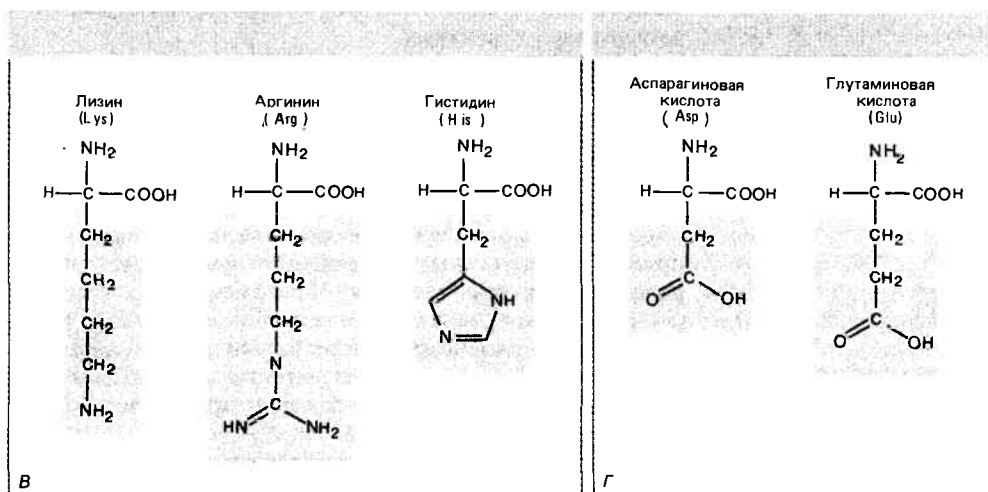
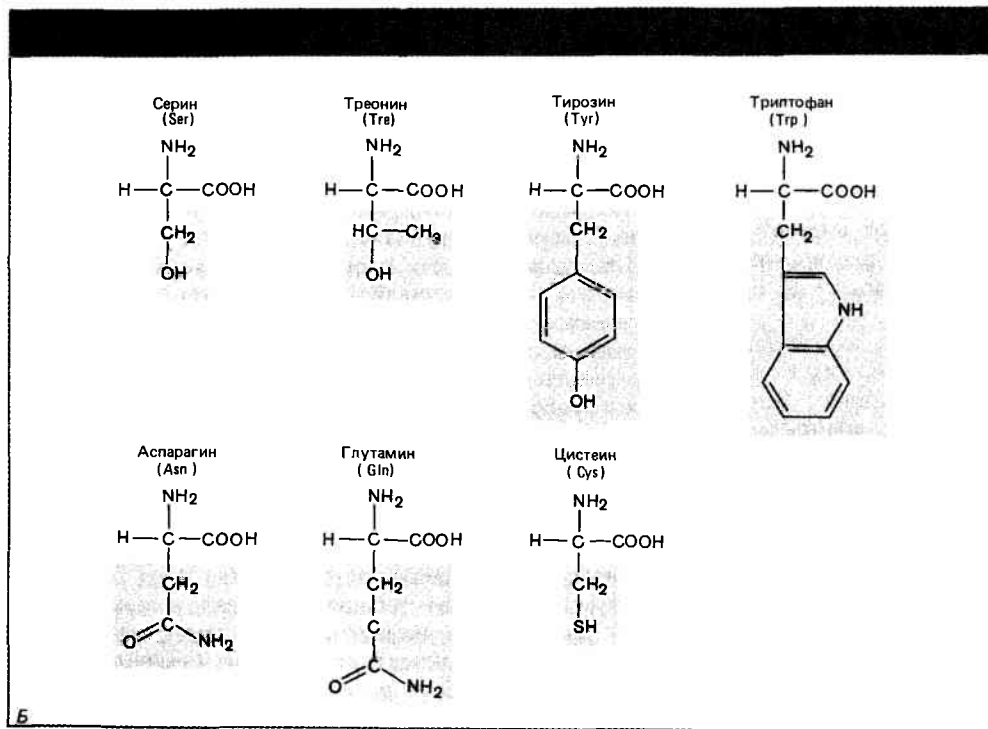
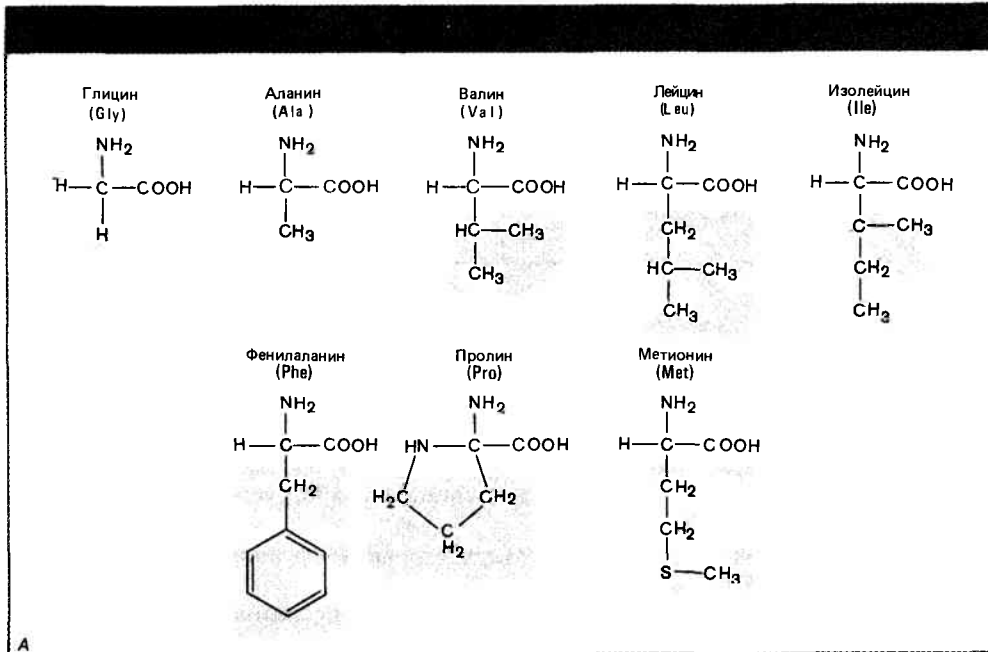


Рис. 4.1. Двадцать аминокислот, из которых построены белки, делятся на несколько классов в зависимости от природы боковых групп (R).
 А – нейтральные гидрофобные аминокислоты, Б – нейтральные полярные аминокислоты, В – основные аминокислоты, Г – Кислые аминокислоты.

рядка? Заложено ли это в полной первичной последовательности? Тогда можно было бы считать, что окончательная структура термодинамически наиболее стабильна. Или же правильная структура образуется только во время исходного синтеза белка? Это давало бы преимущество и менее стабильным структурам, поскольку тип структуры белка определялся бы во время синтеза.

В некоторых случаях первичная структура однозначно определяет структуру более высокого порядка. Впервые это было показано в опытах по денатурации белка рибонуклеазы. Денатурацией здесь называют потерю конформации в результате разрушения S—S-мостиков, вызванного нагреванием или специальной химической обработкой белка. Активная конформация может быть восстановлена в результате процесса, обратного денатурации; следовательно, вся информация, необходимая для ее образования, содержится в первичной структуре. Таким образом, для того чтобы получить активную рибонуклеазу, нужно только правильно синтезировать последовательность аминокислот. (Термин «денатурация» используют также, говоря о ДНК. Тогда это означает, что двухцепочечная структура превращается в одноцепочечную; см. гл. 2.)

В других случаях денатурация необратима. Это означает, что у белка могут быть альтернативные стабильные конформации, образование которых зависит от определенных условий. Возможно, что правильную конформацию белки приобретают *только* во время биосинтеза. При этом специфические области уже синтезированного участка белка вступают во взаимодействие между собой, пока другая часть молекулы еще не синтезирована. Синтез белка происходит только в одном направлении, и это может иметь важное значение для определения структур высшего порядка. Таким образом, само направление синтеза обеспечивает информацию, необходимую для образования правильной конформации. (Тот же принцип используется для некоторых РНК; см. гл. 15.)

Иногда полипептидная цепь может принять нужную конформацию лишь при наличии кофактора, представляющего собой часть белка (такого, например, как железосвязывающая группа гема в цитохромах).

Для белков характерно огромное разнообразие форм, а структура и (или) ферментативная активность каждого белка обусловлены его первичной аминокислотной последовательностью. Иными словами, определяя аминокислотную последовательность каждого белка, гены содержат в себе всю информацию, необходимую для построения активной полипептидной цепи. Таким способом элементарная структурная единица — ген может выражать себя в бесчисленных полипептидных формах.

Генетический код считывается триплетами

Итак, в нуклеотидной последовательности должно быть достаточно кодирующих единиц, чтобы зашифровать 20 аминокислот. Но в ДНК содержится только 4 основания. Это означает, что **кодовое отношение** должно быть больше единицы, т.е. специфичность одной аминокислоты должна быть детерминирована более чем одним основанием. Если бы кодовое число было равно двум и два основания определяли одну аминокислоту, то в ДНК могло бы быть закодировано только 4^2 , т.е. 16 типов аминокислот. Поскольку этого недостаточно, кодовое отношение должно быть не меньше 3.

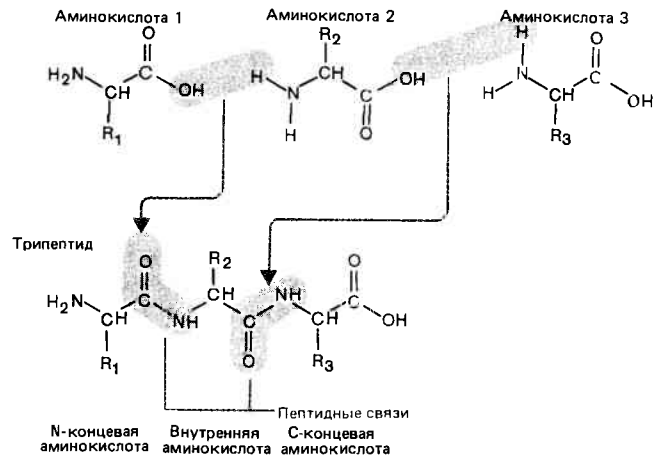


Рис. 4.2. Пептидные связи образуются в результате реакции конденсации.

Если генетический код триплетен, каждой аминокислоте должны соответствовать три расположенных рядом основания. Поскольку число возможных триплетных комбинаций составляет 4^3 , т.е. 64, существование триплетного кода предполагает, что либо не все триплеты участвуют в кодировании аминокислот, либо некоторые аминокислоты кодируются более чем одним кодоном.

Первые идеи о природе генетического кода содержали представления о прямой связи ДНК с белком, что могло бы обеспечиваться стереохимическим соответствием. Еще и теперь высказываются предположения о том, что современный код мог бы в процессе эволюции пройти через стадию, когда действовал именно такой механизм. Однако в ныне существующем виде дешифровка кода происходит с участием сложного аппарата, который связан и с ДНК, и с белком (и участие которого необходимо, если в ДНК заложена существенная информация).

Триплетная природа генетического кода была впервые продемонстрирована в генетических экспериментах. Последовательность из трех нуклеотидов, соответствующая одной аминокислоте, называется **кодоном**. Последовательность кодонов читается непрерывно, начиная с фиксированной стартовой точки на одном конце гена, и заканчивается в точке терминации на другом конце гена. Записывая последовательность нуклеотидов условно в направлении от 5'-конца к 3'-концу, мы увидим, что она соответствует аминокислотной последовательности, записанной в направлении от N-конца к С-концу.

Для генетического анализа использовали делеции, полученные с помощью акридина в *rII*-области фага T4. В 1964 г. Крик (Crick) и его сотрудники показали, что мутации, индуцированные акридинами, существенно отличаются от мутаций, индуцированных заменой оснований, в двух отношениях. В результате замены оснований (примеры см. на рис. 2.17 и 2.18) часто образуются так называемые leaky (лики)-мутанты, у которых сохраняется остаточная функция. У акридиновых мутантов, напротив, функция гена утеряна полностью. Мутации, возникшие в результате замены оснований, могут ревертировать к дикому типу под действием мутагенов, вызывающих замены оснований. Мутации, индуцированные акридинами, ревертируют только после повторного воздействия акридинами.

Мутагенный эффект акридинов обусловлен вставками и выпадениями отдельных нуклеотидов ДНК. Нарушение функции гена под действием акридинов можно объяснить тем, что генетический код читается в виде *неперекрывающихся триплетов с фиксированной стартовой точки*.

Понятие «неперекрывающийся код» означает, что каждый кодон состоит из трех нуклеотидов, а последовательно расположенные кодоны являются последовательно расположенными триплетами нуклеотидов.

Использование фиксированной стартовой точки означает, что синтез белка должен начинаться на одном конце и продолжаться до другого, иными словами, различные части кодирующей последовательности не могут читаться независимо.

Если генетический код читается неперекрывающимися триплетами, возможны три способа трансляции нуклеиновой кислоты в белок, в зависимости от стартовой точки, т.е. **три рамки считывания**. Например, для последовательности

ACGACGACGACGACGACG

возможны три рамки считывания:

ACG ACG ACG ACG ACG ACG
CGA CGA CGA CGA CGA CGA
GAC GAC GAC GAC GAC GAC

Если для считывания используется только одна рамка, то точечная мутация, изменяющая лишь одно основание, изменит только один кодон, в котором находится это основание. При этом в белке изменится всего одна аминокислота, а это не обязательно приведет к полному исчезновению активности белка; активность белка может лишь уменьшиться. Этим и объясняется тот факт, что замены оснований часто приводят к мутациям типа «leaky».

Вставка же или выпадение одного основания, напротив, изменят рамку считывания всей последовательности, расположенной после мутации, т.е. произойдет «сдвиг рамки считывания». Поскольку нуклеотидная последовательность в новой рамке считывания будет полностью отличаться от прежней, последовательность аминокислот в белке изменится сразу после мутировавшего участка. Функция такого белка, вероятно, пропадет совсем, что характерно для мутаций, индуцированных акридинами.

Если акридиновый мутант образован, скажем, в результате добавления одного нуклеотида, он может ревертировать к дикому типу при делеции этого нуклеотида. Однако реверсия может произойти и при делеции другого основания поблизости от мутантного сайта — до или после него. Вторую такую мутацию называют **супрессором**.

На **рис. 4.3** показано, что в этом случае неправильное считывание информации ограничено расстоянием между двумя мутациями. В остальной части белок читается соответственно правильной рамке считывания. Заметим, что в контексте этой работы термин «супрессор» используется в ином смысле, чем обычно (см. гл. 2).

Когда с помощью генетической рекомбинации акридиновые мутации отделяли от их супрессоров, супрессоры (как и сами мутации) проявляли свойства, характерные для мутаций со сдвигом рамки считывания.

В более общем виде все акридиновые мутации можно разделить на два типа, обозначив их (+) и (-). Каждый тип мутаций сам по себе вызывает сдвиг рамки считыва-

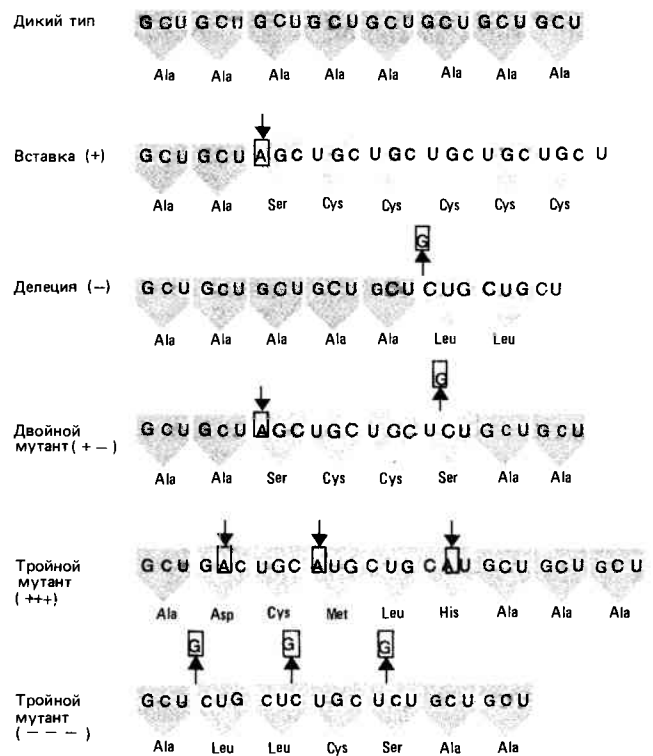


Рис. 4.3. Мутации со сдвигом рамки показывают, что генетический код читается триплетами, начиная с фиксированной стартовой точки.

Основания, вставленные или делетированные из последовательности дикого типа, указаны стрелками или обведены. Отдельные триплеты обозначены пятиугольниками серого цвета для кодонов дикого типа и цветными — для мутантных кодонов. Показаны соответствующие аминокислоты.

ния. Мутации со сдвигом рамки, обусловленные добавлением одного основания, относят к (+)-типу, а мутации, обусловленные выпадением одного основания, — к (-)-типу. Двойные мутанты типа (+ +) или (- -) также обладают мутантным фенотипом. Однако при сочетании мутаций (+ -) или (- +) мутации супрессировали друг друга.

Из этих опытов видно, что генетический код читается как последовательность, у которой рамка считывания фиксирована наличием стартовой точки, так что одиночные вставки и делеции компенсируют друг друга. При двойных же вставках или двойных делециях мутации не компенсируются. Из этого, однако, не ясно, из скольких нуклеотидов состоит кодон. Но если сконструировать тройные мутанты, то комбинации типа (+ + +) и (- - -) будут иметь дикий фенотип, другие же комбинации останутся мутантными. Из этого следует, что код считывается триплетами, так как тройные вставки и тройные делеции добавляют или удаляют только по одной аминокислоте. Измененная часть белка ограничена при этом участком между первым и третьим мутационными сайтами (рис. 4.3).

Предсказание о том, что у двойных и тройных мутантов измененная часть белка ограничивается областью между мутировавшими участками, не удалось проверить на *rII*-системе. К сожалению, тогда этот белок был недоступен. Однако это предсказание было подтверждено на другой системе при анализе аналогичного типа.

Аппарат для последовательного белкового синтеза

Если генетический код читается как серия триплетов, белок синтезируется последовательно с одного конца до другого. Впервые это продемонстрировал Динцис (Dintzis), исследуя белок эритроцитов – гемоглобин. К клеткам, синтезирующим гемоглобин, добавляли радиоактивно меченную аминокислоту. Затем выделяли целый белок, расщепляли его трипсином и исследовали содержание метки в пептидах из разных участков молекулы.

На рис. 4.4 показано, что радиоактивная метка может включаться во все возможные участки белка, в зависимости от той точки, в которой находился синтез данной молекулы в момент добавления метки. Но в целых, *завершенных* молекулах радиоактивность содержится только в том конце, который синтезируется последним. Если увеличить время мечения, то метку можно обнаружить и в начале молекулы. Из этого следует, что радиоактивность включается в обратном порядке – от конца к началу. Согласно этим данным, белок синтезируется в направлении от N-конца к C-концу.

Как мы уже говорили в гл. 3, первая ступень экспрессии гена состоит в образовании мРНК, копирующей одну цепь ДНК. Это совпадает с представлением о том, что любой из двух комплементарных цепей двойной спирали достаточно для кодирования. Цепь, имеющая *ту же* последовательность, что и мРНК (с той лишь разницей, что в ней содержится Т вместо U в РНК), иногда называют *кодирующей цепью*. Другая цепь, которая обеспечивает синтез мРНК путем комплементарного спаривания оснований, оказывается *антикодирующей цепью*. Поскольку генетический код считывается с мРНК, его обычно записывают, используя четыре основания, присутствующие в РНК: U, C, A, G.

Синтез одноцепочечной мРНК на ДНК называют *транскрипцией*. мРНК служит матрицей, по которой аминокислоты собираются в полипептидную цепь. Этот процесс получил название «*трансляция*», поскольку именно на этой стадии нуклеотидные триплеты дешифруются как определенные аминокислоты.

Трансляция выполняется специальными клеточными органеллами – *рибосомами*. Рибосомы – это частицы, состоящие из РНК и белка слегка сплюсненной сферической формы. Как мы увидим в следующей главе, они прикрепляются к мРНК на 5'-конце и передвигаются вдоль полинуклеотидной цепи по направлению к 3'-концу, попутно транслируя каждый триплет в аминокислоту.

Из-за различия структур нуклеотида и аминокислоты сразу возникает вопрос: как каждый кодон подбирает свою определенную аминокислоту? Еще до того, как код был расшифрован, Крик предположил, что посредником в трансляции может служить особая молекула – «адаптор». Ею оказалась *транспортная РНК*, сокращенно *тРНК*, состоящая из 75–85 нуклеотидов и являющаяся самой маленькой РНК в клетке.

Каждая тРНК обладает двумя необходимыми свойствами адаптора. Она способна узнавать только одну аминокислоту, с которой она *ковалентно связывается*, и превращается в *аминоацил-тРНК*. Аминоацил-тРНК в свою очередь способна связываться с рибосомой, выполняя вторую важную функцию тРНК. Каждая тРНК имеет тринуклеотидную последовательность – *антикодон*, который комплементарен кодону, соответствующему данной аминокислоте. Поэтому тРНК узнает кодон путем

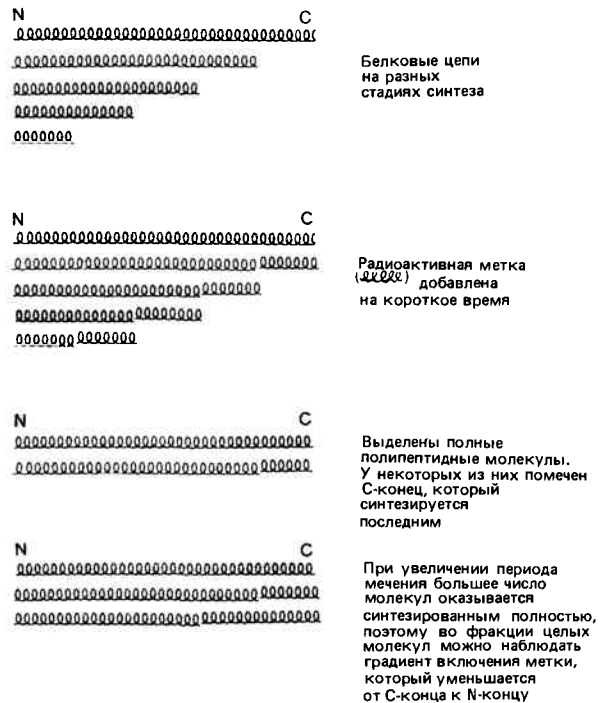


Рис. 4.4. Белки синтезируются в направлении от N-конца к C-концу.

Радиоактивная метка включается во все растущие участки полипептидной цепи одновременно. После выделения полных белковых молекул метку можно обнаружить в тех участках молекул, которые были синтезированы последними. Градиент метки от C-конца к N-концу показывает, что синтез происходит в противоположном направлении.

комплементарного спаривания оснований. В рибосоме одновременно находятся два вида тРНК для двух последовательно расположенных кодонов. Таким образом, между двумя находящимися в рибосоме аминокислотами образуется пептидная связь.

Кодоны, соответствующие аминокислотам

Когда была установлена триплетная природа генетического кода, казалось, что нелегко будет найти подход, позволяющий расшифровать кодоны для всех аминокислот. Но очень скоро было разработано два метода, с помощью которых расшифровали большинство кодонов.

Систему для белкового синтеза можно получить в виде бесклеточного экстракта (см. гл. 6). Исходная процедура состояла в следующем: клетки бактерии *E. coli* разрушали, из полученного экстракта методом центрифугирования удаляли фрагменты оболочек и мембран. ДНК разрушали добавлением ДНКазы (а бактериальная мРНК слишком нестабильна и не могла служить эндогенной матрицей). Полученная в результате надосадочная фракция активно транслирует любую добавленную мРНК, поскольку в этой фракции содержатся необходимые предшественники и источник энергии. Позднее были разработаны другие системы, содержащие более чистые компоненты. Такие системы были созданы из многих типов как прокариотических, так и эукариотических клеток.

Синтетические полинуклеотиды стали применять для синтеза белка после того, как в 1961 г. Ниренберг и Мат-

теи (Nirenberg, Matthaei) сообщили, что poly(U)–полиуридиловая кислота,– может служить матрицей, как и мРНК. При этом *in vitro* из фенилаланина синтезируется полифенилаланин. Из этого следовало, что триплет UUU должен быть кодоном для фенилаланина. Аналогично этому гомополимеры poly(C) и poly(A) стимулировали синтез полипролина и полилизина.

Следующей ступенью в этих исследованиях стало использование *гетерополимеров*, содержавших более одного основания и способных направлять включение в белок нескольких аминокислот. Такие гетерополимеры можно приготовить, используя фермент полинуклеотид-фосфорилазу (обычный фермент *метаболизма* нуклеиновых кислот), которая сшивает вместе нуклеозиддифосфаты. (Эта искусственная ситуация является единственным исключением из правила, согласно которому предшественниками для синтеза нуклеиновой кислоты служат трифосфаты.)

Если отдельные нуклеотиды включаются в полинуклеотид случайно, тогда частота образования случайных триплетных комбинаций должна совпадать с относительным количеством каждой аминокислоты в белке.

Более точный метод разработал Корана (Khorana), который синтезировал полинуклеотиды известной последовательности. Это дало возможность установить точную корреляцию между кодонами и аминокислотами в полипептиде. С помощью этого метода была расшифрована примерно половина всех кодонов.

Ниренберг и Ледер (Nirenberg, Leder) в 1964 г. разработали другой метод расшифровки кодонов. Тринуклеотиды известной последовательности стимулируют связывание специфических молекул аминоацил-тРНК с рибосомами. Это имитирует процесс взаимодействия кодона с антикодоном, происходящий на рибосомах между тРНК и мРНК. Связывая рибосомы на нитроцеллюлозных фильтрах, можно выделить тройные комплексы: тринуклеотид·аминоацил-тРНК·рибосома. Сама аминоацил-тРНК не связывается с нитроцеллюлозными фильтрами, но ее можно выделить в составе комплекса: эффективность специфического связывания измеряют по радиоактивной метке, используя меченую аминокислоту. Определив, какая из 20 меченых аминоацил-тРНК задерживается на фильтре, с помощью этого метода (получившего название **метод связывания рибосом**) установили специфическое значение всех тринуклеотидов.

Два описанных выше метода позволили определить значение 61 из 64 кодонов. Однако в экспериментах обоих типов реакции проводили *in vitro*, поэтому следовало подтвердить, что полученные данные справедливы и *in vivo*. Первая серия доказательств того, что кодоны имеют такое же значение в живых клетках, была получена при анализе мутаций. Мутации, индуцированные заменами оснований, коррелировали с аминокислотными заменами в белке. Сопоставление белка дикого типа и участка белка между двумя мутациями типа «сдвиг рамки» также выявило четкую корреляцию между измененным смыслом кодонов и последовательностью аминокислот в испорченном участке белка.

Эксперименты такого рода подтверждали, что смысл кодонов был определен правильно, но они отнюдь не доказывали, что *in vivo* и *in vitro* кодоны считаются одинаково. Однако позднее благодаря развитию методов определения нуклеотидной последовательности ДНК появилась возможность прямо определять кодоны. Мы уже говорили о том, что теперь можно быстрее узнать последовательность самого гена, чем его продукта. Гены с из-

вестными белковыми продуктами можно прямо читать триплетами, сопоставляя тексты ДНК и белка. Последовательность кодирующей цепи ДНК, прочитанная в направлении 5'–3', состоит из триплетов, соответствующих аминокислотам в направлении от N-конца белка к С-концу. Таким методом была получена полная характеристика генетического кода, которая целиком и во всех деталях подтвердила прежние результаты.

Генетический код изображен на **рис. 4.5**. Отчетливо видно, что код – **вырожденный**: 20 аминокислот представлены 61 кодоном. Почти каждой аминокислоте соответствует несколько **кодонов-синонимов**. Число кодонов для одной аминокислоты достаточно хорошо отражает частоту встречаемости данной аминокислоты в белках. Как видно на **рис. 4.6**, такая корреляция наблюдается для всех аминокислот, за исключением аргинина.

Другой особенностью кода является тенденция к группировке кодонов, соответствующих одной аминокислоте. Часто основание в третьем положении кодона оказывается несущественным для его специфичности. Одна аминокислота может быть представлена четырьмя кодонами, различающимися только по третьему основанию. Иногда различие состоит в предпочтении пурина пиримидину в этом положении. Меньшую специфичность этого положения в кодоне называют **вырожденностью третьего основания**.

Эта особенность, а также тенденция к сходству кодонов для аминокислот одного типа (т.е. полярных, гидрофобных и т.д.) сводят до минимума эффект мутаций. При такой организации кода случайно возникшая замена основания с большей вероятностью (чем при случайном подборе кодонов) приведет к замене на сходную аминокислоту или же замены не произойдет вовсе.

Природа сигналов терминации

Только 61 из 64 триплетов используется для обозначения аминокислот. Какова же функция оставшихся трех кодонов?

Триплеты UAG, UAA и UGA являются кодонами-терминаторами, на которых синтез белка останавливает-

		Второе основание			
		U	C	A	G
Первое основание	U	UUU } Phe UUC } UUA } Leu UUG }	UCU } Ser UCC } UCA } UCG }	UAU } Tyr UAC } UAA } TЕРМ UAG }	UGU } Cys UGC } UGA } TЕРМ UGG } Trp
	C	CUU } Leu CUC } CUA } CUG }	CCU } Phe CCC } CCA } CCG }	CAU } His CAC } CAA } Gln CAG }	CGU } Arg CGC } CGA } CGG }
	A	AUU } Ile AUC } AUA } Met AUG }	ACU } Thr ACC } ACA } ACG }	AAU } Asn AAC } AAA } Lys AAG }	AGU } Ser AGC } AGA } Arg AGG }
	G	GUU } Val GUC } GUA } GUG }	GCU } Ala GCC } GCA } GCG }	GAU } Asp GAC } GAA } Glu GAG }	GGU } Gly GGC } GGA } GGG }

Рис. 4.5. Все триплетные кодоны осмысленны.

Шестьдесят один кодон соответствует аминокислотам, и все аминокислоты, за исключением триптофана и метионина, кодируются несколькими кодонами. Кодоны-синонимы обычно образуют группы, в которых два первых основания в кодоне являются общими, а третье – варьирует. Три кодона вызывают терминацию (TЕРМ). Порядок оснований в кодоне записан, как обычно, в направлении от 5'-конца к 3'-концу.

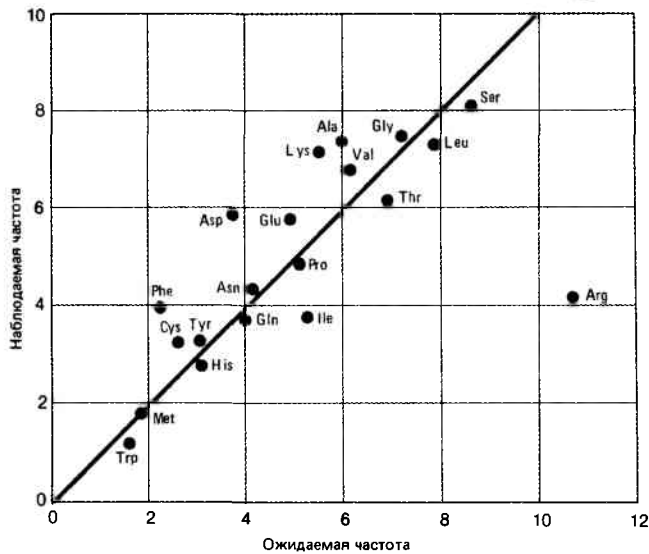


Рис. 4.6. Число кодонов для каждой аминокислоты коррелирует с частотой встречаемости данной аминокислоты в белках. Исключение составляет аргинин, поскольку в эукариотической ДНК дуплет CG встречается редко. Поэтому четыре кодона, соответствующие аргинину, которые начинаются с этого дуплета, встречаются реже, чем следовало бы ожидать, исходя только из состава их оснований.

ся. Если такой кодон находится в синтетическом полинуклеотиде, синтез белка происходит только до места расположения такого кодона. (Это, однако, не очень строго доказано из-за некоторых недостатков системы *in vitro*.)

Как показал анализ мутаций, роль этих кодонов *in vivo* отличается от роли всех остальных кодонов. Точковую мутацию, которая приводит к изменению триплета на кодон для другой аминокислоты, называют **миссенс-мутацией**. Влияние такой мутации на белок зависит от природы аминокислотной замены.

Если же в результате мутации образуется один из трех кодонов-терминаторов, это приводит к **преждевременной терминации** синтеза белка в том месте, где расположен мутантный кодон. При этом нарушится функция белка, так как в мутантной клетке синтезируется только часть белковой молекулы. Такие мутации называют **нонсенс-мутациями**. [Иногда термин «нонсенс-кодон» используют для обозначения кодонов-терминаторов. «Нонсенс» (бессмысленный) — неправильное название для этих кодонов, так как они имеют вполне определенный смысл, хотя и неблагоприятный для мутантного гена.]

Мутации типа «нонсенс» и «миссенс» впервые удалось разграничить благодаря генетическому тесту, использованному Бензером и Чеймпом (Benzer, Champe) в 1961 г. Существует вариант фага T4, у которого делетирован промежуток между цистронами *rIIA* и *rIIB*. В результате оба цистрона соединились воедино, и вместо двух отдельных белков синтезируется один слившийся белок. У этого белка сохраняется активность белка В, несмотря на соединение двух полипептидов. На рис. 4.7 показано, как можно различать нонсенс- и миссенс-мутации в *rIIA*-области. Для этого нужно сконструировать двойной мутант, который кроме исследуемой мутации несет делецию, соединяющую цистроны А и В. Если в *rIIA*-области возникла миссенс-мутация, то активность *rIIB*-цистрона не будет нарушена. Но если возникнет нонсенс-мутация, синтез белка остановится и полипептид В не синтезируется.

У нонсенс-мутаций, отобранных таким методом, была

обнаружена интересная особенность: их поведение зависело от штамма *E. coli*, использованного в качестве клетки-хозяина. Была создана **условно-летальная система**, аналогичная той, которую мы описывали в конце гл. 2. В **непермиссивных** штаммах мутанты фага не размножаются из-за нонсенс-мутации, прекращающей синтез белка. Однако те же самые мутанты размножаются в пермиссивных штаммах бактерий. Это происходит благодаря тому, что **пермиссивные** штаммы бактерий содержат мутации-**супрессоры**, преодолевающие эффект нонсенс-мутаций фага. Супрессором служит мутантная форма аминоксил-тРНК, которая узнает кодон-терминатор и вставляет в этом месте свою аминокислоту. Таким образом, прерванный ранее синтез белка продолжается, проходя мутантный сайт (см. гл. 7).

По чувствительности к различным супрессорам нонсенс-мутации делятся на три класса. Исходный класс нонсенс-мутаций, изолированных у фага T4, был назван **амбер-мутациями**. Все эти мутации оказались чувствительными к одному супрессору *E. coli*. Анализируя способность мутантов фага размножаться на разных штаммах *E. coli*, несущих амбер-супрессоры, обнаружили новый класс нонсенс-мутаций, названный **охра-мутациями**. Мутации типа «охра» не супрессируются амбер-супрессорами, а соответствующие им супрессоры называют охра-супрессорами. Интересно, что охра-супрессоры *способны* супрессировать и амбер- и охра-кодоны, что говорит о возможном сходстве этих типов нонсенс-мутаций. Позднее был обнаружен третий класс нонсенс-мутаций, которых назвали опал-мутациями. **Опал-мутации** не чувствительны ни к охра-, ни к амбер-супрессорам, а их супрессоры не действуют на кодоны-терминаторы типа «охра» и «амбер».

Нуклеотидную последовательность кодонов-терминаторов сначала определили, исходя из особенностей мутаций, вызвавших их образование. Зная, какие аминокислоты находились на месте мутантных кодонов-терминаторов, установили, что амбер-кодону соответствует триплет UAG, охра-кодону — UAA, а опал-кодону — UGA. Позднее эти же значения были получены при определении нуклеотидной последовательности генов, несущих соответствующие нонсенс-мутации.

Во всех проведенных экспериментах наблюдали, что любой из кодонов UAG, UAA или UGA, появившихся в результате нонсенс-мутации, вызывает остановку белкового синтеза. Однако используются ли эти же сигналы для обычной остановки синтеза белка в конце генов? Оказалось, что в конце всех генов с известной нуклеотидной последовательностью всегда находится один (а иногда не менее двух) кодонов-терминаторов, расположенных непосредственно после кодона для С-концевой аминокислоты белка дикого типа. Следовательно, это обычные сигналы, которые действуют и внутри гена, если они там возникают в результате мутации.

Ни для одного из кодонов-терминаторов не найдено соответствующей тРНК. Это исключает возможность механизма терминации с участием специальной тРНК, которая узнает нонсенс-сигнал для прекращения белкового синтеза. Вместо этого существуют сигнальные белковые факторы, которые вступают в действие как раз в тот момент, когда рибосома доходит до кодона-терминатора. Таким образом, терминирующие кодоны являются **знаками пунктуации**, механизм действия которых отличается от механизма действия кодонов, детерминирующих аминокислоты.

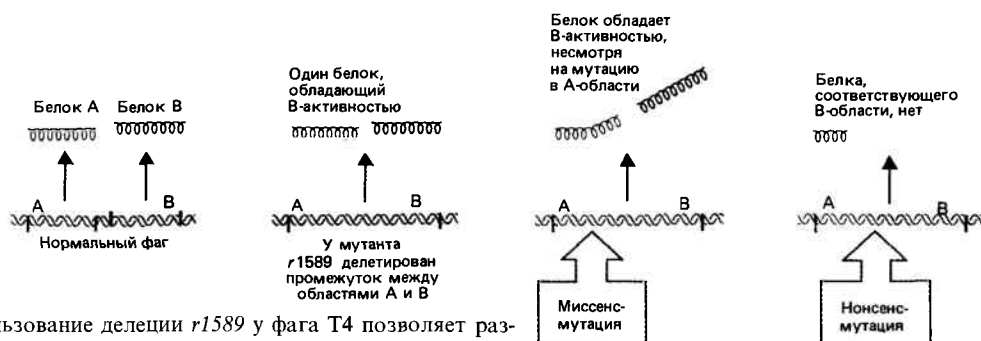


Рис. 4.7. Использование делеции *r1589* у фага Т4 позволяет различить нонсенс- и миссенс-мутации.

Универсален ли код?

Генетический код первоначально был установлен у бактерии *E. coli*. Естественно возник вопрос: таков ли он и у всех других живых организмов?

Известно, что мРНК одного вида может быть правильно прочитана *in vivo* и *in vitro* белок-синтезирующим аппаратом другого вида. Теперь уже разработано много таких гетерологических систем трансляции. Сам факт их успешного действия указывает, что кодоны, используемые в мРНК одного вида, имеют такой же смысл для рибосом и тРНК других видов. Это говорит о постоянстве кода.

Прямое доказательство универсальности кода было получено при сравнении последовательностей ДНК с соответствующими белковыми последовательностями. Оказалось, что во всех бактериальных и эукариотических геномах используются *одни и те же наборы кодовых значений*. Однако состав оснований различных геномов сильно варьирует в противоположность относительному постоянству аминокислотного состава белков. Можно думать поэтому, что различные виды используют различающиеся характерные наборы кодонов-синонимов. Действительно, наблюдаемое постоянство аминокислотного состава можно объяснить только вырожденностью генетического кода.

Первые исключения из универсальности генетического кода были недавно обнаружены в митохондриях некоторых видов. Во всех случаях было замечено одно общее изменение: кодон UGA читался так же, как UGG, и поэтому из кодона-терминатора превратился в кодон для триптофана. Другие изменения были специфическими для каждого вида. У дрожжей кодон CUA и, возможно, все семейство CUN кодирует треонин вместо лейцина. У млекопитающих AUA имеет то же значение, что AUG, и означает метионин вместо изолейцина. Кодоны AGA и AGG вызывают терминацию цепи, а не кодируют аргинин. Некоторые из этих изменений приводят к упрощению кода — в том смысле, что два кодона, которые имели разное значение, замещаются парой кодонов с одинаковым значением.

Универсальность кода (игнорируя эти редкие исключения) свидетельствует о том, что он сложился очень рано в процессе эволюции. Согласно некоторым моделям, первоначально должны были существовать некоторые стереохимические взаимоотношения между аминокислотами и их кодонами. Затем, по мере развития белок-синтезирующей системы, шел отбор на увеличение эффективности и уменьшение числа ошибок.

Возможно, сначала код существовал в примитивном виде, когда малое число кодонов обозначало сравнительно

небольшое число аминокислот; один кодон мог соответствовать любому члену одной группы аминокислот. Более точное значение кодонов и большее число аминокислот могли быть введены позже. Сначала только первые два из трех оснований в кодоне могли быть использованы для узнавания, а роль третьего положения определилась позднее. С этим согласуется тип кодон-антикодонного взаимодействия, который мы наблюдаем в митохондриях. Для кода, используемого в митохондриях, характерна полная вырожденность третьего основания (по сравнению с обычной, частичной вырожденностью), так что третье основание в митохондриях не влияет на смысл кодона.

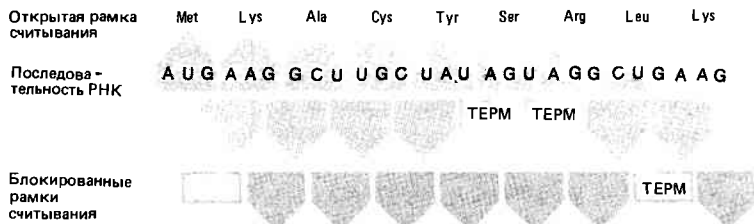
В некоторый момент времени код был «заморожен» в его современном виде, поскольку сложившаяся система кодирования оказалась достаточно изощренной для того, чтобы любые изменения значений кодонов не приводили к критическим изменениям в белке, обусловленным аминокислотной заменой. Универсальность кода означает, что он возник еще в первоначальном наборе примитивных клеток, из которых произошли все ныне живущие организмы.

Почему возникли изменения в коде митохондрий? Поскольку значения митохондриальных кодонов проще (по сравнению с общим кодом), это, возможно, говорит об их более примитивной организации. Примером тому служит использование UGA наряду с UGG для кодирования триптофана. Однако возможно и обратное толкование, т.е. что это упрощение было развито в результате особенностей белкового синтеза в митохондриях (см. гл. 7). Существование видоспецифических изменений говорит о большей гибкости механизмов белкового синтеза в митохондриях по сравнению с целым организмом. Система митохондриального белкового синтеза производит не много белков (всего около 10); поэтому проблемы значительных нарушений из-за перекодировки не так остры. Вероятно, измененные кодоны не часто использовались в тех местах, где аминокислотные замены приводили к серьезным повреждениям.

Трансляция при перекрывающихся рамках считывания

Представление о том, что ген соответствует отдельной полипептидной цепи, оказалось особенно плодотворным после того, как был расшифрован генетический код, так как на его основе была создана концепция, что последовательность белка закодирована в последовательности ДНК в виде серии триплетных кодонов. В этой

Рис. 4.8. Одной открытой рамке считывания обычно соответствуют две рамки считывания, которые находятся в других фазах и блокированы благодаря высокому содержанию кодонов-терминаторов.



концепции предполагалось также, что только одна из трех возможных рамок считывания используется в каждом гене.

Рамка считывания, содержащая последовательную серию кодонов, соответствующих аминокислотной последовательности, называется **открытой**. До сих пор мы говорили о кодовых значениях, имея в виду открытую рамку считывания, начинающуюся с какой-то фиксированной точки. Но в чем состоит стартовый сигнал? Точно так же, как нонсенс-кодоны используются для терминирования белкового синтеза, другой набор кодонов служит для его запуска. Общим сигналом инициации является соответствующий метионину кодон AUG в комбинации с предшествующей ему последовательностью, нужной для прикрепления рибосом к мРНК. В некоторых случаях для инициации используется также кодон GUG, который вопреки коду транслируется как метионин, а не как валин. Таким образом, кодирующая область состоит из **кодона AUG** (или GUG), следующей за ним серии триплетов, составляющих открытую рамку считывания, и оканчивается терминирующими кодонами UAA, UAG и UGA.

Две другие рамки считывания, которые находятся в иной фазе по отношению к открытой рамке считывания, обычно не могут быть использованы для синтеза белка, поскольку в их последовательности слишком часто встречаются кодоны-терминаторы. Такие рамки считывания называют **блокированными**. Типичный пример перекрывания открытой рамки считывания с двумя блокированными рамками считывания показан на **рис. 4.8**. Поскольку давление отбора происходило в пользу открытой рамки считывания, как о том свидетельствует последовательность аминокислот, в двух других фазах считывания шло беспрепятственное накопление кодонов-терминаторов. Возможно даже, что их накопление оказалось благоприятным, для того чтобы избежать образования нежелательных белков. В случайной последовательности ДНК кодоны-терминаторы составляют 3/64, что соответствует 1 кодону-терминатору на 20 триплетов (в зависимости от точного состава оснований). (Поэтому у мутантов со сдвигом рамки синтез полипептидной цепи обычно рано прекращается из-за нонсенс-кодона, оказавшегося во внефазовой рамке считывания.)

Существование открытой рамки считывания обычно рассматривается как свидетельство ее возможной трансляции. Только благодаря давлению отбора против накоп-

ления кодонов-терминаторов смогли образоваться белки длиннее нескольких аминокислотных остатков. Поскольку теперь появилась возможность определять последовательность длинных отрезков ДНК, периодически находят открытые рамки считывания, белковый продукт которых неизвестен. Таким образом, их фактическое использование может быть установлено путем идентификации белка.

Совершенно необычная ситуация была впервые обнаружена у фага ФХ174, где были найдены две перекрывающиеся открытые рамки считывания с достаточно протяженным участком перекрывания. Гипотетический пример такого перекрывания приведен на **рис. 4.9**. Внутри первой открытой рамки считывания находится кодон AUG в *другой рамке считывания*, за которым следует серия кодонов, детерминирующих различные аминокислоты. Таким образом, в одной последовательности ДНК сосуществуют две открытые рамки считывания. В случае фага ФХ174 (и в других с тех пор найденных случаях у вирусов и у митохондрий) были идентифицированы белки, соответствующие альтернативным рамкам считывания. Поскольку белки кодируются альтернативными рамками, они могут начинаться и кончаться в разных местах, так что участок перекрывания может варьировать в широких пределах.

До сих пор мы говорили о генетическом коде, рассматривая только одну цепь ДНК. Если же обратиться к обеим цепям, то окажутся возможными сразу шесть различных рамок считывания, по три в каждой цепи, читающиеся, конечно, в разных направлениях, поскольку цепи антипараллельны. Однако, несмотря на такую возможность, *почти* во всех случаях в любой данной точке только одна цепь служит кодирующей, тогда как другая блокирована во всех потенциальных рамках считывания. В рассмотренном выше случае фага Ф174 (а также в большинстве таких случаев) перекрывание двух рамок считывания происходит в одной цепи ДНК. Но поскольку такие ситуации в принципе допустимы, можно ожидать и перекрывания открытых рамок считывания в разных цепях. Случаи такого перекрывания уже обнаружены.

Использование перекрывающихся рамок считывания означает, что одна последовательность ДНК может кодировать два белка. Это немедленно ограничивает обычную гибкость, присущую кодирующей области, так как изменение основания в третьем положении не повреждает одного белка, но, вероятно, повредит другой белок. Мутации в перекрывающихся участках должны приводить

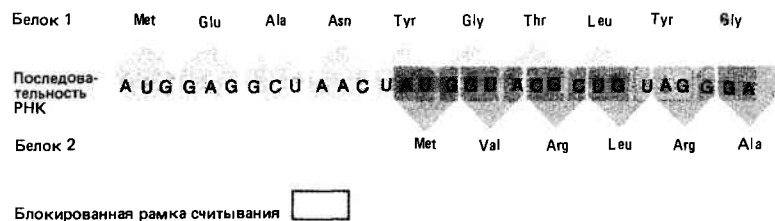


Рис. 4.9. В одной последовательности ДНК одновременно могут существовать две открытые рамки считывания (третья рамка считывания блокирована).

к аминокислотным заменам в одном или двух белках. Мутации, затрагивающие один белок, будут относиться к одной группе комплементации; мутации, затрагивающие другой белок, будут принадлежать к другой, но перекрывающейся группе, а некоторые мутации могут входить в обе группы.

Это редкая ситуация, и возникает она, вероятно, в результате того, что вирусам приходится кодировать необходимое количество белков, не увеличивая размера генома. Труднее объяснить подобные случаи у млекопитающих и в митохондриях. Каким образом развиваются перекрывающиеся рамки считывания и что их ограничивает, не совсем ясно.

Рекомендуемая литература

Открытие генетического кода посвящено бесчисленное количество обзоров. Подробное изложение можно найти в книге Ичас М. (Ycas M., *The biological code*, North Holland, Amsterdam, 1969). [Имеется перевод: Ичас М., *Биологический код*—М.: Мир, 1971.] Краткое, но достаточно полное обсуждение вопроса дано в первой главе книги Льюина (Lewin, *Gene Expression, 1, Bacterial genomes*, Wiley, New-York, 1974). Природа мутаций со сдвигом рамки рассматривается в специальном обзоре Рота (Roth, *Ann. Rev. Genet.*, 8, 319–346, 1974).

Глава 5

ОТ ГЕНА К БЕЛКУ

«Центральная догма [молекулярной биологии] гласит, что, раз проделав свой путь от гена к белку, «информация» не может вернуться в обратном направлении. Информация может быть перенесена от одной нуклеиновой кислоты к другой или от нуклеиновой кислоты к белку, но перенос информации от белка к белку или от белка к нуклеиновой кислоте невозможен. Информацией мы здесь называем точную последовательность оснований в нуклеиновой кислоте или аминокислотных остатков в белке».

Этими словами в 1958 г. Крик (Crick) провозгласил правило, называемое с тех пор **центральной догмой** молекулярной биологии. Генетическая информация хранится в последовательности ДНК (или у некоторых вирусов в РНК). В результате транскрипции с последующей трансляцией она может превратиться в аминокислотные последовательности белков.

Схематическое изображение центральной догмы молекулярной биологии приведено на **рис. 5.1**. До некоторого

времени считалось, что весь этот процесс необратим: если информация уже транскрибирована в РНК, ее невозможно вернуть в исходное состояние для использования в качестве генетического материала; ее можно лишь использовать для трансляции в белок. С тех пор, однако, был открыт процесс **обратной транскрипции**, действующий у РНК-содержащих онкодеродных вирусов. Геном вируса, состоящий из одноцепочечной РНК, может превратиться в двухцепочечную ДНК, встроиться в геном клетки-хозяина и, слившись с ним, наследоваться подобно остальным генам. Были обнаружены также РНК-содержащие вирусы, способные реплицироваться в форме РНК, используя комплементарное спаривание оснований. При этом иногда образуется стабильная двухцепочечная РНК, а иногда — нет. Хотя ни обратная транскрипция, ни репликация РНК обычно, по всей видимости, не используются самой клеткой, сам факт их существования отчетливо указывает, что при наличии подходящих ферментов вполне возможна передача и воспроизведение информации между различными и идентичными формами нуклеиновых кислот.

Синтез белка происходит в рибосомах

Тот факт, что синтез белка происходит в рибосомах, впервые был установлен в опытах с введением крысам меченых аминокислот. Через различные промежутки времени после инъекции у животных удаляли часть печени и исследовали распределение метки в различных фракциях клетки. По истечении нескольких минут после инъекции метку обнаружили во фракции, содержащей рибосомы. Затем она переместилась в клеточный сок (т.е. в общую цитоплазматическую фракцию). Из этого был сделан вывод о том, что аминокислоты собираются в полипептидную цепь на рибосомах и после завершения синтеза отделяются от них.

Рибосомы — это компактные **рибонуклеиновые частицы**, построенные из двух субчастиц. Каждая субчастица в свою очередь построена из нескольких белков, связанных с одной молекулой РНК. Молекулы РНК, входя-

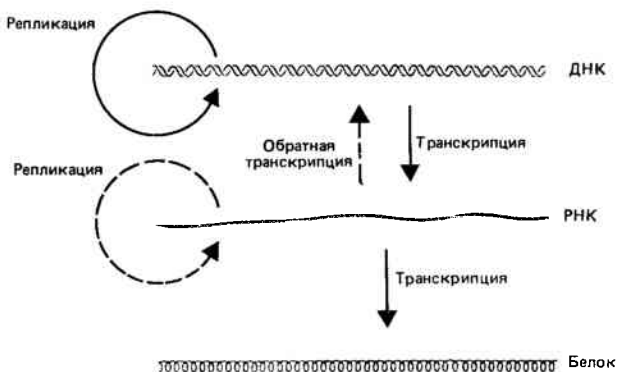


Рис. 5.1. Согласно центральной догме молекулярной биологии, генетическая информация, заключенная в нуклеиновой кислоте, может передаваться в ряду поколений, но при этом перенос информации происходит лишь в направлении нуклеиновая кислота → белок. Передача информации в обратном направлении — от белка к нуклеиновой кислоте — невозможна.

Процессы, происходящие обычно в клетке, указаны сплошными линиями, а те, которые характерны только для вирусных геномов, показаны пунктиром.

щие в состав рибосом, называют **рибосомными РНК** или сокращенно **рРНК**. При понижении концентрации ионов магния рибосомы диссоциируют на субчастицы. Типичная концентрация Mg^{+} для разделения бактериальных рибосом — 1,5 мМ. Рибосомы (и их субчастицы) бактерий отличаются по размеру от цитоплазматических рибосом эукариот. Впервые это было обнаружено по различию в скорости седиментации рибосом этих двух типов. (Скорость седиментации измеряют в единицах Сведберга, сокращенно S. Чем больше масса, тем выше скорость седиментации и выше значение S. Форма частиц также влияет на скорость седиментации, так как более компактные частицы седиментируют быстрее.)

Бактериальные рибосомы обычно седиментируют при 70S, а отдельные субчастицы — при 50S и 30S. Большая субчастица имеет почти сферическую форму и по размеру примерно вдвое больше маленькой асимметричной субчастицы.

Цитоплазматические рибосомы высших эукариот больше бактериальных и обычно седиментируют при 80S. Скорость седиментации отдельных субчастиц составляет 60 и 40S.

Другие формы рибосом обнаружены в митохондриях и хлоропластах. Они широко варьируют по размеру и могут быть значительно меньше бактериальных рибосом (например, размер рибосом митохондрий человека или лягушки — 60S) или занимать промежуточное положение между бактериальными рибосомами и цитоплазматическими рибосомами эукариот. (Размер митохондриальных рибосом у грибов равен примерно 74S.)

Приведенные выше размеры рибосом можно считать «номинальными». Условно размер бактериальных рибосом обозначают «70S», а размер цитоплазматических рибосом эукариот — «80S», однако реальная скорость седиментации рибосом обоих типов варьирует в зависимости от конкретных условий.

Поиски посредника

Расшифровка генетического кода показала, что генетическая информация хранится в виде нуклеотидных триплетов. Однако оставалось неясным, каким образом каждый кодон транслируется в соответствующую аминокислоту. Представление о том, что для реализации информации нужно дешифровать код, развивалось одновременно с идеей об обязательном участии **матрицы** в процессе трансляции. В эукариотической клетке ядро, содержащее генетический материал, и цитоплазма, в которой синтезируется белок, пространственно разобщены. Из этого следует, что ДНК *сама по себе* не может служить матрицей.

Поскольку уже была очевидна связь между количеством РНК и уровнем белкового синтеза в клетке, возникла мысль, что роль посредника между ДНК и белком может выполнять РНК. Сначала считали, что посредник должен, по-видимому, быть составной частью рибосомы, и были недалеки от истины, хотя некоторые тонкости при этом не учитывались. Но поскольку рибосомы содержат и белок, и РНК, думали, что, возможно, рибосомы специализированы благодаря наличию в них различных РНК, используемых в синтезе соответствующих белков. Но, как теперь уже известно, все рибосомы в клетке имеют одинаковый состав — как в отношении белка, так и в отношении РНК. Инструкции о синтезе определенных белков поступают в РНК с молекулами другого типа (**ин-**

формационной, или **матричной**, **мРНК**), которая временно связывается с рибосомой.

Поначалу пытались обнаружить посредника у бактерий, но оказалось, что выделить его довольно трудно. Бактериальная мРНК нестабильна и существует только в течение короткого промежутка времени. Первые данные о существовании посредника были получены при изучении бактерий, зараженных фагом T2. Затем эти данные подтвердились на незараженных бактериях.

При фаговой инфекции не удается заметить каких-либо изменений ни в рибосомах, содержащих около 80% всей клеточной РНК, ни в общем содержании РНК в клетке. Однако сразу после проникновения в клетку фаговой ДНК синтезируется небольшое количество РНК, которая быстро распадается. Состав этой новой РНК больше напоминает ДНК фага, чем ДНК бактерии.

Были проведены эксперименты другого рода — с использованием кратковременного введения радиоактивной метки. Это позволило определить включение предшественников в РНК незараженных бактерий. Необходимость в кратковременной (импульсной) метке была обусловлена быстрым распадом мРНК. При более продолжительном периоде меченая мРНК успевает деградировать прежде, чем удастся измерить включенную в нее метку. При использовании импульсной метки можно успеть произвести измерения раньше, чем РНК деградирует.

И в зараженных, и в незараженных бактериях ново-синтезированная мРНК связывается с рибосомами, но вскоре от них отделяется и деградирует (**рис. 5.2**). В этих опытах продемонстрирован важный факт, что в ходе фаговой инфекции белковый синтез происходит на предсуществовавших рибосомах бактериальной клетки, которые связывают нестабильную РНК как в неинфицированных бактериях, так и в бактериях, инфицированных фагом.

Фаговая мРНК по своей последовательности соответствует последовательности ДНК фага. Были проведены опыты, в которых исследовали способность меченой мРНК гибридизоваться с ДНК фага. Поскольку ДНК фага T2 оказалось возможным денатурировать и разделить на отдельные цепи, удалось показать, что в мРНК воспроизводится последовательность только одной из цепей. С ДНК клетки-хозяина фаговая мРНК не гибридизуется совсем.

Таким методом, не прибегая к выделению мРНК в чистом виде, удалось показать, что существует очень нестабильная РНК, которая по своей последовательности соответствует одной из цепей ДНК и которая связывается с рибосомами. Мы можем предположить, что транскрипция происходит так, как это изображено на **рис. 5.3**. Чтобы спаривание оснований было возможным, ДНК должна быть расплетена в соответствующем участке. Одна из расплетенных цепей при этом используется в качестве матрицы. По ходу транскрипции расплетенный участок передвигается вдоль ДНК. Затем РНК отделяется от ДНК, и прежняя двухцепочечная структура восстанавливается. Все нуклеиновые кислоты синтезируются в направлении от 5'-конца к 3'-концу. Следовательно, транскрипция и трансляция у бактерий происходят в одном направлении. А это означает, что у бактерий трансляция мРНК может начаться раньше, чем завершится транскрипция.

Попытки обнаружить посредника у эукариот также встретились с определенными трудностями. Здесь, как и у бактерий, мРНК составляет только небольшую часть общей клеточной РНК (примерно 3%). Первой изолиро-

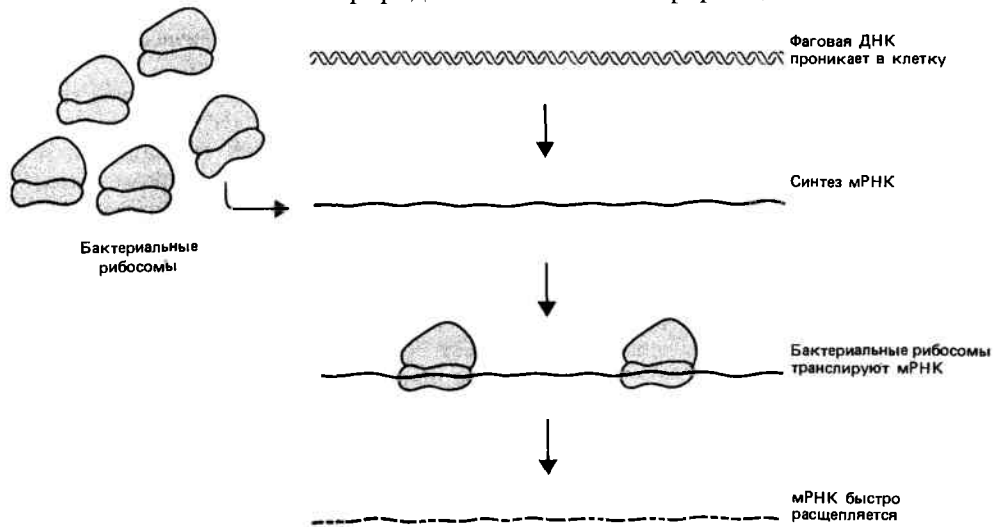


Рис. 5.2. Матричная РНК бактерий быстро деградирует. Когда фаговая ДНК проникает в клетку, на ней начинается транскрипция новых мРНК, которые транскрибируются бактериальными рибосомами,

существовавшими в клетке еще до инфицирования фагом. Однако вскоре мРНК деградирует, а рибосомы могут быть использованы повторно для трансляции других мРНК.

ванной индивидуальной мРНК была глобиновая мРНК эритроцитов. С тех пор было выделено много индивидуальных мРНК, и теперь получение какой-либо определенной мРНК стало обычной процедурой. Эти мРНК, по крайней мере в течение нескольких часов, сохраняют стабильность. Поэтому их можно выделять в интактном состоянии и транскрибировать *in vitro*, если добавить в систе-

му рибосомы и другие необходимые компоненты. При необходимости доказать, что данная мРНК соответствует определенному белку, этот белок следует синтезировать *in vitro* с использованием данной мРНК в качестве матрицы. Затем для полной уверенности можно сравнить нуклеотидную последовательность этой РНК с аминокислотной последовательностью данного белка.

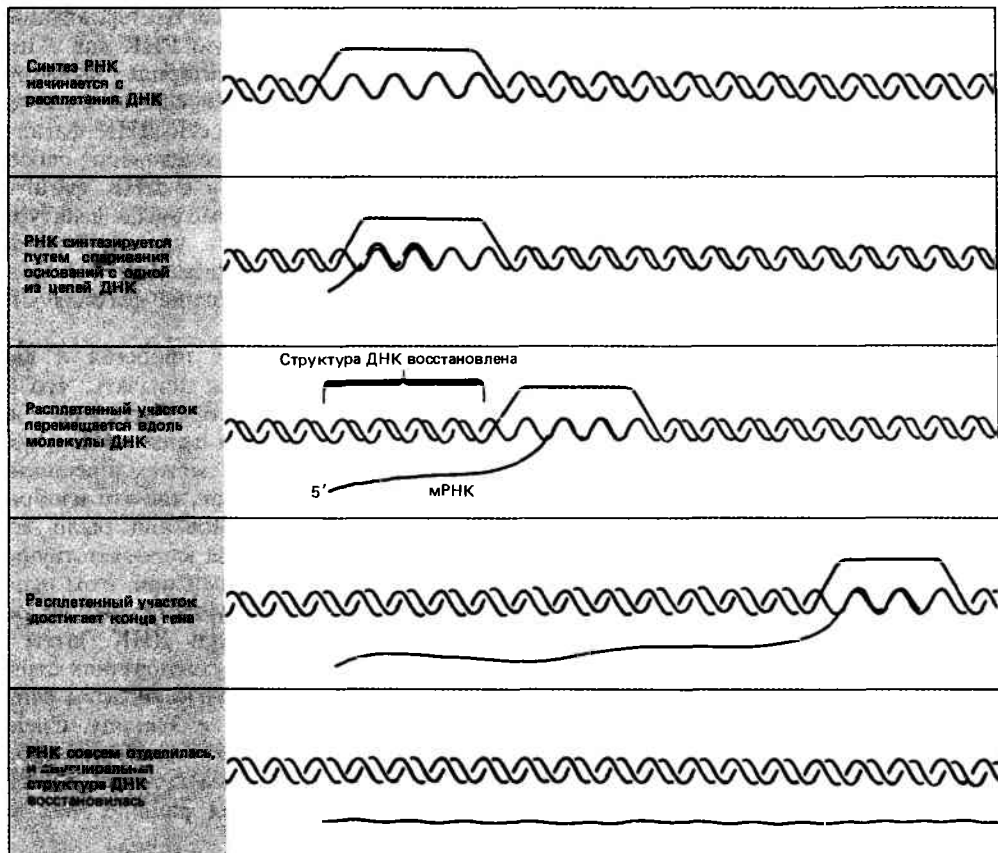


Рис. 5.3. Синтез мРНК происходит путем спаривания оснований на одной из цепей ДНК и сопровождается локальным расплетением двойной спирали.

Расплетанный участок перемещается по ДНК слева направо, и по мере его передвижения происходит восстановление двуспиральной структуры и отделение одноцепочечной мРНК.

Транспортная РНК – адаптор

Каким образом происходит трансляция тринуклеотидной последовательности мРНК в соответствующую аминокислоту белка? Это ключевой момент в процессе передачи информации. Мы уже упоминали о том, что существовавшая ранее идея о возможном стереохимическом взаимодействии участка нуклеиновой кислоты и его белкового продукта была отвергнута и все внимание сосредоточилось на поиске посредников в этом процессе. Оказалось, что транспортная РНК играет роль «адаптора» и выполняет сразу две функции: узнает и кодон, и соответствующую аминокислоту. Первоначально тРНК идентифицировали как фракцию РНК, седиментирующую при 4S, и назвали ее «растворимой», исходя из ее маленького размера. Типичные тРНК состоят из 75–85 нуклеотидов.

Нуклеотидную последовательность каждой тРНК можно изобразить в виде **клеверного листа**, как показано на **рис. 5.4**. Участки молекулы, комплементарные друг другу, спариваются и образуют «**стебли**», а одноцепочечные участки остаются в виде **петель**. Часть молекулы, состоящую из стебля и петли, называют «**шпилькой**».

Каждая тРНК может **присоединять** только ту аминокислоту, которой соответствует ее антикодон. Связывание аминокислоты происходит с образованием эфирной связи, в которой участвуют ее карбоксильная группа и одна из гидроксильных групп рибозы последнего нук-

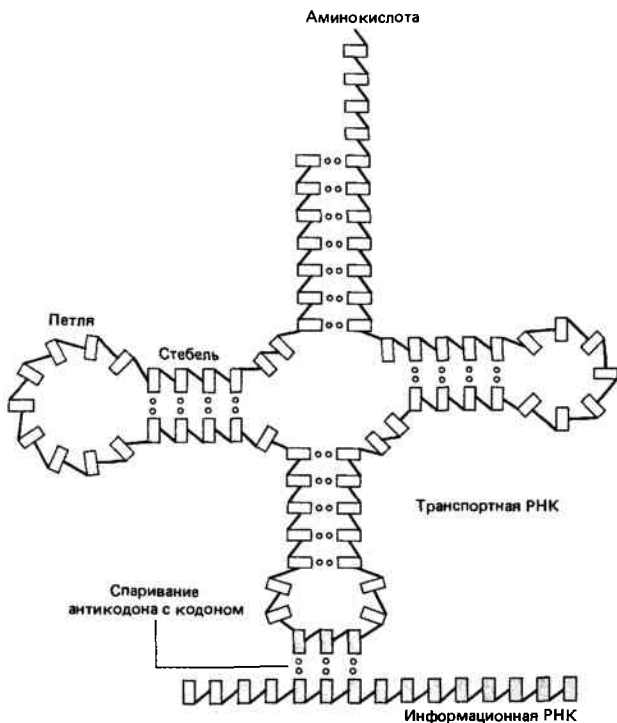


Рис. 5.4. Транспортная РНК обладает двумя основными свойствами адаптора.

Остатки сахара изображены в виде прямоугольников, а соединяющие их линии обозначают фосфодиэфирные связи. Стебли шпильки образованы спариванием комплементарных оснований (показано маленькими светлыми кружками). На конце каждой шпильки находится одноцепочечная петля. В нижней петле расположен триплет, комплементарный кодону той аминокислоты, которую представляет данная тРНК. Этот антикодон узнает кодон путем спаривания оснований. Аминокислота присоединяется к последнему нуклеотиду 3'-конца, который изображен в верхней части рисунка.

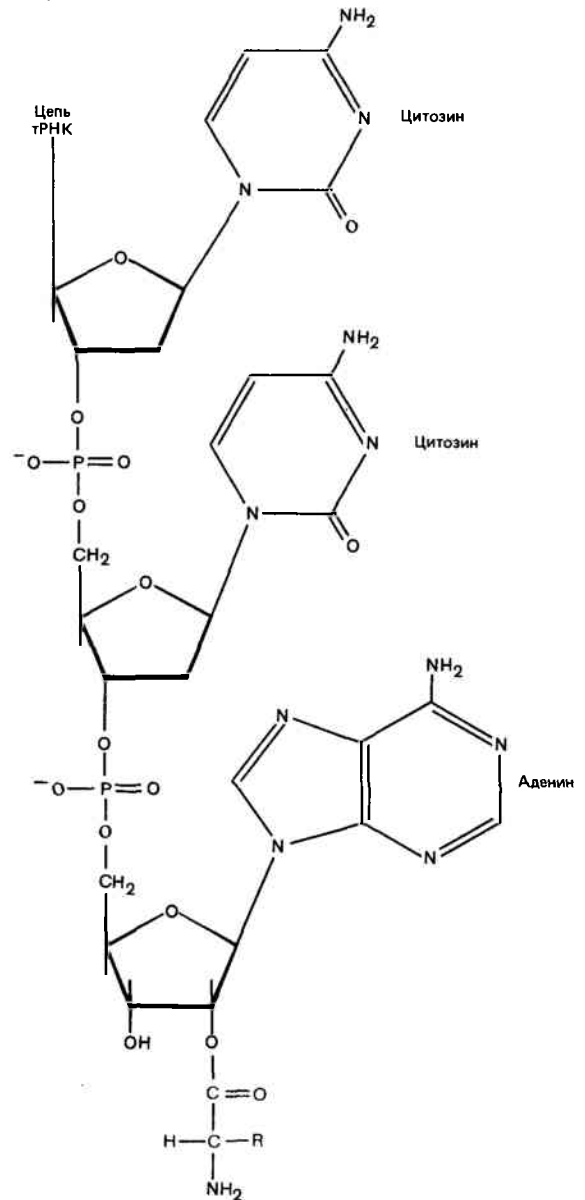


Рис. 5.5. Аминоацил-тРНК состоит из аминокислоты, соединенной (обычно) эфирной связью с концевой 2'-ОН-группой тРНК, на конце которой всегда находится последовательность из трех оснований – ССА.

леотида тРНК (обычно аденилата). Структура, которая при этом образуется, показана на **рис. 5.5**.

Обычно аминокислота присоединяется к 2'-ОН-группе, но в некоторых случаях она может связываться с 3'-ОН-группой. Вопрос о том, какая из гидроксильных групп использована при образовании связи в каждом отдельном случае, остается открытым, так как 2'- и 3'-формы аминоацил-тРНК в растворе находятся в состоянии равновесия. Такое же равновесие, вероятно, существует и в клетке.

Образование аминоацил-тРНК катализируется особым ферментом – **аминоацил-тРНК-синтегазой**. Для каждой аминокислоты имеется свой фермент, узнающий только одну аминокислоту и *все типы РНК*, к которым ее можно присоединить. Таких (изоакцепторных) тРНК может быть много: кроме тех тРНК, которые узнают различные кодоны-синонимы, может быть несколько видов тРНК, реагирующих с одним и тем же кодоном.

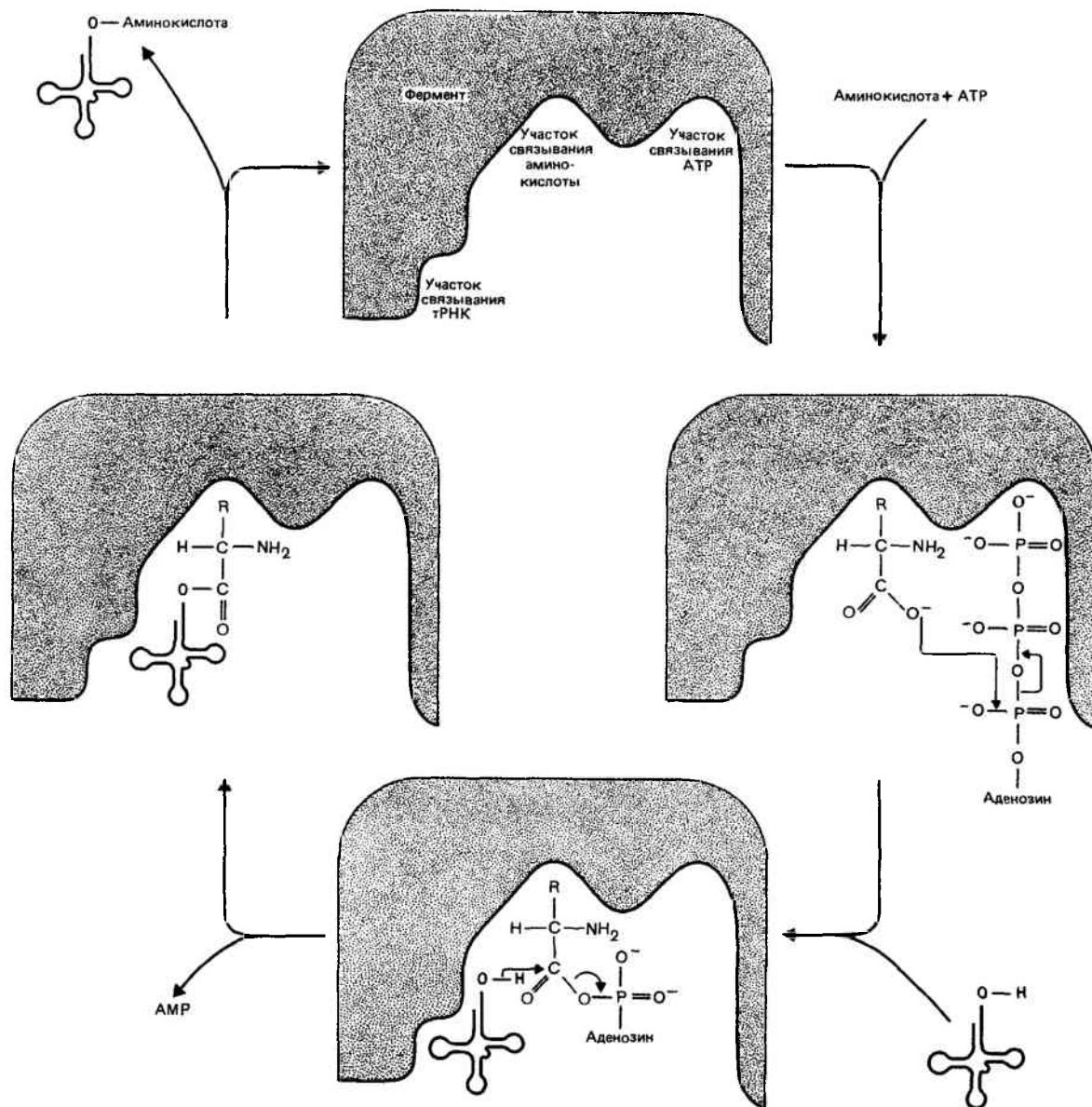


Рис. 5.6. Аминоацил-тРНК—синтетаза присоединяет аминокислоту к тРНК.

В ферменте имеются специальные участки для соединения с тРНК, аминокислотой и АТР. В первой реакции участвуют аминокислота и АТР и образуется аминоксил-АМР. Затем происходит связывание с тРНК

и перенос на нее аминокислоты. После отделения аминоксил-тРНК фермент приступает к следующему циклу. Существуют аминоксил-тРНК—синтетазы для всех аминокислот. Каждый фермент имеет специфические участки для R-группы аминокислоты и соответствующей тРНК.

Процесс аминокислотирования включает реакцию между аминокислотой и АТР с образованием аминоксил-аденилата; затем активированная аминокислота переносится на тРНК. Как показано на рис. 5.6, при этом образуются АМР и пирофосфат. Необходимая для реакции энергия высвобождается при расщеплении АТР. Двухступенчатый механизм этой реакции подтверждается следующими данными: при добавлении в реакционную смесь только фермента, АТР и аминокислоты образуется первый комплекс, а собственно аминокислотирование происходит при добавлении тРНК.

Специфическое узнавание кодона возможно благодаря наличию в тРНК антикодона — тринуклеотидной последовательности, комплементарной кодону. Достаточно ли этого для узнавания кодона? В ранних работах было по-

казано, что в сформированной аминоксил-тРНК антикодон узнает только правильный кодон. Рассмотрим пример, показанный на рис. 5.7. Как видно из рисунка, если подвергнуть восстановительному десульфированию цистеинил-тРНК, то остаток цистеина в этой аминоксил-тРНК превратится в аланин и в результате из цистеинил-тРНК образуется аланил-тРНК_{1^{Cys}}. Для условного наименования тРНК используют трехбуквенное сокращение соответствующей ей аминокислоты (в верхнем индексе) и соответствующую цифру (в нижнем индексе), если есть несколько тРНК для одной аминокислоты. Этот образованный из цистеина остаток аланина, будучи присоединенным к цистеиновой тРНК, включается в белок вместо цистеина. Таким образом, функция аминоксил-тРНК в белковом синтезе точно определяется ее антикодоном.

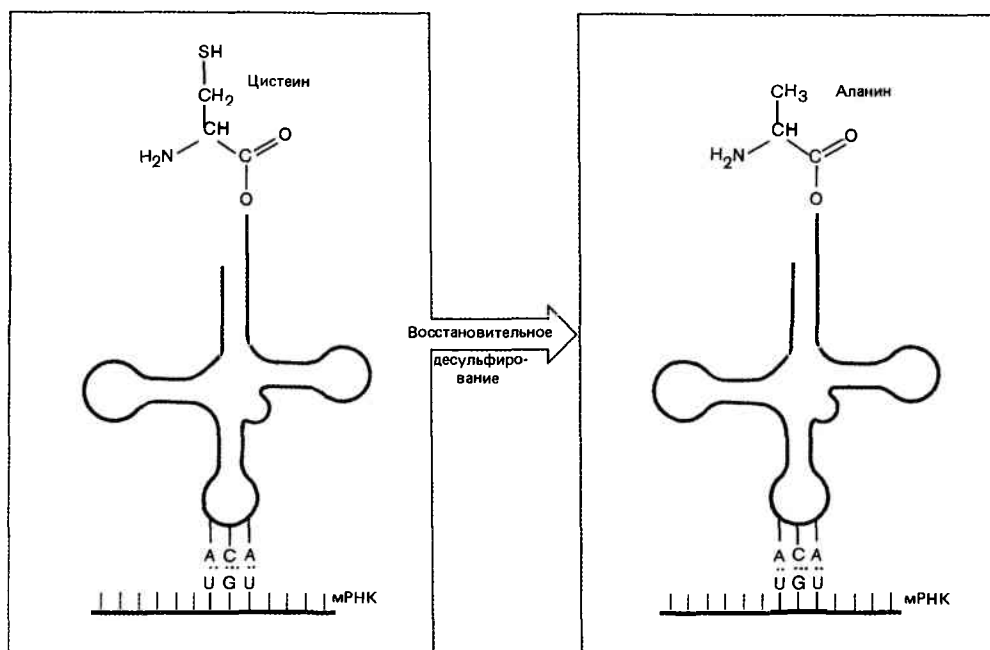


Рис. 5.7. Специфичность тРНК определяется ее антикодоном, а не аминокислотой.

Антикодон у цистеинил-тРНК спаривается с кодоном UGU. После хими-

ческой модификации остаток цистеина превращается в аланин. Однако тРНК по-прежнему узнает кодон UGU, но теперь вместо цистеина в белок включается аланин.

Рибосомы передвигаются как конвоиры

Взаимодействие между тРНК и мРНК происходит с участием рибосом, обеспечивающих условия, необходимые для кодон-антикодонного узнавания. В процессе синтеза белка рибосомы **движутся** вдоль мРНК, последовательно считывая триплеты путем присоединения к ним аминоксил-тРНК, несущих нужные антикодоны. Каждый этап наращивания полипептидной цепи удлиняет ее на одну аминокислоту.

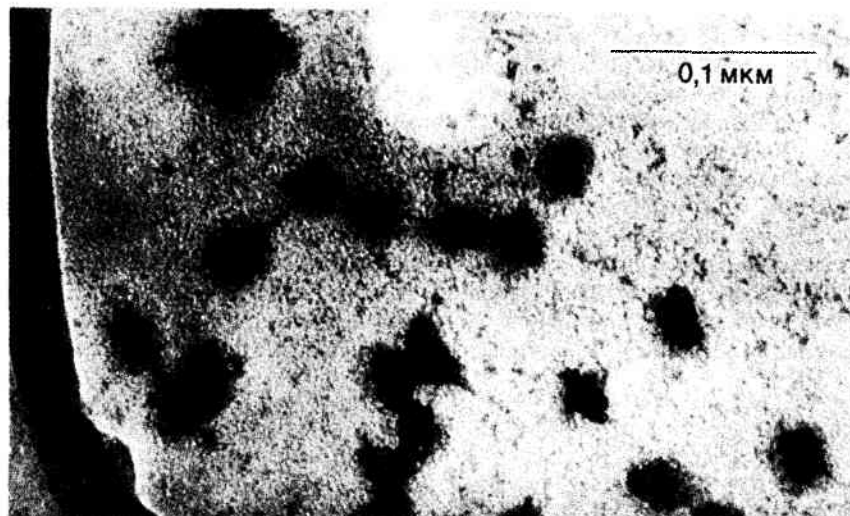
Активная фракция рибосом, связанная с мечеными аминокислотами, осаждается с более высокой константой седиментации, чем обычно, — больше 70S (или 80S). Быстро седиментирующая единица состоит из одной мРНК, связанной с несколькими рибосомами. Эти комплексы называют **полирибосомами** или **полисомами**.

Классическая картина показана на электронной микрофотографии (рис. 5.8). Видна пентасома, образованная в процессе синтеза глобина в эритроцитах. Каждая пентасома состоит из пяти рибосом, связанных с цепью мРНК. Рибосомы находятся в различных положениях на мРНК. Те рибосомы, которые расположены ближе к одному концу, еще только начали синтезировать белок, а те, которые ближе к другому концу, почти закончили синтез полипептида.

Рибосома, движущаяся вдоль мРНК, прикреплена к полипептидной цепи, которая постепенно удлиняется с прибавлением каждой следующей аминокислоты. За время своего путешествия по мРНК каждая рибосома синтезирует только одну полипептидную цепь. Таким образом, по мРНК от 5'-конца к 3'-концу движется серия рибосом, несущих белковые продукты увеличивающегося

Рис. 5.8. Синтез белка происходит в полирибосомах.

На электронных микрофотографиях видны пентасомы, синтезирующие глобин. Рибосомы имеют почти сферическую форму — около 7 нм (70 Å) в диаметре — и связаны с нитью мРНК. Расстояние между отдельными рибосомами в полисомах зависит от условий. Фотография любезно предоставлена А. Ричем (Alex Rich).



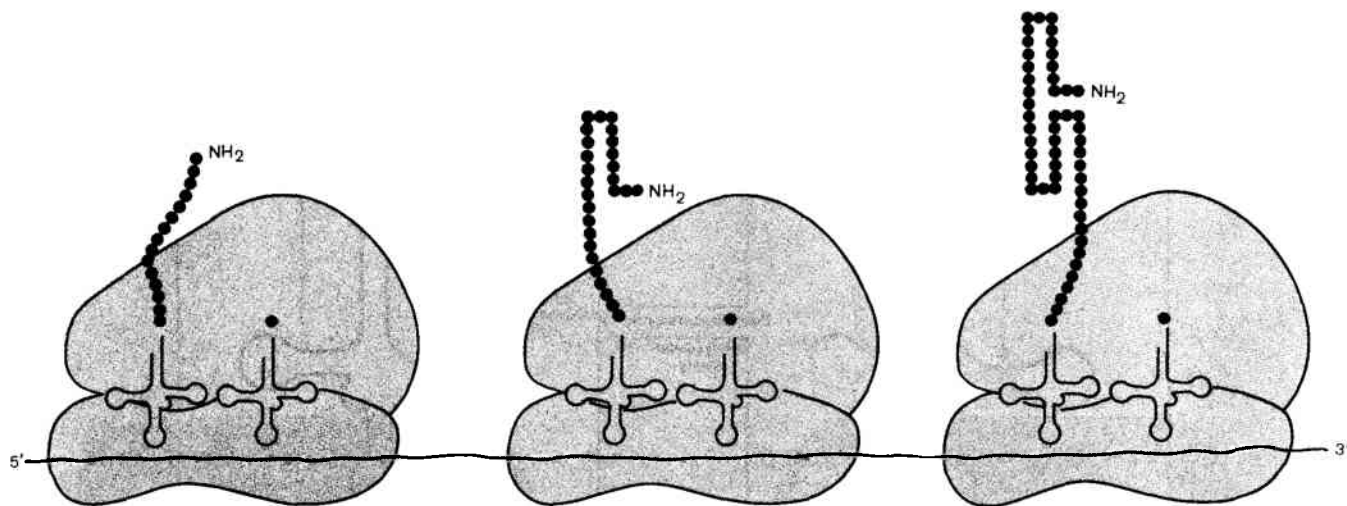


Рис. 5.9. Полирибосомы состоят из нескольких рибосом, транслирующих мРНК, и передвигаются в направлении от 5'-конца к 3'-концу.

В каждой рибосоме находятся две молекулы тРНК. Первая из них со-

единена с синтезированной к этому времени полипептидной цепью, а вторая несет следующую аминокислоту, которую нужно присоединить к полипептиду.

размера (рис. 5.9). Примерно 30–35 аминокислот, присоединенных последними к растущей полипептидной цепи, покрыты рибосомной частицей. По-видимому, часть белка, синтезированная ранее, находится вне рибосомы и начинает свертываться, приобретая характерную конформацию. (Об этом свидетельствует узнавание специфическими антителами белков, еще находящихся в полисомах.)

Размер полисомы зависит от многих факторов. У бактерий они иногда очень большие и состоят из десятков рибосом, одновременно занятых трансляцией. Частично это зависит от большого размера мРНК (которые могут кодировать более одного белка). Частично это связано с эффективностью трансляции мРНК рибосомами. Поскольку рибосомы могут прикрепляться к бактериальной мРНК раньше, чем завершится ее транскрипция, полисомы, вероятно, связаны с ДНК.

В эукариотической клетке мРНК должна обязательно переместиться из ядра в цитоплазму к рибосомам. Полисомы эукариот обычно бывают меньшего размера, чем у бактерий, но и в этом случае их размер зависит от длины мРНК и от того, с какой частотой к ней прикрепляются рибосомы. Считается, что к эукариотическим мРНК одновременно присоединяется менее 10 рибосом.

Ни у бактерий, ни у эукариот число рибосом на каждой молекуле мРНК, синтезирующей белок, точно не предопределено. Оно выражается статистической функцией, зависящей от размеров мРНК и эффективности синтеза.

Общая характеристика содержания различных компонентов бактериальной клетки, имеющих отношение к белковому синтезу, дана в табл. 5.1. Около 20000 рибосом составляют ощутимую долю общей клеточной массы. Молекулы тРНК численно превышают количе-

Таблица 5.1

Основные компоненты бактериальной (*E. coli*) клетки

Компонент	Доля от сухого веса клетки, %	Количество в расчете на одну клетку
Клеточная стенка	10	
Мембрана	10	
ДНК	2	1
мРНК	2	
тРНК	3	160 000
рРНК	21	
Рибосомный белок	9	20 000
Растворимый белок	42	10^6
Низкомолекулярные метаболиты	1	$6,5 \cdot 10^6$

ство рибосом почти в десять раз. Большинство тРНК находится в аминоацилированном состоянии в полной готовности к белковому синтезу. Из-за нестабильности молекул мРНК трудно подсчитать их число. Предположительно в клетке содержится 2000–3000 молекул мРНК на разных стадиях синтеза и деградации.

Рекомендуемая литература

Существует довольно мало обзоров, рассматривающих общие процессы белкового синтеза с современных позиций. (Большинство обзоров посвящено отдельным аспектам, и ссылки на них приводятся в гл. 6–9 этой книги.) Общее описание этого процесса можно найти в гл. 2 книги Льюина, т. 1 (*Lewin, Gene Expression, 1, Bacterial Genomes, Wiley, New York, 1974*).

Часть II

СИНТЕЗ БЕЛКОВ

Результаты [наших] экспериментов с радиоактивной меткой оказались даже более удивительными в свете представлений того времени [1955] о биосинтезе белков и их локализации в клетке. Работа Шоенхеймера (Schoenheimer) [1942] фактически убедила большинство биохимиков в том, что белкам организма присуще «динамическое» состояние, при котором каждая молекула непрерывно разрушается и реконструируется путем обмена аминокислотных остатков. Однако наши эксперименты показывали, что *in vivo* в условиях нормального роста бактерий β -галактозидаза полностью стабильна, как и другие бактериальные белки... Эти исследования получили новое истолкование в свете тех перспектив, которые открылись в биологии незадолго до 1955 г. В 1953 г. Уотсон и Крик (Watson, Crick) предложили свою модель структуры ДНК. Между тем еще годом раньше Сэнгер (Sanger) установил первичную последовательность инсулина, и уже было известно ... что генетическая мутация может вызывать некоторые изменения в структуре белка. В 1954 г. на основе теории генетического кода [было высказано предположение, что] первичная структура белков предопределяется и детерминируется линейной последовательностью нуклеотидов в ДНК.

Жан Моно, 1965

Глава 6

КОНВЕЙЕР ДЛЯ СБОРКИ ПОЛИПЕПТИДНЫХ ЦЕПЕЙ

Синтез белков до некоторой степени напоминает сборочный конвейер, в котором рибосомы все время передвигаются относительно информационной РНК, доставляя аминоксил-тРНК – реальные строительные блоки для сборки белковых молекул. Рибосома представляет собой маленькую фабрику, в которой компактно упакованные белки и тРНК образуют несколько активных центров, способных осуществлять многочисленные каталитические функции. Различные группы дополнительных факторов участвуют в работе рибосомы на каждой из трех стадий белкового синтеза: инициации, элонгации и терминации. Энергия для биосинтеза белка обеспечивается гидролизом GTP.

При **инициации** происходят события, предшествующие образованию пептидной связи между двумя первыми аминокислотами белка. Для инициации необходимо, чтобы рибосома связалась с информационной РНК, образовав иницирующий комплекс совместно с первой аминоксил-тРНК. Это относительно медленная стадия в синтезе белка.

Элонгация включает в себя все реакции – от образования первой пептидной связи до присоединения последней аминокислоты к белковой молекуле. Это наиболее быстрая стадия белкового синтеза, во время которой рибосома перемещается от первого до последнего кодона на информационной РНК.

Терминация состоит из последовательных этапов, необходимых для освобождения синтезированной полипептидной цепи; при этом рибосома отделяется от мРНК. По сравнению со временем, требующимся для включения аминокислоты в полипептидную цепь, диссоциация происходит медленно.

Только в сравнении со скоростью элонгации и можно считать, что стадии инициации и терминации происходят медленно. Синтез белка – это быстрый процесс (хотя скорость в большой степени зависит от температуры). У бактерий при 37°C скорость элонгации варьирует: в растущую полипептидную цепь за 1 с включается от 12 до 17 аминокислот. Конкретная величина скорости элонгации зависит от условий роста клеток. Для синтеза среднего белка размером в 300 аминокислот требуется около 20 с. В синтезе белка одновременно участвует примерно 80% бактериальных рибосом; следовательно, в свободном состоянии находится лишь небольшая их часть. В эукариотических клетках скорость белкового синтеза ниже; так, в ретикулоцитах при 37°C скорость элонгации составляет 2 аминокислоты в 1 с.

Большинство экспериментов по изучению стадий белкового синтеза проводилось в системе *in vitro*, содержащей рибосомы, аминоксил-тРНК, дополнительные факторы и источник энергии. В этих системах скорость белкового синтеза может быть примерно на порядок ниже, чем *in vivo*.

Функциональные участки рибосомы

Информационная РНК взаимодействует с малой субчастицей рибосомы. При этом в связывание вовлечен уча-

сток мРНК длиной из 30 оснований. Но в каждый данный момент только две молекулы тРНК могут разместиться в рибосоме. Поэтому в синтезе белка непосредственно участвуют только два из десяти кодонов, находящихся в рибосоме.

На **рис. 6.1** показано, что каждая тРНК располагается в отдельном участке. Эти два участка характеризуются различными свойствами.

Тот участок, который может быть занят очередной (новопоступающей) аминоксил-тРНК, называется **А-участком** (или **акценторным участком**). До того как в рибосому поступит аминоксил-тРНК, в этот участок входит триплет, кодирующий ту аминокислоту, которая должна быть включена в полипептидную цепь.

Кодон, соответствующий *последней* аминокислоте, уже успевшей включиться в полипептид, располагается в **Р-участке** (или **донорном участке**). Этот участок занят комплексом **пептидил-тРНК**, представляющим собой тРНК, несущую целую полипептидную цепь, синтезированную к данному моменту (на рис. 6.1 стадия 1).

Когда оба участка заняты, происходит образование пептидной связи в результате следующей реакции: полипептидный компонент комплекса пептидил-тРНК переносится на аминокислоту, связанную в аминоксил-тРНК (реакция транспептидации). Эта реакция происходит на *большой* субчастице рибосомы.

Таким образом, тот конец тРНК, который несет аминокислоту, расположен на *большой* субчастице, тогда как другой ее конец, т.е. антикодон, взаимодействует с мРНК, связанной *малой* субчастицей. Следовательно, расположение участков Р и А таково, что в их образовании принимают участие обе рибосомные частицы (рис. 6.1).

В результате переноса полипептида в рибосоме создается следующая ситуация: тРНК, теперь уже лишенная аминокислоты (деацилированная тРНК), продолжает удерживаться в Р-участке, а новая пептид-тРНК сидит в А-участке (стадия 2). Эта пептидил-тРНК на один аминокислотный остаток длиннее, чем та, которая была в Р-участке на стадии 1.

Затем рибосома перемещается на один триплет вдоль информационной РНК. При этом деацилированная тРНК вытесняется из рибосомы, а удлинившаяся на одну аминокислоту пептидил-тРНК поступает в Р-участок (стадия 3). Теперь в А-участке устанавливается следующий кодон, и в результате в рибосому может поступить новая аминоксил-тРНК. Цикл, таким образом, может повториться.

Инициация: специальная иницирующая тРНК

Почему начинается синтез белка, т.е. как распознается первый кодон в гене, являющийся стартовой точкой при трансляции?

Для осуществления процесса инициации информационная РНК должна обладать двумя характерными особенностями. Во-первых, в молекуле мРНК имеется уча-

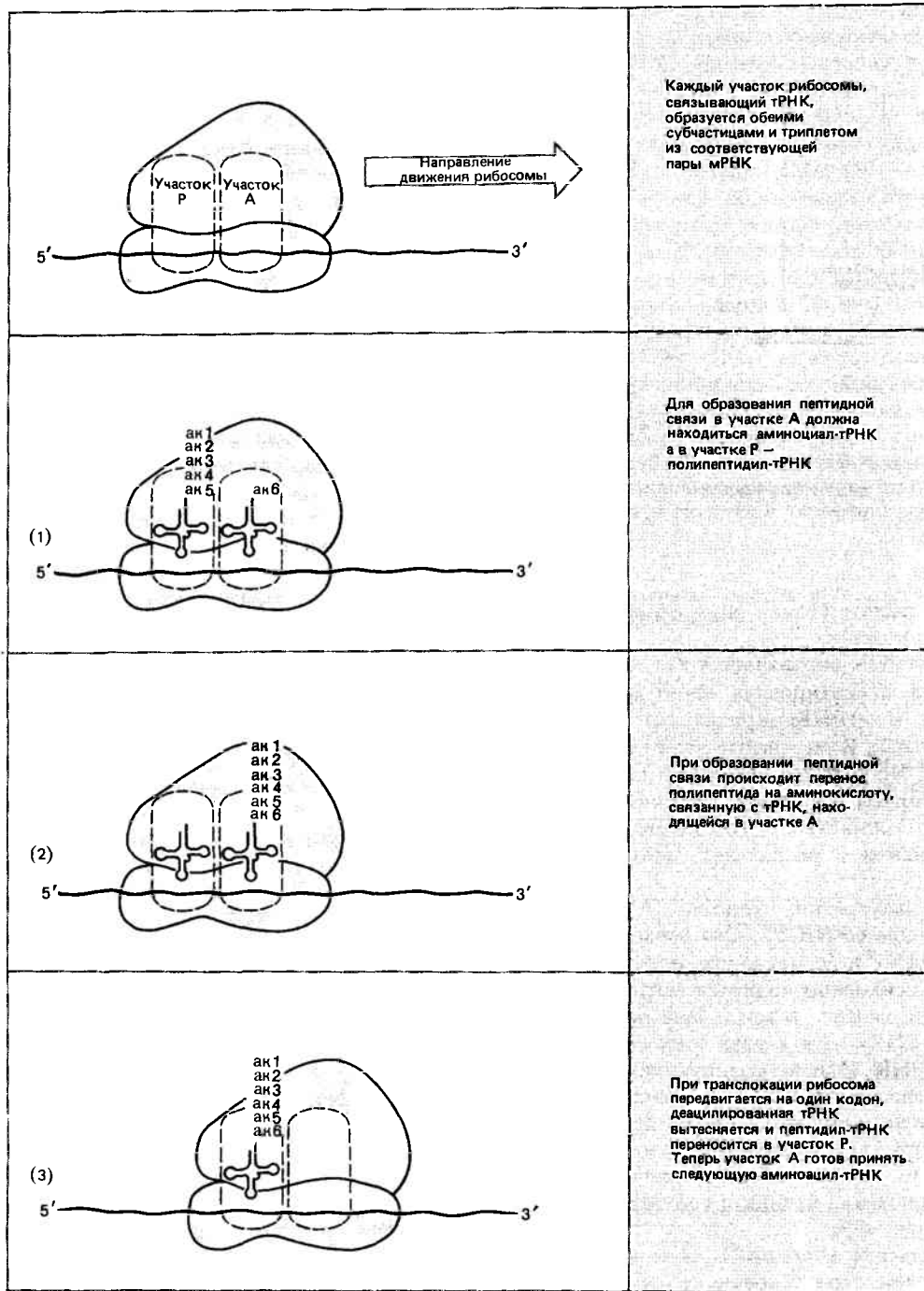


Рис. 6.1 В рибосоме имеется два участка, с которыми связываются молекулы тРНК.

сток, предшествующий началу кодирующей области, который служит сигналом для связывания с рибосомой. Природу таких сигналов мы рассмотрим в гл. 9. Во-вторых, имеется специальный кодон AUG, который является стартовой точкой и определяет рамку считывания (у бактерий вместо AUG иногда используется кодон GUG). Для инициации необходимо одновременное наличие последовательности мРНК, ответственной за связывание с рибосомой, и иницирующего кодона.

Первым указанием на то, что инициация не может происходить в *любой* точке, а требует специального меха-

низма, явились данные о том, что аминокислотный состав N-концевой части белков *E. coli* имеет определенную тенденцию. Почти у 50 белков в качестве первой аминокислоты присутствует метионин. Затем была открыта специальная *N*-формилметионил-тРНК. Оказалось, что она очень подходит на роль инициаторной тРНК, поскольку заблокированная аминогруппа не позволяет ей участвовать в элонгации, но не мешает инициировать синтез белка.

У *E. coli* существует два вида тРНК, способных акцептировать метионин. Один вид используется для инициа-

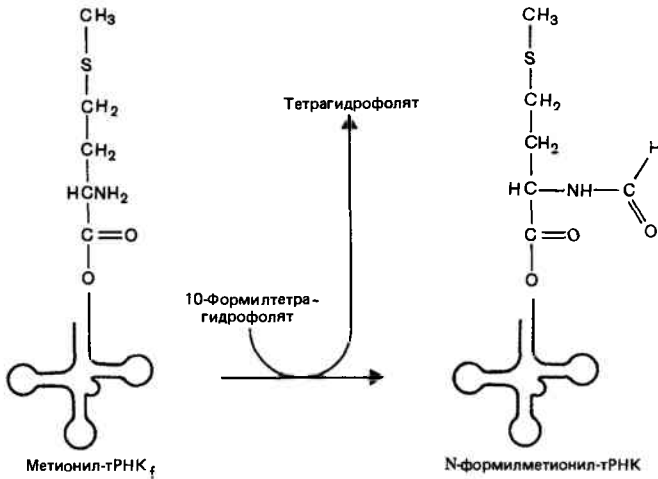


Рис. 6.2. Инициаторная N-формилметионил-тРНК ($fMet$ -тРНК $_f$) образуется в результате формилирования метионил-тРНК $_f$, используя формилтетрагидрофолят в качестве кофактора.

ции; другой — для распознавания внутренних AUG кодонов в процессе элонгации.

Инициаторная тРНК обозначается как $mPНК_f^{Met}$. После того как она акцептировала свою аминокислоту, образуя комплекс Met -тРНК $_f$, происходит реакция формилирования (рис. 6.2). В результате этой реакции блокируется свободная NH_2 -группа. Продукт реакции, обладающий инициаторной функцией, обычно сокращенно обозначается как $fMet$ -тРНК $_f$. Эта тРНК используется только для инициации и распознает либо AUG, либо GUG.

Распознавание внутренних кодонов AUG осуществляется другим видом $mPНК_m^{Met}$. Она узнает только внутренние кодоны AUG, и ее метионин не формилируется.

Триплет AUG неизменно кодирует метионин, независимо от того, расположен ли он в начале или внутри информационной РНК, хотя в двух этих случаях он узнается разными тРНК. Однако кодирующая способность триплета GUG зависит от его локализации. Когда он является иницирующим кодоном, то при инициации прочитывается как формилметионин. Когда же он расположен внутри гена, то узнается обычной валиновой тРНК, т.е. определяет включение валина в соответствии с генетическим кодом (рис. 4.5).

Конкретное значение кодонов AUG и GUG определяется их окружением. При условии что иницирующий кодон находится по соседству с сигналом, ответственным за связывание рибосомы с мРНК, происходит инициация; если же кодон встречается с уже транслирующей рибосомой, то наблюдается обычное считывание (см. следующий раздел).

Иницирование трансляции у эукариот несколько отличается от этого процесса у прокариот. Во-первых, только кодон AUG используется в качестве иницирующего; кодон GUG таковым уже не является. Если в результате мутации кодон AUG замещается кодоном GUG, то инициации не происходит. Во-вторых, иницирующая тРНК, представленная особым видом, обозначается как тРНК $_f^{Met}$, так как ее метионин не формилирован. Таким образом, различия между молекулами Met -тРНК, участвующими в инициации и элонгации, заключаются в самой тРНК. При этом Met -тРНК $_i$ используется при

инициации, а Met -тРНК $_m$ — при элонгации (однако формилаза *E. coli* может использовать *in vitro* Met -тРНК $_i$ в качестве субстрата).

В митохондриях также обнаружены две тРНК, распознающие кодон AUG, причем так же, как и в бактериях, у иницирующей тРНК метионин формилирован. Изменения в специфичности узнавания, по-видимому, позволяют AUA, а иногда AUC и AUU также использоваться в качестве инициаторных кодонов в митохондриях млекопитающих.

Во всех случаях существуют ферментные системы, удаляющие ненужные аминокислотные остатки с N-конца. В бактериальных клетках и в митохондриях формильный остаток отщепляется специальным деформирующим ферментом. В том случае, если в белке N-концевой аминокислотой является метионин, то достаточно только этой модификации. В случае бактериальных и эукариотических белков, имеющих на N-конце другие аминокислоты, метионин отщепляется с помощью аминопептидазы. Когда же возникает необходимость в обеих реакциях, они осуществляются последовательно (рис. 6.3). Реакция (реакции) удаления ненужных аминокислотных остатков происходит довольно быстро — возможно, когда образующаяся полипептидная цепь достигнет в длину 15–30 аминокислот.

В инициации принимают участие 30S-субчастицы и вспомогательные факторы

Рибосомы (70S или 80S), передвигаясь вдоль мРНК, наращивают полипептидную цепь, осуществляя синтез

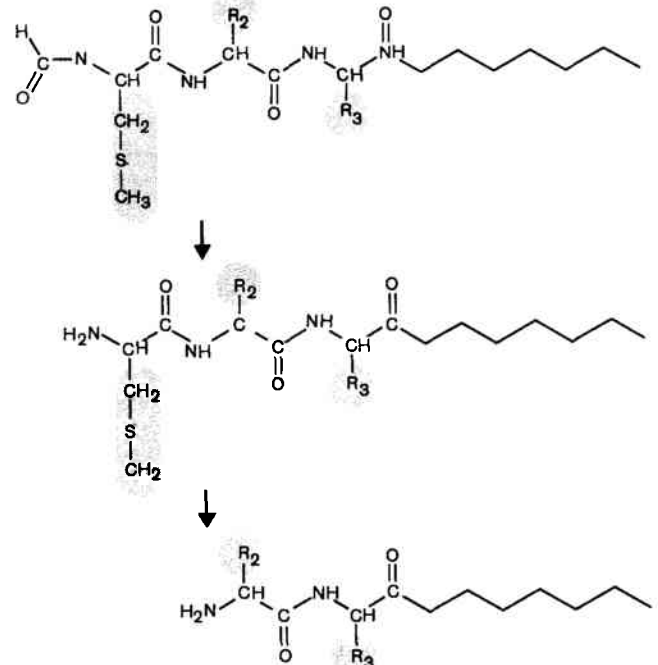


Рис. 6.3. Новосинтезирующиеся бактериальные белки начинаются с формилметионина.

Формильная группа удаляется в результате деформирувания, образуя обычный NH_2 -конец. Затем приблизительно у половины синтезирующихся белков с помощью аминопептидазы удаляется метионин, и в результате аминокислота R_2 превращается в N-концевую. (Исходно эта аминокислота была включена в пептидную цепь второй.)

белка. Но интактная рибосома не может непосредственно взаимодействовать с мРНК и инициировать синтез белка. Первоначально с мРНК связывается малая 30S-субчастица, образуя **иницирующий комплекс** с первым кодоном AUG.

Малая субчастица садится на мРНК таким образом, что инициирующий кодон располагается внутри той части участка Р, которая принадлежит этой субчастице (30S). *Только* инициирующая аминоацил-тРНК обладает уникальным свойством непосредственно входить в Р-участок, распознавая находящийся там кодон. Это показано на **рис. 6.4**.

Затем большая субчастица связывается с комплексом, образуя полную рибосому, в которой инициирующая аминоацил-тРНК расположена в новообразованном интактном Р-участке. В свою очередь А-участок становится способен акцептировать аминоацил-тРНК, комплементарную следующему кодону – второму кодирующему триплету в гене.

В ходе этого процесса инициирующие кодоны взаимодействуют только с инициаторной аминоацил-тРНК (через малую субчастицу), тогда как триплеты, расположенные в *пределах* кодирующей области, взаимодействуют только с обычными аминоацил-тРНК (благодаря целой рибосоме).

Несмотря на то что 30S-субчастица участвует в инициации, самостоятельно она не способна связывать ни мРНК, ни инициаторную тРНК. Неполющенность рибосомных субчастиц в реакции иницирования впервые была обнаружена при описании эффекта «промывки» 30S-субчастиц раствором хлористого аммония. При такой обработке они теряют способность иницировать синтез новой белковой цепи. Причина состоит в том, что активные 30S-субчастицы содержат **факторы инициации (IF)**. Это белки, которые, будучи лабильно связанными с рибосомами, удаляются при промывке раствором хлористого аммония, не оказывающего никакого влияния на истинные или структурные рибосомные белки.

В настоящее время в бактериальных системах известно три фактора инициации, обозначенных как IF1, IF2 и IF3. (Их характеристики даны в табл. 6.1.) Они необхо-

димы на каждой стадии взаимодействия малой субчастицы рибосомы с другими компонентами: мРНК и инициирующей тРНК. На первой стадии происходит специфическое связывание 30S-субчастицы исключительно с инициирующим участком мРНК. В этой реакции принимает участие фактор IF3. Вторая стадия состоит в том, что только инициирующая fMet-тРНК попадает в недостроенный Р-участок, образуемый субчастицей и мРНК. Для обеспечения этой реакции необходим фактор IF2.

Роль фактора IF1 в инициации точно не установлена. Этот белок с мол. массой 9500 дальтон – самый маленький из факторов инициации. Он связывается с 30S-субчастицей рибосомы только как составная часть инициирующего комплекса. IF1 может принимать участие в стабилизации инициирующего комплекса, вероятно вызывая конформационные перестройки в 30S-субчастице, а не играя непосредственную роль в узнавании каких-либо компонентов. Количество IF1 в клетках достаточно только для того, чтобы связаться с 3% общего числа 30S-субчастиц.

Таблица 6.1

Для запуска белкового синтеза у *E. coli* необходимо наличие трех факторов инициации

Фактор	Масса (дальтоны)	Молекулы фактора / Свободные рибосомы	Функции
IF3	23000	25%	Связывание с мРНК и диссоциация субчастиц
IF2	100000	?	Иницирует связывание с тРНК
IF1	9000	15%	Повторение цикла?

Факторы инициации обнаруживаются только на свободных 30S-субчастицах, происходящих из тех самых 20% рибосом, которые не участвуют в данный момент в бел-

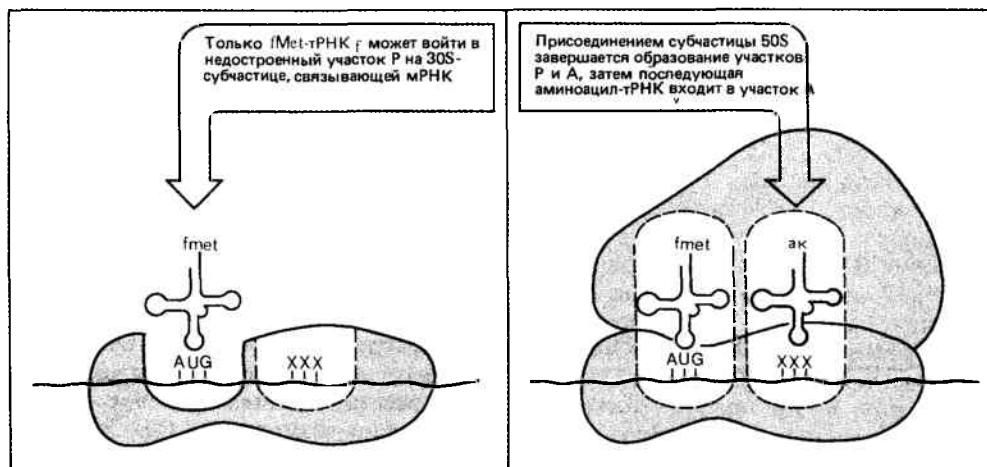


Рис. 6.4. При инициации белкового синтеза 30S-субчастица использует лишь fMet-тРНК_f. Только другие молекулы амино-

ацил-тРНК используются 70S-рибосомами в процессе элонгации.

ковом синтезе. Число копий каждого фактора относительно невелико и достаточно лишь для взаимодействия с малым количеством свободных рибосом. Таким образом, все факторы должны использоваться многократно. При превращении комплекса 30S-мРНК в 70S-рибосому факторы инициации освобождаются. Этим они отличаются от рибосомных структурных белков. Таким образом, факторы инициации участвуют исключительно в формировании иницирующего комплекса. На стадии элонгации они покидают 70S-рибосому и в дальнейшем процессе не принимают никакого участия.

Недолговечная «свобода» 30S-субчастиц

Рибосомы, занятые в синтезе белка, находятся в форме 70S- или 80S-частиц; в таком же виде они диссоциируют от мРНК. Далее они поступают в пул свободных рибосом. Как же образуются необходимые для инициации «свободные» 30S-субчастицы?

У бактерий диссоциации свободных 70S-рибосом на субчастицы способствует белковый фактор, первоначально описанный как фактор диссоциации, но впоследствии идентифицированный как фактор IF3. Ранее этот фактор был обнаружен благодаря той роли, которую он играет при ассоциации 30S-субчастицы с мРНК. Таким образом, он выполняет две функции: сначала используется для образования свободных 30S-субчастиц, а затем для их связывания с мРНК. В этом смысле он является фактором, контролирующим свободное состояние 30S-субчастиц, которое продолжается с момента их диссоциации от свободных 70S-рибосом до их реассоциации с 50S-субчастицей при инициации.

На рис. 6.5 показано, что свободные 70S-рибосомы находятся в динамическом равновесии с 30S- и 50S-субчастицами. Когда IF3 связывается с 30S-субчастицей, равновесие сдвигается, поскольку образуется свободная субчастица, не способная реассоциировать с 50S-субчастицей.

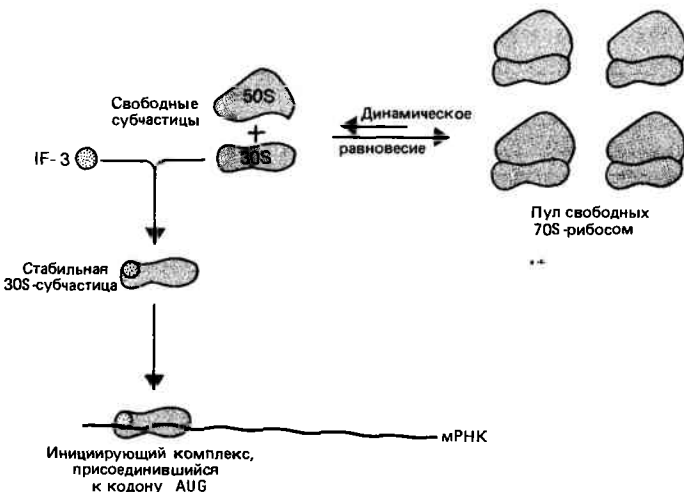


Рис. 6.5. IF3 обуславливает сдвиг равновесия, существующего между свободными 70S-рибосомами и субчастицами, образуя стабильные 30S-субчастицы, способные взаимодействовать с иницирующими сайтами мРНК.

При этом IF3 действует как фактор, препятствующий ассоциации. 30S-субчастица и фактор IF3 взаимодействуют в стехиометрических соотношениях: с одной субъединицей связывается одна молекула IF3 (молекулярная масса одной полипептидной цепи составляет 23 000 дальтон). В клетке количество фактора лимитировано, и поэтому с ним связано только 5% 30S-субчастиц (или около 25% 30S-субчастиц, не участвующих в синтезе белка).

Для того чтобы 30S-субчастица связалась с мРНК, совершенно необходимо присутствие фактора IF3. В результате промывки свободных 30S-субчастиц (раствором хлористого аммония) или при искусственной диссоциации 70S-рибосомы на субчастицы (при пониженной концентрации ионов Mg^{2+}) можно получить малую субчастицу, лишенную этого фактора. Такие субчастицы не способны к образованию иницирующего комплекса с мРНК. Однако специфичность узнавания участков инициации на мРНК всецело определяется рибосомной субчастицей. Таким образом, фактор IF3, по-видимому, не участвует в выборе участка инициации, а необходим только для его взаимодействия с малой субчастицей.

Фактор IF3 освобождается до того, как произойдет объединение малой (30S) и большой (50S) субчастиц. 30S-субчастица не способна одновременно взаимодействовать с фактором IF3 и 50S-субчастицей. Следовательно, 30S-субчастице предоставляется право выбора или находиться в «свободном» состоянии, присоединяясь при этом к IF3, или стать только частью 70S-рибосомы. Из этих превращений и состоит жизненный цикл 30S-субчастицы. Когда рибосома находится в свободном состоянии, то и динамическое равновесие с субчастицами позволяет фактору IF3 заменить 50S-субчастицу. Как только иницирующий комплекс сформируется, фактор IF3 освобождается, давая возможность 30S-субчастице соединиться с 50S-субчастицей. После этого IF3 немедленно поступает в новый цикл, находя другую 30S-субчастицу. Такое быстрое повторное использование характерно для всех вспомогательных факторов.

Освобождение инициаторной тРНК

Роль фактора IF2 в связывании инициаторной fMet-тРНК_f с комплексом 30S·мРНК показана на рис. 6.6. Этот фактор представляет собой белок с мол. массой около 73 000 дальтон, способный специфично связываться с fMet-тРНК_f. Какое свойство инициаторной тРНК позволяет отличать ее от других аминоксил-тРНК, пока еще не ясно, но точно установлено, что для этого узнавания необходимо наличие блокированной NH_2 -группы. In vitro специфичность узнавания проявляется только при низкой концентрации ионов магния.

Образование бинарного комплекса, состоящего из fMet-тРНК_f и IF2, может являться первой стадией, необходимой для попадания этой тРНК в иницирующий комплекс. Далее комплекс 30S·мРНК связывается с IF2·fMet-тРНК_f таким образом, что тРНК находится в недостроенном Р-участке. Однако эта последовательность событий точно не доказана. Не исключено, что IF2 непосредственно связывается с 30S-субчастицей и тем самым дает возможность fMet-тРНК_f взаимодействовать с ним, находясь уже в составе рибосомной субчастицы. При любой последовательности событий на этой стадии белок IF2 остается в составе субчастицы – в дальнейшем ему еще предстоит сыграть свою роль.

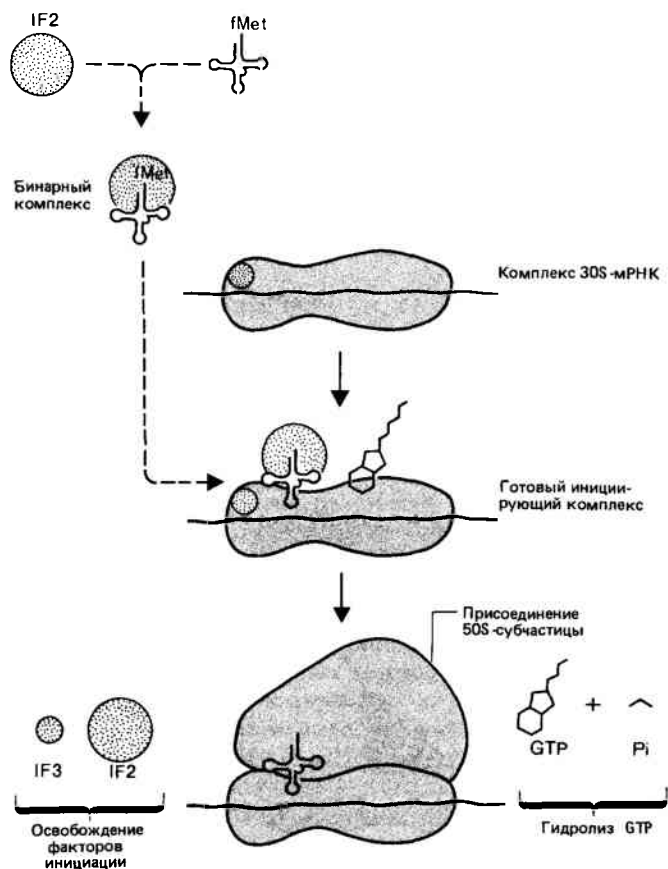


Рис. 6.6. Бактериальные факторы инициации необходимы для связывания 30S-субчастицы с мРНК, а также для взаимодействия fMet-тРНК с комплексом 30S-мРНК; их освобождение сопряжено с гидролизом GTP, сопровождающим присоединение 50S-субчастицы.

Давно известно, что фактор IF2 обладает рибосомозависимой GTP-азной активностью. В присутствии рибосом данный фактор осуществляет гидролиз GTP, высвобождая энергию, запасенную в высокоэнергетической связи. Несмотря на то что GTP является составной частью полностью сформированного иницирующего комплекса, все же точно не ясно, на какой стадии присоединяется нуклеотид. Наиболее вероятно, что это происходит во время или сразу же после связывания бинарного комплекса с 30S-субчастицей. Далее, когда присоединяется 50S-субчастица и образуется полная рибосома, в присутствии IF2 происходит гидролиз GTP. Возможно, фактор IF2 сам не является GTPазой, но активирует какой-то рибосомный белок, выполняющий эту функцию.

Роль фактора IF1 окончательно не выяснена. Возможно, он участвует в возобновлении цикла, способствуя высвобождению фактора IF2 из комплекса.

Для чего нужен гидролиз GTP, точно не известно. Возможно, что он сопряжен с конформационными изменениями в рибосоме, в результате которых независимые субчастицы превращаются в единую функционирующую 70S-рибосому. Вслед за расщеплением GTP P- и A-участки рибосомы принимают свою правильную ориентацию. В P-участке располагается fMet-тРНК, а A-участок готов принять аминоксил-тРНК. Согласно этой модели, гидролизу GTP отведена роль, в чем-то похожая на ту, которую выполняет разрыв высокоэнергетической связи при движении рибосомы (см. ниже).

В инициации у эукариот участвует много факторов

Процесс иницирования трансляции у эукариот в общих чертах аналогичен тому, который наблюдается у *E. coli*. Но здесь обнаружено больше факторов инициации. В ретикулоцитах (незрелые эритроциты), на которых была выполнена основная часть работы, обнаружено по крайней мере восемь факторов, но их конкретная роль недостаточно изучена. Факторы эти обозначаются так же, как и у бактерий, но с приставкой «е», что указывает на их эукариотическое происхождение. К настоящему времени в эритроцитах обнаружены следующие факторы: eIF1, eIF2, eIF4A, eIF4B, eIF4C, eIF4D и eIF5. Препараты eIF2 и eIF3 содержат множество полипептидных цепей, а их названия отражают тот факт, что их роль, по-видимому, аналогична бактериальным факторам IF2 и IF3. Остальные факторы, как правило, состоят из одной полипептидной цепи, а их функции точно не установлены.

Кроме указанных выше различий, процесс инициации у эукариот, по-видимому, в общих чертах аналогичен тому, что происходит у *E. coli*. В ретикулоцитах (незрелых эритроцитах), с которыми проводится основная часть работ, существует значительно больше факторов инициации, по крайней мере, девять обнаружены к настоящему времени. Факторы названы аналогично бактериальным, но с добавлением приставки «е», указывающей на их эукариотическое происхождение. Известный на сегодняшний день перечень факторов приведен в табл. 6.2. Каждая из фракций eIF2 и eIF3 содержат по несколько цепей. Другие факторы преимущественно представлены единичными полипептидами, функции которых пока еще изучены недостаточно полно.

Инициация у эукариот включает образование **тройного комплекса**, содержащего Met-тРНК, eIF2 и GTP. Он образуется в две стадии: сначала GTP связывается с фактором eIF2, увеличивая сродство фактора к Met-тРНК, а это затем приводит к образованию комплекса. На рис. 6.7 показано, что этот комплекс непосредственно взаимодействует со свободной 40S-частью. Процесс стимулируется присутствием факторов eIF3 и eIF4C, которые, возможно, стабилизируют комплекс.

У млекопитающих фактор eIF2 — это белок, состоящий из трех субъединиц — α , β и γ (с мол. массами 35 000, 38 000 и 55 000 дальтон соответственно). Сходной структурой обладает аналогичный фактор из зародыша пшеницы, что свидетельствует об идентичности функции этого фактора у многих эукариот. Свойства фактора eIF2 приведены в табл. 6.3. Точная роль всех субъединиц еще не определена, но, по-видимому, eIF2 α связывается с GTP, тогда как eIF2 γ связывает Met-тРНК; фактор eIF2 β , возможно, является фактором рециклизации.

Взаимодействие 40S- и тройного комплекса с мРНК зависит от фактора eIF3, который, возможно, играет роль, аналогичную фактору IF3, поддерживая 40S-субчастицу в свободном состоянии. Для этого взаимодействия также необходимы некоторые дополнительные факторы (eIF1, eIF4A, eIF4B) и гидролиз высокоэнергетической связи АТФ. Связывание 60S-субчастицы с иницирующим комплексом происходит в присутствии фактора eIF5 и также сопряжено с гидролизом GTP, входившим в состав тройного комплекса.

Если факторы пометить радиоактивными изотопами, то можно непосредственно продемонстрировать наличие eIF2 и eIF3 в составе иницирующего 40S-комплекса.

Таблица 6.2

Ретикулоциты содержат по крайней мере девять факторов инициации

Фактор	Структура	Функция
eIF3	> 500 000 мультимер	Связывание мРНК
eIF1 eIF4B eIF4A	15 000 мономер мономер мономер	Помогает связывать мРНК Помогает связывать мРНК и связывает АТФ
eIF6	23 000 мономер	Предотвращает объединение 40S–60S-субчастиц
eIF5 eIF4C	150 000 мономер мономер	Освобождение eIF2 и eIF3. Связывание 60S-субчастицы
eIF2	Тример	Связывание Met-тРНК (см. табл. 6.3)
eIF4D	Мономер	Не известна

Другие факторы не обнаруживаются с помощью такого подхода, но надо отметить, что ингибитор белкового синтеза, эдеин, останавливающий процесс на стадии инициации, приводит к тому, что факторы eIF4A, eIF4B и eIF4C остаются в составе комплекса 40S-мРНК. Вероятно, когда малая частица соединяется с большой, образуя полную рибосому, все эукариотические факторы инициации освобождаются.

В связи с характеристикой иницирующих факторов необходимо учитывать, что в данном контексте термин «эукариотические» имеет отношение только к небольшому количеству систем. В других случаях могут использоваться способы инициации, отличающиеся в деталях от описанных выше. Точно так же в бактериях, отличных от *E. coli*, обнаруживаются некоторые особенности в способах инициации.

Важная роль фактора eIF2 в синтезе белка

Предположение о том, что фактор eIF2 взаимодействует с элементами, контролирующими белковый синтез, вытекает из двух фактов. В обоих случаях малая субъединица фактора eIF2 α фосфорилируется, причем эта модификация каким-то непонятным образом приводит к нарушению процесса инициации.

Таблица 6.3

eIF2 состоит из трех субъединиц

Субъединица	Масса (дальтоны)	Функция в инициации
α	35 000	Связывает GTP Контролируется с помощью механизма фосфорилирования
β	38 000	Возможно, является фактором, отвечающим за возобновление цикла
γ	55 000	Связывает Met-тРНК

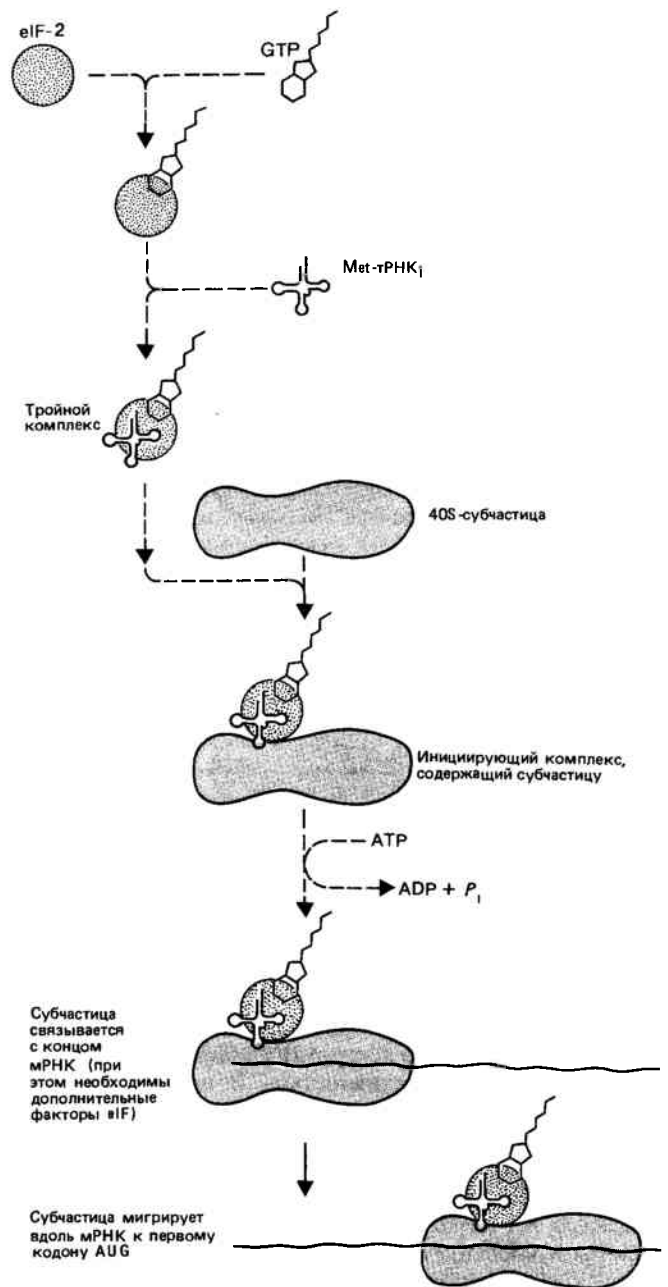


Рис. 6.7. При инициации синтеза белка у эукариот eIF2 образует тройной комплекс с Met-тРНК₁, который связывается с 40S-субчастицей; затем иницирующий комплекс взаимодействует с 5'-концом мРНК и начинает мигрировать к иницирующему кодону AUG (гл. 9).

Первая ситуация возникает в ретикулоцитах, когда те лишаются гемина, кофактора, необходимого для сборки гемоглобина – основного белка эритроцитов. В отсутствие гемина инициация белкового синтеза прекращается. Это связано с тем, что падение концентрации кофактора до уровня ниже критического активирует ингибитор белкового синтеза, т.е. протеинкиназу (фермент, катализирующий перенос фосфатной группы на белки), а фактор eIF2 α является субстратом этого фермента.

Фосфорилирование фактора eIF2 α ингибирует его активность непрямым путем. После того как он выполнил свою роль в реакции инициации, фактор eIF2 освобо-

ждается в форме комплекса GDP, образовавшимся при гидролизе GTP, входящего в состав тройного комплекса. До тех пор пока GDP не будет снова заменен на GTP, фактор не может выполнять свою функцию (рис. 6.8). Реакция замещения требует дополнительного фактора, известного под разными названиями, но в настоящее время обозначаемого как eIF2B. Этот фактор не может функционировать в том случае, если α -субъединица фосфорилирована.

Фосфорилирование хотя бы 20–30% eIF2 достаточно, чтобы остановить инициацию белкового синтеза. По всей вероятности, фосфорилированный eIF2 связывает весь eIF2B (содержание молекул фактора eIF2B составляет всего лишь 10–20% от содержания молекул фактора eIF2) и таким образом делает его недоступным для взаимодействия с нефосфорилированным eIF2. Фосфорилирование eIF2-субъединицы непосредственно не мешает фактору eIF2 принять участие в синтезе белка один раз, однако оно препятствует его взаимодействию с каким-то другим фактором. Последний, по-видимому, принимает участие в регенерировании фактора eIF2-GTP из eIF2-GDP, освобождающегося после инициации. Это превращение необходимо для повторных циклов инициации.

Иная ситуация встречается во множестве других клеток, в том числе и в ретикулоцитах. В присутствии двухцепочечной РНК (дцРНК) активируется новая протеинкиназа, которая точно так же фосфорилирует eIF2 α . Синтез киназы может стимулироваться интерфероном, для которого дцРНК служит мощным индуктором. Интерферон избирательно подавляет синтез вирусспецифических белков, не влияя при этом на синтез белков клетки-хозяина. Однако не ясно, каким образом фосфорилирование eIF2 связано со специфичностью ингибирования, так как eIF2 сам по себе в равной степени участвует в инициации белкового синтеза как на вирусных, так и на клеточных мРНК.

Последовательность событий в прокариотах и эукариотах

Не ясно, существует ли строгая очередность при связывании мРНК и инициаторной тРНК с бактериальной 30S-субчастицей. Согласно существующей точке зрения, как та, так и другая может присоединиться первой. В противоположность этому, как показано на рис. 6.7, эукариотическая 40S-субчастица сначала связывает инициаторную тРНК. Следовательно, эта реакция не должна зависеть от взаимодействия кодон–антикодон. Для последующего присоединения мРНК необходимо присутствие инициаторной тРНК.

30S-субчастица бактерий и 40S-субчастица эукариот по-разному находят на мРНК связывающие участки. В бактерии инициаторный комплекс формируется на последовательности, непосредственно соседствующей с иницирующим кодоном AUG. У эукариот малые субчастицы сначала узнают 5'-конец-мРНК, а затем продвигаются к иницирующему кодону, где они соединяются с большими субчастицами.

Бактериальные мРНК часто содержат несколько кодирующих областей, каждая из которых обладает собственной инициаторной последовательностью, способной непосредственно связывать 30S-субчастицу. Иногда эти последовательности могут быть маскированы вторичной структурой мРНК, но, как правило, их фун-

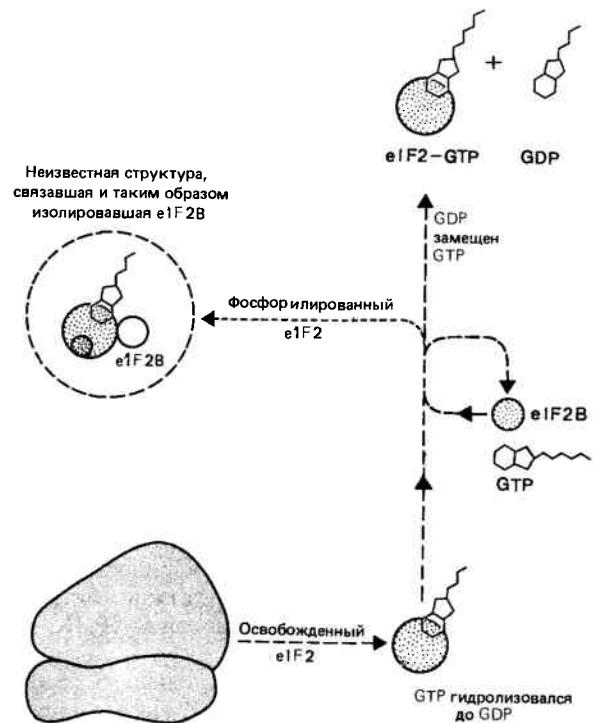


Рис. 6.8. У эукариот фактор eIF2 освобождается в реакции инициации в виде бинарного комплекса, содержащего GDP; фактор eIF2B нужен для того, чтобы заменить GDP на GTP для повторного использования eIF2.

кция не зависит от местоположения в молекуле. Следовательно, близость к 5'-концу мРНК не сказывается на распознавании участка инициации.

Каждая эукариотическая мРНК обычно кодирует только один полипептид. При этом 5'-конец мРНК способен взаимодействовать с 40S-субчастицей, которая не может непосредственно связываться с внутренними участками молекулы. Часто иницирующий кодон AUG не располагается в непосредственной близости от 5'-конца. Из этого следует, что 40S-субчастица может передвигаться вдоль мРНК до тех пор, пока не достигнет иницирующей последовательности. Для этого передвижения требуется гидролиз АТФ. Такой процесс можно представить как некоторый вид «скольжения» вдоль последовательности мРНК в поиске кодона AUG. К этому вопросу мы еще вернемся в гл. 9.

Элонгация: поступление аминоксил-тРНК в А-участок

Как только на иницирующем кодоне собралась «полная» рибосома, стадия инициации переходит в цикл, в ходе которого аминоксил-тРНК попадает в А-участок рибосомы, Р-участок которой уже занят пептидил-тРНК. Для образования первой пептидной связи инициаторная тРНК, связавшаяся в участке Р, должна выступать в качестве аналога пептидил-тРНК. Нам уже известно, что любая аминоксил-тРНК, за исключением инициаторной, может входить в А-участок. Это происходит не спонтанно, а при участии фактора элонгации (EF). Так же как и его аналог при инициации, этот фактор при поступле-

нии аминоксил-тРНК в А-участок связывается с рибосомой. Для продолжения синтеза белка он должен диссоциировать. Наличие такого циклического связывания с рибосомой и отделения от нее — характерная особенность дополнительных факторов.

Факторы, имеющие отношение к поступлению аминоксил-тРНК в рибосому, первоначально были получены из *E. coli* в виде индивидуальной фракции, названной фактором Т (название фактор Т происходит от англ. transfer factor — фактор переноса). Позже эту фракцию разделили на два компонента, названные *EF-Tu* и *EF-Ts*. Данные обозначения отражают следующие характерные особенности факторов: фактор *EF-Tu* неустойчив к нагреванию (*unstable*), тогда как *EF-Ts* — устойчив (*stable*).

Аминоксил-тРНК может связываться с рибосомой одним из двух способов. Некоторое взаимодействие, называемое **неферментативным взаимодействием**, наблюдается даже в отсутствие соответствующих факторов. Оно не подвержено такому контролю, какое имеет место при синтезе белка в клетке. Это взаимодействие отражает присущее аминоксил-тРНК сродство к рибосоме. Истинное связывание наблюдается в присутствии фактора *EF-Tu*, обеспечивающего доставку аминоксил-тРНК в участок А.

На рис. 6.9 показана последовательность событий, в результате которой аминоксил-тРНК связывается с рибосомой. Белок *EF-Tu* связывает гуаниновый нуклеотид. Активная форма *EF-Tu* содержит *GTP*. Поэтому **бинарный комплекс** *EF-Tu·GTP*, взаимодействующий с аминоксил-тРНК, образует **тройной комплекс**, состоящий из аминоксил-тРНК·*EF-Tu·GTP*. Последний связывается только с участком А рибосомы, у которой участок

Таблица 6.4

Для функционирования аминоксил-тРНК требуется два белковых фактора

Фактор	Кодируется геном	Молекулярная масса	Число молекул на клетку	Функция	Вещество, ингибирующее активность фактора
<i>EF-Tu</i>	<i>tufA</i> <i>tufB</i>	43 225	70 000	Связывает аминоксил-тРНК и <i>GTP</i>	Кирромицин
<i>EF-Ts</i>	<i>tsf</i>	74 000	10 000	Связывает <i>EF-Tu</i> , вытесняя <i>GTP</i>	

Р уже занят пептидил-тРНК. Эта стадия является решающей в обеспечении правильной ориентации аминоксил-тРНК и пептидил-тРНК при образовании пептидной связи.

Связавшись с рибосомой, *GTP* гидролизуются с высвобождением бинарного комплекса *EF-Tu·GDP*. В этой форме *EF-Tu* не связывает аминоксил-тРНК. Возможно, функция гуанилового нуклеотида в этой реакции состоит в том, что *EF-Tu* принимает нужную для связывания аминоксил-тРНК конформацию только в присутствии трифосфата. Поэтому *GTP*, а не *GDP* поступает в рибосому для последующего гидролиза высокоэнергетической связи.

Активная форма *EF-Tu·GTP* регенерируется из использованной формы *EF-Tu·GDP* при участии фактора *EF-Ts*. Он способен вытеснить *GDP* из комплекса с *EF-Tu*, образуя объединенный фактор *EF-Tu·EF-Ts* (который является подлинным «полным» фактором Т). Затем *EF-Ts* в свою очередь замещается *GTP*, что снова приводит к образованию *EF-Tu·GTP*. Этот бинарный комплекс способен вновь связать аминоксил-тРНК, а освободившийся *EF-Ts* может поступить в следующий цикл.

Ассоциация между *EF-Ts*, *EF-Tu* и *GTP* *in vitro* обратима — в отличие от реакции с аминоксил-тРНК, являющейся той стадией, после которой события разворачиваются только в одном направлении. Превращения, происходящие с факторами элонгации, заключаются в следующем: *EF-Ts* совершает цикл, в котором он то связан с *Tu*, то находится в свободном состоянии, тогда как *EF-Tu* либо связывается с *GDP* или с *GTP* в бинарном или тройном комплексе, либо связывается с *EF-Ts*, образуя фактор Т.

В табл. 6.4 приведены некоторые свойства факторов *EF-Tu* и *EF-Ts*. Количество фактора *EF-Tu* равно примерно числу молекул аминоксил-тРНК. Из этого следует, что большая часть молекул аминоксил-тРНК, очевидно, находится в составе тройного комплекса, а не в свободном состоянии. На долю фактора *EF-Tu*, кодируемого двумя генами, приходится около 5% суммарного клеточного белка. Большая часть *EF-Tu*, по-видимому, должна находиться в составе бинарного или тройного комплексов, так как количество молекул *EF-Ts* значительно меньше, будучи равным приблизительно числу рибосом.

Белок *EF-Tu* представляет собой одну полипептидную цепь с мол. массой 43 225, состоящую из 393 аминокис-

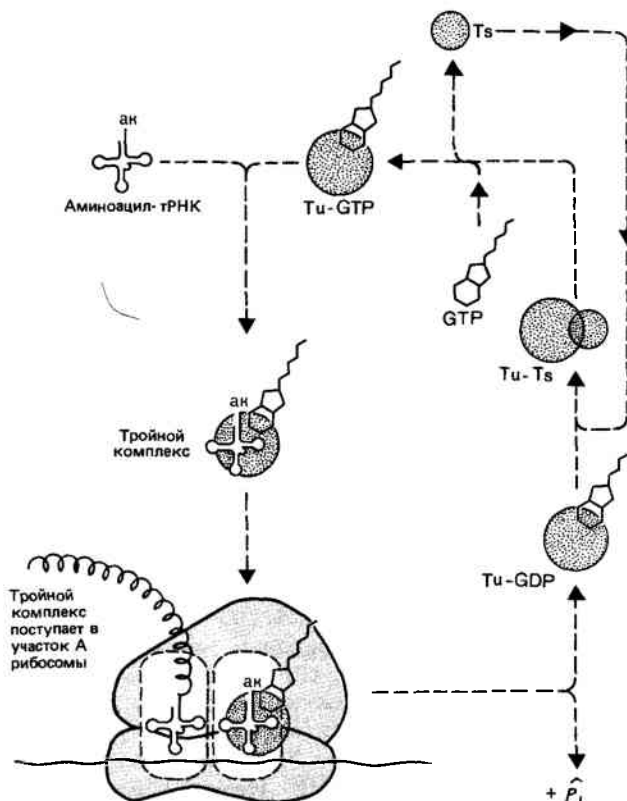


Рис. 6.9. *EF-Tu·GTP*, доставляющий аминоксил-тРНК в рибосому, освобождается в виде *EF-Tu·GDP*, а затем под действием *EF-Ts* регенерирует с образованием *EF-Tu·GTP*.

лот. У *E. coli* имеется два гена — *tufA* и *tufB*, кодирующих EF-Tu. Это один из основных белков *E. coli*. В клетке насчитывается около 70 000 его молекул, что составляет 5% от общего содержания белка. Этого достаточно для связывания всего клеточного пула аминоацил-tРНК, количество которых примерно в десять раз превышает число рибосом. Содержание EF-Ts соответствует числу рибосом. Следовательно, большинство EF-Tu в клетке входит в состав бинарного и тройного комплексов.

При формировании тройного комплекса EF-Tu·GTP наиболее важной особенностью, по-видимому, является структура акцепторного стебля tРНК (двухцепочечная последовательность, имеющая на конце АССА-ОН, к которой присоединяется аминокислота). В частности, изменения в структуре концевой аденозина могут препятствовать узнаванию аминоацил-tРНК. Комплекс EF-Tu·GTP способен узнавать аминоацил-tРНК, содержащую аминокислоту либо в 2'-, либо в 3'-положении сахарного остатка. Следовательно, положение, к которому присоединяется аминокислота в сахарном кольце, существенно для данного взаимодействия. Однако на определенной стадии, перед тем как связаться с рибосомой, тройной комплекс стабилизирует эфирную связь аминоацил-tРНК в 2'-положении.

fMet-tРНК_f — это единственная аминоацил-tРНК, которая не может быть узнана бинарным комплексом EF-Tu·GTP. Именно данное свойство не позволяет ей реагировать с внутренними кодонами AUG и GUG. Это является гарантией того, что данные кодоны узнаются только соответствующими видами аминоацил-tРНК (tРНК_m^{Met} и tРНК^{Val}). Одной из причин отсутствия взаимодействия между tРНК_f^{Met} и бинарным комплексом может быть структура ее акцепторного стебля. Перед концевой АССА-последовательностью, акцептирующей аминокислоту, в стебле располагается некомплементарная пара нуклеотидов С·А вместо комплементарной пары, присутствующей во всех остальных tРНК. Кроме того, могут быть и другие причины, связанные с особенностями структуры инициаторной tРНК. В частности, немаловажное значение имеет блокирование NH₂-группы, так как молекулы аминоацил-tРНК, обладающие такими концами, не способны связываться с EF-Tu·GTP. (Следует отметить, что упомянутые особенности узнавания инициаторной tРНК могут быть специфичны для бактерий и иметь существенные отличия в случае эукариот.)

У эукариот, как показано на примере ретикулоцитов, имеется два фактора элонгации — eEF1 и eEF2, которые эквивалентны соответственно бактериальным факторам EF-T и EF-G (см. далее).

Роль eEF1 заключается в том, чтобы доставлять аминоацил-tРНК в рибосому. И в этой реакции происходит гидролиз высокоэнергетической связи с образованием GDP из GTP. В состав активного фактора входит несколько полипептидных цепей различного размера. Как происходит регенерация GTP после его расщепления, в деталях не ясно, но вполне возможно, что фактор содержит компоненты, аналогичные EF-Tu и EF-Ts (поэтому eEF1 характеризуется как аналог составного фактора EF-T). Однако среди этих полипептидов могут быть компоненты, относящиеся не только к eEF1. Поэтому проводить аналогии между компонентами, входящими в состав этих факторов, пока еще преждевременно. В количественном отношении ситуация у прокариот и эукариот, возможно, сходная: eEF1 (так же как и EF-Tu) является наиболее обильно представленным белком клетки.

Не исключено, что в митохондриях и хлоропластах существуют независимые факторы элонгации, ответственные за связывание аминоацил-tРНК с рибосомой, но они еще не охарактеризованы.

Гидролиз GTP происходит после присоединения аминоацил-tРНК

Гидролиз GTP происходит через некоторое время, после того как комплекс аминоацил-tРНК·EF-Tu·GTP связывается с рибосомой. Роль GTP была изучена с помощью его негидролизующих аналогов. В соединении 5'-гуанилилметилendifосфонат (GMP-МДФ) имеется структура, в которой метиленовая группа заменяет кислород, связывающий β- и γ-фосфаты в GTP. В присутствии GMP-МДФ возможно образование тройного комплекса, доставляющего аминоацил-tРНК в рибосому. Но при этом пептидная связь образоваться не может. Из этого следует, что для связывания аминоацил-tРНК с А-участком необходимо присутствие GTP, а не его гидролиз, происходящий позже. Количественные эксперименты показывают, что при поступлении каждой аминоацил-tРНК в рибосому гидролизуется одна молекула GTP.

У бактерий большое количество информации об отдельных этапах синтеза белка было получено при использовании антибиотиков, ингибирующих процесс на определенных стадиях. Например, **кирромицин** связывается с EF-Tu и мешает его функционированию. Одним из последствий этого связывания является активация GTPазной активности, не зависящая от рибосом. Это означает, что гидролиз GTP, происходящий вслед за связыванием тройного комплекса с рибосомой, скорее осуществляется фактором EF-Tu, чем рибосомным белком (см. ниже).

Кирромицин блокирует синтез белка еще до стадии образования пептидной связи. Фактор EF-Tu, связанный с этим антибиотиком, способен доставлять аминоацил-tРНК в А-участок. Затем может произойти единственный цикл гидролиза GTP. Но комплекс EF-Tu·GDP не способен покинуть рибосому. Это препятствует образованию пептидной связи между пептидил-tРНК и аминоацил-tРНК; в результате рибосома начинает «буксовать» на мРНК. Если хотя бы одна рибосома не может перемещаться по матрице, это мешает трансляции, осуществляемой другими рибосомами, и приводит к тому, что синтез белков прекращается. Воздействие данного антибиотика показывает, что нарушение одной из стадий синтеза белка приводит к блокированию последующих стадий. В данном случае правильная конформация рибосомы, необходимая для образования пептидной связи, достигается при высвобождении EF-Tu.

Устойчивость к кирромицину можно получить с помощью мутаций, препятствующих действию антибиотика на активность фактора EF-Tu. Поскольку для данного белка существует два гена, то они оба должны быть мутантными, иначе присутствие в клетке даже небольшого количества чувствительного EF-Tu позволит кирромицину блокировать трансляцию.

За образование пептидной связи ответственна рибосома

Растущая полипептидная цепь удлиняется в результате переноса полипептида, связанного с tРНК, находящейся в рибосомном Р-участке, на аминоацил-tРНК, распо-

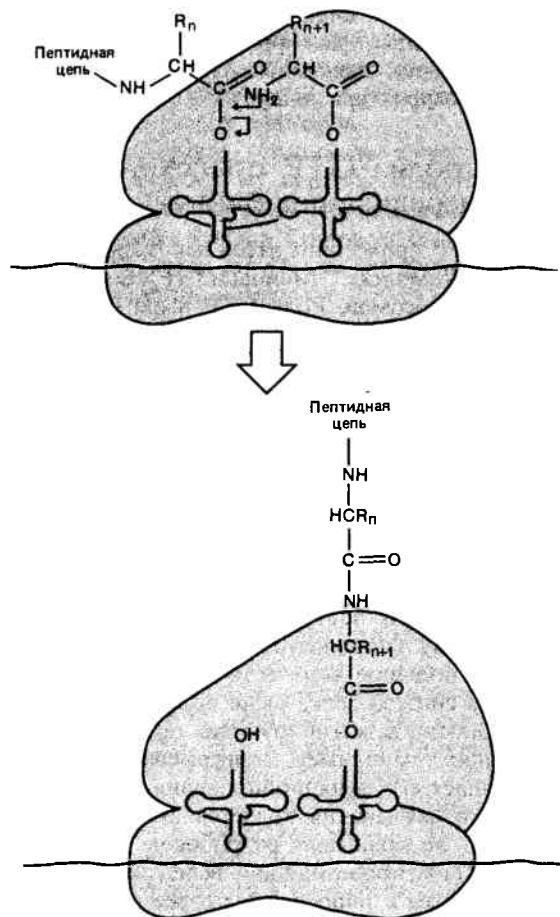


Рис. 6.10. Образование пептидной связи происходит благодаря переносу полипептида от тРНК, находящейся в Р-участке, на aminoacyл-тРНК, расположенную в А-участке.

ложенную в А-участке (рис. 6.10). Синтез полипептидной связи осуществляется **пептидилтрансферазой**, предпочтительным субстратом которой служит aminoacyл-3'-тРНК. Возможно, перенос аминокислоты из 2'-положения в 3'-положение рибозного кольца может предшествовать образованию пептидной связи.

Пептидилтрансферазная активность является функцией 50S- (или 60S-) субчастицы. Хотя активный центр фермента расположен полностью на большой субчастице, обычно активность обнаруживается только тогда, когда субчастица входит в состав «полной» 70S- (или 80S-) рибосомы. Возможно, это ограничение необходимо для того, чтобы большая субчастица, находясь в свободном состоянии, не могла беспорядочно образовывать связь между молекулами aminoacyл-тРНК. Однако и у прокариот, и у эукариот пептидилтрансфераза больших субчастиц может быть активирована *in vitro* различными воздействиями, например такими, как добавление соответствующих количеств этанола.

Пептидилтрансфераза, возможно, является частью того участка рибосомы, в котором концы пептидил-тРНК и aminoacyл-тРНК приходят в тесный контакт. Ведь один из еще не решенных вопросов белкового синтеза сводится к следующему: как две довольно громоздкие молекулы тРНК могут комплементарно спариваться с соседними триплетами на мРНК при том, что их акцепторные стебли должны быть расположены достаточно близко, чтобы могла образоваться пептидная связь? Как преодолевается это стерическое препятствие, не ясно.

На рис. 6.11 изображена стадия **транслокации**, которой завершается процесс включения аминокислот в растущую полипептидную цепь. При этом рибосома передвигается на три нуклеотида вдоль мРНК.

После образования пептидной связи в Р-участке рибосомы находится деацилированная тРНК, а в А-участке — пептидил-тРНК. Как составная часть одного и того же механизма, при транслокации из рибосомы одновременно выталкивается деацилированная тРНК, а пептидил-тРНК передвигается в Р-участок. В результате в рибосоме образуется свободный А-участок, готовый принять следующую aminoacyл-тРНК, требующуюся для возобновления цикла.

Необходимость в дополнительных факторах для транслокации была обнаружена в самых первых экспериментах, выявивших наличие двух участков — А и Р. При этом использовали антибиотик **пуромицин** (рис. 6.12). Как видно из рисунка, этот антибиотик во многом напоминает аминокислоту, присоединенную к концевому аденозину тРНК. Атом азота (N) в пуромицине, соответствующий атому кислорода (O) в aminoacyл-тРНК, предотвращает расщепление. Рибосома, очевидно, воспринимает пуромицин как aminoacyл-тРНК. Это сходство обнаруживается и при образовании пептидной связи. При этом происходит освобождение растущего полипептида в форме полипептидилпуромицина. Преждевременная остановка белкового синтеза объясняет летальное действие антибиотика на бактериальные клетки.

Поскольку пуромицин попадает в рибосому тем же путем, что и aminoacyл-тРНК, он может взаимодействовать с пептидил-тРНК только в том случае, если Р-участок занят, а А-участок свободен. (Эта ситуация изображена в нижней части рис. 6.10.) Когда образование пептидной связи уже завершено, но транслокация еще не произошла, пуромицин не способен проникнуть в рибосому, так как А-участок остается занятым пептидил-тРНК (такая ситуация показана в верхней части рис. 6.10). На этой стадии рибосомы могут стать чувствительными к пуромицину (так как произошла транслокация) только при наличии GTP и фактора элонгации EF-G. Название этого белка отражает тот факт, что исходно он был описан как GTP-зависимый фактор. EF-G является одним из основных компонентов клетки (табл. 6.5). На каждую рибосому приходится примерно одна молекула этого фактора.

Транслокация состоит из согласованной последовательности событий, начинающейся с ассоциации EF-G с рибосомой. Процесс зависит от гидролиза GTP. Фактор EF-G является одним из основных белков клетки. Его молекулярная масса равна 72 000 дальтон, и на его долю приходится 2% растворимого клеточного белка. Фактор кодируется одним-единственным геном *fus*. Он присутствует в количествах, приблизительно эквивалентных с рибосомными белками.

Фактор связывается с рибосомой для того, чтобы способствовать транслокации, а далее после гидролиза GTP он освобождается в среду. Для этого связывания необходимо присутствие GTP, но оно также может произойти и при использовании GMP-МДФ. Из этого следует, что для ассоциации фактора с рибосомой не нужен гидролиз GTP, но транслокация при этом произойти не может. Возможный комплекс, который EF-G образует с GTP (или GDP), трудно охарактеризовать вследствие их нестабильности.

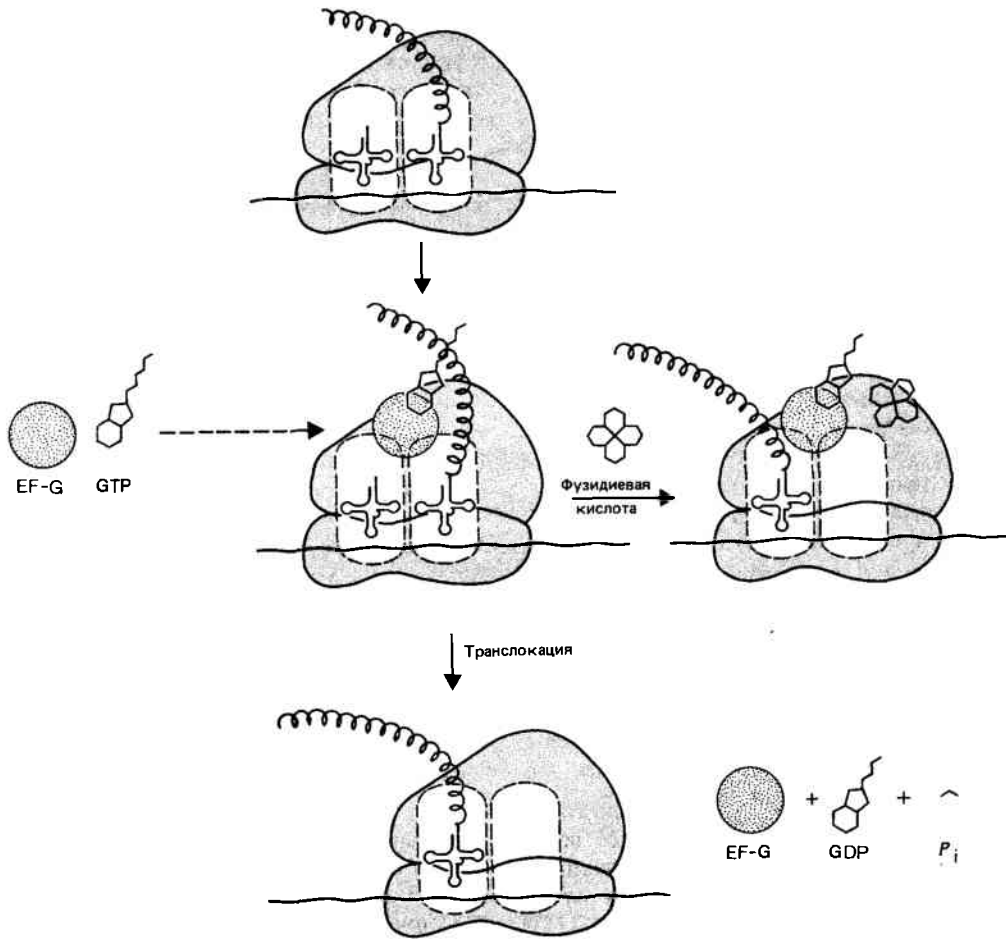


Рис. 6.11. После образования пептидной связи для осуществления транслокации необходимы EF-G и GTP, которые затем от-

деляются от рибосомы (если только не присутствует фузидиевая кислота, стабилизирующая посттрансляционную стадию).

В связи с этим мы не знаем, попадает ли молекула GTP в рибосому в комплексе с EF-G или взаимодействует с ней непосредственно. Гидролиз GTP катализируется не фактором EF-G, а рибосомой.

Рибосома не может связать EF-Tu и EF-G одновременно, поэтому в процессе трансляционного цикла эти факторы попеременно взаимодействуют с частицей. Фактор EF-G может связаться с рибосомой, как комплекс EF-Tu·GDP покинул рибосому. Затем EF-G должен уступить место аминоацил-tPHK·EF-Tu·GTP.

При изучении механизма действия антибиотика стероидной природы — фузидиевой кислоты, — которая «консервирует» рибосому в транслоцированном состоянии, было обнаружено, что для успешного протекания синтеза белка необходимо освобождение EF-G. В присутствии фузидиевой кислоты происходит один цикл транслокации: EF-G связывается с рибосомой, GTP гидролизуется и рибосома передвигается на три нуклеотида. Но далее, вместо того чтобы покинуть рибосому, EF-G и GDP остаются в ней (см. рис. 6.11). Такое действие этого антибиотика, возможно, объясняется тем, что фузидиевая кислота стабилизирует комплекс рибосомы EF-G·GDP. Из-за того что освобождения фактора EF-G не происходит, прерывается дальнейший ход событий, так как А-участок рибосомы уже не способен акцептировать тройной комплекс, состоящий из аминоацил-tPHK·EF-Tu·GTP. В результате дальнейший рост полипептидной цепи прекращается.

Нам не известно, чем определяется взаимоисключающее взаимодействие факторов элонгации между собой. Возможно, что оно обусловлено только конформационным эффектом рибосомы или же непосредственной конкуренцией за одни и те же или частично перекрывающиеся участки связывания (лекарственное вещество тиострептон препятствует связыванию как EF-Tu, так и EF-G). Таким образом, необходимость того, чтобы каждый фактор удался из рибосомы, с тем чтобы следующий мог войти в нее, убеждает нас в существовании строго установленной очередности событий в процессе белкового синтеза.

Оба фактора используют GTP для того, чтобы связаться с рибосомой, хотя гидролиз его осуществляется

Таблица 6.5
Для транслокации необходим фактор элонгации

Фактор	Кодируется геном	Молекулярная масса	Количество молекул на клетку	Функция	Вещество, ингибирующее активность фактора
EF-G	<i>fus</i>	77 444	20 000	Взаимодействует с рибосомой и GTP	Фузидиевая кислота

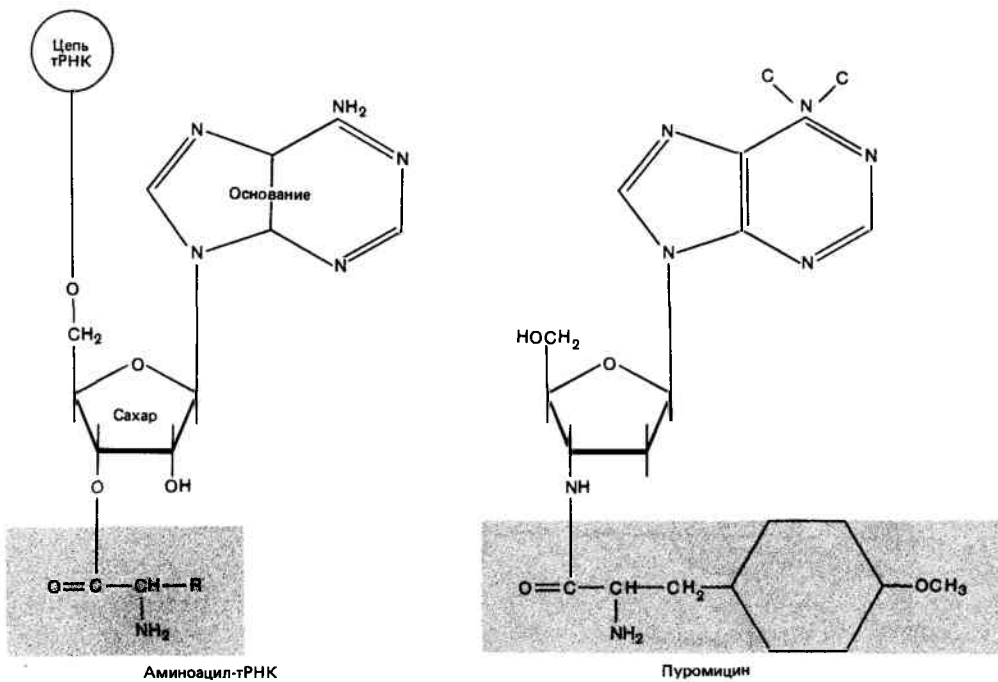


Рис. 6.12. Пуромицин по своему строению напоминает аминоацил-тРНК, так как он имеет сходство с ароматической аминокислотой (затемненная часть), соединенной с нуклеозидной частью.

позднее. Поскольку ни один из них не может использовать для связывания с рибосомой GDP, то вероятнее всего, что GTP требуется факторам для принятия правильной конформации. Благодаря такому механизму факторы получают доступ в рибосому только в присутствии GTP, который в дальнейшем они используют для выполнения своей функции. При гидролизе GTP высвобождается энергия, которая может быть затрачена на изменение конформации рибосомы. В присутствии EF-Tu она нужна для активации аминоацил-тРНК в А-участке. При наличии фактора EF-G энергия расходуется на перемещение рибосомы.

Гидролиз GTP, сопряженный с действием EF-G, не катализируется самим фактором, а является функцией рибосомы. Это свойство присуще 50S-субчастице (гл. 8). Как мы уже отмечали, вопрос о том, расщепляется ли GTP, входящий в тройной комплекс тРНК·EF-Tu·GTP, под действием EF-Tu или рибосомного белка, остается открытым.

Расстояние в три основания, которое проходит рибосома, вероятно, определяется взаимодействием между кодоном мРНК и антикодоном тРНК. Некоторые мутантные тРНК узнают антикодоны из четырех оснований, что позволяет рибосомам продвигаться на четыре основания в каждом транслокационном акте. Поэтому, исходя из стереохимии взаимодействия кодон—антикодон в А-участке, можно определять расстояние, на которое перемещается рибосома (гл. 8).

У эукариот аналогом EF-G является белок eEF2, функционирующий, по-видимому, сходным образом, как транслоказа, зависящая от гидролиза GTP. Действие этого фактора также ингибируется фузидиевой кислотой, хотя и в меньшей степени, чем в случае бактериальных рибосом. Фактор eEF2 является крупным белком с мол. массой около 100 000 дальтон. Это один из самых многочисленных компонентов клетки, на долю которого приходится, вероятно, около 1% растворимого белка. В отли-

чие от бактериального фактора удается выделить стабильный комплекс eEF2 с GTP. Комплекс может связываться с рибосомами с последующим гидролизом входящего в его состав GTP.

Уникальным свойством eEF2 является его чувствительность к дифтерийному токсину. Действие токсина, использующего NAD (никотинамидадениндинуклеотид) в качестве кофактора, заключается в переносе аденозиндифосфатрибозильного компонента (ADPR) на eEF2. Образующийся ADPR-eEF2 неактивен в синтезе белка. Субстратом для этой модифицирующей реакции служит необычная аминокислота, которая образуется в eEF2 в результате изменения гистидинового остатка. Она обнаружена в составе eEF2, выделенного из представителей многих видов. Такое ADP-рибозилирование и обуславливает летальный эффект дифтерийного токсина. Реакция чрезвычайно эффективна: одна молекула токсина способна модифицировать такое количество фактора eEF2, которое вызывает гибель клетки. Существуют мутантные линии клеток (но не целые организмы), в которых eEF2 устойчив к ADP-рибозилированию.

Извлечение энергии, необходимой для работы рибосомы

Для включения одной аминокислоты в полипептидную цепь в сумме затрачивается три высокоэнергетические связи. Во время присоединения аминокислоты к тРНК расщепляется АТФ. Гидролиз GTP происходит после включения аминоацил-тРНК (в составе тройного комплекса) в А-участок. Еще одна молекула GTP гидролизует, когда рибосома передвигается на один триплет по цепи мРНК (см. также гл. 8).

Молекулы GTP, участвующие в работе рибосомы, вероятно, выполняют две функции. Они поддерживают в со-

ответствующей конформации факторы элонгации, с которыми ассоциированы. Когда фактор элонгации связан с GDP, он не способен выполнять свои функции. GTP также обеспечивает энергией соответствующие этапы в синтезе белка, в каждом случае, вероятно, вызывая конформационные изменения частицы.

При поступлении аминоацил-тРНК в рибосому энергия расходуется на придание последней нужной конформации, в которой аминоацил-тРНК располагается в А-участке и становится доступной для дальнейшего взаимодействия. Все эти события не простое следствие связывания с рибосомой. На стадии транслокации в результате гидролиза GTP высвобождается энергия, необходимая для передвижения рибосомы. В каждом случае характерным признаком реакции является высвобождение фактора элонгации из рибосомы.

В эукариотических системах провести точный подсчет энергетических затрат представляется более трудной задачей, но, вероятно, и здесь потребность в энергии столь же велика, как и в случае бактериальных рибосом.

Терминация: завершение синтеза белка

Окончание белкового синтеза — это необычная реакция, так как при этом происходит непосредственное узнавание терминирующего кодона белковым фактором. Поскольку эта реакция совершенно отличается от взаимодействия кодон — антикодон, характерного для инициации и элонгации, то, по-видимому, совсем не обязательно, чтобы и в терминации участвовала последовательность, состоящая из трех нуклеотидов. Очевидно, этот факт отражает некоторые аспекты эволюции генетического кода.

Только 61 триплет кодирует аминокислоты. Остальные три триплета являются **терминирующими кодонами**, функция которых — прекращать белковый синтез. Им даны случайные названия, отражающие историю их открытия. Триплет UAG называется амбер-кодоном; UAA — охра-кодоном и UGA иногда называемый опал-кодоном.

Природа этих триплетов первоначально была установлена с помощью генетических тестов, благодаря которым было выделено два типа точковых мутаций.

У *E. coli* обнаружено два белка, участвующих в терминации. Они получили название **факторы освобождения (RF)**. Эти факторы заставляют рибосому связываться с нуклеотидными триплетами; RF1 узнает UAA и UAG, а RF2 — UGA и UAA соответственно. Для реакции освобождения полипептидной цепи необходимо, чтобы в Р-участке присутствовала полипептидил-тРНК. Факторы RF, по-видимому, взаимодействуют с участком А, так как все мутантные тРНК, способные узнавать терминирующие кодоны, могут конкурировать с ними за рибосому (гл. 7). Фактор RF3 стимулирует действие двух первых.

В эукариотических системах был обнаружен только один фактор освобождения — eRF. Для его связывания с рибосомой, по-видимому, требуется GTP (в случае бактерий этого не нужно). Возможно, гидролиз GTP происходит после завершения терминации; это, вероятно, необходимо для диссоциации eRF от рибосомы.

Реакция терминации включает в себя освобождение полипептида из связи с последней тРНК, вытеснение тРНК из рибосомы и отделение рибосомы от мРНК. Отделение полипептида и тРНК с последующим их освобождением в клетку происходит в результате реакций, аналогичных тем, которые рассматривались в случае элонгации цепи, но модифицированных для нужд терминации. Мы, к сожалению, еще не знаем механизмов, обеспечивающих отделение рибосомы от мРНК. Диссоциация, возможно, обусловлена конформационными изменениями, вызываемыми белком RF. Не исключено, что в этом процессе может участвовать какой-либо белковый фактор или факторы.

Рекомендуемая литература

В большинстве современных обзоров внимание концентрируется на факторах, контролирующих стадии белкового синтеза. Среди них статьи Вейсбаха, Очоа (*Weissbach, Ochoa, Ann. Rev. Biochem.*, **45**, 191–216, 1976); Лодиша (*Lodish, Ann. Rev. Biochem.*, **45**, 39–72, 1976); Очоа, де Хара (*Ochoa, de Haro, Ann. Rev. Biochem.*, **48**, 549–580, 1979). Самые последние данные по инициации и регуляции синтеза белка у эукариот приведены в статье Джагуса, Андерсона, Сафера (*Jagus, Anderson, Safer, Prog. Nucleic Acid. Res.*, **25**, 127–185, 1981).

Глава 7

ТРАНСПОРТНАЯ РНК: ТРАНСЛЯЦИОННЫЙ ПОСРЕДНИК

Транспортная РНК занимает центральное место в биосинтезе белка, выполняя функцию адапторной молекулы, ответственной за перевод (трансляцию) последовательности нуклеотидных триплетов в последовательность аминокислот. По выражению Крика (Crick), тРНК возникла в результате попытки природы наделить нуклеиновую кислоту свойствами, обычно присущими белку. Поскольку эти РНК принимают участие во множестве различных реакций, они должны быть похожи по одним свойствам и различаться по другим. Это требование удо-

влетворяется благодаря способности молекул тРНК складываться, образуя при этом характерную третичную структуру.

Рассматривая участие тРНК в биосинтезе белка, начиная с заключительных стадий этого процесса, можно заметить, что все тРНК способны располагаться в рибосомных участках А и Р. При этом одним концом в результате кодон-антикодонного узнавания они взаимодействуют с мРНК, тогда как другой конец участвует в перемещении полипептида. Чтобы участки А и Р были

доступны для всех молекул тРНК, конфигурация этих молекул по размеру и форме полностью должна соответствовать одному и тому же шаблону.

Аналогичным образом все тРНК, за исключением инициаторных, обладают способностью взаимодействовать с факторами белкового синтеза (EF-Tu или eEF1) для последующего связывания с рибосомой. Инициаторные тРНК узнаются факторами IF2 или eIF2. Таким образом, для взаимодействия с факторами элонгации молекулы тРНК должны обладать определенными общими чертами, но у инициаторных тРНК эти структурные особенности отсутствуют или же изменены и служат для других целей.

В то же время должны существовать значительные различия между отдельными группами тРНК. В большинстве случаев одной аминокислоте соответствует несколько тРНК. Такие тРНК называются **изоакцепторными тРНК**. Они узнаются единственной аминоацил-тРНК-синтетазой, специфичной только для данной аминокислоты. Поэтому изоакцепторные тРНК должны обладать общими свойствами, дающими возможность ферменту выбирать среди различных тРНК ту, которая соответствует данной аминокислоте. Таким образом, полный набор тРНК поделен на 20 изоакцепторных групп, что соответствует количеству аминокислот, встречающихся в белках, причем каждая группа обладает уникальным свойством, благодаря которому она и узнается специфической синтетазой.

Что же собой представляют те характерные особенности в структуре тРНК, которые узнаются рибосомой и факторами элонгации? Чем различаются между собой группы изоакцепторных тРНК? Первоначальные попытки выявить характерные особенности в структуре тРНК были направлены на поиски коротких нуклеотидных последовательностей, общих для всех или для определенной группы тРНК. Несмотря на то что некоторые участки первичной последовательности тРНК высококонсервативны и необходимы, по-видимому, для выполнения ряда функций, тем не менее попытки связать отдельный участок или последовательность с определенной реакцией белок-нуклеинового узнавания не увенчались успехом. Эти общие (или различные) особенности в строении тРНК могут выйти за рамки первичной и вторичной структур и выразиться, по крайней мере частично, на уровне пространственной организации молекулы. Для детального анализа этого предположения необходимо сравнение нескольких (возможно, всех) представителей семейства тРНК. Однако это займет много времени, поскольку сейчас третичная структура известна только для малого числа индивидуальных тРНК.

Универсальная структура клеверного листа

В настоящее время определены первичные последовательности нескольких сот тРНК, полученных из различных источников. Все эти последовательности обладают общей вторичной структурой, предложенной на основе анализа последовательности первой из расшифрованных тРНК. Оказалось, что последовательность любой тРНК может быть изображена в форме **клеверного листа**, образуемого в результате спаривания оснований между короткими комплементарными участками (см. рис. 5.4). Реальность такой конформации, а не какой-либо из других возможных впервые была продемонстрирована в опы-

тах, в которых препараты тРНК обрабатывали нуклеазами. При этом было установлено, что если основания расположены в неспаренных участках клеверного листа, то как и предполагалось, этот сайт наиболее чувствителен к действию нуклеаз.

На **рис. 7.1** изображена наиболее часто встречающаяся форма клеверного листа. Как видно из рисунка, в составе тРНК можно выделить четыре основные **ветви**, обозначаемые в соответствии с их структурой или функцией. **Акцепторная ветвь** состоит из двухцепочечного стебля, который заканчивается одноцепочечной последовательностью, свободная 2'/3'-ОН группа которого подвергается аминоацилированию. Остальные ветви состоят из двухцепочечных стеблей, на конце которых находятся неспаренные участки, образующие петли. **ТψС-ветвь** названа так по имени входящего в нее триплета ТψС. **Антикодоновая ветвь** всегда содержит антикодоновый триплет в центре петли. D-петля получила свое название благодаря тому, что содержит дигидроуридин (ψ обозначает псевдоуридин, а D—дигидроуридин; эти два «необычных» основания, встречающиеся в тРНК, рассматриваются ниже).

Как видно из рис. 7.1, каждый остаток в тРНК имеет свой определенный номер, что говорит о постоянстве структуры этих молекул. Наиболее стандартная структура тРНК содержит 76 остатков; они пронумерованы в направлении от 5'-конца к 3'-концу. Размеры молекул тРНК колеблются в пределах от 74 до 95 оснований. Эти отклонения обусловлены вариациями только в двух ветвях.

В D-петле вариации в размере петли могут достигать четырех оснований. Дополнительные нуклеотиды обозначаются как 17:1 (дополнительный нуклеотид расположен между 17 и 18 положениями), 20:1 и 20:2 (дополнительные нуклеотиды находятся между 20 и 21 положениями). Однако в наименьшей из известных D-петель могут отсутствовать как остаток под номером 17, так и эти три дополнительных нуклеотида.

Наибольшую вариабельность в структуру тРНК вносит так называемая **дополнительная ветвь**, находящаяся между антикодоновой и ТψС-петлями. В зависимости от размеров дополнительной ветви молекулы тРНК можно подразделить на два класса. **тРНК класса 1** имеют небольшую дополнительную петлю, состоящую только из 3–5 оснований. Такие тРНК составляют 75% от общего числа молекул. **К тРНК класса 2** относятся тРНК, имеющие большие дополнительные ветви (они могут даже являться наиболее длинными в тРНК), состоящие из петли размером от 13 до 21 основания и двухцепочечного стебля размером около 5 пар оснований. Дополнительные основания имеют номера от 47:1 до 47:16.

Комплементарные основания, поддерживающие структуру тРНК, как правило, консервативны. Двигаясь по часовой стрелке вдоль структуры клеверного листа, всегда можно обнаружить 7 комплементарных пар оснований в акцепторном стебле, 5—в ТψС-ветви, 5—в антикодоновой ветви и обычно 3 (иногда 4) в D-ветви. В пределах данной тРНК большинство комплементарных пар образуются в результате взаимодействия А—U и G—C, но иногда встречаются пары G—U, G—ψ или А—ψ. Дополнительные типы спаривания менее стабильны, чем обычные пары, но и они способствуют образованию двуспиральной структуры в РНК.

При сравнении первичных последовательностей тРНК можно видеть, что основания, находящиеся в некоторых позициях, являются **консервативными** (или **инвариантными**).

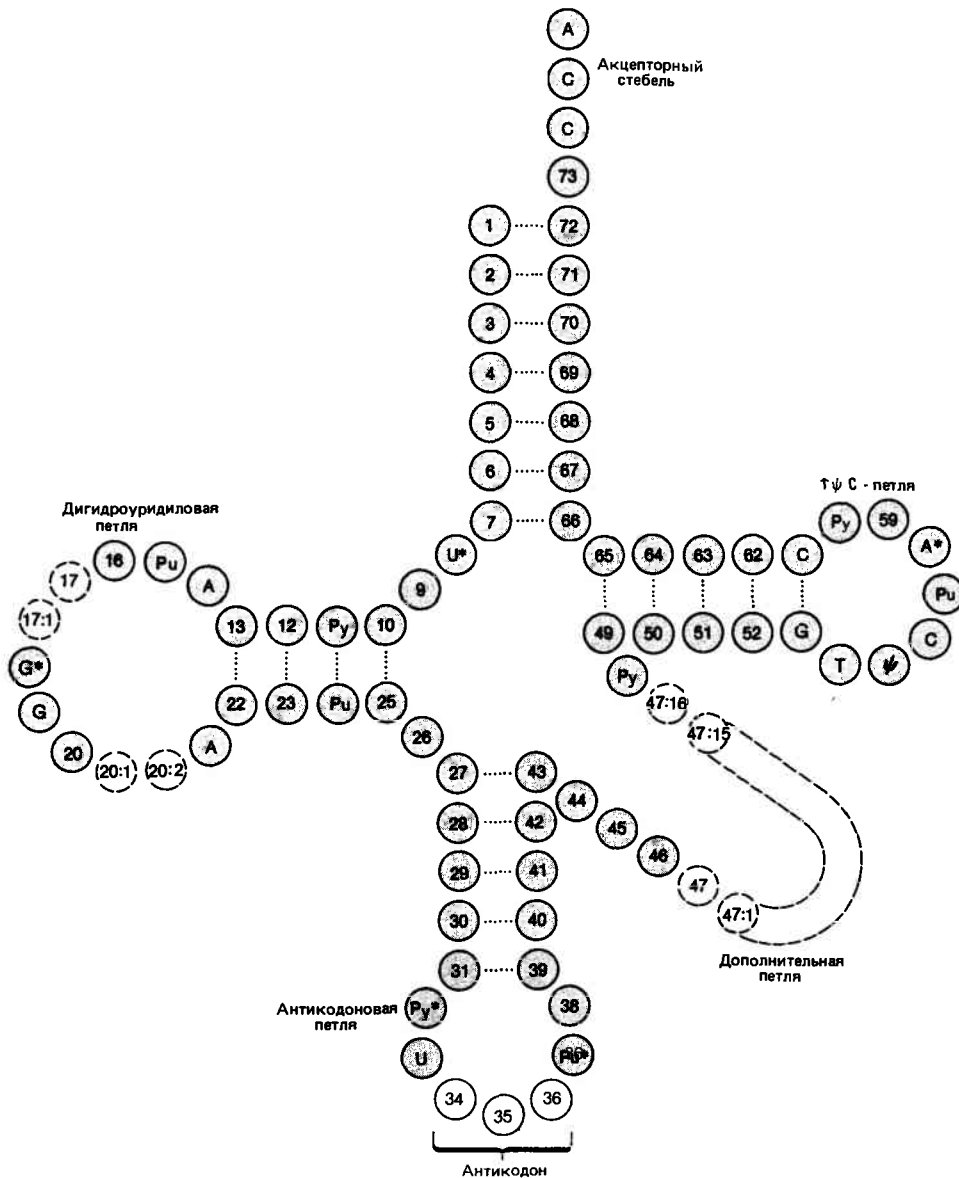


Рис. 7.1. Транспортная РНК образует клеверный лист.

Кружками показано положение оснований. Номера в закрашенных кружках обозначают такие положения, которые всегда имеются в молекулах тРНК, но в которых может находиться любое основание. Если в определенном положении всегда находится одно и то же основание, то оно указано в соответствующем кружке. Для консервативных оснований ука-

заны их конкретные значения, для полуконсервативных оснований Рy обозначает любой пиримидин, а Рu – любой пуриин. Наличие звездочки говорит о том, что основание модифицировано, но конкретная форма модификации может быть различной. Светлыми кружками показаны положения, имеющиеся не у всех молекул тРНК. Точки указывают на наличие спаренных оснований.

ми), т.е. в этом положении почти всегда обнаруживается определенное основание. На самом деле при анализе большого количества тРНК оказалось, что в позициях, которые раньше казались полностью инвариантными, иногда встречаются замены. Поэтому для практических целей надо иметь в виду, что если какая-либо позиция считается консервативной, то это означает, что данное основание присутствует в более чем 95% тРНК с известной последовательностью оснований. Иногда замены оснований носят индивидуальный характер; иногда их можно разделить на группы, отражающие особенности данной клетки.

Некоторые позиции являются **полуконсервативными** (или **полуинвариантными**), так как в них обнаруживается только один тип оснований (пуриин или пиримидин), но при этом любое основание данного типа может присутствовать в этом положении.

тРНК содержит много модифицированных оснований

Одна из характерных особенностей транспортных РНК, отличающих их от других нуклеиновых кислот, – это наличие в молекуле необычных оснований. Необычным основанием называется любое пуриновое или пиримидиновое кольцо, за исключением А, G, C или U, из которых синтезируются все РНК. Все другие основания образуются в результате **модификаций** одного из четырех названных оснований; модификация происходит после того, как основание будет включено в полирибонуклеотидную цепь. Во всех классах РНК обнаруживается некоторая степень модификации оснований, но во всех остальных случаях, за исключением тРНК, эти модификации сводятся к простым реакциям, таким, как введение метильных групп. Только в случае тРНК обнаруживают-

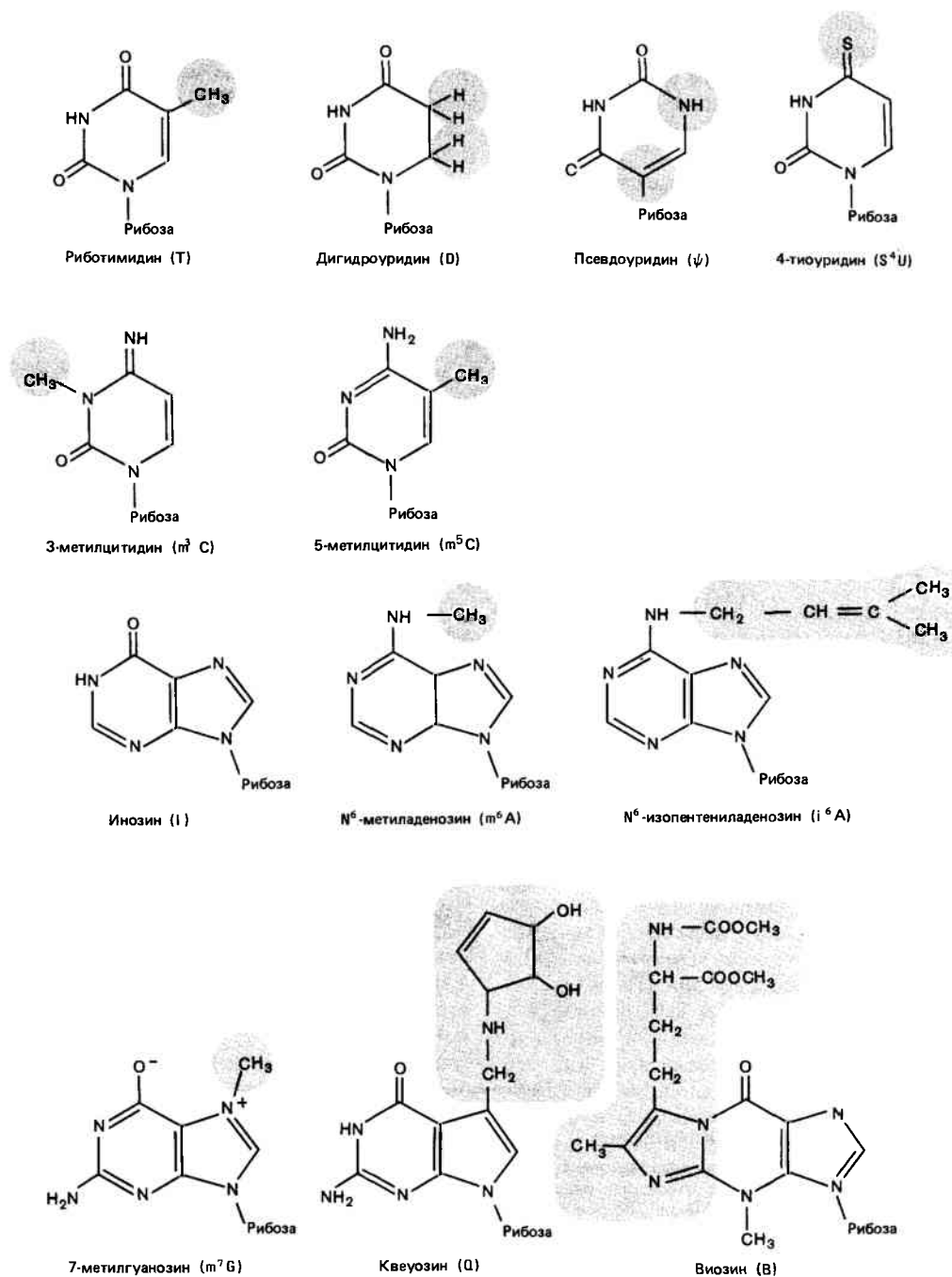


Рис. 7.2. Некоторые модифицированные основания, обнаруженные в тРНК (модификации показаны оттенением).

ся разнообразные модификации, начиная от простого метилирования до крупномасштабного перестраивания пуриновых колец. Они обуславливают более широкое структурное многообразие тРНК, которое, по-видимому, существенно для выполнения их различных функций.

На рис. 7.2 изображены некоторые из наиболее часто встречающихся модифицированных оснований. Наблюдаемая тенденция состоит в том, что у пиримидинов (С и U) модификации менее сложны, чем у пуринов (А и G). Некоторые из них затрагивают основания, принимающие участие в комплементарном взаимодействии, и поэтому приводят к изменению стабильности пар или даже к сме-

не партнера, с которым эта пара была образована. Прочие модифицированные основания могут располагаться в любых других местах, не влияя на способность к образованию комплементарных пар.

Простая, наиболее часто встречающаяся модификация затрагивает уридин. Метилирование в пятом положении кольца приводит к появлению риботимидина (Т). Такое же основание обнаруживается и в ДНК, но если в РНК тимидин соединен с дезоксирибозой, то в случае тРНК он связан с рибозой. Тимин, присутствующий в РНК, является необычным основанием, возникающим в результате модификации U (в ДНК же урацил будет считаться

необычным основанием, иногда возникающим при воздействии мутагеном, как это уже рассматривалось в гл. 2).

В молекуле дигидроуридина (D) имеется насыщенная двойная связь, изменяющая структуру кольца. В псевдоуридине (Ψ) кольцо повернуто таким образом, что атомы N и C меняются местами. В 4-тиоуридине на месте кислорода находится атом серы. Другие превращения, которые могут происходить с уридином, состоят из сочетания описанных выше изменений или же заключаются в введении более сложных групп. Например, оксиуксусная кислота может быть включена в пятое положение кольца.

Модификации цитидина чаще всего сводятся к метилированию, а также ацетилированию 4-го положения и образованию 2-тиоцитидина.

Инозин представляет собой основание, которое обычно обнаруживается в клетке в качестве промежуточного продукта при биосинтезе пуринов. Тем не менее он непосредственно не включается в РНК; его появление связано с дезаминированием А. Другие модификации А включают присоединение различных сложных групп.

При модификации G образуются две сложные серии нуклеотидов. Наиболее просто устроенным основанием, относящимся к типу Q, является **квеуозин**, содержащий дополнительное пентозное кольцо, присоединенное с помощью NH_2 -связи к метильной группе 7-метилгуанозина. Пентозное кольцо может содержать различные дополнительные группы. Среди оснований типа Y простейшим является **виозин**. Это основание содержит дополнительное кольцо, слитое воедино с пуриновым кольцом и соединенное с длинной углеродной цепью, к которой в свою очередь присоединены различные дополнительные группы.

Кроме описанных модификаций, затрагивающих структуру оснований, можно также обнаружить изменения, связанные с метилированием рибозного кольца в 2'-О-положении.

Во всех случаях, за исключением одного, суть реакции модификации состоит в изменении или в добавлении определенной группы к основаниям, уже включенным в состав тРНК. Исключение представляет синтез основания типа Q. В этом случае был обнаружен фермент, способный обменивать свободный квеуозин на гуаниновый остаток, включенный в тРНК. Реакция состоит из разрыва и последующего восстановления связей по обе стороны от нуклеозида. Такое превращение является беспрецедентным в биосинтезе нуклеиновых кислот.

Перечень модифицированных нуклеозидов, обнаруженных в настоящее время в тРНК, включает более 40 названий. Они синтезируются особыми ферментами, модифицирующими тРНК. Нуклеозид, присутствовавший исходно в каком-то положении тРНК, может быть идентифицирован путем сравнения последовательностей тРНК и кодирующего ее гена. Эту задачу возможно также решить, если выделить молекулы-предшественники тРНК, не подвергшиеся модификациям. С помощью последнего подхода было показано, что различные модификации в тРНК вносятся на разных стадиях ее образования из предшественника.

Каким образом тРНК-модифицирующие ферменты узнают свои субстраты, неизвестно. В некоторых случаях данная модификация в тРНК обнаруживается более чем в одной позиции. Например, псевдоуридин может находиться в нескольких местах. Не исключено, что для синтеза псевдоуридина, присутствующего в различных поло-

жениях, необходимы и различные модифицирующие ферменты. Нам не известно, чем определяется специфичность этих ферментов, но очевидно, что существует большое количество ферментов с различной специфичностью. Некоторые ферменты могут осуществлять строго определенные реакции с индивидуальными тРНК; другие, напротив, обладают широкой субстратной специфичностью. В некоторых реакциях для проведения нужных модификаций может участвовать более одного фермента.

Ряд модифицированных оснований присутствует во всех молекулах тРНК — например, дигидроуридин, давший название D-петле, и псевдоуридин, находящийся в Т Ψ U-последовательности. За антикодоном (со стороны 3'-конца) всегда находится модифицированный пурин, хотя в каждом конкретном случае характер этой модификации может широко варьировать.

Таким образом, модификации затрагивают все участки молекулы тРНК. В характере модификаций наблюдается как некоторая видоспецифичность, так и некоторая специфичность, определяемая самой разновидностью тРНК. Например, основание виозин характерно для тРНК^{Phe}, выделяемой из бактерии, дрожжей и млекопитающих.

В настоящее время мы недостаточно хорошо понимаем функциональное значение модификаций. Наиболее непосредственное значение этого явления наблюдается в антикодоне, когда модификация может влиять на специфичность взаимодействия с кодоном, определяя тем самым конкретное значение данной тРНК. Другие модификации, находящиеся вблизи от антикодона, также могут сказываться на специфичности этого взаимодействия.

L-образная пространственная структура тРНК

Форма клеверного листа, изображающая вторичную структуру тРНК, не позволяет составить правильное представление о третичной структуре этой молекулы, которая в действительности является компактной. Для исследования третичной структуры тРНК необходимо получить ее кристаллы, пригодные для рентгеноструктурного анализа. Сделать это технически трудно, и в настоящее время только несколько тРНК, полученные в основном из дрожжей, исследованы подробно. Тем не менее сходство структур и природы взаимодействий, поддерживающих каждую исследованную пространственную структуру, свидетельствуют о том, что, возможно, имеется общая основа третичной структуры, свойственная всем тРНК, хотя каждой и присущи свои особенности.

Двухцепочечные участки вторичной структуры сохраняются и в третичной структуре, но при этом их расположение в пространстве, по существу, создает две двойные спирали, ориентированные под прямым углом по отношению друг к другу, так, как это изображено на **рис. 7.3**. Акцепторный и Т Ψ С-стебли образуют одну двойную спираль с одноцепочечным интервалом. Антикодоновый и D-стебли формируют другую двойную спираль, в которой также имеется брешь. Область между двойными спиралями, в которой происходит изгиб, содержит Т Ψ С-и D-петли. Таким образом, аминокислотный остаток находится на одном конце буквы L, в то время как антикодоновая петля располагается на противоположном конце молекулы.

Данная модель в деталях была разработана для дрож-

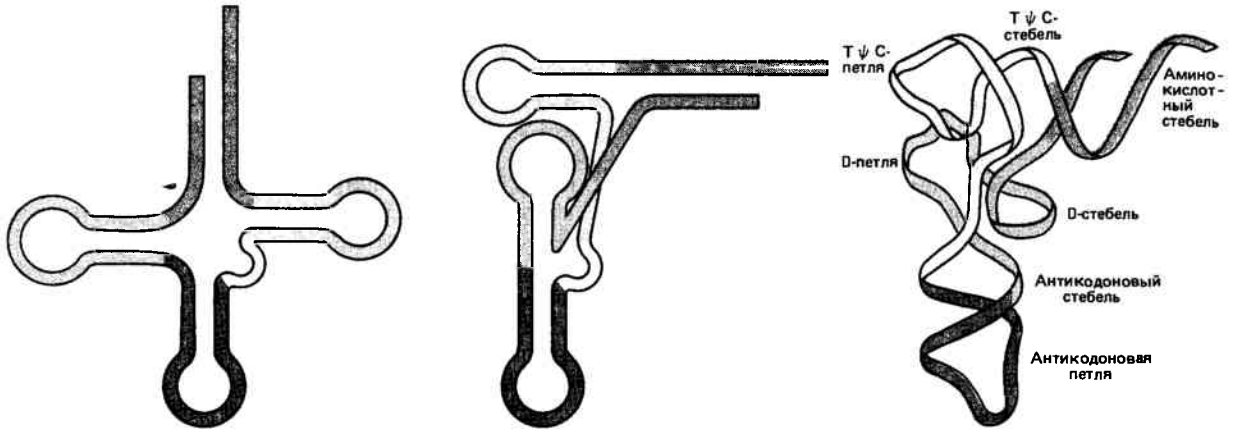


Рис. 7.3. Транспортная РНК сложена в компактную L-образную третичную структуру.

На левом рисунке показан код, используемый для обозначения различных участков клеверного листа. В центре изображена двумерная про-

екция, показывающая, как накладываются две двойные спиральные области, образуя прямые углы. Справа приведена схема укладки молекулы, построенная с учетом комплементарных и неспаренных областей. В данном случае изображена структура дрожжевой тРНК^{Phe}. Другие тРНК имеют аналогичную структуру.

жевой тРНК^{Phe} – первой молекулы тРНК, которую удалось получить в кристаллическом виде. Ее молекулярная модель изображена на рис. 7.4. В случае тРНК^{Asp} угол между двумя спиралями немного больше, поэтому молекула обладает сходной, но несколько более открытой конформацией. В других тРНК также обнаруживаются небольшие различия, затрагивающие ту или иную область. Эти данные проливают свет на ранее непонятное обстоятельство: с одной стороны, все тРНК должны обладать сходной формой, а с другой – должна существовать возможность различить их.

Кроме тех пар оснований, которые уже существуют во вторичной структуре, при образовании третичной структуры формируются дополнительные водородные связи. Водородные связи, поддерживающие структуру клеверного листа, обозначаются как **вторичные водородные связи**; дополнительные связи, обуславливающие формирование третичной структуры, называются **третичными водородными связями**. Многие консервативные и полуконсервативные основания принимают участие в образовании третичных водородных связей. Это, по-видимому, лежит в основе их консервативности и свидетельствует в пользу

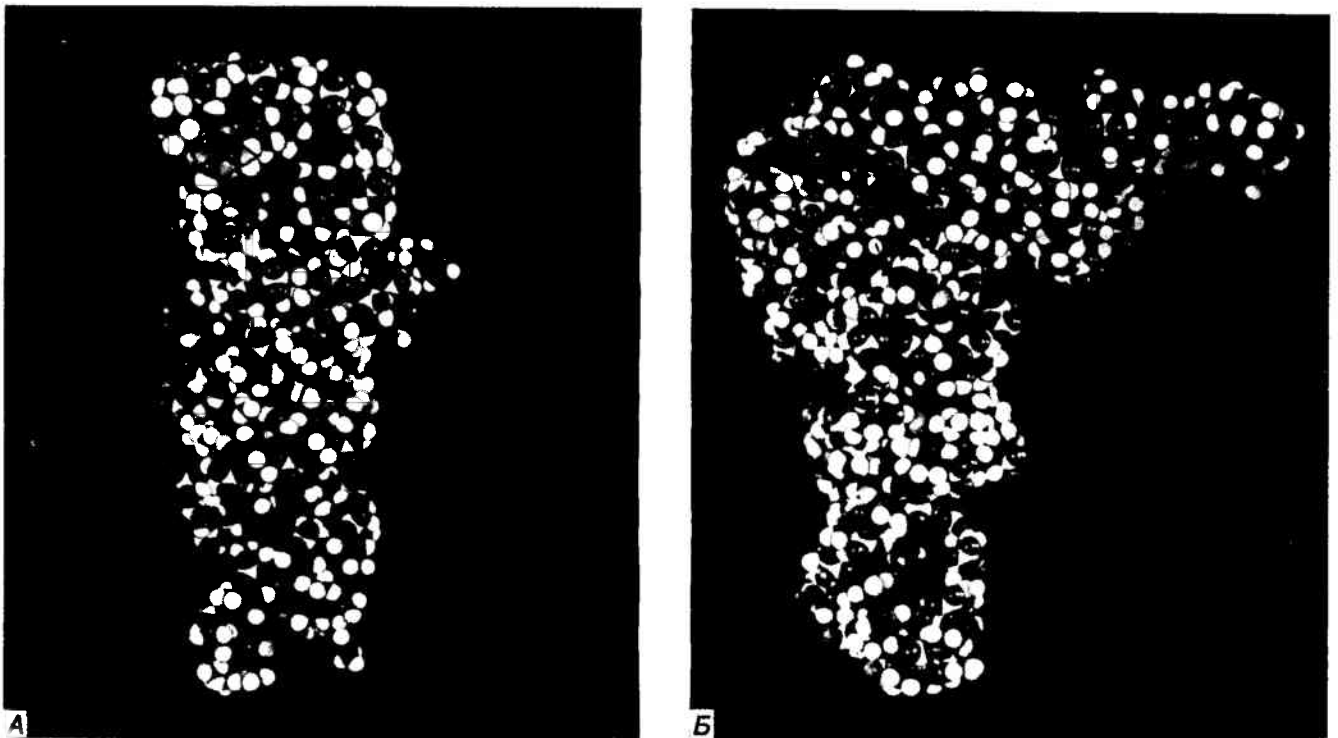


Рис. 7.4. Пространственная модель, показывающая, что третичная структура тРНК компактна.

Модель дрожжевой тРНК^{Phe}, изображенная в двух проекциях, отличаю-

щихся на 90°. Проекция в правой части соответствует двумерной проекции, приведенной на предыдущем рисунке. (Фотография любезно предоставлена S. H. Kim.)

того, что для всех тРНК может существовать единый план строения третичной структуры. На рис. 7.5 изображены дополнительные водородные связи, образующиеся в случае тРНК^{Phe}. Можно видеть, что в формировании некоторых из этих связей принимают участие необычные основания или же эти связи образуются за счет взаимодействия между тремя основаниями. Большая часть третичных взаимодействий возникает между D- и ТΨС-петлями. Не в каждой тРНК формируются все возможные третичные связи. Например, некоторые из них отсутствуют в тРНК^{Asp}. Но очевидно, что такие взаимодействия являются универсальным способом создания пространственной структуры тРНК.

Рентгеноструктурный анализ показал, что молекула тРНК обладает стабильной структурой. Некоторая гибкость у молекулы появляется при формировании комплекса с белками или в растворе. При этом, возможно, изменяется ее конформация, но не известно, как это отражается на структуре кристалла.

Исходя из структуры молекулы тРНК, можно высказать следующие соображения о ее функционировании. Во-первых, участки, ответственные за выполнение различных функций, максимально разобщены. Аминокислота удалена от антикодона на максимально возможное расстояние. Это согласуется с тем, что аминоацильная группа находится на большой субчастице рибосомы рядом с пептидилтрансферазным участком, тогда как взаимодействие антикодона с мРНК происходит на малой субчастице (см. гл. 8). Последовательность ТΨС лежит в углу L-подобной структуры; вероятно, такое ее положение имеет критическое значение для осуществления контроля за изменениями в третичной структуре. (В гл. 8 бу-

дет рассмотрено предположение, согласно которому такие изменения могут происходить в процессе синтеза белка.) Таким образом, с помощью нескольких различных доменов, аналогично тому, как это происходит у белков, структура тРНК может приспосабливаться для выполнения разнообразных функций, присущих тРНК.

Синтетазы ответственны за подбор соответствующих друг другу аминокислот и тРНК

Аминокислоты вовлекаются в белковый синтез с помощью аминоацил-тРНК—синтетаз. Эти ферменты обеспечивают взаимодействие аминокислот с тРНК. Синтетазы являются очень интересными ферментами, способными сортировать тРНК и аминокислоты в соответствии с их специфичностью. У бактерий имеется по одной аминоацил-тРНК—синтетазе для каждой аминокислоты, причем этот единственный фермент распознает все изоакцепторные тРНК, обслуживающие данную аминокислоту. Вероятно, похожая ситуация наблюдается в цитоплазме эукариот, хотя возможны исключения, и в некоторых случаях имеется более одного фермента, способного присоединять к молекулам тРНК соответствующую аминокислоту. В свою очередь органеллы обладают собственным набором этих ферментов.

Назначение аминоацил-тРНК—синтетазы хорошо известно. Ее функция состоит в том, чтобы обеспечить взаимодействие трех субстратов: тРНК, аминокислоты и АТР. Энергия гидролиза высокоэнергетической связи АТР используется для присоединения аминокислоты

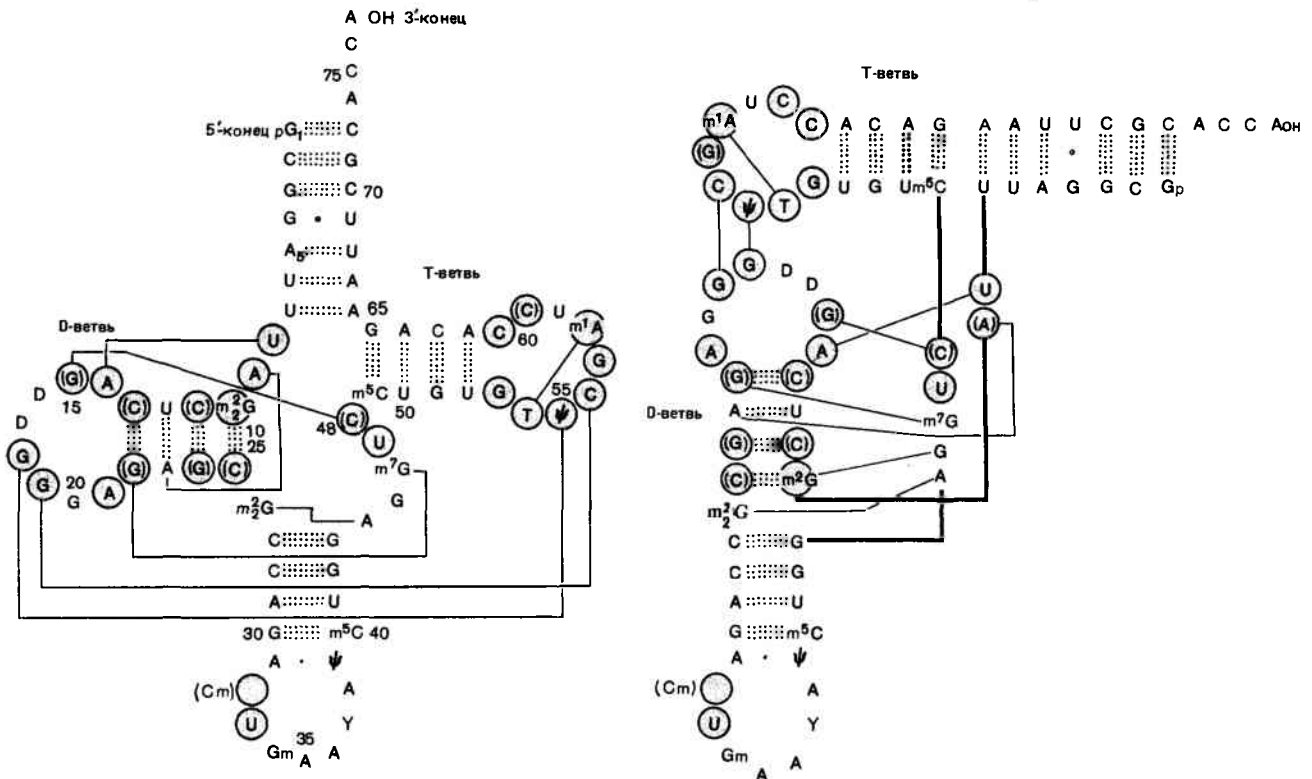


Рис. 7.5. Большинство третичных взаимодействий в структуре тРНК образуется с участием консервативных и полуконсервативных оснований.

Молекула изображена в виде клеверного листа (слева) или в виде L-об-

разной структуры (справа). Третичные водородные связи показаны тонкими линиями. Толстые линии обозначают ковалентные связи, соединяющие две части L-образной структуры. Закрашенные кружки соответствуют консервативным или полуконсервативным основаниям.

к молекуле тРНК. Промежуточным продуктом этой реакции является комплекс АМР—аминокислота. Обобщая сведения, приведенные на рис. 5.6, можно заключить, что, по-видимому, тот же самый двухстадийный процесс характерен для всех синтезас.

Однако строение этих ферментов очень разнообразно. В настоящее время среди них выделяют четыре класса, обозначаемых соответственно: α , α_2 , α_4 и $\alpha_2\beta_2$, где α и β представляют два типа полипептидных цепей. Таким образом, синтезасы могут существовать в форме мономеров (α) с мол. массой от 50 000 до 120 000 дальтон или мультимеров, состоящих из одного или двух типов полипептидных цепей с мол. массой до 300 000 дальтон. Следовательно, одни и те же функции, выполняемые аминокцил-тРНК—синтезасой, могут быть осуществлены несколькими различными путями.

Как синтезаса узнает определенную группу тРНК и соответствующую ей аминокислоту, является одним из основных вопросов при исследовании белок-нуклеинового взаимодействия. Многочисленные тРНК, соответствующие определенной аминокислоте, могут быть сходны в двух отношениях. Во-первых, могут существовать тРНК, обладающие *одинаковыми антикодонами*, но различающиеся по другим участкам молекулы. Такие тРНК будут однозначно отвечать на один и тот же кодон. Во-вторых, возможно существование тРНК с *разными антикодонами*, узнающими различные синонимические триплеты, кодирующие одну и ту же аминокислоту. Как правило, такие тРНК имеют и другие различия в своей структуре. Молекулы тРНК, одинаково хорошо узнающиеся одной и той же синтезасой, называются **изоакцепторными**, «своими», специфическими.

При сравнении многих последовательностей изоакцепторных тРНК, а также из опытов по индукции химических модификаций, влияющих на специфичность акцентирования аминокислот, возникло представление о том, что местоположение оснований, участвующих в узнавании синтезас, не ограничивается каким-либо единственным участком первичной или вторичной структуры молекулы.

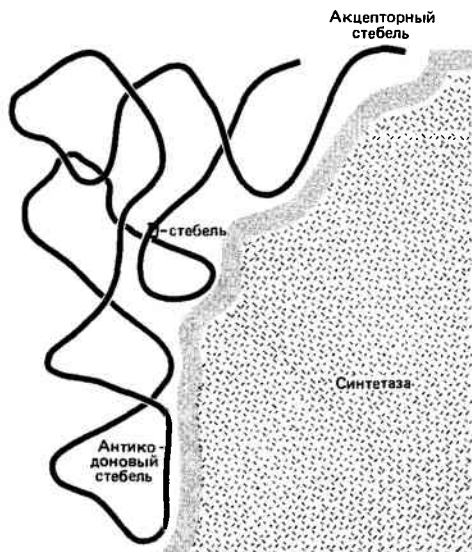


Рис. 7.6. Аминокцил-тРНК—синтезасы взаимодействуют с тРНК, контактируя со внутренней стороной L-образной структуры, преимущественно с акцепторным стеблем, D-стеблем и антикодоновым стеблем.

Не менее трех участков молекулы тРНК взаимодействуют с синтезасой. Несомненно, что в этот процесс вовлечен акцепторный стебель, так как аминокислота присоединена к 3'-концу тРНК. Участие D-стебля в этом взаимодействии продемонстрировано путем образования фотохимической сшивки между ферментом и тРНК. Это также доказывает, что данная область служит точкой контакта между синтезасой и тРНК. С ферментом часто сшивается также антикодоновый стебель. Эти данные свидетельствуют в пользу существования модели, изображенной на рис. 7.6. Согласно этой модели, фермент взаимодействует с тРНК, контактируя при этом со внутренней стороной L-образной структуры.

Это взаимодействие является чрезвычайно точным. Точковые мутации в тРНК могут изменять специфичность ее узнавания, нарушая взаимодействие с собственной синтезасой и позволяя реагировать с другими синтезасами. Мутации с такими свойствами локализуются в акцепторном стебле и в антикодоновой ветви. Таким образом, различия в свойствах тРНК, обуславливающие возможность их узнавания, могут быть очень тонкими; одна тРНК может отличаться от другой всего лишь одним основанием. (Конечно, не все мутации обладают таким действием; например, супрессорные мутации, рассматриваемые далее в этой главе, изменяют основание в антикодоне и тем самым кодоны, на которые тРНК реагирует, но не влияют на тип аминокислоты, соединяющейся с тРНК.) Так же были получены мутации в гене, кодирующем аминокцил-тРНК—синтезасу, влияющие на способность фермента распознавать тРНК. Эти результаты позволяют сделать общий вывод о том, что узнавание определяется взаимодействием между несколькими аминокислотами, формирующими активный центр фермента, и несколькими контактными точками в тРНК, сосредоточенными на оконечностях молекулы и в D-стебле.

Стадия активации

Основные закономерности белок-нуклеинового взаимодействия можно понять, изучая отдельные этапы, лежащие в основе экспрессии генов. Как синтезаса узнает нужные ей тРНК и аминокислоты? Как рибосомы узнают только ту тРНК, кодон которой находится в участке А? Как ДНК- и РНК-полимеразы выбирают основание, комплементарное матрице? В каждом из этих случаев возникает одна и та же проблема: как из целого набора элементов выбрать один нужный, при условии что все они обладают одинаковыми основными характеристиками.

Можно предположить, что существует равновероятный процесс, при котором любой компонент первоначально может проконтактировать с активным центром фермента, затем несоответствующие компоненты удаляются, и в реакции участвует только один нужный. Такой компонент всегда находится в меньшинстве (1 из 20 аминокислот, 1 из приблизительно 40 тРНК и 1 из 4 оснований), поэтому критерии его отбора должны быть строгими. Возможно существование двух основных способов, с помощью которых принимается решение о пригодности данного компонента для включения.

Во-первых, поступление, проверка правильности выбора, включение или удаление данного компонента могут представлять собой единый акт, объединяющий все перечисленные реакции. Это предположение равносильно тому, что в связывающий участок может попасть только

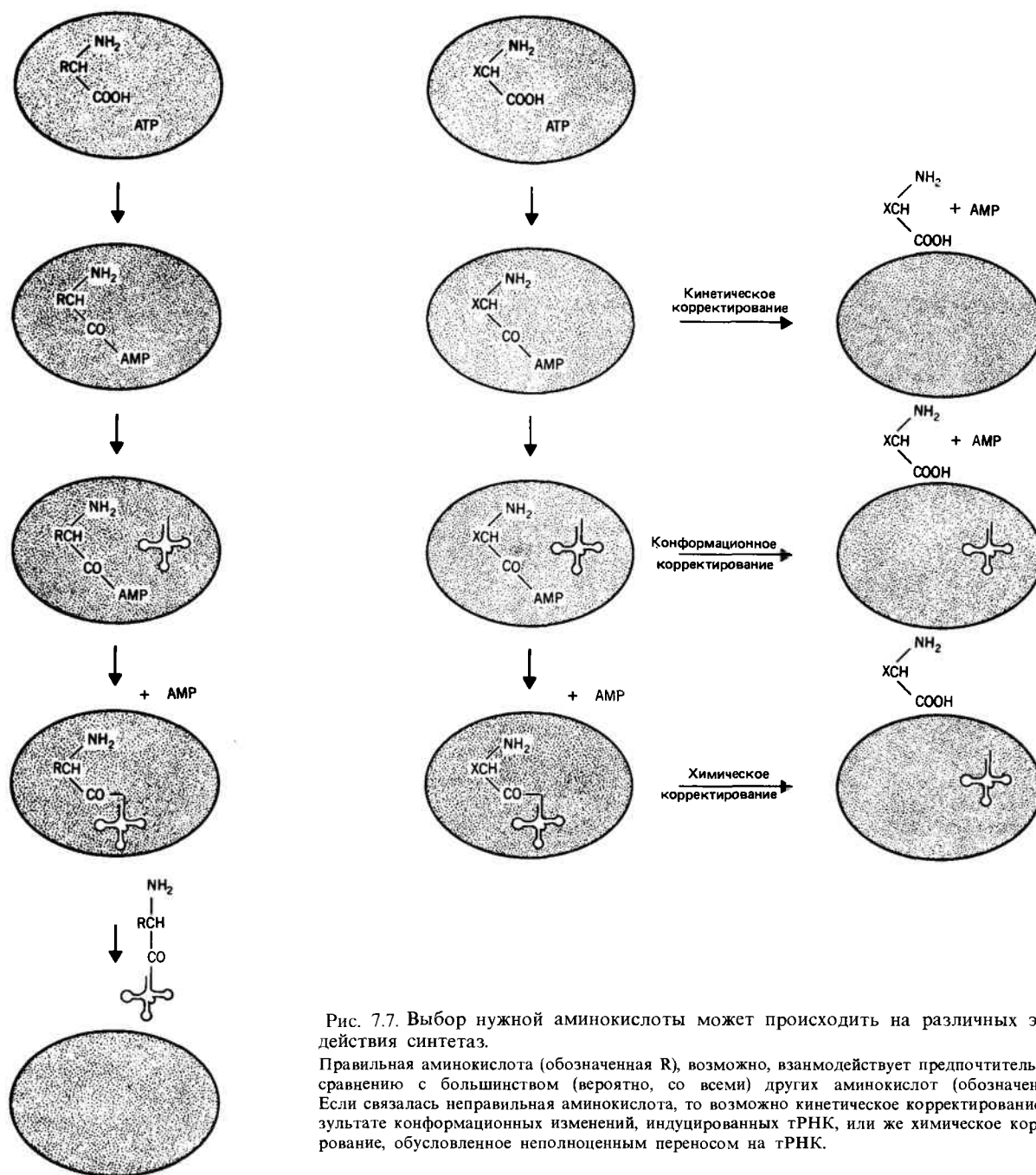


Рис. 7.7. Выбор нужной аминокислоты может происходить на различных этапах действия синтетаз.

Правильная аминокислота (обозначенная R), возможно, взаимодействует предпочтительнее по сравнению с большинством (вероятно, со всеми) других аминокислот (обозначены X). Если связалась неправильная аминокислота, то возможно кинетическое корректирование в результате конформационных изменений, индуцированных тРНК, или же химическое корректирование, обусловленное неполноценным переносом на тРНК.

нужный компонент. В случае синтетаз это означает, что только изоакцепторные молекулы тРНК обладают способностью стабильно прикрепляться к активному центру.

Другая возможность состоит в том, что реакция может пройти ряд стадий, прежде чем будет принято решение о правильности присутствующего компонента. Если компонент удовлетворяет предъявляемым критериям, то реакция продолжается до завершения. В противном случае реакция протекает в обратном или в обходном направлении, и в результате неподходящий компонент удаляется. Такой способ проверки, осуществляемый уже после того, как произошло связывание, принято называть **коррекцией**. В случае синтетаз это означает, что реакция может продолжаться далее, проходя через ряд стадий, даже если в системе присутствуют несоответствующие друг другу тРНК и аминокислота.

Способность синтетаз выбирать нужную тРНК и аминокислоту характеризуется некоторой вариабельностью. Об этом свидетельствует тот факт, что у различных ферментов существуют свои особые стадии, в большей или меньшей степени определяющие специфичность процесса. Например, синтетазы по-разному осуществляют первоначальное взаимодействие с похожими аминокислотами. Некоторые синтетазы обладают способностью активировать только определенную аминокислоту, тогда как другие могут это делать с целым рядом похожих по строению аминокислот. Эти аминокислоты могут быть превращены в аденилатную форму, но ни в одном из известных случаев неверно активированная аминокислота не образует стабильной аминоацил-тРНК.

На какой же стадии выбраковывается неправильный аминоациладенилат? На **рис. 7.7** показаны различные

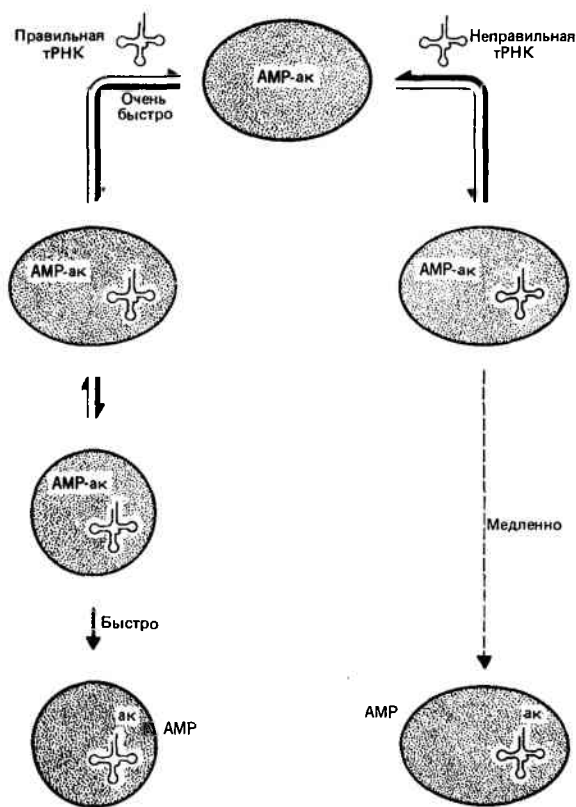


Рис. 7.8. Процесс узнавания правильной тРНК синтетазой контролируется в два этапа.

Во-первых, фермент обладает повышенным сродством к своим изоакцепторным тРНК, стабилизирующим связывание в результате конформационных изменений в ферменте. Во-вторых, аминоацилирование неправильной тРНК происходит очень медленно (V_{\max} падает в 10^4 раз по сравнению с изоакцепторными тРНК), что также увеличивает вероятность отделения такой тРНК.

способы, благодаря которым гипотетический фермент может удалять неправильную аминокислоту. Одна из возможностей состоит в том, что тщательной проверке подвергается аминокислота, а не аминокислота как таковая. При этом неправильный аминокислота будет гидролизован. Такая дополнительная проверка может непосредственно осуществляться в процессе активации аминокислоты (**кинетическая коррекция**), или же связывание тРНК с ферментом может индуцировать конформационные изменения (**конформационная коррекция**). Однако существуют доводы в пользу той точки зрения, что селекция осуществляется на последующем этапе, путем **химической коррекции**. При этом неправильная аминокислота переносится на тРНК, которая затем, находясь в тРНК-связывающем центре фермента, узнается как неправильная по своей структуре, гидролизуется и освобождается. Такой механизм коррекции подразумевает наличие постоянно действующего цикла, состоящего из связывания и последующего гидролиза, продолжающегося до тех пор, пока правильная аминокислота не присоединится к тРНК.

Классическим примером, показывающим, что для различения аминокислот необходимо присутствие тРНК, является работа Phe-tRNA —синтетазы *E. coli*. Этот фермент способен активировать валин, образуя валиладенилат, но после присоединения тРНК^{Ile} гидролизует ранее образовавшийся аденилат. Конечная ошибка при образовании аминокислота-тРНК определяется специфичностью

индивидуальных стадий. В рассмотренном выше примере неправильная активация валина происходит с частотой, оцениваемой как $1/225$. В свою очередь химическая коррекция позволяет образовываться только $1/270$ от общего числа молекул $\text{Val-tRNA}^{\text{Ile}}$. Таким образом, полная ошибка составляет $1/60\,000 = 1,5 \times 10^{-5}$. Эта величина сравнима со значениями, полученными при измерении частоты замены валина на изолейцин в глобине кролика (от $2 \cdot 10^{-4}$ до $5 \cdot 10^{-4}$). Следовательно, можно заключить, что неправильное активирование тРНК составляет только небольшую долю ошибок, происходящих на самом деле.

Взаимодействие тРНК с синтетазой включает в себя этапы, изображенные на рис. 7.8. Во-первых, это этап быстрого связывания, которое при условии поступления правильной тРНК стабилизируется путем конформационных изменений фермента. После этого события происходит аминоацилирование. Если же поступает неправильная тРНК, то реакция протекает гораздо медленнее и увеличивается вероятность того, что отделение тРНК произойдет раньше, чем она будет аминоацилирована. Возможно, решающим моментом, от которого зависит более медленное протекание реакции, является расположение 3'-конца тРНК, который может занять необходимое для аминоацилирования положение только при условии образования всех надлежащих контактов.

Поэтому у синтетаз для узнавания субстрата, возможно, существует двухкомпонентная система, в которой вначале достигается стабильное связывание, зависящее от правильности возникших контактов с тРНК, а затем сама протекающая каталитическая реакция определяется связавшейся тРНК.

Кодон-антикодоновое узнавание и гипотеза неоднозначного соответствия

Задача тРНК в синтезе белка состоит в том, чтобы, попав в А-участок рибосомы, узнать нужный кодон. Взаимодействие между антикодоном и кодоном происходит путем комплементарного спаривания оснований, но по правилам, которые значительно отличаются от обычных, ограничивающих контакты только парами G—C и A—U.

Имеющиеся экспериментальные данные не позволяют окончательно решить вопрос о силах, лежащих в основе кодон-антикодонового взаимодействия. Существуют даже некоторые сомнения о том, какую роль играет образование пар G—C в стабилизации взаимодействия по сравнению с парами A—U. Параметры взаимодействия, измеренные в растворе, могут отличаться от истинных, так как структура тРНК и сам участок А играют определенную роль в создании условий для ассоциации тРНК с мРНК.

Поэтому необходимо установить основные правила, управляющие взаимодействием между антикодонами и соответствующими кодонами. Способность любой тРНК узнавать определенный кодон выявляется с помощью теста на специфическое связывание с триплектидом или использованием тРНК в системе белкового синтеза *in vitro*. (Первоначально этот прием был использован для расшифровки генетического кода, как это уже было описано в гл. 4.)

Сама структура генетического кода позволяет сделать ряд важных заключений о процессе узнавания кодона. На рис. 7.9 приведены примеры, показывающие, что код

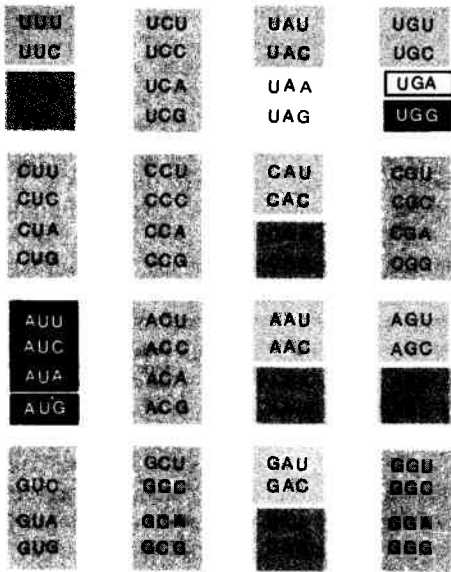


Рис. 7.9. Основания в третьем положении менее всего влияют на значения кодонов.

Код представлен в виде прямоугольников, содержащих группы кодонов, среди которых вырожденность основания в третьем положении обеспечивает их одинаковое значение. Возможно выделить четыре класса. Третье основание не играет роли в определении смысла кодона.

любой из U, C, A, G кодирует одинаковые аминокислоты: 32 таких кодона подразделяются на 8 семейств.

Различия проводятся только между пуринами и пиримидинами.

■ или U, или C кодирует одинаковую аминокислоту: 14 таких кодонов образуют 7 пар.

■ или A, или G кодирует одинаковую аминокислоту: 12 таких кодонов образуют 6 пар (включая терминирующие кодоны UAA/UAG).

Три из четырех

■ любое из U, C или A имеют одинаковое значение: существует только три таких кодона (AUC).

Уникальные значения

■ возможно присутствие только G: имеется только два таких кодона (AUG = Met, UGG = Trp).

□ возможно присутствие только A: таким кодоном является UGA (не узнающийся тРНК).

является вырожденным в том смысле, что третье основание в кодоне может быть неоднозначным: либо третье основание является некомплементарным, либо различия проводятся только между пуринами и пиримидинами. Существует восемь семейств кодонов. Все четыре кодона в каждом из этих семейств имеют одинаковые первые два основания и кодируют одну и ту же аминокислоту. Следовательно, в этих случаях третье основание не играет роли в определении смысла кодона. Существует также семь кодоновых пар, значения которых не зависят от того, какой из пиримидинов присутствует в третьем положении. Кроме этого имеется еще пять пар кодонов, в которых любой из пуринов может находиться в третьем положении, не изменяя при этом смысла кодона (пара такого типа представлена двумя терминирующими кодонами).

Теперь, возможно, имеет смысл рассмотреть генетический код в другом аспекте. Известно только три случая, когда основание, находясь в третьем положении, придает уникальное значение кодону: AUG (кодирует метионин), UGG (кодирует триптофан) и UGA (является нонсенс-кодоном). Это означает, что если в третьем положении находится C или U, то кодон не имеет однозначного соответствия. В случае когда в третьем положении находится A, это также не приведет к тому, что кодон будет соответствовать единственной аминокислоте.

Поскольку антикодон комплементарен кодону, то первое основание в антикодоновой последовательности пишется согласно существующим правилам – в направлении от 5'-конца к 3'-концу. При написании кодоновой последовательности по этим правилам третье основание кодона взаимодействует с первым основанием антикодона. Поэтому комбинация

Кодон	5'ACG	3'
Антикодон	3'UGC	5'

обычно записывается как кодон ACG5' антикодон CGU. Для того чтобы быть комплементарной кодону, последовательность антикодона должна считываться в обратном направлении. Поэтому во избежание путаницы мы будем записывать все последовательности, как обычно, в направлении от 5' к 3', но при этом последовательность антикодона обозначается с помощью стрелки, направленной в обратную сторону, для напоминания о имеющейся взаимосвязи с кодоном. Так, рассмотренная выше пара кодон/антикодон будет соответственно записана как ACG и $\overleftarrow{\text{CGU}}$.

Существует ли для каждого кодона своя тРНК с комплементарным антикодомом? Или одна тРНК может взаимодействовать со всеми или по крайней мере с несколькими представителями кодонового семейства? Для ряда высокоочищенных тРНК, выделенных из различных источников, было показано, что они способны узнавать несколько различных кодонов. Это означает, что основание в первом положении антикодона должно обладать способностью образовывать пару с несколькими различными основаниями, находящимися в третьем положении соответствующих кодонов. Другими словами, взаимодействие оснований, находящихся в этом положении, не ограничено образованием канонических пар G—C и A—U.

Правила, согласно которым происходит взаимодействие кодон-антикодон, суммированы в гипотезе неоднозначного соответствия (гипотеза качаний, Wobble-гипотеза). Эта гипотеза постулирует, что образование пары кодон-антикодон в двух первых положениях кодона всегда происходит по каноническим правилам, но в третьем положении возможно «качание» (неоднозначное соответствие). Объясняется это тем, что конформация антикодовой петли тРНК допускает значительную подвижность первого основания антикодона. Некоторые пары оснований, предусмотренные гипотезой неоднозначного соответствия, приведены на рис. 7.10.

Эти правила в дополнение к обычному взаимодействию разрешают образование пар между G и U. Как следствие этого A, находясь в третьем положении, уже более не имеет однозначного соответствия (так как U, который узнает его, должно также узнавать и G). Аналогично C также не имеет однозначного соответствия (так как G, узнающее его, должно узнавать и U).

Возможные варианты спаривания перечислены ниже:

U в антикодоне узнает как A, так и G в кодоне,
C в антикодоне узнает только G в кодоне,
A в антикодоне узнает только U в кодоне,
G в антикодоне узнает C или U в кодоне.

Поэтому однозначно узнавать уникальные кодоны возможно только тогда, когда третьим основанием является G или U. К таким кодонам относятся UGG и AUG, хотя аналогичных кодонов в случае U не существует.

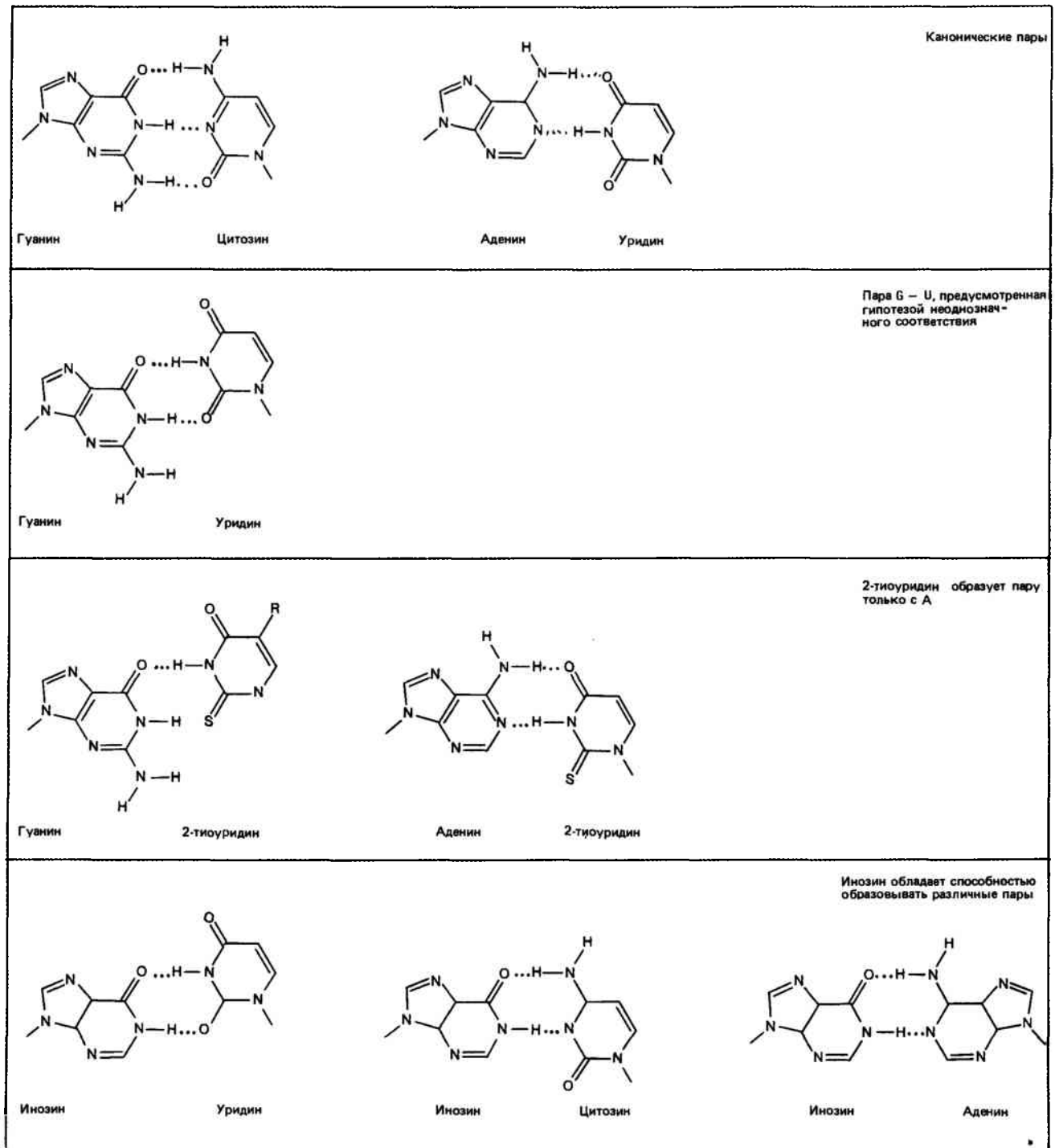


Рис. 7.10. Неоднозначное соответствие при образовании кодон-антикодоновой пары позволяет некоторым основаниям в первом положении антикодона узнавать более чем одно основание в третьем положении кодона.

Ряд неоднозначных взаимодействий, в которых участвуют обычные

и модифицированные основания, сравнивается с каноническим взаимодействием G–C, A–U. Основание G может взаимодействовать с U, но превращение в 2-тиоуридин допускает взаимодействие модифицированного U только с A, так как с G возможно образование только одной водородной связи. Инозин способен взаимодействовать с U, C и A.

Модификация оснований может контролировать узнавание кодона

В связи с тем что перечень оснований, входящих в состав тРНК, не ограничивается только четырьмя стандартными пуринами и пиримидинами (A, G и U, C), образующими пары согласно обычным правилам и гипотезе неоднозначного соответствия, возможно образование и

других, дополнительных пар оснований. В действительности некоторые обычные варианты пар не встречаются в силу того, что ряд оснований всегда модифицирован. Например, в первом положении антикодона не используются U и A.

Как правило, U в этом положении превращается в модифицированную форму с видоизмененной способностью образовывать пары. Среди бактериальных и эукариотиче-

ских тРНК имеется единственное исключение из этого правила. Эта необычная тРНК использует антикодон \overline{UAG} для узнавания всех шести лейциновых кодонов!

В первом положении антикодона А никогда не обнаруживается, так как обычно он превращается в инозин (I), способный при этом образовывать пары с тремя основаниями – U, C и A. Надо отметить, что это особенно важно для изолейциновых кодонов, когда AUA кодируют изолейцин, несмотря на то что AUG кодирует метионин. Когда А находится в третьем положении кодона, при помощи обычных оснований невозможно образовать однозначную пару, так как любая тРНК, у которой имеется в антикодоне U, будет узнавать AUG так же, как и AUA.

Вместе с тем кодон AUA должен узнаваться наряду с кодонами AAU и AUC. Эта проблема разрешилась благодаря существованию тРНК, содержащей I в первом положении антикодона.

Два других модифицированных нуклеозида, способных образовывать пары с тремя основаниями, находящимися в третьем положении кодона, – это уридин-5-оксиуксусная кислота и 5-метоксиуридин. Они эффективно узнают А и G и менее эффективно U.

Другой пример неоднозначного соответствия обнаруживается в ряду **квеузина** и его производных. Это модифицированное основание G сохраняет способность узнавать как C, так и U, но предпочтительно взаимодействует с U.

Ограничение, не достигаемое при использовании обычных правил, возможно при наличии в антикодоне 2-тиоуридина. Эта модификация не нарушает способности основания образовывать пару с А, но не допускает его участия во взаимодействии с G, возможным по правилам гипотезы неоднозначного соответствия.

Рассмотрение кодон-антикодонных взаимодействий позволяет сделать вывод о том, что в принципе существует несколько способов для создания набора молекул тРНК, способных узнавать все 61 кодон. В любом данном организме ни один из этих способов не имеет преимуществ, хотя отсутствие определенного пути модификации может помешать использованию некоторых способов узнавания. Так, в различных клетках определенное семейство кодонов может быть прочитано молекулами тРНК, обладающими различными антикодонами. Часто встречаются молекулы тРНК, чьи способности к узнаванию перекрываются. Вследствие этого определенный кодон может быть прочитан более чем одной тРНК. При этом возможны различия в эффективности альтернативных узнаваний. И кроме того, в наборе адапторных молекул часто встречается несколько тРНК, способных отвечать на одинаковые кодоны.

Почти для всех тРНК предсказания гипотезы неоднозначного соответствия хорошо согласуются с имеющимися данными. Но существует ряд исключений, при которых тРНК образует пары, не предусмотренные правилами гипотезы, или же когда невозможно спаривание с одним из кодонов, хотя гипотезой это разрешено. Вероятно, эти эффекты обусловлены влиянием соседних оснований или (и) конформационными изменениями в антикодовой петле. Действительно, самой идее, заложенной в основу гипотезы неоднозначного соответствия, присуще представление о важности структуры антикодовой петли. Дальнейшее исследование этого вопроса подтвердило и то, что ближайшее окружение влияет на способность антикодона узнавать кодоны. Были получены спонтанные мутанты, несущие замены оснований в других, не антико-

довых областях молекулы, изменяющие взаимодействие тРНК с кодоном (см. ниже).

Другой пример взаимодействия, не предсказанный гипотезой, заключается в том, что бактериальная инициаторная fMet-тРНК способна узнавать как кодон AUG, так и кодон GUG. Это вынуждает третье основание антикодона взаимодействовать не по правилам.

Митохондрии содержат минимальный набор тРНК

Согласно гипотезе неоднозначного соответствия, для узнавания всех существующих 61 кодона необходимо минимум 31 тРНК (без учета инициаторных тРНК). Эту цифру получают, исходя из того, что по крайней мере, 2 тРНК необходимы для узнавания каждого кодового семейства и 1 тРНК для кодовой пары или единичного кодона. Но в некоторых митохондриях (выделенных из грибов и млекопитающих) наблюдается загадочная ситуация: они содержат не более чем 23–24 различных тРНК. Как же этот неполный набор тРНК обслуживает все кодоны?

Определяющим моментом при ответе на этот вопрос является упрощение кодон-антикодонного взаимодействия, в результате которого одна тРНК становится способной узнавать все четыре представителя кодового семейства. Это позволяет свести минимальное число тРНК, необходимых для узнавания внутренних кодонов, к 23. У тРНК, обслуживающих все восемь кодовых семейств, в первом положении антикодона находится немодифицированный U.

Оставшиеся кодоны распределяются по парам, в которых кодоны, заканчивающиеся пиримидинами, прочитываются основанием G, находящимся в антикодоне, а все кодоны, заканчивающиеся пуринами, прочитываются основанием U в антикодоне, согласно предсказаниям гипотезы неоднозначного соответствия. Сложность с уникальными кодонами UGG и AUG в митохондриях млекопитающих полностью устраняется в результате изменений в генетическом коде, обусловленных действующими там правилами считывания: кодон UAG прочитывается наряду с UGG как триптофан, тогда как AUA больше не кодирует изолейцина и узнается наряду с AUG как метионин (см. гл. 4). В результате все не вошедшие в семейства кодоны подразделяются на 14 пар по способности взаимодействовать с молекулами тРНК.

Поэтому 23 идентифицированных гена для тРНК кодируют соответственно инициаторную тРНК, 14 молекул тРНК, узнающих пары кодонов, и 8 молекул тРНК, узнающих семейства кодонов. В результате два обычных нонсенс-кодона UAG и UAA не узнаются тРНК, так же как и аргининовая пара AG_G^A , которая тоже может быть использована для терминации (дополнительное изменение в коде). Сходные результаты обнаружены при исследовании митохондрий грибов. Отличие состоит только в том, что AUA остается изолейциновым кодоном (для этого необходимо, чтобы G, находящийся в антикодоне и, по-видимому, модифицированный, мог взаимодействовать с U, C или A), а C в антикодоне обеспечивает специфическое узнавание AUG как метионинового кодона.

Теперь необходимо ответить на вопрос, как U, находящийся в антикодоне, используется для узнавания и кодовых семейств, и кодовых пар, оканчивающихся пуринами. По-видимому, эта трудность преодолевается за счет модификации U только в тех молекулах тРНК,

которые узнают кодоновые пары. В других молекулах тРНК немодифицированные U способны образовывать пары с U и C по правилам гипотезы «качаний» так же хорошо, как с A и G (рис. 7.10). Если такой дуализм тРНК является наследственной способностью, то это может объяснить, почему U всегда модифицирован в первом положении антикодонов у тРНК, выделяемых из бактерий и цитоплазмы эукариотических организмов (хотя мы и не знаем, почему при этом U никогда не используется для узнавания целых кодоновых семейств).

Мутантные тРНК способны прочитывать различные кодоны

Мутантные тРНК явились одним из наиболее мощных инструментов при исследовании кодон-антикодонного взаимодействия, а также при выяснении роли, которую играют различные участки молекулы тРНК в этом взаимодействии.

Клоны, содержащие мутантные тРНК, выделяют благодаря способности последних подавлять эффекты мутаций, возникающих в генах, кодирующих белки. Мы уже упоминали, что мутации, способные устранять эффекты, вызванные первичными мутациями, получили название **супрессорные** (см. гл. 4). В системах, в которых изучают супрессорные тРНК, первичная мутация, изменяющая кодон в мРНК, инактивирует белковый продукт. Вторичная супрессорная мутация в свою очередь изменяет антикодон тРНК таким образом, что он узнает мутантный кодон в мРНК вместо (или так же хорошо) своего исходного кодона. В результате функция белка восстанавливается. В зависимости от природы первоначальной мутации супрессоры получили названия **нонсенс-** или **миссенс-**супрессоры (nonsense- или missense-супрессоры).

Преждевременная терминация белкового синтеза, вызываемая нонсенс-мутацией, может быть супрессирована благодаря тому, что в тРНК возникают изменения, позволяющие ей узнать терминирующий кодон как смысловой. В нормальной клетке терминирующий кодон узнается только фактором терминации. Следовательно, мутация в гене тРНК, приводящая к узнаванию терминирующего кодона, придает новое свойство трансляционной системе. Как показано на рис. 7.11, благодаря этому восстанавливается способность включать аминокислоту в ответ на мутантный кодон. В результате синтезируется белок нужной длины. Если аминокислота, включенная в результате супрессии, отличается от аминокислоты, которая исходно присутствовала в белке дикого типа, то его активность может быть частично снижена.

Нонсенс-супрессоры подразделяются на три класса — по одному на каждый терминирующий кодон. В табл. 7.1 приведены свойства некоторых наиболее хорошо изученных супрессоров.

Относительно легко были охарактеризованы **амбер-супрессоры**. У *E. coli* по крайней мере 6 тРНК в результате мутаций начинают узнавать кодон UAG. У всех амбер-супрессорных тРНК антикодоном служит триплет CUA. В каждом случае такой триплет возникает в результате единичной замены основания в кодоне дикого типа. Как видно из сравнения последовательностей *supD*, *supE* и *supF*, мутация может затрагивать любое из трех оснований в антикодоне. Теперь вместо своего прежнего кодона(ов) каждая супрессорная тРНК узнает только кодон UAG. При этом в белок включаются серин, глутамин или тирозин, т.е. те аминокислоты, которые присоединяются к тРНК дикого типа.

Таблица 7.1

Нонсенс-супрессорные тРНК возникают в результате мутаций в антикодоне¹

Локус	тРНК	Дикий тип			Супрессор	
		узнаваемый кодон	антикодон	антикодон ²	узнаваемый кодон	
<i>supD(su1)</i>	Ser	UCG	CGA	CUA	UAG	
<i>supE(su2)</i>	Gln	CAG	CGU	CUA	UAG	
<i>supF(su3)</i>	Tyr	UA ^C _U	GUA	CUA	UAG	
<i>supC(su4)</i>	Tyr	UA ^C _U	GUA	UUA	UAG ^A	
<i>supG(su5)</i>	Lys	AA ^A _G	UUU	UUA	UAG ^A	
<i>supU(su7)</i>	Trp	UGG	CCA	UCA	UG ^A _G	

¹ Используемые обозначения: исходно локусы получили обозначения *su1*, *su2* и так далее, как указано в скобках. *su⁺* обозначает супрессорную мутантную форму гена, а *su⁻* — дикый тип. Теперь локусы переименованы как *sup* с соответствующей буквой от A до V. Все локусы, приведенные в этой таблице, являются структурными генами для молекул тРНК.

² Цветом выделены основания в антикодоне, которые мутировали.

Возникновение **охра-супрессора** также связано с мутацией в антикодоне. Наиболее хорошо изученными супрессорами такого типа являются *supC* и *supG*, включающие тирозин или лейцин в ответ на кодоны охра (UAA) и амбер (UAC). Такой тип супрессии полностью согласуется с предсказанием гипотезы неоднозначного соответствия о том, что кодон UAA не может узнаваться однозначно. Действительно, все охра-супрессоры узнают как охра-, так и амбер-кодона. Это было наглядно продемонстрировано при введении дополнительных мутаций в штамм, уже содержащий амбер-супрессоры *supE* или *supF*. Если при этом амбер-кодона мутировали от CUA к UUA, то молекулы тРНК превращаются в охра-супрессорные, взаимодействующие и с охра-, и с амбер-кодонами. Амбер-супрессоры узнают только амбер-кодона. (В первом положении антикодона мы пишем U, но надо помнить, что на самом деле он может быть модифицирован.)

Очень интересны супрессоры кодонов UGA. Одна их разновидность (*supU*) возникает в результате мутации в гене, кодирующем триптофановую тРНК. Антикодоном триптофановой тРНК дикого типа является триплет CCA, узнающий только кодон UGG. Супрессорная тРНК содержит мутантный антикодон UCA, который узнает как свой исходный кодон, так и терминирующий кодон UGA.

Другая тРНК, супрессирующая кодон UGA, обладает совершенно непредвиденным свойством. Она происходит от той же триптофановой тРНК, но только в результате единичной мутации, заменяющей A в 24-м положении на G. Это приводит к замене пары G—U в D-стебле на A—U, что повышает стабильность спирали. Последовательность антикодона остается той же самой, что и в тРНК дикого типа, CCA. Таким образом, мутация в D-стебле должна каким-то образом изменять конформацию в антикодоновой петле, позволяя CCA взаимодействовать с UGA. Это происходит в результате образования пары между C и A, что не предусмотрено гипотезой однозначного соответствия. Данная супрессорная тРНК сохраняет способность узнавать свой обычный кодон UGG.

На самом деле способность триптофановой тРНК узнавать терминирующий кодон UGA, возможно, не

является уникальной для супрессорной формы, так как UGA-мутанты всегда отчасти «лики» (leaky). Это означает, что даже в отсутствие мутантной супрессорной тРНК наблюдается некоторый уровень остаточной супрессии. Следовательно, вероятней всего, что тРНК^{Tyr} дикого типа взаимодействует как с UGA, так и с UGG (хотя и хуже). Поэтому мутация в положении 24 скорее может усиливать эффект этой супрессии, чем создавать его заново. Таким образом, если кодирующая область заканчивается кодоном UGA, то полноценно функционировать в качестве терминатора он не будет. Поэтому некоторые рибосомы смогут продолжать трансляцию до следующего терминирующего кодона. В результате образуются два белка с одинаковой N-концевой областью, но один из них обладает более длинной C-концевой частью. Такая неполная супрессия нонсенс-кодонов, возможно, используется в качестве естественного механизма в процессе возникновения сходных белков.

Аналогичное кодон-антикодоновое взаимодействие обнаруживается в случае одной из эукариотических тРНК. Бычья печень содержит тРНК, аминоацелированную серином. Ее антикодоном является триплет ^mCCA, который, согласно гипотезе неоднозначного соответствия, может отвечать на триптофановый кодон UGG. Но фактически он узнает только терминирующий кодон UGA. Таким образом, ^mC специфически взаимодействует с А. Поэтому в рассмотренном случае использование триплета UGA в качестве терминирующего кодона, возможно, является определенным видом супрессии.

Главный вывод, который можно сделать на основе этих данных, состоит в том, что кодон-антикодоновое узнавание не может быть предсказано только на основе известных последовательностей триплетов. Это связано с тем, что на данное взаимодействие могут влиять другие особенности молекул — как в случае тРНК дикого, так и в случае мутантного типа.

Все нонсенс-супрессоры мутаций возникают в результате точковых мутаций в структурных генах, кодирующих тРНК. Эти изменения могут возникать как в антикодоне, так и в других местах молекулы. Однако в принципе возможно возникновение супрессорных тРНК и в результате нарушения процесса модификации тРНК, который оказывает влияние на специфичность узнавания. Это, вероятно, может происходить в результате мутаций в генах, кодирующих ферменты модификации.

Миссенс-мутации, изменяющие смысл кодона, приводят к замене одной аминокислоты на другую, не способную функционировать в белке вместо исходной. Формально любая замена аминокислоты в белке является миссенс-мутацией, но на практике мутации обнаруживаются только в том случае, если они приводят к образованию неактивного белка. Эти мутации супрессируются в результате включения или исходной, или какой-либо другой аминокислоты, не нарушающей функционирования белка. На рис. 7.12 показано, что это осуществляется таким же образом, как и супрессия нонсенс-кодонов. В результате мутации в антикодоне какой-либо тРНК, несущей подходящую аминокислоту, тРНК становится способной узнавать мутантный кодон. Таким образом, суть миссенс-супрессии заключается в изменении смысла кодона.

Миссенс-супрессоры были получены с помощью ряда тРНК, включающих глицин в ответ на один из кодонов GGN. В одном случае мутация в триптофан-синтазном гене *E. coli* приводила к замене глицинового кодона на

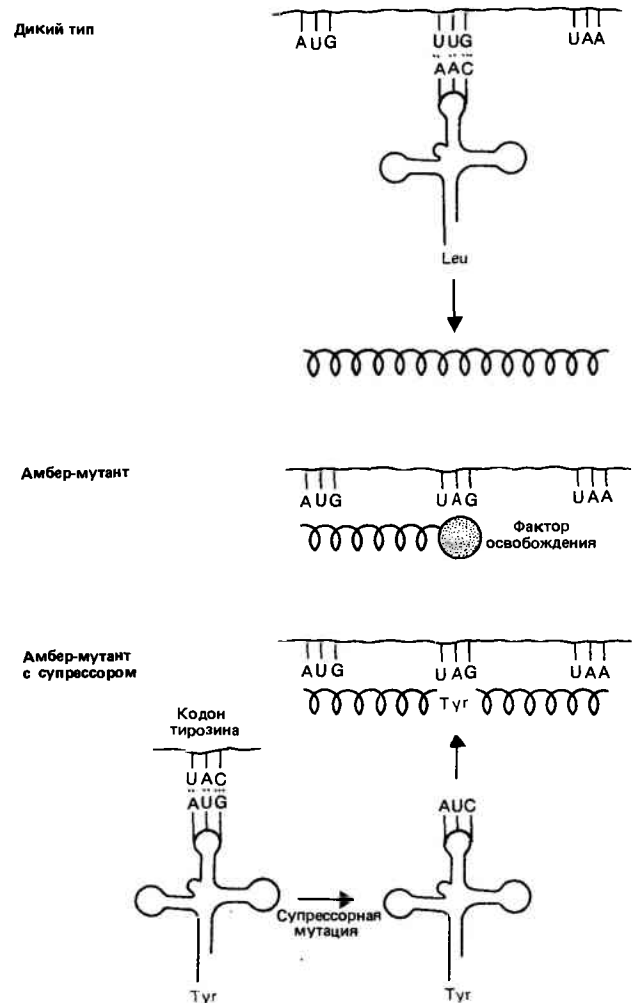


Рис. 7.11. Нонсенс-мутации могут быть супрессированы с помощью тРНК с мутантным антикодоном.

Ген дикого типа содержит кодон UUG, детерминирующий включение лейцина. Этот кодон узнается тРНК с антикодоном CAA. Мутация превращает исходный кодон в UAG, с которым взаимодействует фактор освобождения. Мутация в гене, кодирующем тирозиновую тРНК, превращает ее антикодон из GUA в CUA, позволяя тем самым взаимодействовать ей с амбер-кодоном. В результате синтезируется белок обычной длины, в котором исходный лейцин заменен на тирозин.

аргининовый. Супрессорные мутации, возникающие в тРНК^{Gly}₁ или в тРНК^{Gly}₂ (кодируемые соответственно локусами *supT* или *supA*) заменяют первоначальные антикодоны CCC или UCC (узнающие GGG и GG^A_G) на антикодон UCU, который теперь узнает аргининовые кодоны AG^A_G.

Другой путь возникновения миссенс-супрессии состоит в получении мутаций в генах, кодирующих тРНК (или синтетазу), вызывающих замену аминокислоты, которую акцентирует данная тРНК. Вследствие этого тРНК будет узнавать тот же кодон, что и ранее, но при этом включать другую аминокислоту, ответственную, возможно, за супрессию. Редко, но все же возможно получить такие мутации из *supF*-супрессоров, включающих тирозин в ответ на амбер-кодон. Ряд дополнительных мутаций в акцепторном стебле такой супрессорной тРНК позволяют ей включать глутамин в ответ на кодон UAG. При этом обычно сохраняется некоторая способность включать тирозин. Следовательно, эти мутации приводят к тому, что тРНК узнается Gln-тРНК-синтетазой, а также родственной Tyr-тРНК-синтетазой.

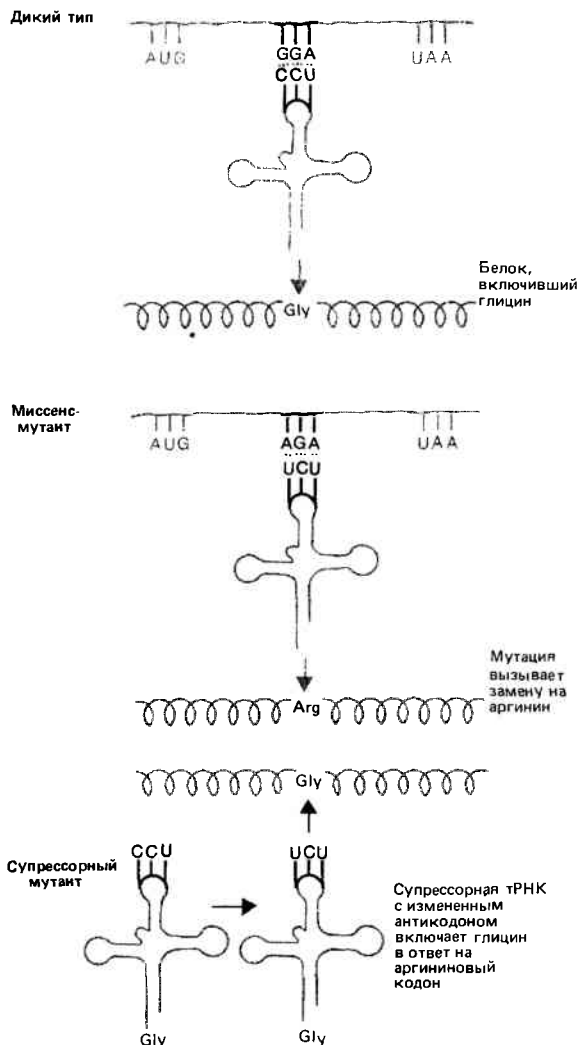


Рис. 7.12. Миссенс-супрессия состоит в изменении антикодона тРНК, которая в результате считывает «неправильный» кодон как смысловой.

В результате мутации GGA превращаются в AGA, что приводит к замене Glu в белке дикого типа на Arg в мутантном белке. Если антикодон тРНК^{Gly} мутирует от \overline{CC} в \overline{CU} , то он включает глицин в ответ на кодон AGA, кодирующий аргинин. Заметьте, что тРНК^{Arg} и супрессорная тРНК^{Gly} будут отвечать на кодон AGA; поэтому эффективность супрессии будет неполной.

Все рассмотренные случаи супрессии были исследованы на примере *E. coli*. У других бактерий (преимущественно у *S. typhimurium*) также были выделены похожие мутанты, и это свидетельствует о сходстве ситуаций во всех изученных случаях. Значительно меньше известно о распространенности и возможности супрессии нонсенс- и миссенс-мутаций у эукариот. Супрессоры охра- и амбер-мутаций, включающие тирозин, серин или лейцин, были выделены у дрожжей, причем каждый супрессор узнает только свой кодон. Возможно, это достигается благодаря использованию модифицированных оснований в антикодонах охра-супрессоров.

Конкуренция между супрессорными и обычными тРНК

Существуют серьезные различия между ситуациями, когда кодон узнается своей собственной аминокислотированной тРНК и когда мутация позволяет супрессорной тРНК узнать новый кодон. В клетке дикого типа опреде-

ленному кодону может быть приписано только одно значение, которое соответствует или конкретной аминокислоте, или сигналу терминации. Но в клетках, содержащих супрессорную мутацию, у мутантного кодона имеется две возможности: либо быть узнаваемым супрессорной тРНК, либо быть прочитанным согласно своему обычному значению.

Следовательно, нонсенс-супрессорные тРНК должны конкурировать с освобождающими факторами за терминирующие кодоны. В свою очередь миссенс-супрессорные тРНК должны конкурировать с тРНК, которые обычно узнают данный кодон. Степень конкуренции будет влиять на эффективность супрессии, приводя к зависимости этого процесса не только от степени сродства между антикодомом и комплементарным кодоном, но и от концентрации молекул-супрессоров в клетке, а также от подобных параметров, влияющих на протекание конкурирующих реакций терминации или включения. Такая зависимость может объяснить, почему эффективность супрессии никогда не превышает 50%. При этом эффективность супрессоров различного типа неодинакова. Из этого, очевидно, можно сделать вывод, что эффективность супрессии обусловлена характерными особенностями каждого класса супрессоров.

При выделении нонсенс-супрессоров использовали их способность узнавать мутантные нонсенс-кодоны, которые в силу особенностей своего расположения в гене оказывали летальное действие. Между тем один из нормальных терминирующих кодонов имеет такую же последовательность оснований как и супрессируемый нонсенс-кодон. Следовательно, мутантная тРНК, супрессирующая нонсенс-мутацию, в принципе должна быть способна супрессировать и нормальные терминирующие кодоны в конце генов. Тогда в результате **сквозного прочитывания** текста произойдет образование более длинного белка с дополнительным С-концевым пептидом. Таким образом, вероятнее всего, что эффективная супрессия терминации окажется летальной для клетки, так как при этом появятся более длинные белки с измененной функцией.

Аналогичным образом миссенс-супрессор одного гена скорее всего будет мутатором для другого. Например, если супрессор исправляет эффект мутации, заменяя одну аминокислоту на другую в мутантном сайте, то он будет при этом вводить новую аминокислоту взамен обычной в другом сайте. Поэтому перед клеткой возникает дилемма: с одной стороны, она должна супрессировать летальный кодон, с другой — не допустить слишком сильных изменений нормальных значений кодона в других местах.

Эффективность супрессии амбер-мутаций весьма высока и составляет в зависимости от конкретной системы от 10 до 50%. Допустим, что клетка не может быть толерантной к такому высокому уровню супрессии природных терминаторов, тогда очевидно, что амбер (amber) кодоны используются в конце генов с меньшей частотой, чем другие нонсенс-триплеты.

Выделить супрессоры охра-мутаций всегда трудно, причем для них характерна более низкая эффективность — обычно ниже 10%. Скорость роста клеток, содержащих охра-супрессоры, снижена. Из этого следует, что одновременная супрессия кодонов UAA и UAG неблагоприятна для клетки, возможно, в силу того, что охра-кодоны наиболее часто используются в качестве природных сигналов терминации. Отсутствие сколько-нибудь

сильных миссенс-супрессоров объясняется аналогичным образом: эффективная замена в одном гене оказывает вредное действие на другие гены.

На эффективность считывания любого кодона может оказывать влияние его локализация в гене, так что эффективность супрессирования нонсенс-кодона молекулами определенной тРНК изменяется в широких пределах в зависимости от соседствующих оснований. В действительности мы не знаем механизмов, при помощи которых **окружение** оказывает влияние на частоту узнавания кодона определенной тРНК. Но в зависимости от локализации эта частота может изменяться примерно на порядок. Таким образом, несмотря на то что в целом амбер-кодоны прочитываются более эффективно, чем охра-кодоны, коэффициент их полезного использования может варьировать в зависимости от локализации. Вероятно, основание, примыкающее к 3'-концу кодона, оказывает особенно сильное влияние на эти различия.

Мутация, вызывающая образование супрессорной тРНК, может иметь два последствия. Во-первых, она дает возможность тРНК узнать новый кодон. Во-вторых, в определенных случаях она не позволяет тРНК взаимодействовать со своими старыми кодонами. Существенно, что все высокоэффективные амбер-супрессоры возникают в результате мутации в гене, кодирующем одну из копий тРНК, образующихся в клетке в избытке. Другими словами, клетка располагает несколькими видами тРНК, способными реагировать с тем исходным кодоном, с которым реагирует тРНК дикого типа. Следовательно, супрессорная мутация не упраздняет узнавания старых кодонов, которые продолжают по-прежнему обслуживаться избыточными копиями тРНК.

В противоположность этому мутация в гене, кодирующем другую амбер-супрессорную тРНК, является **рецессивно-летальной**. Это означает, что тРНК может действовать как супрессор лишь в тех случаях, когда клетка содержит две копии гена: один — дикого типа, а другой мутантный. Но если имеется только единственная мутантная копия, то клетка гибнет. Причина состоит в том, что ген-супрессор (*supU*) является единственным структурным геном для триптофановой тРНК у *E. coli*. Мутация приводит к замене антикодона \overline{CCA} на \overline{CUA} . Но в результате клетка лишается единственной тРНК, способной распознавать кодон UGG, что приводит к летальному исходу (как было описано ранее, в другом случае ген этой тРНК *может* мутировать так, что она становится способной распознавать как UGA, так и UGG). Ограничениями, накладываемыми этими обстоятельствами, можно объяснить, почему не удается получить все возможные нонсенс-супрессоры — производные тРНК, узнающие такие кодоны, которые отличаются только одним основанием от терминирующего кодона

Транспортная РНК может изменить рамку считывания

Как правило, у мРНК имеется постоянная рамка считывания. Трансляция начинается с кодона AUG и последовательно продолжается до терминирующего кодона. При этом процесс происходит таким образом, что считывается все подряд. Поэтому вставка или делеция основания вызывают мутацию со **сдвигом рамки** считывания. В результате после сайта, в котором произошла мутация, трансляция происходит уже в другой рамке считывания. Рибосомы и тРНК, продолжая узнавать три-

плет за триплетом, синтезируют при этом совершенно другую последовательность аминокислот.

В результате мутаций, **супрессирующих мутации со сдвигом рамки**, восстанавливается исходная рамка считывания. В гл. 4 мы уже обсуждали, что для этого необходимо компенсировать делеции и вставки, возникшие в гене. Но оказалось, что наряду с этим можно получить ряд *внегенных* супрессоров, которые не действуют на терминирующие кодоны и не вызывают замены одной аминокислоты на другую.

У бактерий описано несколько таких супрессирующих мутаций. В одном случае супрессор исправляет рамку считывания, когда мутация вызвана вставкой дополнительного основания в участок, состоящий из нескольких идентичных остатков, — например, вставка в участок из нескольких оснований G. Супрессором такой мутации служит тРНК^{Gly}, содержащая дополнительное основание в антикодонной петле. В результате антикодон превращается из обычной триплетной последовательности \overline{CCC} в квадриплетную последовательность \overline{CCCC} . Поэтому в некоторых случаях такая тРНК узнает квадриплетный кодон и с вариациями в последнем положении, согласно правилу неоднозначного соответствия.

Сходный тип супрессии был обнаружен в митохондриях дрожжей. При этом супрессируются мутации, вызванные вставками или делециями одного остатка T на участке из пяти T.

Все мутации в этом случае были лики (*leaky*). Возможный механизм супрессии состоит в том, что тРНК^{Phe}, узнающая кодон UUU, иногда допускает проскальзывание рибосомой основания в том или ином направлении, тем самым супрессируя мутацию. Такое проскальзывание можно рассматривать как нормальное, хотя и не слишком частое явление (менее чем в 5% случаев).

Другой пример эпизодически происходящего сдвига рамки обнаруживается при нормальном развитии РНК-содержащего бактериофага MS2. Механизм, обуславливающий сдвиг рамки считывания, не известен, но о его наличии свидетельствует тот факт, что в новой рамке считывания рибосомы производят преждевременную терминацию.

Это позволяет начать трансляцию с другого гена — гена лизиса. Иницирующий кодон этого гена лежит впереди по ходу трансляции. Таким образом, терминация трансляции, связанная со сдвигом рамки считывания, необходима для последующего взаимодействия рибосомы с иницирующим кодоном гена, ответственного за лизис. Поэтому редко происходящие ошибки считывания на самом деле могут являться важным компонентом естественного процесса трансляции.

Некоторые тРНК, супрессирующие мутации со сдвигом рамки, могут взаимодействовать более чем с одним квадриплетным «кодоном». Например, бактериальная супрессорная тРНК^{Lys} способна узнавать кодоны AAAA и AAUU вместо обычного кодона AAA. Другой супрессор может прочитывать любой квадриплетный «кодон», содержащий в трех первых положениях ACC (природа следующего основания в данном случае не влияет на процесс узнавания). Альтернативные основания, находящиеся в четвертом положении удлиненного «кодона», не подчиняются обычным правилам гипотезы «качаний». Вероятный механизм этой супрессии состоит в том, что в действительности тРНК узнает триплетный кодон, но по какой-то причине — возможно, из-за стерических помех — соседнее основание заблокировано. Это заставляет

одно основание переместиться, прежде чем следующая тРНК сможет провзаимодействовать с кодоном.

Сходное явление, способное пролить свет на механизм только что описанной супрессии, наблюдается в обычной системе трансляции *in vitro*. При определенных условиях в этой системе можно вызвать сдвиг фазы считывания. Такими условиями являются или большой избыток, или нехватка одной из аминоксил-тРНК. Возможно, длительная пауза во время узнавания молекулой тРНК триплетного кодона, находящегося в участке А, позволяет другой тРНК, комплементарной перекрывающемуся (находящемуся в другой фазе) кодону, провзаимодействовать с ним.

Эти данные свидетельствуют о существовании универсального механизма, с помощью которого рибосома контролирует свое перемещение. Передвигаясь, рибосома сдвигает тРНК, находящуюся в участке А, таким образом, что теперь она занимает участок Р. Когда тРНК связывается с «кодоном», состоящим из 4 оснований, или когда она в силу стерических изменений растягивается за пределы триплетного кодона, следующая аминоксил-тРНК взаимодействует с триплетом, находящимся в участке А, но уже в другой фазе считывания. В процессе транслокации рибосома перемещает эту тРНК в участок Р, так что следующая аминоксил-тРНК в участке А взаимодействует с триплетным кодоном уже обычным образом. Следовательно, геометрия взаимодействия тРНК с мРНК используется для измерения расстояния, на которое рибосома должна передвинуться.

Рекомендуемая литература

Литература по тРНК хорошо подобрана в ряде обзорных книг. В сборнике под редакцией Алтмана (Altman (Ed.), *Transfer RNA* (MIT Press, Cambridge, 1978).

Первичная и вторичная структуры рассмотрены в обзоре Кларка (Clark, стр. 14–47) и третичная структура – в обзоре Кима (Kim, стр. 248–293). Модифицированные основания рассматриваются Нишимура (Nishimura, стр. 168–195). Одна из точек зрения на аминоксил-тРНК – синтетазу приведена в обзоре Иглуа и Крамера (Igloi, Cramer, стр. 294–340), другую можно найти у Шиммела и Солла (Schimmel, Soll, *Ann. Rev. Biochem.*, **48**, 601–648, 1979). Обширные представления о тРНК дает двухтомное издание, включающее обзоры и экспериментальные статьи. Первый том – под редакцией Шиммела, Солла и Абелсона (Schimmel, Soll, Abelson, *Transfer RNA: Structure, Properties and Recognition*); второй том – под редакцией Солла, Абелсона и Шиммела (Soll, Abelson, Schimmel, *Transfer RNA: Biological Aspects*, Cold Spring Harbor Laboratory, New York, 1979). Кодирование свойства тРНК и их мутационные изменения рассматриваются у Льюина (Lewin, *Gene Expression, I, Bacterial Genomes* (гл. 5), Wiley, New York, 1974). Супрессия рассматривается у Корнера, Файнштейна и Алтмана (Korner, Feinstein, Altman, стр. 105–135 в сборнике под редакцией Алтмана (Altman), а последние результаты – в этой области в недавно появившемся коротком сообщении Рота (Roth, Cell, **24**, 601–602, 1981).

Глава 8

РИБОСОМЫ КАК ФАБРИКИ БЕЛКОВОГО СИНТЕЗА

Первоначально рибосомы выделяли в виде нескольких различных структур; их описывали по-разному – как *клеточные органеллы*, как *микросомные частицы*, как *рибонуклеопротеины* или просто как *белоксинтезирующие машины*. На деле же это макромолекулярные комплексы с определенной структурой, которые совместно с некоторыми вспомогательными факторами обладают определенными ферментативными активностями, необходимыми для различных этапов белкового синтеза. Вкратце рибосому можно представить как маленькую передвижную фабрику, которая, передвигаясь вдоль матрицы, осуществляет синтез пептидных связей. При этом с необычайно большой скоростью в рибосому входят (и выходят из нее) молекулы тРНК, несущие аминокислоты; кроме того, с рибосомой циклически связываются (и отделяются от нее) факторы элонгации.

Рибосомы представляют собой весьма важный клеточный компонент. В быстрорастущих бактериях содержится около 20 000 рибосом в расчете на один геном (см. табл. 5.1). В состав рибосом входит 10% суммарного клеточного белка, а рРНК составляет приблизительно 80% всей клеточной РНК. В эукариотических клетках на долю рибосом приходится относительно меньшая часть клеточного белка, чем в прокариотических клетках, но абсолютное число рибосом у эукариот больше. Что касается РНК, то у эукариот, как и у прокариот, на долю рибосом

приходится большая часть клеточной РНК. Количество рибосом в прокариотах и эукариотах прямо связано с белоксинтезирующей активностью их клеток.

У бактерий рибосомы прикреплены к молекулам мРНК, которые сами все еще связаны с ДНК-матрицей. В цитоплазме эукариотических клеток рибосомы, как правило, ассоциированы с цитоскелетом – фибриллярным матриксом. В ряде эукариотических клеток некоторая часть рибосом ассоциирована с мембранами эндоплазматического ретикулума и может быть выделена в виде микросомной фракции. Всем рибосомам присуще одно общее свойство: будучи заняты в синтезе белков, они находятся в клетке не в свободном состоянии, а обязательно бывают прикреплены прямо или опосредованно к клеточным структурам.

Все рибосомы при диссоциации распадаются на две неравные субчастицы; большая из этих частиц по размеру в два раза превышает меньшую. Каждая субчастица содержит высокомолекулярную рРНК и определенное количество различных белков; почти все из них представлены в рибосоме только одной копией. Большая субчастица помимо основной рРНК может содержать еще и молекулы малых РНК.

Несмотря на наличие в клетке нескольких копий гена, кодирующего рРНК, последовательность рРНК консервативна, что свидетельствует о ее важном значении. Из

этого следует, что давление отбора препятствует накоплению мутационных изменений. Действительно, мутации, затрагивающие рРНК, могут влиять на активность рибосом.

Роль высокомолекулярных рРНК в функционировании рибосом до конца не выяснена. Одна из ее непосредственных функций, возможно, заключается в связывании мРНК и тРНК (путем образования комплементарных пар). Но очевидно, что в основном рРНК играет структурную роль. Белки связываются с рРНК в особых участках и в определенной последовательности, необходимой для сборки субчастиц. В действительности интерес к рибосомам объясняется не только их способностью транслировать мРНК, но также присущей им способностью к самосборке из составляющих их РНК и белковых молекул.

Рибосомные белки совместно с вспомогательными факторами обеспечивают ферментативные активности, необходимые для биосинтеза белка. При этом одни функции обусловлены индивидуальными белками, для проявления других активностей необходимо наличие нескольких белков. В рибосомах имеется несколько функциональных центров, каждый из которых построен из определенной группы белков. В этом смысле рибосомы представляют собой набор определенных ферментов, активных лишь в том случае, если они находятся в составе частиц. В результате их совместного координированного действия осуществляется акт трансляции.

Только у небольшой части рибосомных белков идентифицированы определенные ферментативные или структурные функции, необходимые для синтеза белка. В то же время у бактерий получены мутации, затрагивающие большинство генов, кодирующих рибосомные белки. Все это дало возможность идентифицировать соответствующие гены; кроме того, было показано, что функции изучаемых белков являются незаменимыми. Вероятно, многие из этих белков (а также рРНК) необходимы для формирования полной структуры рибосомы, обуславливающей правильное взаиморасположение различных активных участков и передвижение рибосом. Отнюдь не обязательно, чтобы они непосредственно участвовали в реакциях синтеза.

Рибосомы – компактные рибонуклеопротеиновые частицы

Рибосомы, выделенные из бактерий или из цитоплазмы и органелл эукариотических клеток, различаются как по размеру, так и по относительному содержанию РНК и белка. Следовательно, рибосомы из различных источников могут иметь разную форму и по-разному быть сконструированы. В табл. 8.1 перечислены компоненты бактериальных рибосом, выделенных из *E. coli*.

У этих рибосом определена структура всех компонентов. Известны нуклеотидные последовательности молекул рРНК, а также аминокислотные последовательности большинства белков. Молекулярные массы двух частиц рибосомы равны соответственно $0,93 \cdot 10^6$ и $1,59 \cdot 10^6$. Доля РНК составляет 60% в малой субчастице и 70% – в большой субчастице. Из 52 рибосомных белков 21 входит в состав малой субчастицы (обозначаются S1–S21), а 31 – в состав большой субчастицы (обозначаются L1–L34). (Нумерация доходит до 34, поскольку ранее была допущена ошибка.)

Рибосомные белки не имеют никакого сходства ме-

Таблица 8.1
Состав рибосом *E. coli*

	Рибосома	Малая суб- частица	Большая суб- частица
Коэффициент седиментации	70 S	30 S	50 S
Масса, дальтон	2 520 000	930 000	1 590 000
РНК			
Большая		16S = 1541 оснований	23S = 2904 оснований 5S = 120 оснований
Малая			
Масса РНК		560 000	1 104 000
Доля РНК	66%	60%	70%
Белки		21 полипептид	31 полипептид
Масса белков		370 000	487 000
Доля белков	34%	40%	30%

жду собой. Они различаются как по аминокислотным последовательностям, так и по иммунологическим свойствам. Из этого правила есть два исключения. Белки S20 и L26 идентичны. Более чем 80% этого белка входит в состав малой субчастицы, но некоторая его часть связана с большой субчастицей. Возможно, это отражает тот факт, что данный белок расположен на стыке двух субчастиц. Белок L7 отличается от белка L12 только наличием ацетильной группы на N-конце (некоторые другие рибосомные белки также ацетилированы). Эти белки, получившие название «L7/L12», в растворе образуют димеры. В каждой рибосоме содержится два таких димера. Белок S6 присутствует в двух копиях. Все другие белки, в том числе S20/S26, представлены одной молекулой на рибосому. Из этого следует, что все бактериальные рибосомы имеют одинаковый состав; они представляют собой комплекс из молекул рибосомных РНК и специфического набора уникальных белков.

Цитоплазматические рибосомы эукариот крупнее бактериальных; молекулярная масса их субчастиц составляет $1,4 \cdot 10^6$ и $28 \cdot 10^6$. В табл. 8.2 приведены параметры для 80S-рибосом млекопитающих. Абсолютное содержание

Таблица 8.2
Состав рибосом, полученных из цитоплазмы печени крыс

	Рибосома	Малая суб- частица	Большая суб- частица
Коэффициент седиментации	80 S	40 S	60 S
Масса, дальтон	4 420 000	1 400 000	2 820 000
РНК			
Большая		18S = 1900 оснований	28S = 4700 оснований 5,8S = 160 оснований
Малая			
Масса РНК		700 000	5S = 120 оснований 1 820 000
Доля РНК	60%	50%	65%
Белки		33 полипептида	49 полипептидов
Масса белков		700 000	1 000 000
Доля белков	40%	50%	35%

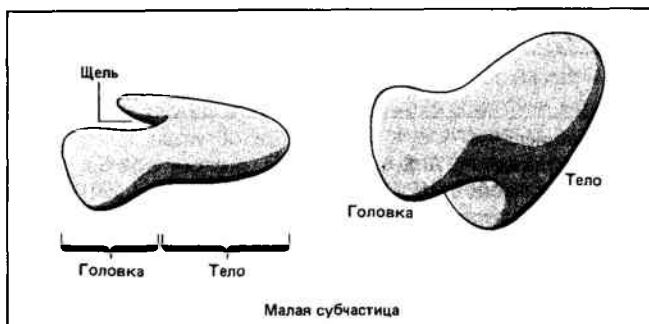


Рис. 8.1. 30S-субчастица обладает вытянутой и асимметричной формой с уплощенной областью и с щелью, отделяющей головку от тела.

Субчастица показана в двух взаимно перпендикулярных проекциях.

РНК и белка в них больше, чем в бактериальных рибосомах. Это объясняется следующими причинами: молекулы РНК длиннее, и больше белковых молекул (~82) входит в состав рибосом. Доля РНК в 80S-рибосомах меньше, чем в 70S-рибосомах и составляет 50 и 65% от масс соответственно малой и большой субчастиц (как говорилось в гл. 6, намного сложнее, чем у бактерий, организованы и вспомогательные факторы, принимающие участие в процессах инициации и элонгации). Возможно, большинство, а может быть, и все белки входят в состав эукариотических рибосом в стехеометрических соотношениях, однако точно это не доказано. Сходства между рибосомными белками бактерий и эукариот выявить не удалось, но за одним исключением: L7/L12 могут заменить аналогичные эукариотические белки в составе 80S-рибосом.

Размеры рибосом, выделенных из различных оргanelл, варьируют. Митохондриальные рибосомы низших эукариот, таких, как грибы, имеют несколько большие размеры, чем рибосомы *E. coli*. В митохондриях растений рибосомы лишь ненамного меньше, чем в окружающей цитоплазме. Однако в митохондриях млекопитающих и амфибий их размеры еще меньше и составляют 60S. Эти рибосомы характеризуются более низким содержанием РНК (25–31%). Рибосомы хлоропластов приблизительно такого же размера, как и рибосомы бактерий, хотя с более высоким содержанием РНК.

По своей форме бактериальные рибосомы и цитоплазматические рибосомы эукариот в общих чертах сходны

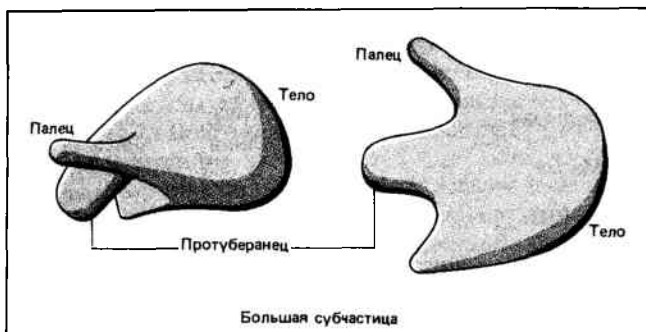


Рис. 8.2. 50S-субчастица имеет весьма компактное тело, из которого выпячиваются протуберанец и тело.

Субчастица показана в двух взаимно перпендикулярных проекциях.

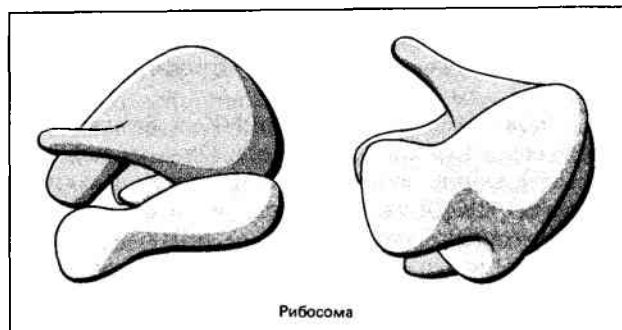


Рис. 8.3. Образование 70S-рибосом, возможно, происходит путем ассоциации между несколькими отдельными участками двух субчастиц.

Субчастица показана в двух взаимно перпендикулярных проекциях.

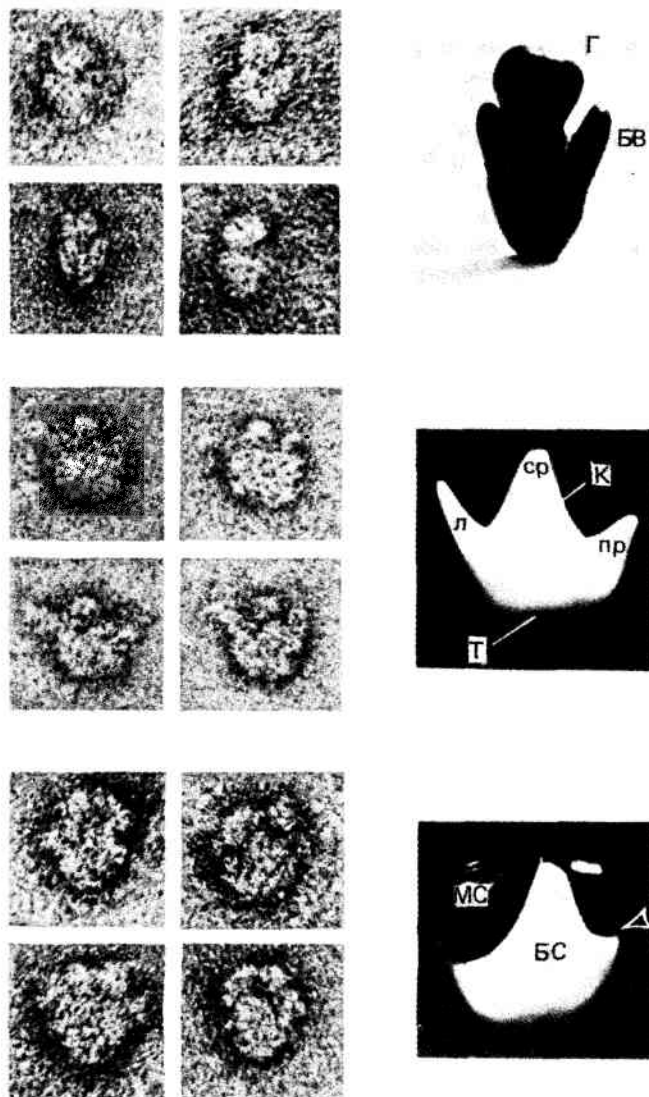


Рис. 8.4. Электронно-микроскопические фотографии рибосом и субчастиц, позволяющие представить их форму.

Для каждой исследованной структуры приведено по четыре фотографии, на которых они изображены в одинаковых ориентациях. МС – малая субчастица; БС – большая субчастица; Т – тело большой субчастицы; К – «корона» большой субчастицы, л, ср, пр – левая сторона, средняя и правая сторона; Г – головка малой субчастицы, БВ – боковой выступ. (Фотографии любезно предоставлены Miloslav Voublik.)

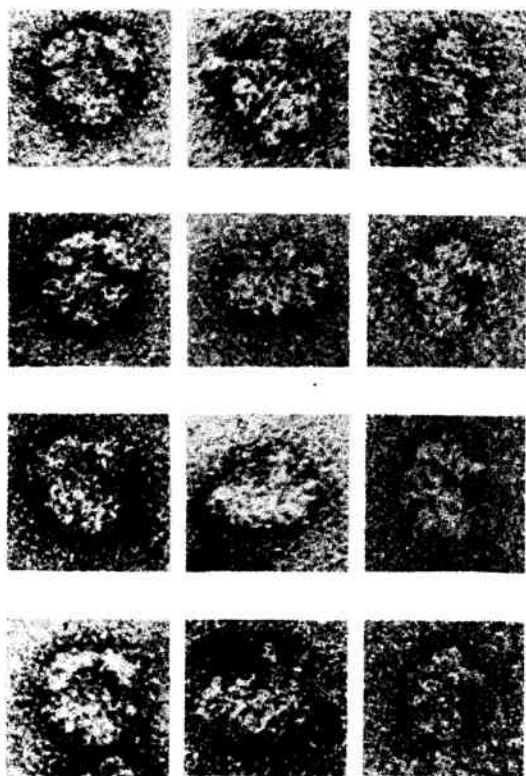


Рис. 8.5. Электронные микрофотографии 80S-рибосом и их субчастиц. Видна консервативность структуры среди эволюционно различных видов и взаимосвязь с бактериальными рибосомами. В левой колонке показаны целые рибосомы, в средней – большие субчастицы (60S), в правой – маленькие субчастицы (40S). Исследованные препараты рибосом выделены из рачка *Artemia salina* (верхний ряд), миксомицета *Dictyostelium discoideum* (второй ряд), ретикулоцитов кролика (третий ряд) и клеток (человека) *HeLa* (четвертый ряд). (Фотографии любезно предоставлены Miloslav Boublik.)

(рибосомы из различных органелл эукариотической клетки в этом отношении не исследовались). На рис. 8.1 схематически изображена структура малой бактериальной субчастицы, которую можно описать как сплюснутый эллипсоид с несколько уплощенной формой и размерами $55 \times 220 \times 220 \text{ \AA}$. Как показано на рис. 8.2, большая субчастица имеет более сферическую форму размерами $150 \times 200 \times 200 \text{ \AA}$.

Полная 70S-рибосома не охарактеризована так подробно, как субчастицы; одна из моделей ее строения изображена на рис. 8.3. Согласно этой модели, две субчастицы соединены в результате взаимодействия между щелью, расположенной на 30S-субчастице, и протуберанцем на 50S-субчастице. Между субчастицами, возможно, образуется пространство в виде тоннеля.

Несколько электронно-микроскопических фотографий субчастиц и целых бактериальных рибосом приведено на рис. 8.4; одновременно приведены их модели, расположенные в соответствующей ориентации.

Цитоплазматические рибосомы млекопитающих имеют сходную структуру, но отличаются большими размерами. Некоторые электронные микрофотографии субчастиц и целых 80S-рибосом приведены на рис. 8.5. Детальной модели, описывающей структурную организацию 80S-рибосом млекопитающих, пока еще нет. Однако взаимное расположение субчастиц представляется похожим на то, которое было обнаружено в случае 70S-рибосом. Примечательно, что малые субчастицы бактериальных и эукариотических рибосом сходны. Наиболь-

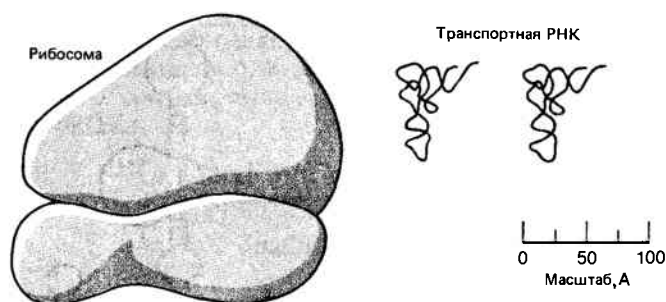


Рис. 8.6. Сравнительные размеры, показывающие, что рибосомы достаточно большие, чтобы связать две молекулы тРНК (так же как и 40 оснований мРНК). Масштаб приведен в ангстремах.

шие различия, по-видимому, наблюдаются в структуре большой субчастицы.

Очень много усилий было затрачено на выявление локализации мРНК. Согласно одной из распространенных гипотез, мРНК располагается между субъединицами или в непосредственной близости от места контакта субчастиц – возможно, в «тоннеле», проходящем между ними. В некоторых моделях мРНК изображена в виде «нитки», протодетой через рибосому, но настолько же вероятно, что она расположена на поверхности частицы. Как видно из рис. 8.6, две молекулы тРНК вполне соизмеримы с рибосомой. Возможно, что они входят в открывающиеся на поверхности частиц большие щели, соответствующие им по своей геометрии.

Структура рибосомной РНК

Молекулы рибосомной РНК транскрибируются в виде единого предшественника, содержащего последовательности для обеих рРНК. У эукариот этот предшественник достаточно стабилен, и его можно выделить. В случае бактерий это сделать значительно труднее. Необходимо создать специальные условия, при которых процессинг предшественника будет нарушен, поскольку у бактерий разделение молекул под действием нуклеазы происходит раньше, чем заканчивается транскрипция (гл. 25). В результате образуются молекулы, которые несколько длиннее, чем зрелые рРНК. Например, первый тестируемый предшественник рРНК, обозначаемый как р16, на 94 нуклеотида длиннее, чем 16S-рРНК, входящий в состав малой субчастицы. Молекула р16 образует комплекс с рибосомными белками, а ее избыточная последовательность удаляется позднее.

Рибосомная РНК метилируется. В клетках млекопитающих 18S- и 28S-рРНК содержат соответственно 43 и 74 метильные группы. Таким образом, около 2% ее нуклеотидов метилировано. Приблизительно 80% метильных групп находится при С-2' рибозы, остальные приходятся на модифицированные основания (преимущественно на аденин). В бактериальной 16S-рРНК обнаруживается 10 метилированных нуклеотидов и, вероятно, около 20 – в 23S-рРНК, что составляет 0,7% от их общего количества. В 16S-рРНК метильные группы встречаются главным образом в нуклеотидах, расположенных вблизи 3'-конца молекулы. Среди этих нуклеотидов обнаруживаются диметиладенин, метиладенин и метилгуанин. Существует модификация другого типа – превращение некоторых уридинов в псевдоуридин.

При сравнении последовательностей рРНК, выделенных из различных бактерий, прослеживается определенная степень консервативности; кроме того, обнаружи-

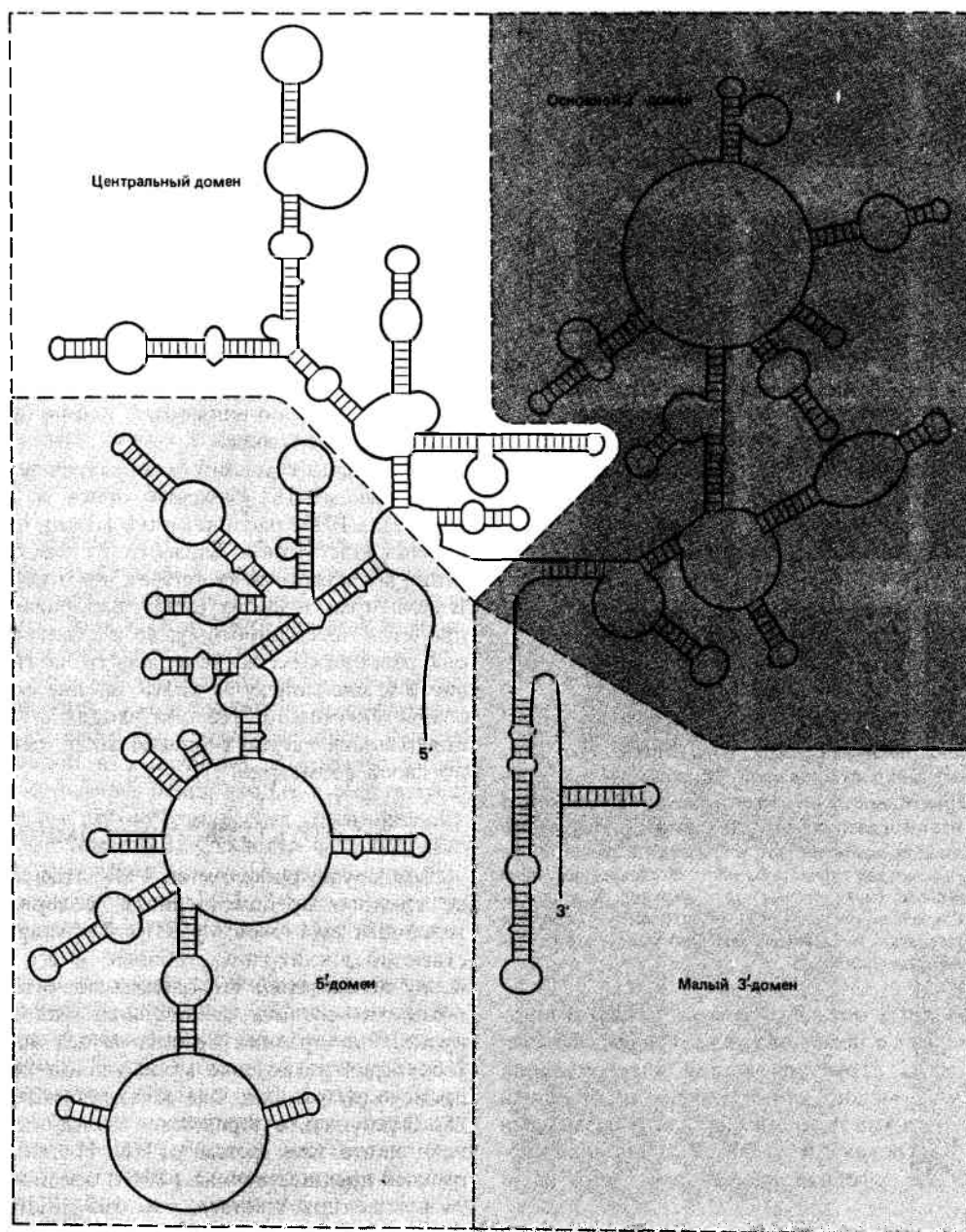


Рис. 8.7. Во вторичной структуре 16S-рРНК *E. coli* содержится много коротких двухцепочечных участков, образующих четыре основных домена.

вадается некоторая гомология с рРНК органелл эукариотической клетки. Каково значение этого родства, пока еще не ясно. Среди эукариотических рРНК не прослеживается значительной консервативности, за исключением участков, расположенных в непосредственной близости от метилированных оснований. Это характерно даже для таких отдаленных видов, как дрожжи и млекопитающие.

Зная последовательность рРНК, можно предсказать образование двухцепочечных участков. Однако в молекулах с большими двухцепочечными участками возможно образование нескольких различных конформаций. Заранее нельзя предсказать, какая из них окажется более предпочтительной, если даже допустить, что в бактериальной клетке молекула совершенно свободна и находится как бы в растворе. Вероятные модели вторичной структуры рРНК в принципе могут быть различны в свя-

зи с различной относительной стабильностью определенных двухцепочечных участков, но получение подобных сведений во многом ограничено.

Наиболее информативный подход к выяснению вторичной структуры — это сравнение последовательностей рРНК, выделенных из родственных организмов (последовательности рРНК определяются непосредственно по результатам анализа последовательности генов, кодирующих соответствующие РНК). Участки, имеющие важное значение для поддержания определенной вторичной структуры, консервативны, так как благодаря своему значению соответствующие пары должны образовываться в каждой рРНК. На основе такого подхода была предложена модель для 16S- и 23S-рРНК.

Наиболее распространенная модель для 16S-РНК *E. coli* схематично изображена на рис. 8.7. Около половины

последовательности оснований образует двухцепочечные структуры. Двойные спирали, возникающие во вторичной структуре РНК, преимущественно короткие (менее восьми пар оснований). Часто двойной участок несовершенен и содержит петли из неспаренных оснований. В целом молекула образует четыре основных домена.

Соответствует ли структура свободной рРНК той, которая на самом деле существует в составе рибосомной субчастицы? Вероятно, что взаимодействие с рибосомными белками может сказаться на способности образовывать двухцепочечные участки. Для исследования структуры РНК, находящейся в составе субчастиц, были использованы различные методы. Например, определяли основания, способные реагировать с кетоксалем (вещество, взаимодействующее с неспаренным гуанином). Выявляли области рРНК, образующие под действием псоралена поперечные сшивки. Локализовали участки, предпочтительно расщепляемые нуклеазами. Наконец, определяли, какие участки рРНК и с какими белками можно ковалентно сшить. По мере проведения этих исследований накапливались результаты, неплохо согласовавшиеся с моделью, предложенной на основе изучения свободной рРНК в растворе. И тем не менее рРНК, связанная в субчастице, по-видимому, все же отличается от свободной рРНК. Возможное различие, вероятно, состоит в том, что у рРНК, находящейся в составе субчастицы, меньше двойных спиралей, чем у свободной формы. В структуре 16S-РНК в зависимости от того, находится ли она в составе 30S-субчастицы или 70S-рибосомы, обнаруживаются некоторые различия. Из этого, очевидно, следует, что объединение субчастиц может вызывать конформационные изменения в рРНК.

Помимо двух основных рРНК, образующих, как это принято говорить, скелет субчастиц, большая субчастица содержит еще молекулу 5S-РНК. (Это относится ко всем рибосомам, за исключением митохондриальных.) Молекулы 5S-РНК прокариотических организмов обнаруживают некоторую консервативность нуклеотидной последовательности, особенно в тех участках, которые взаимодействуют с рибосомными белками. Обнаруживается консервативность и среди 5S-РНК эукариот. Например, у млекопитающих преимущественно встречается одна и та же последовательность. Однако между 5S-РНК прокариот и эукариот не обнаруживается ничего общего при сравнении их первичных последовательностей.

Во всех молекулах 5S-РНК имеется много комплементарных последовательностей, потенциально способных к образованию двухцепочечных участков. Тем не менее при выборе единственно правильной модели из всех возможных (а их предложено более 20) возникает ряд трудностей. Объясняется это существованием значительной степени неопределенности при выборе возможных вариантов комплементарного спаривания даже в совсем небольших молекулах — если не имеется дополнительного источника информации (как, например, общее соответствие структуре клеверного листа, обнаруживаемое у молекул тРНК). На рис. 8.8 изображена наиболее распространенная модель, образующая структуру с четырьмя спаренными участками, которая удовлетворяет всем известным прокариотическим последовательностям для 5S-РНК. В аналогичной модели для эукариотической 5S-РНК имеется пять двухцепочечных участков. Так же как и в случае тРНК, исходя из вторичной структуры, нельзя точно представить пространственную форму молекулы.

В составе большой субчастицы эукариотических рибо-

сом обнаруживается другая короткая РНК с коэффициентом седиментации 5,8S. Ее последовательность, по-видимому, соответствует 5'-концу прокариотической 23S-рРНК.

Каждый рибосомный белок характеризуется специфической локализацией

Для анализа строения рибосомных субчастиц используются два подхода: изучают сборку рибосомы (процесс, с которым мы познакомимся позже) и непосредственно исследуют ее структуру.

Биохимические исследования сосредоточены главным образом на идентификации местоположения индивидуальных рибосомных белков с помощью метода иммунной электронной микроскопии, на определении взаиморасположения соседствующих белков путем образования сшивок между ними и на анализе взаимодействий между определенными рибосомными белками и рРНК.

При использовании метода в иммунной электронной микроскопии получают антитела к индивидуальным рибосомным белкам. Этими антителами обрабатывают интактные субчастицы и затем с помощью электронного микроскопа определяют места связывания. Таким способом была определена локализация всех белков 30S-субчастицы, а также многих белков 50S-субчастицы.

Используя для образования ковалентных сшивок соединения с различной длиной молекулы, можно определить близость расположения двух данных белков. Суть метода состоит в следующем: нативные субчастицы обрабатывают реагентами, образующими ковалентные сшивки, с последующим анализом белков, вошедших в состав конгломератов. (Однако возможность этого метода ограничена анализом белков.) Ковалентные сшивки белков с рРНК позволяют определить места их связывания с нуклеиновой кислотой. Другие методы, с помощью которых были получены сравнимые результаты о местоположении белков, — это использование нейтронного рассеивания субчастиц, в которых определенные белки дейтерированы, и измерение энергии переноса между флуоресцентно мечеными парами рибосомных белков.

Применение различных методов в конечном счете дало сходные результаты. Появилась возможность локализовать каждый белок рибосомы в определенном месте субчастицы. Более того, некоторые белки уже сейчас могут быть сопоставлены с определенными особенностями формы рибосомы. По-видимому, организация каждой субчастицы высокоспецифична. В настоящее время вырисовывается подробная картина для 30S-субчастицы. Сходные результаты получаются для 50S-субчастицы, хотя в этом случае из-за большого размера исследования оказываются более трудоемкими.

Взаимодействие рибосомных белков и рРНК

Рибосомные РНК по массе составляют большую часть бактериальных 30S- и 50S-субчастиц. Они присутствуют повсеместно, и, вероятно, большинство или даже все рибосомные белки связаны с рРНК. Поэтому укладка рРНК определяет структуру каждой субчастицы. Сложность в определении конформации молекул рРНК и способа связывания ими рибосомных белков состоит в том,

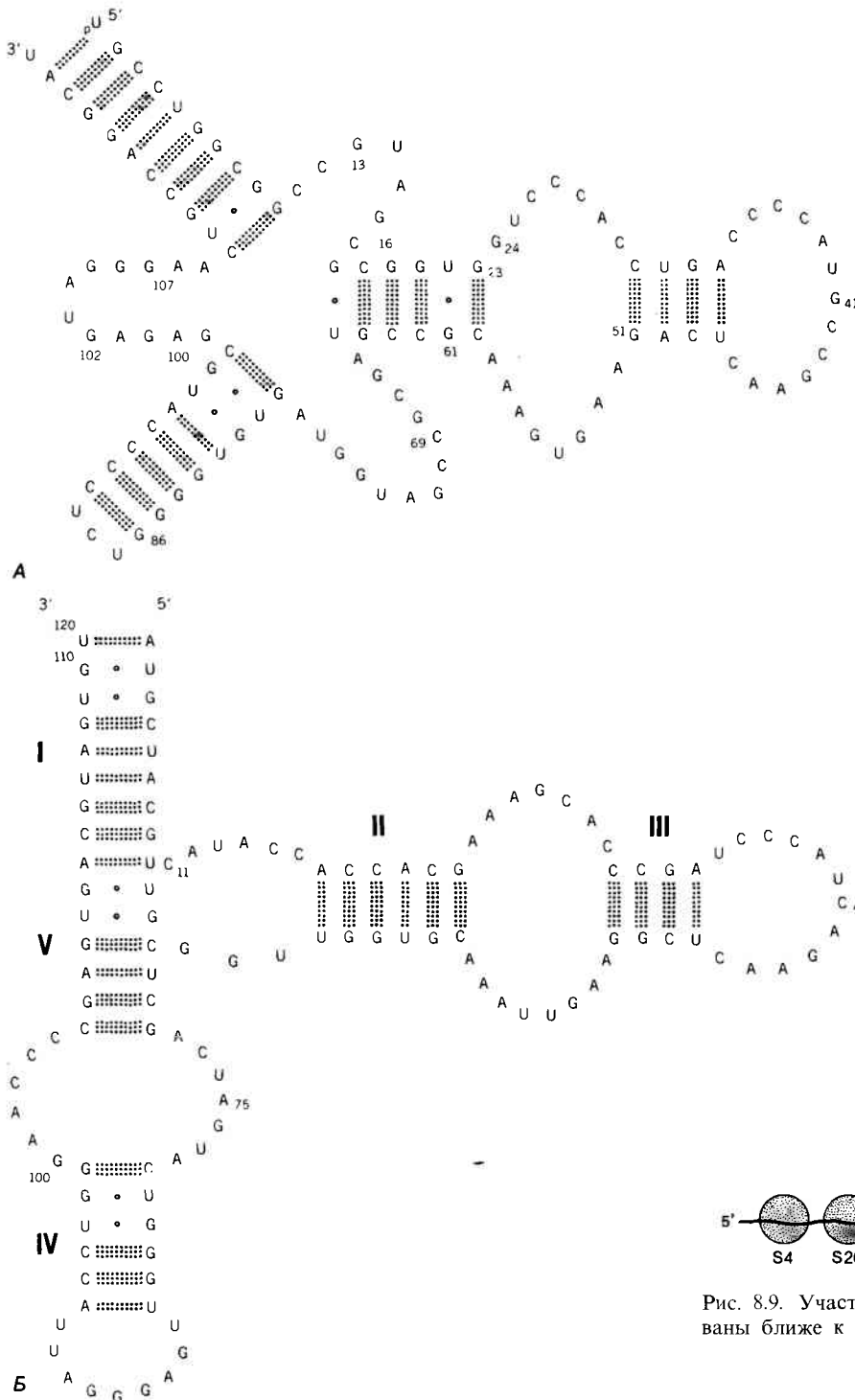


Рис. 8.8. Во вторичной структуре прокариотической 5S-РНК имеется 4 спиральных участка, тогда как эукариотическая 5S-РНК образует 5 спиральных участков. А. Модель бактериальной (*E. coli*) 5S-РНК. Б. Модель эукариотической (*Chlorella*) 5S-РНК.

что способность белков взаимодействовать с рРНК зависит от условий реакции.

Первые исследования по связыванию очищенных рибосомных белков с рРНК показали, что некоторые из них прочно присоединены к нуклеиновой кислоте. К таким белкам относятся и те, которые в процессе самосборки *in vitro* первыми входят в состав частиц. Об этом будет рассказано ниже. Участки рРНК, связывающие эти белки, могут быть охарактеризованы путем выделения фрагментов нуклеиновой кислоты, защищенных от расщепления нуклеазами. Таким путем была построена карта, показывающая расположение участков связывания рибосомных белков на рРНК. Как видно из рис. 8.9,

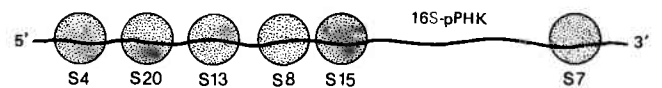


Рис. 8.9. Участки связывания рибосомных белков сгруппированы ближе к 5'-концу 16S-рРНК.

в случае 16S-рРНК места связывания сгруппированы преимущественно на 5'-конце рРНК.

Можно было думать, что остальные белки непосредственно не связаны с РНК, но оказалось, что полученный результат обусловлен самим способом выделения рРНК. В дальнейшем было показано, что, используя различные методы выделения, можно выявить еще ряд белков, способных достаточно прочно связываться с рРНК. Из этого следует, что конформация рРНК (которая зависит от метода выделения) непосредственно определяет возможность существования участков связывания для некоторых белков. По мере связывания определенных групп белков происходят конформационные изменения в рРНК, сказывающиеся на доступности других участков.

вающиеся на последующем взаимодействии с белками, которые до этого не были способны прикрепиться к рРНК. Для участков связывания, обнаруженных на рРНК, характерно наличие вторичной структуры, часто имеющей форму шпилек, чьи двухцепочечные стебли содержат неспаренные области. В отношении 30S-субчастицы *E. coli* установлено, что почти все или даже просто все рибосомные белки способны связываться с 16S-рРНК, хотя и в различной степени. При исследовании 50S-субчастицы получены сходные результаты, но в данном случае ситуация изучена менее детально.

Диссоциация и реконструкция рибосомных субчастиц

При центрифугировании рибосомных субчастиц в градиенте плотности CsCl от каждой субчастицы отделяется определенная группа белков (дополнительные **рибосоמוобразующие** белки). При этом из 30S- и 50S-субчастиц образуются соответственно 23S- и 42S-**рибонуклеопротеиновые** частицы. При обработке рибосом возрастающими концентрациями CsCl или LiCl происходит ступенчатая диссоциация определенных белков. Из этого можно сделать вывод о том, что в составе рибосомы имеются группы белков, взаимодействующих кооперативно. В результате диссоциация одного из членов группы вызывает диссоциацию остальных.

Реконструированные субчастицы обладают функциональной активностью. Оптимальные условия самосборки, первоначально разработанные для 30S-субчастиц *E. coli*, сейчас используются применительно и к 50S-субчастицам. Процесс протекает *in vitro* и состоит из двух стадий, одна из которых характеризуется температурозависимой реорганизацией частиц. Это схематично показано на **рис. 8.10**.

На первой стадии самосборки 30S-субчастиц происходит объединение 16S-рРНК с группой из 15 белков. Эта стадия протекает при пониженной температуре и приводит к образованию *RI-частиц*. Перед тем как произойдет присоединение остальных 6 белков, *RI-частицы* необходимо прогреть. При этом происходит молекулярная перестройка, приводящая к образованию *RI*-частиц*.

На природу такой перестройки проливает свет действие ЭДТА на 30S-субчастицы. Это хелатообразующее соединение связывает ионы магния, вызывая тем самым разворачивание субчастицы и диссоциацию ряда белков. Получающиеся частицы сходны с *RI-частицами* в том отношении, что они не способны снова присоединять недостающие белки с образованием функционально активных субчастиц до тех пор, пока не будет прогреты. Таким образом, лимитирующим этапом при самосборке, очевидно, является температурозависимая перестройка, приводящая к более компактной организации формирующейся частицы.

Мутации, влияющие на самосборку рибосомы

Поскольку при реконструкции *in vitro* рибосомные субчастицы необходимо инкубировать при повышенной температуре, очевидно, что можно получить **холодочувствительные мутации**, нарушающие процесс самосборки. Такие мутации отбирают среди бактериальных клеток, в которых синтез белка протекает нормально при обычной температуре, но прекращается при пониженной тем-

пературе, так как при этом не могут образовываться полноценные рибосомы. Используя такой подход, Номура (Nomura) и его сотрудники выделили несколько типов *sad*-мутаций (нарушающих сборку субчастиц). Оказалось, что ряд мутаций, полученных ранее на основе других свойств (таких, как устойчивость к антибиотикам, взаимодействующим с рибосомой), также обладают подобными свойствами.

Если инкубацию бактериальных клеток, содержащих такие мутации, проводят при пониженной температуре, образование функционально-активных субчастиц прекращается. Некоторые мутации вызывают изменения в определенных рибосомных белках; другие могут локализоваться в генах, продукты которых каким-то образом влияют на процесс самосборки.

При наличии *sad*-мутаций в клетке накапливаются частицы, которые по размеру меньше, чем функционально активные рибосомные субчастицы, и являются их предшественниками. Такие промежуточные частицы содержат рРНК, соединенную с некоторыми, но не со всеми белками, входящими в состав субчастиц. В каждом случае отсутствие конкретного белка блокирует самосборку на определенной стадии. Для каждой субчастицы охарактеризованы основные промежуточные продукты, образующиеся *in vivo*.

В случае 50S-субчастицы предшественниками являются частицы с коэффициентами седиментации 43S и 32S, содержащие 23S-рРНК и некоторые из белков, входящих в состав 50S-субчастицы. 32S-частица – первый промежуточный продукт, идентифицируемый при самосборке. 43S-частица представляет собой структуру, образующуюся на последующей стадии; она содержит все белки, входящие в состав 32S-частицы, а также несколько других белков.

В случае 30S-субчастицы обнаружены предшественники, обладающие в зависимости от условий выделения коэффициентом седиментации 26S или 21S. Они также

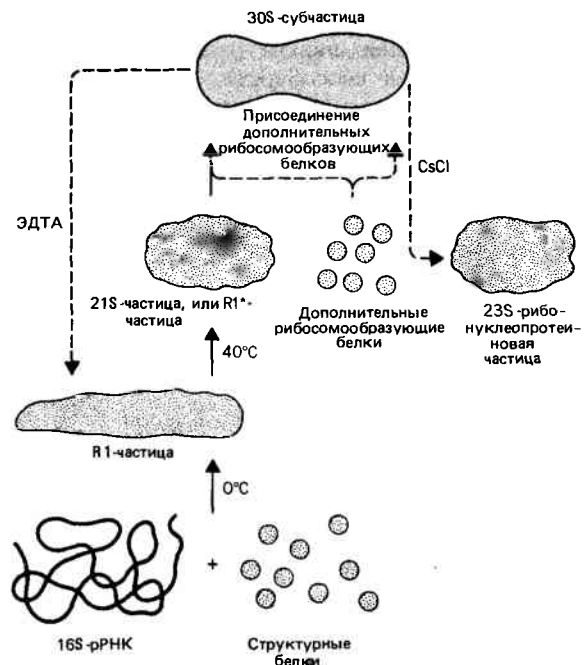


Рис. 8.10. 30S-субчастицы могут быть диссоциированы на более мелкие частицы, из которых при самосборке образуются промежуточные комплексы.

были охарактеризованы как минорные промежуточные компоненты, образующиеся при формировании рибосом в бактериях дикого типа.

Порядок самосборки определяется пространственной организацией субчастиц

Как видно из рис. 8.10, в случае малой субчастицы предшественник, обнаруживаемый *in vivo*, и R1-частицы, реконструированные *in vitro*, содержат очень похожие или даже одинаковые наборы белков. При сравнении же РНК этих частиц выявляются некоторые различия. В состав предшественника, обнаруживаемого *in vivo*, входит более длинная рибосомная РНК, тогда как частицы, сформированные *in vitro*, содержат зрелую 16S-РНК. Тот факт, что зрелая рРНК может быть использована для самосборки, заслуживает особого внимания.

Из данных, приведенных на рис. 8.11, видно, что белки, входящие в предшественники и в R1-частицы, также имеют много общего с белками, которые в присутствии LiCl или CsCl диссоциируют в последнюю очередь. Следовательно, можно предположить, что пространственная организация 30S-субчастицы отражает процесс ее самосборки. Те белки, которые входят в состав предшественника (сюда относятся все белки, наиболее прочно связанные с рРНК), как правило, являются самыми постоянными компонентами 30S-субчастиц и с трудом отделяются от рРНК даже при жестких условиях обработки.

Роль конкретного белка в процессе самосборки рибосомной субчастицы возможно проанализировать, если попытаться осуществить реконструкцию рибосомы в отсутствие этого белка. Таким путем было установлено, что при исключении некоторых белков из реакционной смеси процесс полностью прекращается. Следовательно, именно эти белки необходимы для самосборки. Другие белки *необязательны*; субчастицы, образующиеся в их отсутствие, по физическим параметрам могут быть нормальными, но по их способности осуществлять синтез белка они оказываются дефектными по ряду функций.

Влияет ли отсутствие одного белка на способность других белков к самосборке? Для изучения этого вопроса используется частичная реконструкция. В этом случае необходимо устранить из среды определенный белок, а затем проследить, будут ли оставшиеся белки способны формировать частицы. Такой подход возможно осуществить независимо от того, образуется ли в результате частичной самосборки функционально активная или неактивная частица.

Последовательность самосборки *in vitro*, изображенная на рис. 8.12, была установлена в результате опытов, в которых исследовалось, какие белки должны предвари-

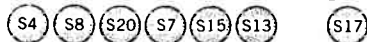
тельно связаться с 16S-РНК, для того чтобы определенный белок мог быть вовлечен в самосборку. Установлено, что рибосомные белки должны присоединяться к формирующейся частице в определенной последовательности. Некоторые белки, характеризующиеся по данным, приведенным на рис. 8.12, тесной взаимосвязью при самосборке, действительно располагаются рядом. Это соответствует представлению о том, что последовательность самосборки во многом определяется непосредственными физическими взаимодействиями. Разумеется, пока еще многое в процессе самосборки нам неясно. Так, присоединение одного белка может в принципе привести к формированию участка связывания для другого белка в любом ином участке рРНК. Все белки, необходимые для самосборки, располагаются в «ранней» части карты, т.е. связываются с рРНК первыми.

Функцию некоторых белков можно определить, проводя самосборку в их отсутствие. Очевидно, что этот прием можно использовать только в отношении белков, участвующих в завершающих стадиях самосборки, а не тех, которые определяют возможность связывания других белков. Таким образом, было установлено, что определенные белки 30S-субчастицы принимают участие в определенных стадиях белкового синтеза, так как сформировавшиеся частицы были лишены некоторых ферментативных активностей. В составе 50S-субчастицы имеется белок L24. Он одним из первых включается в состав первого предшественника и прочно удерживается, связываясь с несколькими участками, находящимися на 5'-конце рРНК. Этот белок можно отделить от субчастицы в присутствии LiCl. При этом субчастицы, лишенные L24, полностью сохраняют активность при синтезе белка. Из этого следует, что L24 необходим для процесса самосборки, но не играет никакой роли в последующем функционировании рибосомы.

Сборка 30S-субчастиц, вероятно, начинается еще во время транскрипции рРНК. Этим, возможно, и объясняется тот факт, что наибольшее число участков, прочно связывающих белки, находятся на 5'-конце транскрипта, образующегося первым. Вероятно, что в процессе разрезания предшественника освобождается рРНК, уже связанная с рядом рибосомных белков (рис. 8.13). Далее белки продолжают присоединяться к рРНК в последовательности, определяемой процессом самосборки.

Единственный промежуточный рибонуклеопротеин, который можно выделить, представляет собой отдельную частицу, содержащую большинство белков. Мы уже упоминали, что в состав такой частицы входит 16S-РНК, представленная незрелой формой, имеющей несколько дополнительных нуклеотидов. На этой стадии рРНК метилирована только частично. Вероятно, конформационные изменения рРНК сопряжены с удалением дополнительных последовательностей и с метилированием; только после этого оставшиеся белки могут войти в со-

Белки, всегда связывающиеся с 16S-рРНК



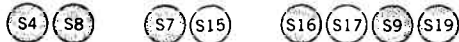
Белки, имеющиеся в составе 21S-частиц



Белки, присутствующие в составе R1-частиц, образующихся *in vitro*



Белки, необходимые для самосборки *in vitro*



Белки, которые не удаляются при обработке CsCl/LiCl



Рис. 8.11. Схожие наборы рибосомных белков входят в состав промежуточных частиц, образующихся в результате самосборки как *in vivo*, так и *in vitro*, а также в состав рибонуклеопротеиновых частиц.

Белки, присутствующие во всех случаях, обозначены затемненными кружками; те же, которые встречаются не всегда, — светлыми кружками.

став частиц. *In vitro* не удается метилировать свободные 16S-rРНК, но промежуточные рибонуклеопротеиновые частицы служат субстратом для метилаз, что согласуется с рассмотренными выше представлениями.

Мутационные изменения могут затрагивать все компоненты рибосомы

Поскольку рибосомы играют центральную роль в жизни клетки, для получения мутантов по каждому из компонентов рибосом понадобилось много времени. Но сейчас уже удалось получить мутации почти для каждого рибосомного белка, что дало возможность картировать соответствующие гены в хромосоме. Некоторые мутации были выделены по их способности подавлять в небольшой степени синтез белка, но мутации по другим белкам удалось получить только в виде условно летальных (гл. 2).

Механизм действия некоторых антибиотиков состоит в блокировании синтеза белка. Можно выделить мутантов, устойчивых к таким антибиотикам. Во многих случаях эти мутации вызывают изменения рибосомных белков (или ферментов, принимающих участие в синтезе белка). Устойчивость к антибиотикам обусловлена мутациями, затрагивающими 5 белков 30S-субчастицы и 4 белка 50S-субчастицы. Используя эти мутантные варианты, можно исследовать функцию данного белка в рибосоме.

Непосредственные попытки выделить мутации, влияющие на синтез белка, были связаны с опытами по получению изменений, сказывающихся на точности работы белоксинтезирующего аппарата. Если в каком-нибудь гене, кодирующем белок, возникла мутация, она может быть супрессирована (как об этом уже говорилось в гл. 7) мутацией в гене, детерминирующем структуру rРНК; другой тип супрессии возникает в результате рибосомных мутаций. Был выделен ряд рибосомных мутаций, вызывающих реверсию первичных мутаций в различных генах. Иногда при этом мутация затрагивает какой-либо из известных этапов трансляции — тогда удается выяснить роль конкретного белка в исследуемом процессе. Таким образом, были получены мутации в генах, кодирующих шесть рибосомных белков.

Мы уже отмечали, что могут быть получены холодочувствительные мутации, влияющие на сборку рибосом. Такой же фенотип свойствен некоторым мутантам, устойчивым к антибиотикам. Таким образом, было установлено, что четыре белка 30S-субчастицы и один белок 50S-субчастицы принимают участие в самосборке рибосом. Поскольку в отсутствие рибосом клетка погибает, данный вид мутаций является условно-летальным.

Ряд рибосомных белков был идентифицирован благодаря изменениям в их структуре у температурочувствительных мутантов. Температурочувствительные мутации приводят к тому, что при повышенной температуре рост бактерий замедляется — эффект, обусловленный каким-то повреждением в рибосомном белке, функция которого при непермиссивной температуре, по-видимому, нарушается. Общий характер таких мутаций не дает ключа к выяснению и функции данного белка.

Все же остается еще несколько рибосомных белков, в генах которых не удалось получить мутаций. Это, возможно, связано с тем, что данные белки не являются обязательными для сборки и функционирования рибосом.

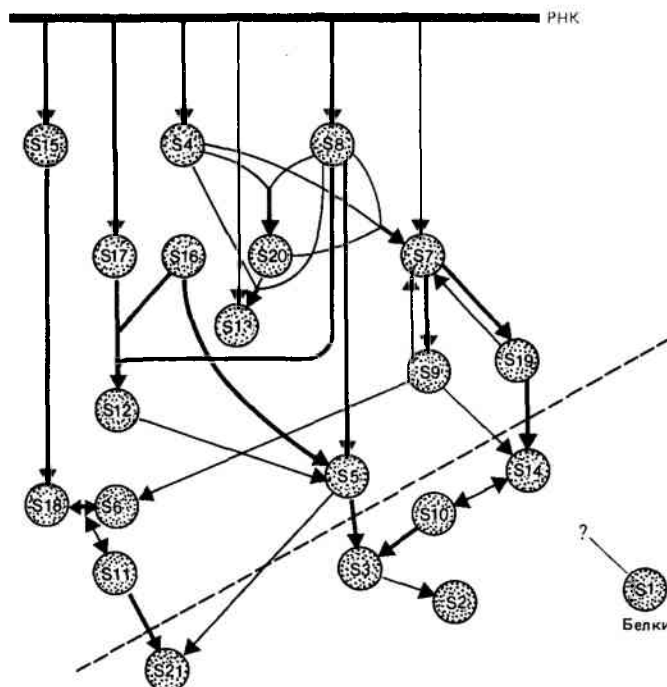


Рис. 8.12. Самосборка рибосом включает ряд последовательных реакций.

Стрелки между белками показывают, что один белок способствует присоединению другого (белок, на который направлена стрелка, может связаться с частицей только в присутствии партнера, от которого идет эта стрелка; двойные стрелки означают, что эффект взаимный). Жирными стрелками показана сильно выраженная зависимость, а тонкими — слабая зависимость. Если два белка оказывают независимое действие на какой-либо другой белок, то этот эффект обозначается индивидуальными стрелками; когда же эффект является кооперативным, то используется объединенная стрелка. Стрелками, идущими от rРНК, указаны белки, наиболее хорошо связывающиеся с rРНК; однако другие белки также могут с ней взаимодействовать.

Белки, расположенные ближе всего к rРНК, встречаются в частицах-предшественниках наиболее часто; белки, расположенные на схеме дальше от rРНК, участвуют в самосборке на более поздних стадиях и диссоциируют при более низких концентрациях соли. Пунктирная линия отделяет белки, часто встречающиеся в частицах-предшественниках, от тех, которые никогда в них не обнаруживаются. Эта последняя группа белков эквивалентна дополнительным рибосомообразующим белкам.

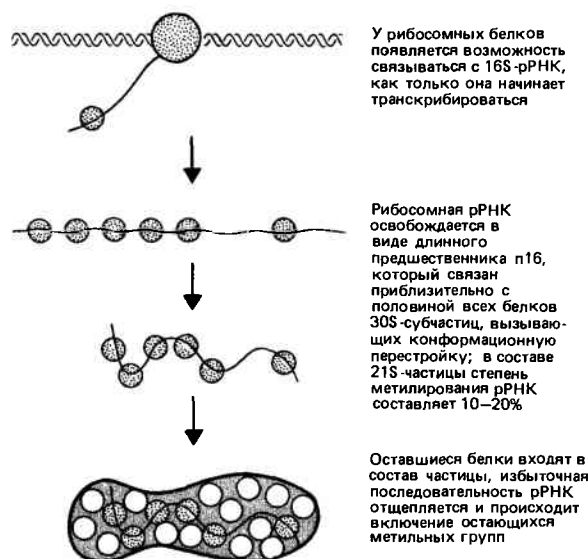


Рис. 8.13. Самосборка рибосомной субчастицы является последовательным процессом, в котором присоединение каждого белка вызывает конформационные перестройки.

Обычно мутации, затрагивающие бактериальные рибосомы, обнаруживаются в результате образования измененных белков. В ряде случаев (пока что в шести) исследуемый ген, возможно, делетирован. Рибосомы таких мутантов особенно полезны при изучении вопроса о необходимости данного белка для функционирования рибосомы.

Путем выявления мутантных белков по их электрофоретической подвижности были выделены мутации, затрагивающие все рибосомные белки. В случае этих мутаций осталось неизвестным, влияют ли мутации на функцию рибосомного белка.

Исследуя мутантов, устойчивых к касугамицину, удалось показать, что 16S-рРНК принимает непосредственное участие в белковом синтезе. Касугамицин блокирует иницирование трансляции. Мутанты, устойчивые к этому антибиотику и относящиеся к типу *KsgA*, лишены метилазы, обеспечивающей присоединение четырех метильных групп к двум соседним аденинам, расположенным недалеко от 3'-конца 16S-рРНК. При этом образуется высококонсервативная последовательность G-ш²A-m²A, обнаруживающаяся как в эукариотических, так и в прокариотических 16S-рРНК.

Из этого следует, что наличие четырех метильных групп в 16S-рРНК рибосом дикого типа делает их чувствительными к действию касугамицина. Почему же отсутствие метильных групп придает устойчивость к этому препарату? 30S-субчастицы как дикого, так и мутантного типов одинаково чувствительны к ингибирующему действию касугамицина, которое состоит в нарушении связывания fMet-тРНК_f. Но в присутствии 50S-субчастицы fMet-тРНК_f освобождается из чувствительных (дикого типа) к действию антибиотика рибосом, тогда как в мутантных рибосомах эта тРНК сохраняется, что позволяет продолжить синтез белка. Следовательно, можно предположить, что метильные группы участвуют во взаимодействии 30S- и 50S-субчастиц, а также в удержании инициаторной тРНК. О влиянии большой субчастицы на этот процесс свидетельствует тот факт, что мутации по белку 50S-субчастицы приводят к зависимости рибосомы от касугамицина. Такие мутации картируются в генах, кодирующих некоторые белки 50S-субчастицы.

Другим указанием на важное значение 3'-конца 16S-рРНК является чувствительность белкового синтеза к колицину E3. Этот белок продуцируется некоторыми бактериями и разрезает 16S-РНК *E. coli* на расстоянии около 50 нуклеотидов от ее 3'-конца. В результате нарушается инициация белкового синтеза.

Некоторые мутационные изменения большой рРНК митохондриальных рибосом придают устойчивость к хлорамфениколу. Этот антибиотик останавливает синтез белка, осуществляемый бактериальными и митохондриальными рибосомами, блокируя пептидилтрансферазную реакцию. При этом у бактериальных мутантов, устойчивых к хлорамфениколу, изменены белки 50S-субчастицы. В митохондриях мутации затрагивают единственное основание, находящееся на расстоянии 243 нуклеотидов от 3'-конца большой рРНК.

Рибосомы содержат несколько активных центров

Исходя из функций и строения рибосомы, в ее составе можно выделить ряд каталитических центров. 30S-субчастица связывает мРНК и комплекс, состоящий из ини-

цирующей тРНК и фактора инициации; после этого происходит присоединение 50S-субчастицы. У 70S-рибосом есть функционально различные участки Р и А, с которыми связывается тРНК; на 50S-субчастице находится пептидилтрансферазный центр. В упрощенном виде взаиморасположение этих участков изображено на рис. 8.14. Как же эти участки соотносятся с реальной топологией рибосомы? Какие рибосомные компоненты принимают участие в реализации различных функций? Несколько подходов было использовано для анализа взаимосвязи между структурой и функцией рибосомы.

Один из способов идентификации участков, выполняющих определенную функцию, основан на использовании аффинной метки (метки по сродству). При этом используют аналоги компонентов, которые связываются с рибосомой. Эти соединения либо сами по себе химически активны, либо их можно активировать *in situ*. Например, можно использовать тРНК с подходящими модификациями; таким путем можно идентифицировать рибосомные компоненты, на которые переносится метка. Сходный метод сводится к использованию соединений, способных в результате фотохимического активирования вызывать сшивки, что позволяет идентифицировать взаимодействующие компоненты.

Производные пептидил-тРНК (содержащие метку в аминокцепторной области) взаимодействуют с белками L2 и L27. Следовательно, можно думать, что эти белки входят в состав участка Р. Используя более протяженные производные, можно пометить белки L14, L18, L24 и L33. Таким путем было показано, что данные белки располагаются вблизи участка Р или тоже входят в его состав. Эти данные показывают, что конец тРНК, связанный с пептидной цепью, располагается на большой субчастице в области, ограниченной несколькими определенными белками.

Часто применяется модификация целых субчастиц. Один из подходов состоит в следующем: субчастицы обрабатывают реагентами, повреждающими белки или РНК, с последующей попыткой увязать возникшее повреждение с потерей определенной функции. Например, кетоксаль был использован для введения модификаций в состав рРНК, а тетранитрометан — для нитрования белков. Проблема состоит в том, чтобы ограничить степень возникающих повреждений, поскольку таким путем можно установить корреляции между определенной модификацией и потерей данной функции.

Обработка рибосом кетоксалем препятствует их связыванию с аминоацил-тРНК, хотя при этом частицы сохраняют способность взаимодействовать с мРНК. (Это доказывает, что кетоксаль не вызывает необратимой инактивации всех функций частицы.) При этом модифицируются шесть гуаниновых остатков 16S-рРНК. Следовательно, 16S-РНК является составной частью участка А или по крайней мере оказывает влияние на его структуру. Более прямые доказательства были получены при использовании фотохимических бифункциональных реагентов, при помощи которых удалось показать, что пептидил-тРНК можно ковалентно «пришить» к участку, находящемуся поблизости от 3'-конца 16S-рРНК. Из этого следует, что рРНК, очевидно, также входит в состав участка Р. Однако не удается ковалентно сшить рРНК и аминоацил-тРНК, находящуюся в участке А.

Включение 5S-РНК в состав 50S-субчастиц, реконструированных *in vitro*, зависит от наличия трех белков: L5, L8 и L25, образующих с этой РНК стехиометрический комплекс. Комплекс способен связываться

с 23S-рНК, хотя по отдельности ни один из его компонентов не обладает такой способностью. Вероятно, этот комплекс располагается поблизости от участков А и Р.

Консервативная последовательность в ТψС-петле тРНК комплементарна определенной последовательности в 5S-рНК. Было высказано предположение, что комплементарное взаимодействие между 5S-рНК и тРНК играет определенную роль в стабилизации аминокил-тРНК при ее попадании в участок А. Однако опыты, выполненные *in vitro*, показали, что удаление соответствующего фрагмента 5S-рНК не мешает рибосоме принимать участие в трансляции. Таким образом, данное взаимодействие, по-видимому, не является обязательным для синтеза белка.

Эксперименты по образованию межмолекулярных швов с использованием EF-Tu показали, что в состав участка А входят белки L1, L5, L7/L12, L20, L30 и L33. К сожалению, мы плохо себе представляем, как эти и другие белки взаимодействуют при образовании тРНК-связывающего участка.

Группа белков, принимающих участие в пептидилтрансферазной реакции, включает L2, L3, L4, L15 и L16. Попытки приписать какому-либо из этих белков ферментативную активность не увенчались успехом. Двумя наиболее возможными кандидатами являются белки L2 и L16.

При отделении 50S-субчастицы от L7/L12 оставшиеся частицы не могут осуществлять гидролиз GTP в присутствии EF-Tu, EF-G или IF2. Однако это еще не означает, что L7/L12 является GTPазой. Возможно, что данный белок (белки) необходим для проявления GTPазной активности другим белком. Ситуация с EF-Tu до конца не выяснена, так как существуют факты, свидетельствующие о том, что этот белок обладает собственной GTPазной активностью, и, следовательно, нет необходимости полагаться на рибосомную GTPазу (гл. 6).

Транслокационный фактор EF-G связывается с 50S-субчастицей; но процесс транслкации представляет собой протергивание мРНК через 30S-субчастицу. Участок связывания для EF-G расположен в непосредственной близости от S12—одного из белков, ответственных за связывание мРНК 30S-субчастицей. При этом EF-G оказывается между субчастицами вблизи димеров L7/L12.

Участки Р и А должны располагаться рядом, поскольку расположенные в них молекулы тРНК взаимодействуют с соседними триплетами мРНК. Между этими участками существует также и функциональная связь: участок Р должен быть занят еще до того, как аминокил-тРНК сможет войти в участок А. При попытке изобразить, как две молекулы тРНК должны располагаться в рибосоме по отношению друг к другу, всегда возникает определенное затруднение. Дело в том, что расстояние между антикодонами не может превышать 10 Å, а между тем диаметр тРНК составляет приблизительно 20 Å.

Молекулы тРНК, находящиеся в участках Р и А, по-видимому, имеют одинаковую конформацию. При этом отнюдь не доказано, что эта конформация такая же, как у молекул тРНК, находящихся в форме кристаллов. Одним из решений этой стереохимической проблемы могло бы быть образование излома или узла в мРНК между кодонами, так чтобы две молекулы тРНК могли подойти к матрице под разными углами. Тогда передвижение рибосомы могло бы сводиться к перебросу тРНК из участка А в расположенный рядом участок Р. Это означало бы, что передвижение мРНК функционально зави-

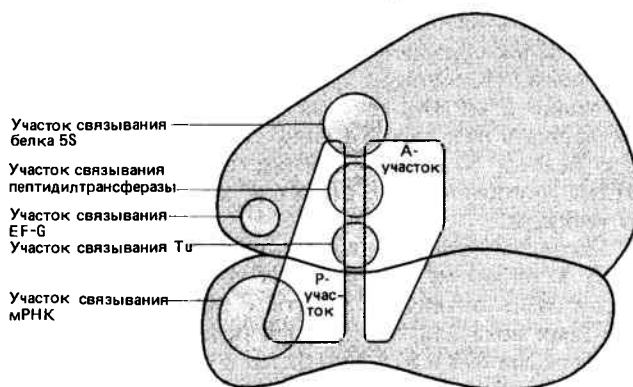


Рис. 8.14. Каталитические центры располагаются на рибосоме в нескольких участках.

Местоположение различных участков обозначено предельно схематично и не может быть связано с трехмерной структурой рибосомы.

Участок связывающий мРНК, расположен на 30S-субчастице; в его состав входят белки S1, S18 и S21, а также S3, S4, S5, S12 и 3'-концевая область 16S-рНК; с этим участком связывается также фактор IF3. Этот участок расположен в непосредственной близости от участка Р.

Участок Р преимущественно расположен на 30S-субчастице и способен связывать инициаторную тРНК. 16S-рНК также входит в состав этого участка благодаря области, расположенной весьма близко к ее 3'-концу, но не ближе, чем 50 оснований, включающих в себя участок связывания мРНК. Со стороны 50S-субчастицы в участок Р входят белки L2 и L27, а также L14, L18, L24 и L33.

Участок А, должно быть, располагается рядом с участком Р, но точные его координаты не известны. 16S-рНК оказывает влияние на участок А и, возможно, входит в его состав. Большинство функций, приписываемых участку А, связано с 50S-субчастицей. В состав этого участка входят белки L1, L5, L7/L12, L20, L30 и L33.

Пептидилтрансферазный участок должен располагаться в области, непосредственно контактирующей с участками А и Р, вблизи от конца тРНК. Белки L2, L3, L4, L15 и L16 необходимы для проявления этой активности. 5S-рНК располагается рядом с пептидилтрансферазным участком; она взаимодействует с белками L5, L8 и L25.

Участок Tu состоит из нескольких расположенных рядом белков (L5, L1, L20 и L7/L12), способных образовывать попеременные швы с EF-Tu. **Участок L7/L12** представляет собою «палец», изображенный на рис. 8.2; здесь он не показан, но «палец» состоит из двух димеров L7/L12. **Участок EF-G**, возможно, расположен на большой субчастице рядом с контактирующей поверхностью, обращенной к малой субчастице, недалеко от белка L12.

сит от перемещения тРНК; при этом кодон-антикодонное спаривание определяет расстояние, на которое должна передвинуться матрица при транслкации, о чем подробно рассказывалось в гл. 7.

Функциональные взаимосвязи между различными участками рибосомы окончательно не установлены. Рибосома, по-видимому, имеет очень сложную организацию, так что изменения в одном центре могут сильно отразиться на работе другого, расположенного поодаль.

Связывание 30S-субчастиц с мРНК

Для первоначального взаимодействия мРНК с 30S-субчастицей необходимо присутствие белка S1. Значение этого белка для структуры рибосомы является предметом дискуссий. В процессе самосборки рибосомы белок S1 последним входит в состав частиц. Этот белок легко отделяется от рибосом и при определенных условиях может быть обнаружен в несвязанном состоянии в избыточном количестве. В связи с этим возникает следующий вопрос: следует ли рассматривать этот белок как фактор трансляции или как составную часть рибосомы? Однако при обычных условиях этот белок присутствует в количестве одной копии на частицу и не отделяется от рибосомы

в процессе синтеза белка. В связи с этим теперь принято рассматривать белок S1 как компонент субчастицы.

Этот белок обладает повышенным сродством к одноцепочечным нуклеиновым кислотам, что способствует его связыванию с мРНК, которая в результате этого поддерживается в развернутом виде. По-видимому, связывание с белком S1 может препятствовать образованию в мРНК двухцепочечных структур, мешающих связыванию рибосом.

Сложность, возникающая при исследовании локализации S1 в рибосоме, обусловлена формой этого белка, которая очень вытянута. Опыты с бифункциональными реагентами показали, что данный белок расположен рядом с белками S18 и S21. Эти белки в опытах с использованием метки по сродству реагируют с тРНК, связанной с кодоном AUG. Они располагаются в передней части малой субчастицы. Исходя из этого, можно предположить то, что рассматриваемые три белка образуют относительно небольшой домен, ответственный как за первичное связывание мРНК, так и за связывание инициаторной тРНК.

Рядом с этим доменом располагается также 3'-конец 16S-рРНК, участвующий в процессе первоначального связывания. На 3'-конце 16S-рРНК имеется относительно короткий участок, комплементарный последовательности, расположенной непосредственно перед каждым иницирующим кодоном во всех бактериальных мРНК. Возможно, что комплементарное взаимодействие между этими последовательностями ответственно за первоначальное связывание рибосом с иницирующими кодонами мРНК (гл. 9).

Фактор инициации IF3 связывается с той же областью рибосомы. Этот фактор можно ковалентно «пришить» к 3'-концу рРНК, так же как и ряд рибосомных белков, включая и те, которые принимают участие в связывании мРНК. При условии что рассматриваемая область 30S-субчастицы также принимает участие в связывании 50S-субчастицы, вероятная роль IF3 может заключаться в стабилизации взаимодействия между 30S-субчастицей и мРНК. В результате последующего присоединения 50S-субчастицы этот фактор вытесняется.

Механизм, обеспечивающий объединение двух субчастиц, изучен плохо. Существует частичная комплементарность между участком 23S-рРНК и (опять) 3'-областью 16S-рРНК. Одна из возможностей состоит в том, что соединение субчастиц обусловлено комплементарным взаимодействием оснований. В этом случае связывание IF3 с 16S-рРНК должно препятствовать ассоциации субчастиц.

Точность трансляции

При анализе первичной последовательности белка не удается обнаружить каких-либо вариаций. Из этого следует, что синтез белка — процесс в высшей степени точный. Очень небольшое число происходящих ошибок составляют замены одной аминокислоты на другую. В процессе синтеза белка ошибки могут возникать на двух стадиях.

Очевидно, что решающим моментом для безошибочного синтеза является активирование тРНК правильной аминокислотой. Как мы увидели в гл. 7, этот процесс осуществляется аминоацил-тРНК-синтетазами. Вероятно, что частота ошибки варьирует у разных ферментов, но усредненная ошибка составляет менее чем 1 на 10^5 актов аминоацилирования.

Критическое значение для безошибочного синтеза белка имеет также правильное кодон-антикодонное взаимодействие. Для различных кодон-антикодонных взаимодействий значения констант связывания могут различаться, поскольку пары образуются между различными основаниями. Поэтому частота неправильного спаривания индивидуальна для каждой кодон-антикодонной пары. Например, замена аргинина на цистеин в бактериальном белке флагеллине происходит с частотой приблизительно 1 на 10^4 триплетов, кодирующих аргинин. (Все неправильные включения в рассматриваемом примере происходят, по всей видимости, из-за ошибок при распознавании кодона, так как неверное ацилирование тРНК в данном случае маловероятно.)

Точность белкового синтеза является одним из наиболее загадочных сторон этого процесса. В растворе свободная тРНК взаимодействует с триплетным кодоном относительно слабо, и похожие, но неправильные триплеты (содержащие только два из трех нужных оснований) узнаются в 10–100 раз менее эффективно, чем правильные триплеты. Таким образом, кодон-антикодонное взаимодействие не настолько специфично, чтобы обеспечить такой низкий уровень ошибки, как 10^{-4} .

Следовательно, можно предположить, что рибосома способна непосредственно или опосредованно осуществлять «корректирующее прочтение», различая правильные и неправильные кодон-антикодонные пары, усиливая в результате относительно слабые первоначальные различия между разными кодон-антикодонными парами. Предположим теперь, что не существует специфичности при первоначальном контакте между тройным комплексом аминоацил-тРНК·EF-Tu·GTP и рибосомой. При условии что любой комплекс независимо от природы входящей в его состав тРНК способен попасть в участок А, количество неправильных взаимодействий должно намного превышать количество правильных. Поэтому необходимо существование некоего механизма, ответственного за стабилизацию правильной аминоацил-тРНК, позволяющего соответствующей аминокислоте выступать в роли акцептора для присоединения полипептидной цепи в реакции транспептидации.

Неправильные контакты должны быстро разрушаться; следовательно, комплекс покидает рибосому, не образовав пептидной связи. Каким же образом рибосома оценивает правильность кодон-антикодонного взаимодействия, происходящего в участке А?

Впервые способность рибосом оказывать влияние на точность процесса трансляции была обнаружена при изучении мутаций, вызывающих устойчивость к стрептомицину. Один из эффектов влияния стрептомицина на синтез белка — это увеличение уровня ошибочного прочтения пиримидиновых оснований U и C. (При этом, как правило, один из пиримидинов включается вместо другого; иногда может включаться A.) Мишенью, чувствительной к действию стрептомицина, является белок S12. В стрептомицинустойчивых мутантах первичная последовательность этого белка изменена. Рибосомы с мутантным белком S12 в отличие от рибосом дикого типа характеризуются пониженным уровнем возникновения ошибок при трансляции. В результате это компенсирует эффект, обусловленный действием стрептомицина на процесс ошибочного считывания.

Мутации в двух других локусах, кодирующих белки S4 и S5, также влияют на процесс неправильного считывания, так как для каждого локуса могут быть выделены

ревертанты, характеризующиеся обычным уровнем ошибок при считывании. Таким образом, точность процесса трансляции контролируется в результате взаимодействия этих трех белков. Как степень ошибочного считывания, так и характер ответа на стрептомицин зависят от набора присутствующих белков. При определенных их комбинациях правильная работа рибосомы даже становится **зависимой** от стрептомицина.

Непосредственно точность процесса трансляции могла бы быть обусловлена стереохимией участка А. Его геометрия могла бы изменяться таким образом, чтобы определять степень свободы, необходимую для правильного кодон-антикодонного взаимодействия. При интерпретации первоначальных данных об эффекте стрептомицина исходили из топологических особенностей рибосом, предполагая, что кодон-антикодонное взаимодействие тщательно выверяется и может служить более или менее точным критерием для акцептирования аминоксил-тРНК.

Другая возможность состоит в том, что наблюдаемый эффект стрептомицина может быть опосредованным. Скорость, с которой функционирует рибосома, должна лимитировать время, имеющееся у нее в распоряжении для идентификации тРНК, что и определяет эффективность процесса. Основываясь на таком представлении, действие стрептомицина можно объяснить его влиянием на кинетические параметры процесса элонгации. Таким параметром является скорость работы рибосомы относительно времени, необходимого для осуществления выбора правильной аминоксил-тРНК. По мере возрастания скорости синтеза пептидной связи увеличивается вероятность захвата неправильных молекул аминоксил-тРНК и включения ошибочной аминоксилоты в пептидную цепь по сравнению с вероятностью удаления такой аминоксил-тРНК. При уменьшении скорости белкового синтеза предоставляется больше времени для исправления ошибок. Имеется ряд доказательств того, что скорость синтеза полипептидной цепи связана с уровнем ошибочного чтения.

Существует предположение, что установление правильного контакта между кодоном и антикодоном может служить сигналом для возникновения конформационных изменений на другом конце тРНК. Это, вероятно, требуется для соответствующего ориентирования аминоксилоты, необходимого для акцептирования полипептида, донором которого служит пептидил-тРНК.

Вопрос, имеющий важное значение при подсчете энергетических затрат на синтез белка, заключается в следующем: на какой стадии процесса принимается решение о возможности акцептирования тРНК. Если это происходит сразу же после присоединения комплекса аминоксил-тРНК · EF-Tu · GTP, который затем, не претерпев никаких изменений, освобождается в случае несоответствия тРНК, то дополнительные затраты энергии минимальны. Но если такое решение происходит после гидролиза GTP, то всегда, когда присоединяется не соответствующая кодону тРНК, расходуется энергия высокоэнергетической связи. Это приведет к увеличению энергетических затрат сверх тех трех высокоэнергетических связей, которые, как известно, расходуются при включении каждой аминоксилоты в полипептидную цепь. Существует ряд фактов, свидетельствующих о том, что *in vitro* на каждую аминоксилоту расходуется на 3–4 молекулы GTP больше, чем это предполагалось.

Рекомендуемая литература

Основной монографией по рибосомам является книга под редакцией Номуры, Тиссьера и Ленджила (*Nomura, Tissieres, Lengyel, Ribosomes, Cold Spring Harbor Laboratory, New York, 1974*). Хотя кое-что в этой книге устарело, она дает хорошее представление о структуре и функции рибосом. Структуре рибосом посвящен ряд обзоров; бактериальным рибосомам — обзор Курланда (*Kurland, Ann. Rev. Biochem., 46, 173–200, 1977*) и Бримакомба, Стофлера и Витманн (*Brimacombe, Stoffler, Wittmann, Ann. Rev. Biochem., 47, 217–250, 1978*), а эукариотическим рибосомам — обзор Вула (*Wool, Ann. Rev. Biochem., 48, 719–754, 1979*). Рибосомной РНК посвящена работа Ноллера и Войса (*Noller, Woese, Science, 212, 403–411, 1981*). Структура и функция рибосомных участков рассматривается Понгсом (*Pongs, In: Altman (Ed.), Transfer RNA, MIT Press, Cambridge, 1978, p. 78–104*). Самосборка 50S-субчастиц рассматривается, включая ранние работы в сообщении Рола и Нейрхауза (*Rohl, Nierhaus, Proc. Nat. Acad. Sci. USA, 79, 729–737, 1982*). Последнее сообщение об отсутствии взаимодействия между 5S-РНК и тРНК поступило от Пэйса и др. (*Pace et al., Proc. Nat. Acad. Sci. USA, 79, 36–40, 1982*). Факторы, влияющие на точность трансляции, вкратце рассмотрены в обзоре Курланда (*Kurland, Cell, 28, 201–202, 1982*).

Глава 9

ИНФОРМАЦИОННАЯ РНК В КАЧЕСТВЕ МАТРИЦЫ ДЛЯ СИНТЕЗА БЕЛКА

В основе высказанного в свое время предположения о существовании мРНК лежал тот очевидный факт, что в эукариотических клетках должен существовать некий посредник, способный переносить генетическую информацию от ядра, где находится ДНК, к цитоплазме, где происходит синтез белков. Данные о том, что этот посредник служит матрицей, на которой аминоксилоты собираются в полипептидную цепь, впервые были получены при исследовании бактерий. В отношении рибосом было установлено, что они являются тем местом, в ко-

тором происходит синтез белка; возможность использования их в качестве матрицы была исключена. Затем удалось показать, что именно информационная РНК (мРНК) является той разновидностью макромолекул, которая, прикрепляясь к рибосоме, транслируется с образованием белков (см. гл. 5).

Недолговечность существования бактериальной мРНК препятствовала выделению этих молекул; тем не менее их свойства были предсказаны довольно детально на основе особенностей процесса трансляции. Прошло

некоторое время, пока мРНК была выделена из эукариот; оказалось, что она является более стабильным цитоплазматическим компонентом, чем бактериальная мРНК и, подобно последней, выделяется как составная часть полирибосом. Поскольку эукариотическая мРНК относительно стабильна, ее можно выделить без особых затруднений, и сейчас уже в принципе можно получить мРНК для любого белка. Таким образом, опять же в силу сложившихся обстоятельств наиболее детально была охарактеризована мРНК именно эукариотических клеток — именно в этих клетках, согласно первоначальному предположению, она и должна была бы быть обнаружена.

Во всех живых клетках перед мРНК стоит одна и та же задача: с помощью генетического кода перевести последовательность нуклеотидов в определенную последовательность аминокислот. Однако между прокариотическими и эукариотическими мРНК существуют различия, связанные с особенностями основных моментов синтеза и структуры. Наиболее очевидное из них состоит в том, что эукариотическая мРНК синтезируется в виде большого предшественника. Далее происходит процесс созревания, как правило включающий значительное уменьшение размеров, а также другие модификации, после чего мРНК поступает в цитоплазму для трансляции. Таким образом, синтез и экспрессия эукариотических мРНК происходят совершенно независимо в разных частях клетки. У бактерий же процесс транскрипции и процесс трансляции мРНК сопряжены. Они настолько тесно связаны между собой, что могут происходить одновременно на единой матрице. Принципиальное различие в процессе трансляции у прокариот и эукариот состоит в том, что у бактерий мРНК может кодировать несколько белков, тогда как с эукариотической мРНК считывается только одна полипептидная цепь.

Недолговечность бактериальных мРНК

Бактериальные мРНК крайне недолговечны. В этом можно убедиться двумя путями. В обоих случаях вызывают остановку транскрипции и прослеживают судьбу мРНК, уже существовавшей в клетке.

Функциональное время полужизни мРНК определяется ее способностью служить матрицей для синтеза белкового продукта. Обычно оно составляет около 2 мин. Другими словами, каждые 2 мин количество белка, новосинтезирующегося на индивидуальной мРНК, уменьшается наполовину.

Химическое время полужизни мРНК измеряется по уменьшению количества мРНК, способной гибридизоваться с соответствующей ДНК. Эта величина сравнима с функциональным временем полужизни, но, как правило, химический распад несколько запаздывает по сравнению с функциональным распадом.

Из этого следует, что на первом этапе деградации мРНК ее использование в качестве матрицы, очевидно, предотвращается, но этих структурных изменений оказывается недостаточно, чтобы они препятствовали гибридизации. Например, единственный разрыв в молекуле мРНК не скажется на способности гибридизоваться с ДНК, но может нарушить ее трансляцию.

На втором этапе происходит собственно деградация мРНК — до составляющих ее нуклеотидов. Детали молекулярных событий, ответственных за деградацию, с трудом поддаются изучению, но есть основания думать, что распад молекулы, очевидно, происходит более или менее

последовательно от 5'-конца к 3'-концу мРНК — в том же самом направлении, в котором мРНК транскрибируется и транслируется.

У бактерий транскрипция и трансляция взаимосвязаны. Конкретные значения скорости процессов зависят от температуры, но обычно они согласуются друг с другом. Например, при 37°C транскрипция мРНК происходит со скоростью 2500 нуклеотидов в 1 мин, что соответствует образованию 14 кодонов в 1 с. Полученное значение очень хорошо совпадает со скоростью белкового синтеза, составляющего приблизительно 15 аминокислот в 1 с. При индукции экспрессии нового гена соответствующая мРНК появляется обычно в клетке через 2,5 мин, а соответствующий белок можно обнаружить через 0,5 мин.

Совпадение рассмотренных значений и соответствие между временем первоначального появления мРНК и соответствующего белка, очевидно, свидетельствуют о том, что транскрипция и трансляция происходят одновременно. Экспрессия гена начинается, когда фермент РНК-полимераза связывается с ДНК и инициирует синтез мРНК. Рибосомы прикрепляются к 5'-концу мРНК и начинают трансляцию еще раньше, чем заканчивается синтез информационной молекулы. В результате за РНК-полимеразой непосредственно следует «связка» рибосом.

Деградация мРНК происходит по мере ее трансляции. Возможно, деградация начинается через 1 мин после того, как начинается транскрипция. Более того, деградация 5'-конца мРНК может начаться раньше, чем закончится синтез или трансляция 3'-конца молекулы. Деградация, по-видимому, следует за последней рибосомой, передвигающейся в «конвое» вдоль молекулы. Но деградация мРНК происходит более медленно, составляя, вероятно, половину от скорости транскрипции или трансляции. В результате дистальная 3'-часть мРНК существует более длительное время, чем расположенная вначале 5'-область.

Очевидно, что такая последовательность событий возможна лишь благодаря тому, что транскрипция, трансляция и деградация осуществляются в одном и том же направлении. **Рис. 9.1** дает некоторое представление о временной очередности событий, лежащих в основе экспрессии типичного бактериального гена.

Мы не должны забывать, что время полужизни — понятие статистическое и что для каждой мРНК в любой момент времени существует вероятность подвергнуться деградации. Так, вновь синтезированная мРНК может быть деградирована с такой же вероятностью, как и долгоживущая мРНК. Поэтому некоторые копии будут транслироваться много раз, тогда как другие едва ли вообще будут функционировать. Такой статистический процесс характерен для молекул мРНК как прокариотического, так и эукариотического происхождения, но при этом общее количество белка, транслируемого с некой информационной последовательности, является вполне определенной предсказуемой величиной.

Строение бактериальной мРНК

Поскольку бактериальные мРНК нестабильны, их не часто удается выделить неповрежденными. Тем не менее некоторые сведения об их строении могут быть получены на основе данных о структуре генов, с которых они транскрибируются. Область ДНК, соответствующую мРНК, можно обнаружить при помощи гибридизационного анализа, определяя способность радиоактивно меченой

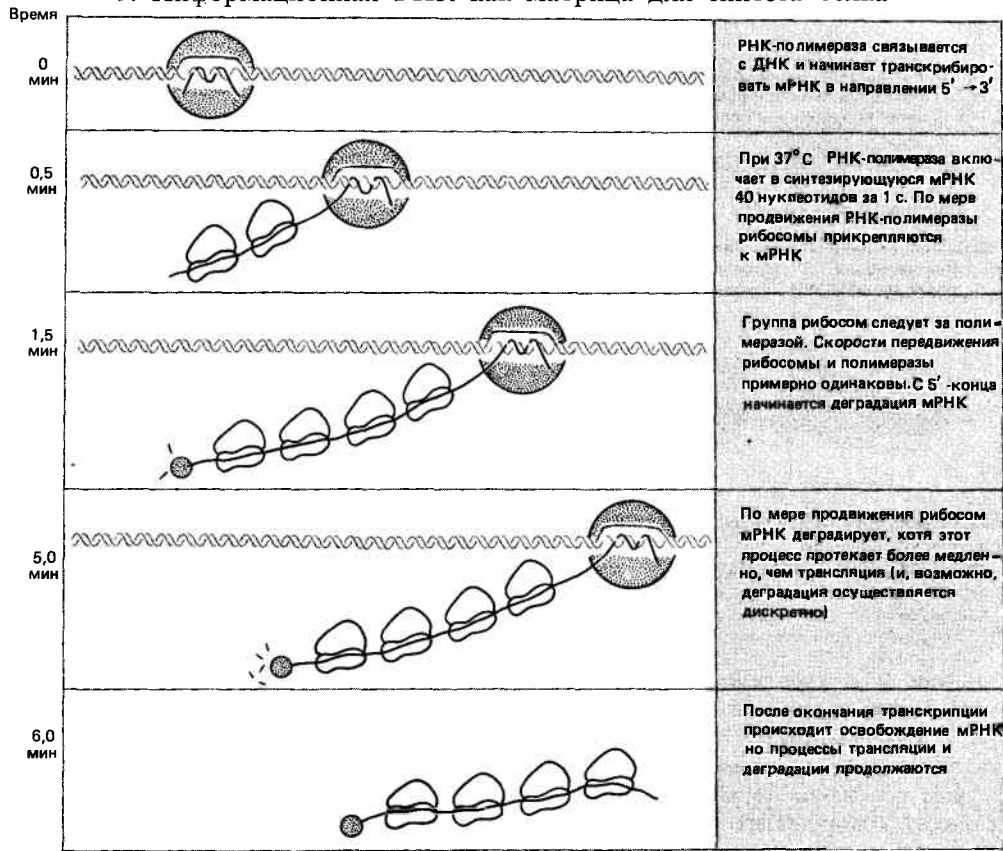


Рис. 9.1. Транскрипция, трансляция и деградация мРНК у бактерий происходит одновременно. На рисунке изображена гипотетическая транскрипционная единица раз-

мером 12 000 п.н. Ее экспрессия начинается в момент времени 0. Для простоты транскрипционная единица изображена так, как будто она содержит один ген.

мРНК гибридизоваться с фрагментами ДНК, представляющими определенные части гена. Поскольку последовательность ДНК определить относительно легко, строение мРНК можно установить, исходя из нуклеотидной последовательности соответствующего гена. С помощью такого подхода без непосредственного исследования бактериальной мРНК была создана подробная картина строения и функционирования этих молекул.

Удобным исключением с точки зрения выделения мРНК является фаговая инфекция. В ряде случаев мРНК более стабильна в клетках, инфицированных фагом, чем в интактной бактерии-хозяине, что дает возможность физически охарактеризовать эту молекулу. Превосходный образец для таких исследований представляют РНК-содержащие фаги, у которых мРНК служит и геномом и поэтому стабильна и может быть получена в больших количествах.

Важным подходом к непосредственному исследованию бактериальной мРНК является использование бесклеточной системы транскрипции и трансляции. При наличии подходящей матрицы мРНК может быть синтезирована *in vitro*. В качестве такой матрицы обычно используют клонированную копию гена. В присутствии рибосом *E. coli*, набора аминоксил-тРНК и прочих необходимых компонентов синтезированную мРНК можно транслировать в белковый продукт. Таким образом, можно исследовать и структуру, и функцию мРНК.

Бактериальные мРНК значительно различаются по количеству кодируемых ими белков. Некоторые молекулы соответствуют только одному гену — это **моноци-**

стронные мРНК. Другие (таких большинство) содержат последовательности, кодирующие несколько белков, — это **полицистронные** мРНК. В этом случае единая мРНК транскрибируется с группы генов, расположенных рядом (как мы увидим в гл. 14, такой кластер генов представляет оперон, который контролируется как единая генетическая единица).

В составе всех мРНК можно выделить участки двух типов. **Кодирующий участок** состоит из набора кодонов, соответствующих аминокислотной последовательности белка: он начинается обычно с кодона AUG и заканчивается терминирующим кодоном. Но мРНК всегда оказывается длиннее, чем кодирующий участок. В моноцистронной мРНК на обоих концах могут находиться дополнительные участки. Дополнительная последовательность на 5'-конце, предшествующая началу кодирующего участка, называется **лидерной**. Дополнительная последовательность, следующая за терминирующим сигналом и образующая 3'-конец, обозначается как **концевая (трейлер)**. Несмотря на то что эти последовательности входят в состав транскрипционной единицы, они не используются для кодирования белков.

На **рис. 9.2** изображена структура полицистронной мРНК. **Междистронные области**, располагающиеся между разными кодирующими участками, значительно различаются по размеру. В случае фаговых РНК они могут быть достаточно протяженными, достигая примерно 100 оснований. В некоторых бактериальных мРНК они состоят из нескольких (до 30) нуклеотидов, но возможно существование еще более коротких последовательно-

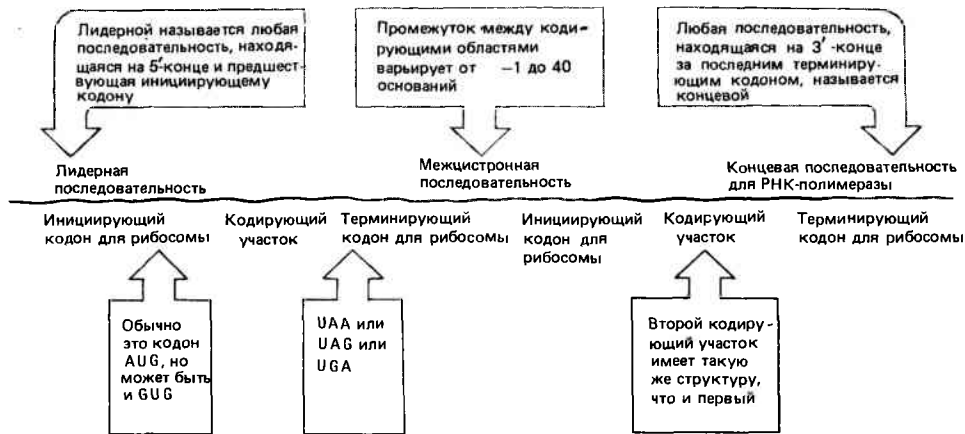


Рис. 9.2. В состав бактериальной мРНК входят как нетранслируемая, так и транслируемая области. Каждая кодирующая область обладает своим собственным инициирующим

и терминирующим сигналами. Обычно мРНК может содержать несколько кодирующих областей.

стей – вплоть до 1 или 2 нуклеотидов, разделяющих терминирующий кодон одного гена от инициирующего кодона следующего. В крайнем случае последовательности двух генов могут практически перекрываться таким образом, что последнее основание терминирующего кодона UGA в конце одного кодирующего участка является одновременно первым основанием инициирующего кодона AUG в начале следующего гена.

Трансляция полицистронной мРНК

Различные кодирующие участки полицистронной мРНК транслируются независимо или же согласованно? Одинаков механизм инициации для всех цистронов или различен, или же инициация первого цистрона отличается от инициации остальных цистронов?

В случае полицистронных бактериальных мРНК очередность событий показывает, что трансляция должна происходить последовательно – цистрон за цистроном. В тот момент, когда рибосомы прикрепляются к первому кодирующему участку, последующая кодирующая область, возможно, еще даже не транскрибировалась. К тому времени, когда второй сайт, связывающий рибосому, только становится доступным, в первом цистроне трансляция уже достигает значительных размеров.

Но что же происходит в межцистронных областях? Это пока еще окончательно не выяснено; возможно, это зависит от конкретной мРНК. Вероятно, в большинстве случаев взаимодействие рибосом с началом каждого цистрона происходит независимо друг от друга. В пользу этого предположения свидетельствуют данные о том, что инициация внутренних цистронов также чувствительна к касугамицину (специфическому ингибитору инициации), как и инициация первого цистрона.

В верхней части рис. 9.3 показан наиболее вероятный ход событий. Трансляция первого цистрона заканчивается обычным образом; при этом рибосомы диссоциируют на субчастицы и покидают мРНК. Затем новая 30S-субчастица должна прикрепиться к следующему инициирующему кодону, соединиться с 50S-субчастицей и начать трансляцию следующего цистрона.

Но эта последовательность событий не говорит о том, что трансляция первого цистрона никак не может влиять

на трансляцию следующего за ним цистрона. Известны случаи, когда мутация в одном гене предотвращает экспрессию другого гена, расположенного за ним в составе той же полицистронной мРНК. Это явление получило название **эффект полярности**. Его основная причина заключается в опосредованном действии мутации на транскрипцию, которая прекращается вскоре после участка, несущего полярную мутацию (см. гл. 13). Однако полярность также может возникать в результате взаимосвязанной трансляции двух цистронов, входящих в состав одной мРНК.

Одна из форм такой взаимосвязи может наблюдаться, когда промежуток между кодирующими последовательностями достаточно мал. Находясь на мРНК, рибосома экранирует около 35 оснований; следовательно, она одновременно может взаимодействовать с терминирующим кодоном и следующим за ним инициирующим кодоном, если они разделены несколькими основаниями. Конечно, такое взаимодействие практически всегда встречается в случае непосредственно соседствующих или перекрывающихся нуклеотидных последовательностей. Такое перекрывание позволяет не замечать некоторые обычно встречающиеся межцистронные сигналы. Например, 30S-субчастица терминирующей рибосомы может не отделяться от мРНК. С большой вероятностью она останется прикрепленной к матрице, так как фактически присоединение к инициирующему сайту уже произошло. Как показано в нижней части рис. 9.3, это означает, что 50S-субчастица и образовавшаяся полипептидная цепь освободятся, а 30S-субчастица должна остаться на месте для реиницирования трансляции следующего цистрона. Не исключено, что существуют ситуации, при которых вся 70S-рибосома остается связанной с матрицей, хотя, по-видимому, это менее вероятно.

Другой вид взаимосвязи между цистронами в полицистронной мРНК опосредован вторичной структурой молекулы. В бактериальной мРНК такая ситуация обычно не встречается, так как рибосомы следуют сразу за РНК-полимеразой и поэтому в области, разделяющей их, комплементарные основания не могут образовать стабильный двухцепочечный участок. Однако, когда рибосомы отделяются от мРНК, дойдя до нонсенс-кодона, расположенного в начале цистрона, последующая

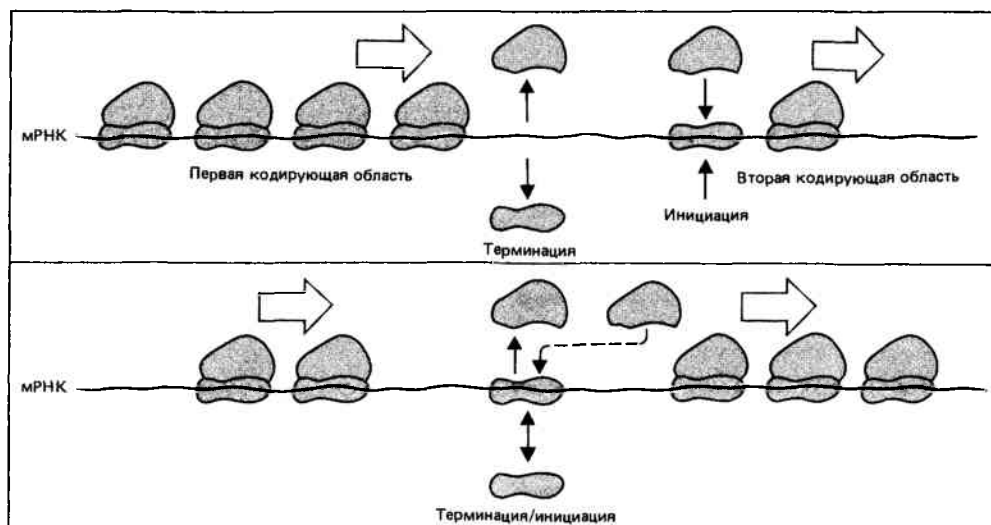


Рис. 9.3. Способ реинициации внутри полицистронных мРНК может зависеть от размеров межцистронных областей.

Вверху. Когда межцистронные промежутки длиннее, чем область, контактирующая с рибосомой, за диссоциацией на конечном сайте независи-

мо происходит инициация на следующем цистроне. Внизу. Если межцистронные промежутки короткие, 30S-субчастица может отделяться на короткое время или даже остаться связанной с мРНК в процессе термации и реинициации.

область мРНК может сформировать вторичную структуру. Поэтому комплементарное взаимодействие какой-либо области с иницирующим кодоном может препятствовать его связыванию с 30S-субчастицей. Возможно, такой механизм обуславливает полярность на уровне трансляции.

Явление той же природы обычно обнаруживается при трансляции фаговых РНК, цистроны которых всегда экспрессируются в определенной последовательности. Причина состоит в том, что фаговая РНК обладает вторичной структурой, позволяющей только одной иницирующей последовательности взаимодействовать с рибосомой. Рибосомы не могут присоединиться к остальным иницирующим последовательностям, так как те образуют пары с другими участками РНК. Однако в процессе трансляции первого цистрона происходит разрушение вторичной структуры, что позволяет рибосомам присоединиться к иницирующему кодону следующего цистрона. Таким образом, в этой мРНК вторичная структура контролирует способность определенных цистронов транслироваться.

Число рибосом, принимающих участие в трансляции определенной мРНК в какой-либо момент времени, зависит от эффективности узнавания иницирующих последовательностей. В случае триптофановых генов *E. coli*, кинетика экспрессии которых изучена хорошо, как правило, в каждый момент времени около 15 молекул РНК-полимеразы находится на транскрипционной единице, состоящей из 7000 пар оснований. С момента транскрипции до момента деградации каждую мРНК, вероятно, успевают протранслировать около 30 рибосом. Поэтому, если каждую минуту происходят пять актов иницирования транскрипции, то это позволяет в данный промежуток времени синтезироваться около 150 молекулам белка, при условии что клетка находится в стационарной фазе роста.

С помощью электронной микроскопии в клетках *E. coli* были обнаружены транскрипционные единицы, изображенные на рис. 9.4. Видно, что одновременно синтезируется несколько молекул мРНК, к которым прикреплены транслирующие рибосомы.

Функциональное определение эукариотической мРНК

Полирибосомы представляют стабильную фракцию эукариотической цитоплазмы, и их можно выделить при центрифугировании клеточных компонентов. Но поскольку на долю мРНК приходится лишь незначительная часть от общей массы РНК, входящей в полирибосомную фракцию (основное количество представлено рибосомной РНК), то ее нельзя выделить, используя только стандартные приемы фракционирования. Попытки избирательно пометить мРНК с помощью радиоактивных предшественников нуклеотидов (подход, использованный применительно к бактериям) не увенчались успехом, так как при этом одновременно метились некоторые виды цитоплазматических РНК, включающие метку так же быстро и эффективно. Все это привело к тому, что выделение эукариотических мРНК было осуществлено позднее, чем выделение бактериальных мРНК.

В основе первого методического приема, позволившего отделить мРНК от других видов РНК, лежала обработка полисом хелатообразующим агентом ЭДТА. Основной эффект этого соединения состоит в связывании ионов Mg^{2+} и в удалении их из состава рибосом. Это вызывает диссоциацию полисом на отдельные рибосомы. В результате освобождается мРНК, которая обнаруживается во фракции с коэффициентом седиментации около 18S. При этом мРНК ассоциирована с белками и находится в форме рибонуклеопротеиновых частиц (мРНКП). Загрязняющий рибонуклеопротеиновый материал, не участвующий в трансляции, не разрушается в присутствии ЭДТА и поэтому по-прежнему седиментирует быстро с коэффициентом седиментации около 200S.

Несмотря на значительные успехи, достигнутые с тех пор, эта методика по-прежнему имеет большое значение по двум причинам. Во-первых, это функциональный тест, позволяющий идентифицировать мРНК, непосредственно транслируемую рибосомами. Поэтому обнаружение мРНК в этой фракции может рассматриваться как дока-



Рис. 9.4. У бактерий визуально можно наблюдать транскрипцию, сопряженную с трансляцией.

ДНК выглядит в виде тонкой линии, проходящей в центре снимка. Отходящие от нее нити — это транскрируемые молекулы мРНК. Приблизитель-

но 15 молекул мРНК синтезируются на одной единице транскрипции. По мере продвижения слева направо, по направлению синтеза, возрастает длина образующихся молекул. Каждая РНК покрыта рибосомами. (Фотография любезно предоставлена Oscar Miller.)

зательство того, что она на самом деле используется для синтеза белка в тех клетках, из которых была выделена. Во-вторых, эта методика, вероятно, позволяет получить мРНК в ее наиболее естественной форме в виде рибонуклеопротеиновых частиц.

Как правило, лишь очень небольшое число белков обнаруживается в составе мРНК. В клетках млекопитающих почти во всех исследованных случаях было обнаружено два основных белка с мол. массой около 52 000 и 78 000 дальтон. Обычно не более двух или трех других белков присутствует в таких же количествах. Остальные белки являются минорными компонентами. В наиболее хорошо изученном случае глобиновой мРНК, полученной из эритроцитов, количество белков с молекулярной массой 52 000 и 78 000 составляет от одного до двух на каждую молекулу мРНК.

Функциональная значимость этих белков непонятна. Они могут принимать участие в транспорте мРНК из ядра в цитоплазму или же оказывать влияние на их трансляцию.

Сопутствующая рибонуклеопротеиновая фракция рассматривается как «загрязняющая» только в том смысле, что она является нежелательным компонентом при выделении полисомной фракции. В действительности РНП, вероятно, представляют собою полноценные клеточные компоненты. Молекулы РНК, входящие в состав этой фракции, обладают свойствами, во многом характерными для мРНК. Но при этом данные молекулы РНК, полученные при обработке РНП частицы ЭДТА, не участвуют в трансляции. Их взаимосвязь с мРНК неясна, но, возможно, они содержат истинные молекулы мРНК (неактивные в данный момент), которые будут вовлечены в трансляцию на более поздних стадиях или при изменившихся условиях.

Некоторые эмбриональные клетки содержат «запасенные» молекулы мРНК. Они существуют в виде рибонуклеопротеиновых частиц, не принимающих участия в трансляции. Но содержащаяся в этих частицах мРНК непосредственно используется для синтеза белка на более поздних стадиях эмбриогенеза. Хотя мы и не знаем, встречается ли аналогичный способ сохранения мРНК в дифференцированных клетках, тем не менее сейчас довольно точно установлено (особенно в случае морских организмов), что значительная часть новообразованной мРНК, синтезированной на ранних стадиях эмбриогенеза,

не транслируется сразу же, а запасается для использования на более поздних стадиях.

Информационная РНК в цитоплазме эукариот относительно стабильна. При измерении ее стабильности обнаруживается несколько дискретных компонентов. Обычно около половины мРНК в культуре клеток млекопитающих имеет период полужизни около 6 ч, тогда как оставшаяся мРНК характеризуется стабильностью, соизмеримой с продолжительностью клеточного цикла, составляющей 24 ч. В дифференцированных клетках, специализированных на синтезе определенных белков, некоторые мРНК могут быть еще более стабильными.

3'-конец эукариотических мРНК может быть полиаденилирован

Большинство эукариотических мРНК содержат на 3'-конце последовательность из полиадениловой кислоты. Этот концевой участок из остатков А принято обозначать как *poly(A)-конец*, а мРНК, имеющую такую структуру, как *poly(A)⁺*.

Poly(A)-последовательность не закодирована в ДНК; она присоединяется к РНК после того, как произойдет транскрипция в ядре. Присоединение poly(A) катализируется специальным ферментом poly(A)-полимеразой, узнающей свободный 3'-ОН конец мРНК и добавляющей к нему еще около 200 остатков А (как подробно изучено на клетках млекопитающих). Этот добавляемый участок состоит исключительно из остатков А; каким образом контролируется его длина, к сожалению, пока не известно.

Когда мРНК только поступает в цитоплазму, ее poly(A)-конец имеет примерно ту же самую длину, что и в ядре. Однако полиаденилированный конец постоянно укорачивается, возможно отдельными частями, в результате эндонуклеолитических разрывов. Таким образом, клеточная популяция мРНК содержит как «новые», так и «старые» молекулы, имеющие относительно длинные и более короткие poly(A)-участки. Однако длина poly(A)-конца в данной молекуле, по-видимому, не влияет ни на ее способность к трансляции, ни на ее стабильность в цитоплазме.

Когда некоторые индивидуальные мРНК инъецировались в овоциты *Xenopus laevis*, было замечено, что ста-

бильность этих молекул увеличивалась с длиной poly(A)-конца; однако в случае других мРНК такой корреляции не наблюдалось. В отношении аденовирусных мРНК было показано, что блокирование добавления к ним poly(A) также уменьшает стабильность этих молекул в клетках-хозяевах. В цитоплазме других клеток (не млекопитающих) протяженность poly(A)-участка может быть несколько короче, но в целом прослеживается та же закономерность: уменьшение длины poly(A) с «возрастом» молекулы. Poly(A) имеется в цитоплазматической РНК всех исследованных эукариот, а также в митохондриальной.

В мРНК клеток млекопитающих poly(A) ассоциирована с одним и тем же белком в 78 000 дальтон, который является преобладающим компонентом мРНК. Из этого, очевидно, следует, что 3'-конец мРНК состоит из poly(A)-участка, связанного примерно с равной массой белка (3'-концевая локализация этого белка в мРНК объясняет, почему этот компонент не создает помех при трансляции рибосом).

После того как было обнаружено существование (A)-последовательностей, первоначально предполагали, что эта структура имеется во всех клеточных мРНК. Хотя poly(A)-содержащая фракция мРНК всегда составляла менее 100%, это несоответствие, как правило, объясняли тем, что в процессе выделения происходит частичная деградация. Один-единственный разрыв в poly(A)-мРНК мог бы привести к образованию 5'-концевой части молекулы, не содержащей poly(A), и такие фрагменты могли бы ошибочно быть приняты за подлинные мРНК. Однако даже при самой тщательной методике выделения мРНК, при которой возможность разрывов сведена к минимуму, все же остается доля молекул, не содержащих poly(A). Обычно эта доля составляет одну треть от общего количества мРНК.

Основной компонент этой poly(A)⁻-фракции представлен молекулами гистоновых мРНК, кодирующих гистоновые белки хромосом. Эти мРНК имеют небольшие размеры и синтезируются только в определенные периоды клеточного цикла. Остальная часть poly(A)⁻-фракции (примерно две трети) идентична фракции poly(A)⁺-молекул как по структурному, так и по функциональному параметрам, отличаясь от нее лишь отсутствием poly(A).

Таким образом, по размерам, стабильности, эффективности трансляции, нуклеоцитоплазматическому транспорту и во всех остальных отношениях оказывается, что poly(A)⁻-фракция не отличается от poly(A)⁺-мРНК.

Разные или одинаковые белки кодируются poly(A)⁺- и poly(A)⁻-мРНК? Гистоновые мРНК специфичны для poly(A)⁻-фракции, поскольку ни одна из гистоновых мРНК обычно не обнаруживается в poly(A)⁺-фракции. Однако другие компоненты poly(A)⁻-фракции значительно перекрываются с молекулами из poly(A)⁺-фракции. Возможно, что все эти мРНК также существуют и в полиаденилированной форме. Таким образом, отдельный ген может быть представлен в транскриптах, часть которых полиаденилирована, а часть — нет. Вероятно, существуют различия и в степени полиаденилирования каждой конкретной молекулы мРНК. Однако в настоящее время пока еще трудно сделать какие-либо общие выводы о значении полиаденилирования (или же его отсутствия). По-видимому, в некоторых случаях наличие poly(A) действительно влияет на стабильность мРНК; других эффектов пока не обнаружено.

Выделение мРНК с использованием poly(A)-конца

Присутствие poly(A) в молекулах мРНК имеет чрезвычайно важное практическое значение. Poly(A)-последовательность мРНК может комплементарно взаимодействовать с олиго (U) или олиго (dT); эта реакция используется для выделения poly(A)-мРНК. Наиболее удобным методом является иммобилизация олиго (U или dT) на твердом носителе, например путем химического присоединения к сефарозе. Когда популяция молекул РНК пропускают через такую колонку, то задерживается только poly(A)⁺-фракция. Эта фракция может быть получена при обработке колонки раствором с низкой ионной силой, в котором разрушаются водородные связи и освобождается РНК, как это схематически изображено на рис. 9.5.

Единственный недостаток данного метода состоит в том, что при этом выделяются все молекулы РНК, содержащие poly(A). Так, например, если используется суммарная клеточная РНК, то на колонке будет задерживаться как ядерная, так и цитоплазматическая poly(A)⁺-РНК. Если же используются полисомы (полученные по обычной методике), то большая часть выделенной poly(A)⁺-РНК представляет собой трансляционно-активные мРНК; однако в эту фракцию попадут и некоторые молекулы РНК из РНК-частиц, обычно загрязняющие препараты полисом, и также содержащие poly(A). Поэтому, когда важно выделить именно активную популяцию мРНК, необходимо применять обработку ЭДТА в качестве функционального теста.

Другой подход, применяемый для очистки мРНК, также основан на использовании poly(A)-конца. Суть данно-

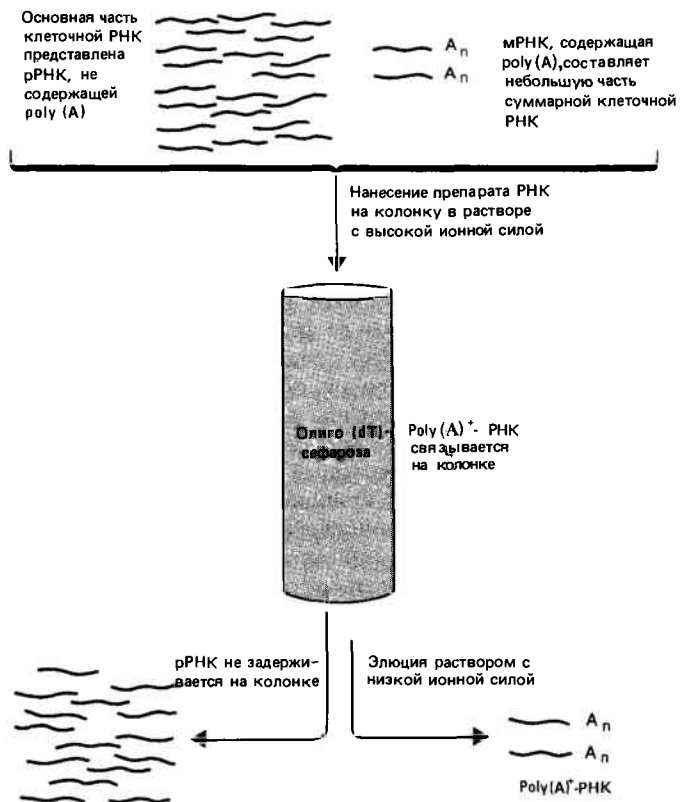


Рис. 9.5. Poly(A)⁺-РНК можно отделить от других видов РНК с помощью фракционирования на oligo(dT)-сефарозе.

го метода состоит в отжиге poly(A)-участка с короткими олиго-(dT)-фрагментами, которые функционируют в качестве стартовой точки, или затравки, для **обратной транскриптазы** (ревертазы), фермента, который копирует мРНК, образуя при этом цепь комплементарной ДНК (обозначенную как **кДНК**). Затем эту кДНК можно использовать в качестве матрицы для синтеза второй цепи ДНК, которая идентична первоначальной последовательности мРНК. Продуктом этих реакций является двухцепочечная ДНК, последовательность которой полностью соответствует последовательности мРНК. Такая ДНК может быть легко «клонирована» с помощью хорошо разработанной технологии (гл. 19).

Любая клонированная ДНК может быть получена в огромных количествах, что дает возможность выделить соответствующую мРНК, используя для этого метод гибридизации. С помощью такого методического приема можно выделить мРНК, представленные всего лишь несколькими копиями в клетке, тогда как непосредственно, без применения техники клонирования, могут быть получены лишь те мРНК, которые присутствуют в относительно больших количествах.

Эукариотические мРНК имеют метилированный «кэп» на 5'-конце

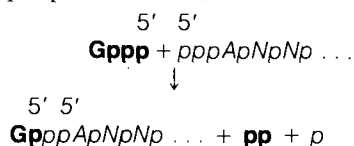
Как и в случае других нуклеиновых кислот, при транскрипции в мРНК включаются только четыре обычных рибонуклеотида. Затем с помощью модифицирующих ферментов в специфические сайты молекулы вводятся дополнительные группы. 5'-конец цитоплазматической мРНК эукариот (но не митохондриальной или хлоропластной) подвергается модификациям двух типов. Возможно, что эти реакции характерны для всех эукариотических клеток.

Транскрипция начинается с нуклеотидтрифосфата (обычно с пурина, А или G). Первый нуклеотид сохраняет 5'-трифосфатную группу и образует обычную фосфодиэфирную связь между своим С-3' и С-5' следующего нуклеотида. Таким образом, изначальная последовательность транскрипта может быть представлена в следующем виде:



Однако если зрелую мРНК обработать *in vitro* ферментами, гидролизующими ее до отдельных нуклеотидов, то ее 5'-конец не образует предполагаемого нуклеозидтрифосфата. Вместо этого он образует два нуклеотида, содержащих метильные группы и связанных между собой *трифосфатной связью* 5'-5'. Концевое основание всегда представлено гуанином, который добавляется к нативной молекуле РНК посттранскрипционно.

Присоединение 5'-концевого G катализируется ядерным ферментом — гуанилилтрансферазой. Эта реакция происходит сразу же после того, как началась транскрипция, так что в ядерной РНК можно обнаружить только следовые количества первоначальных 5'-трифосфатных концов. Суммарная реакция может быть представлена как взаимодействие между GTP и нативным 5' трифосфатным концом РНК:



Реакция может протекать в две стадии. На первой стадии фермент ковалентно связывается с донорным остатком GMP (пирофосфат при этом теряется). Затем GMP присоединяется к РНК, которая при этом теряет γ -фосфатную группу. В результате оказывается, что добавляемый остаток G находится на конце РНК в обратной ориентации по отношению к остальным нуклеотидам.

Эта структура в мРНК получила название «кэп». Она служит субстратом для реакций метилирования, происходящих в определенных положениях и в определенном порядке. Окончательная структура кэпа после всех возможных актов метилирования изображена на **рис. 9.6**.

Первый этап метилирования происходит у всех эукариот; он состоит в добавлении метильной группы в 7-е положение концевого гуанина. Фермент, осуществляющий эту реакцию, гуанил-7-метил-трансфераза, находится в цитоплазме. Как и все остальные метилазы, он использует в качестве кофактора S-аденозилметионин, который служит донором метильных групп. Кэп, имеющий только одну метильную группу, обозначается как **кэп 0**. Такая структура кэпа встречается у одноклеточных эукариот.

Затем происходит добавление еще одной метильной группы к 2'-0-положению предпоследнего основания (т.е. фактически к первому нуклеотиду первоначального, еще не модифицированного транскрипта. Эта реакция катализируется другим ферментом (2'-0-метил-трансферазой). Кэп, который имеет две метильные группы, называется **кэпом 1**. Это основная форма кэпа у всех эукариот, кроме одноклеточных организмов.

В редких случаях у высших эукариот ко второму основанию добавляется еще одна метильная группа. Это происходит только тогда, когда во втором положении находится аденин, и в результате реакции в N'-положение добавляется вторая метильная группа. Фермент, катализирующий эту реакцию, 2'-0-метиладенозин-N⁶-метил-трансфераза, воздействует на аденозиновый субстрат только в том случае, если у него уже имеется метильная группа в положении 2'-0.

У некоторых видов метильная группа может добавляться к третьему основанию кэпированной мРНК. Субстратом для этой реакции служит мРНК с кэпом 1, которая уже имеет две метильные группы. Модификация третьего основания всегда заключается в метилировании 2'-0-положения рибозы. В результате образуется структура, называемая **кэпом 2**. Если это происходит, то кэпы такого типа обычно составляют 10–15% от суммарной кэпированной популяции молекул.

В популяции эукариотических мРНК каждая молекула имеет кэп. Соотношение разных типов кэпов является характерной особенностью данного организма. Остается ли структура молекулы мРНК неизменной, или же в процессе ее функционирования может происходить смена разных типов кэпов? На основе имеющихся в настоящее время данных нельзя еще дать ответа на этот вопрос.

Кроме метилирования при формировании кэпов в мРНК только высших эукариот происходит, хотя и с низкой частотой, метилирование внутренних нуклеотидов. В результате этого процесса образуется модифицированный остаток N⁶-метиладенина, встречающийся примерно один раз на тысячу оснований. мРНК высших эукариот может содержать 1–2 метиладениновых остатка, хотя существование этих нуклеотидов вовсе не обязательно, поскольку в некоторых мРНК (например, глобиновых) их нет совсем.

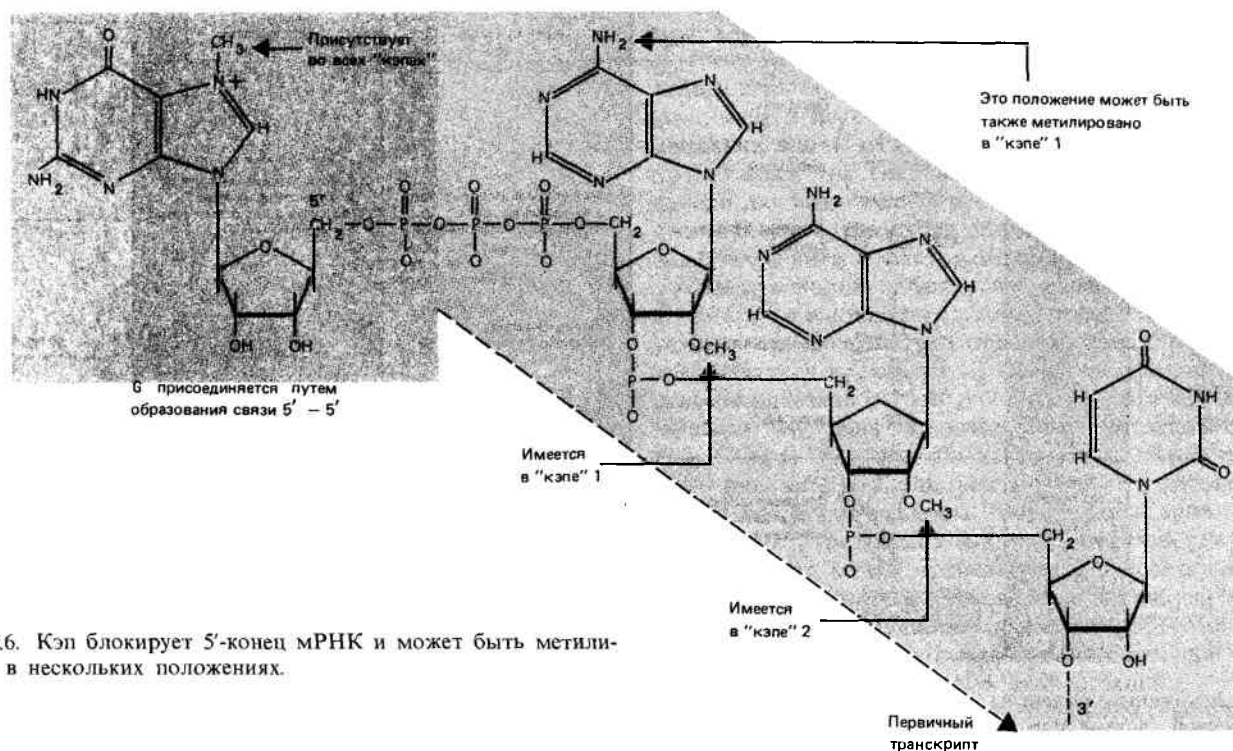


Рис. 9.6. Кэп блокирует 5'-конец мРНК и может быть метилирован в нескольких положениях.

Возможности трансляционных систем *in vitro*

Опыты по транслированию мРНК *in vitro* сыграли решающую роль в изучении процесса трансляции. В-первых, можно получить определенные данные о кодирующих функциях любой конкретной мРНК, доказав, что образующийся продукт обладает теми же свойствами, что и белок, синтезирующийся *in vivo*. (Обычно об этом судят по электрофоретической подвижности.) [Часто это можно проанализировать более детально, установив нуклеотидную последовательность мРНК (как правило, используя для этого ДНК-копию) и обнаружив, что она содержит соответствующую кодирующую область. В ряде случаев, конечно, используя нуклеотидную последовательность мРНК, можно предсказать аминокислотную последовательность неизвестного белка.]

Системы *in vitro* (как в случае транскрипции, так и в случае трансляции) дают уникальную возможность исследовать механизм самого процесса. В обоих случаях определяющим этапом является инициация. С помощью трансляционной системы можно исследовать характерные особенности рибосом и мРНК, необходимые для их взаимодействия.

Хотя в общих чертах трансляционная система *in vitro* во многом напоминает процесс, происходящий *in vivo*, известно два случая, когда различия между продуктами трансляции *in vitro* и *in vivo* дают нам интересную информацию для размышлений.

Иногда бывает возможно выделить мРНК, которые транслируются *in vitro*, хотя соответствующие белки не синтезируются в тех клетках, из которых данные мРНК были получены. Способность мРНК транслироваться *in vitro* доказывает, что в принципе эти молекулы могут функционировать в качестве матрицы. Следовательно, неспособность таких РНК функционировать *in vivo* объясняется, очевидно, наличием контроля на уровне трансля-

ции. Должно быть, *in vivo* функционирует некий механизм, препятствующий трансляции. «Законсервированные» РНК-частицы эмбрионов морских животных, о которых мы упоминали ранее, — наиболее хорошо изученный пример такого явления. (Конечно, такое заключение правомерно лишь в том случае, если синтезирующийся *in vitro* продукт является полноценным белком, а не артефактом, возникшим в результате неправильной трансляции. Поэтому необходимо продемонстрировать, что при других условиях данные белки действительно синтезируются *in vivo*.)

Иногда продукт трансляции, образующийся *in vitro*, родствен подлинному белку, синтезирующемуся *in vivo*, но содержит на N-конце небольшую последовательность дополнительных аминокислот. Эти различия объясняются тем, что белок, обнаруживаемый *in vivo*, не является первичным продуктом трансляции, а появляется в результате расщепления предшественника, содержащего дополнительные аминокислотные остатки. Обычно *in vivo* процессинг такого предшественника происходит быстро, и поэтому их трудно обнаружить, если не применять ингибиторов, блокирующих реакцию расщепления.

Очищенные молекулы мРНК исследуют с использованием трансляционных систем двух типов. Они изображены на рис. 9.7.

В состав реконструированных бесклеточных систем входят рибосомы, факторы белкового синтеза и тРНК. Это традиционный подход, который мы уже описывали. Существует несколько таких систем: наиболее известны системы, получаемые из проростков пшеницы, ретикулоцитов кролика и асцитных клеток мыши. Все они в определенной степени неэффективны, так как каждая мРНК транслируется меньшее количество раз и со значительно меньшей скоростью, чем *in vivo*. Наилучшие системы функционируют на протяжении 90–120 мин. Во всех случаях обнаруживается некоторый остаточный уровень трансляции, который обусловлен неудаленной эндогенной мРНК и который у разных систем может варьировать.



		Преимущества	Недостатки
Бесклеточные системы	 <p>Содержат:</p> <ul style="list-style-type: none"> рибосомы факторы белкового синтеза ак аминоацил-тРНК АТФ энергию 	<p>Может быть использована для изучения роли отдельных компонентов</p> <p>Легко очистить синтезированные белки</p> <p>Возможно выделить белки-предшественники</p>	<p>Работает более медленно и менее эффективно (по сравнению с синтезом белка <i>in vivo</i>)</p>
Ооциты млекопитающих	 <p>Введение мРНК с помощью микропипетки</p> <p>Ооцит</p>	<p>Очень эффективна, работает так же хорошо, как и <i>in vivo</i></p>	<p>Более трудно очистить синтезированный белок</p>

Рис. 9.7. Экзогенные молекулы мРНК можно транслировать в бесклеточной системе или инъецируя их в ооциты *X. laevis*.

Другой трансляционной системой являются интактные ооциты африканской шпорцевой лягушки *Xenopus laevis*. Инъецированные молекулы мРНК транслируются в них природным белоксинтезирующим аппаратом. Система эффективна и использует инъецированные молекулы мРНК так, как если бы они были эндогенными. Поэтому они вовлекаются в повторные циклы трансляции. Единственное ограничение состоит в том, что слишком большое количество мРНК может насытить трансляционную систему (которая, конечно, находится в большом избытке по отношению к эндогенным молекулам мРНК). Как правило, система продолжает быть активной в течение 24–48 часов.

В рассмотренных системах не обнаруживается тканевой или видовой специфичности. Это указывает на то, что мРНК и белоксинтезирующий аппарат (возможно) любой цитоплазмы взаимозаменяемы. Следовательно, можно сделать вывод, что в данных системах не существует контроля на уровне трансляции со стороны каких-то рибосомных факторов, которые были бы способны функционировать с одним типом молекул мРНК, а не с другим. Трансляционный аппарат заранее не запрограммирован, а способен использовать любую мРНК в качестве матрицы. Таким образом, любой контроль на уровне трансляции должен реализоваться в форме, препятствующей молекулам мРНК взаимодействовать с белоксинтезирующим аппаратом.

Наиболее вероятно, это достигается изоляцией молекул мРНК путем их перевода в форму, физически недоступную для белкового синтеза.

Белоксинтезирующая система ооцитов способна осуществлять процессинг по крайней мере некоторых белков, причем сопряженное с процессингом расщепление происходит либо еще во время синтеза белка, либо вскоре после его окончания. В ряде случаев эта система даже может обеспечить проникновение белковых продуктов в необходимый компартмент клетки. Таким образом, сиг-

налы процессинга могут быть общими для различных типов и видов клетки. В качестве практической рекомендации можно отметить, что получение белков-предшественников должно осуществляться в бесклеточной системе трансляции.

Для инициации, по-видимому, необходимо комплементарное взаимодействие между мРНК и рРНК

Участки мРНК, взаимодействующие с рибосомами при инициации белкового синтеза, могут быть идентифицированы по образованию стабильного комплекса между рибосомами и мРНК в условиях, препятствующих элонгации полипептидной цепи. При добавлении рибонуклеазы к иницирующему комплексу вся мРНК, находящаяся за пределами рибосомы, деградирует, и остается только участок, экранированный частицей (рис. 9.8). Защищенный фрагмент можно выделить и исследовать. Такие эксперименты были выполнены с многими очищенными мРНК как бактериального, так и эукариотического происхождения.

Бактериальные рибосомы экранируют иницирующую последовательность размером 35–40 оснований. В состав этой последовательности всегда входит иницирующий кодон AUC (или GUG), располагающийся на расстоянии двух третей от начала защищенного фрагмента.

Наблюдается очень незначительная гомология между последовательностями, образующими участки связывания рибосом у различных бактериальных мРНК. Обнаруживается лишь очень короткая общая последовательность, комплементарная определенной области, расположенной недалеко от 3'-конца 16S-рРНК.

3'-концевой участок бактериальной 16S-рРНК высококонсервативен. Так, он почти идентичен с 3'-концом со-

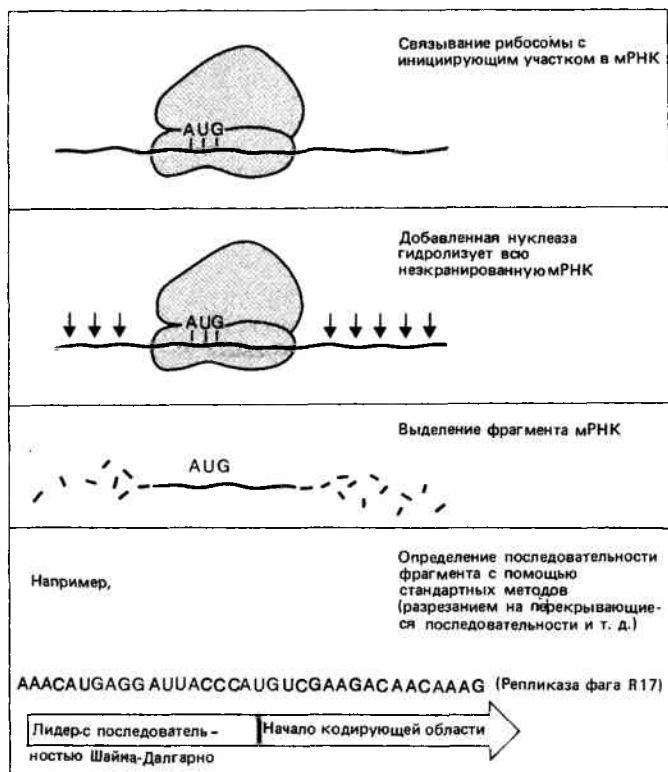


Рис. 9.8. Участок мРНК, связывающий рибосому, может быть получен из состава иницирующего комплекса.

ответствующей рРНК из хлоропластов кукурузы. 3'-конец рРНК обладает двумя особенностями, которые могут иметь важное значение для синтеза белка. Во-первых, последовательность рассматриваемого участка является самокомплементарной и может образовывать шпильку в результате комплементарного взаимодействия оснований (рис. 9.9). Во-вторых, этот участок содержит специально обозначенную на рисунке последовательность, которая, будучи записанной в обратном порядке, выглядит как

3'...UCCUCC...5'.

За одним лишь исключением, в сайтах инициации всех известных мРНК *E. coli* существует участок, комплементарный по крайней мере трем нуклеотидам этой области, а чаще—4–5 нуклеотидам. Таким образом, бактериальная мРНК содержит или часть, или весь олигонуклеотид

5'...AGGAGG...3'.

Этот полипуриновый участок часто называют последовательностью Шайна–Дальгарно. Она располагается на расстоянии 4–7 оснований перед кодоном AUG. Происходит ли взаимодействие между последовательностью Шайна–Дальгарно и комплементарной ей областью в рРНК в процессе связывания матрицы с рибосомой?

Комплементарная последовательность в рРНК не может одновременно быть частью внутримолекулярной шпильки и взаимодействовать с мРНК. Но эти два варианта спаривания могут существовать как альтернативные. Тогда иницирование может включать в себя разрушение концевой шпильки, и комплементарное спаривание между рРНК и мРНК станет в результате воз-

можным. Вслед за этим дуплекс между рРНК и мРНК может быть нарушен, а структура шпильки восстановится. Этот механизм может привести в соответствие необходимость образования стабильного иницирующего комплекса и потребность рибосомы в дальнейшем продвижении вдоль мРНК.

Тот факт, что 3'-концевая область рРНК является мишенью для антибиотика касугамицина, говорит о ее участии в синтезе белка. Эта область также может быть ковалентно сшита с факторами инициации. Основным же доводом в пользу участия последовательности Шайна–Дальгарно в инициации служит тот факт, что эта последовательность повсеместно обнаруживается в иницирующих сайтах у прокариот. В одном случае это непосредственно доказано (мРНК для гена 0,3 фага T7). При этом мутация, ведущая к изменению последовательности GAGC на GAAG, дестабилизирует связывание рибосомы с фаговой мРНК. Единственная мРНК, в которой отсутствует последовательность Шайна–Дальгарно,—это лямбовский P_m -транскрипт фага λ , несущий на 5'-конце рррAUG. Эта мРНК транслируется относительно плохо. Возможно, необычное местоположение кодона AUG—непосредственно на 5'-конце—позволяет рибосоме узнавать его. Однако не известно, сам ли иницирующий кодон узнается малой субчастицей в этой реакции.

При связывании с мРНК 40S-субчастица экранирует область размером до 60 оснований. Когда 60S-субчастица присоединяется к комплексу, размеры защищенной области несколько сокращаются. Участок прочного связывания для 80S-рибосом, как и у прокариот, захватывает область размером в 30–40 оснований. Кодон AUG обычно располагается в центре этого отрезка. Уменьшение размеров экранируемой области, 80S-рибосомами можно объяснить конформационными изменениями в малой субчастице, возникающими при ее взаимодействии с большой субчастицей. С другой стороны, это может быть связано с потерей факторов инициации, способных самостоятельно экранировать некоторую область мРНК.

Сравнение 3'-концевой последовательности рРНК у прокариот и эукариот показывает, что эта последовательность высококонсервативна (рис. 9.10). В последовательности из 20 нуклеотидов, расположенной на отрезке между соседними дважды метилированными аденинами и 3'-концом, имеется только два значительных различия. У бактерий в определенном положении имеется два остатка U, тогда как у высших эукариот здесь же находятся два остатка A (в случае низших эукариот обнаруживается промежуточная по составу нуклеотидов последовательность AU). Кроме того, у всех эукариот обнаруживается делеция последовательности CCUCC длиной пять пар оснований, которая и является областью, комплементарной последовательности Шайна–Дальгарно. (Это означает, что если связывание с последовательностью Шайна–Дальгарно необходимо для инициации, то эукариотические рибосомы не способны иницировать трансляцию на бактериальных мРНК).

Внутри высококонсервативной 3'-концевой области (имеющейся как в бактериальной, так и в эукариотической рРНК) находится последовательность, обогащенная пуринами,

3'...UAGGAAGGCGU...5',

которая у эукариот удалена от 3'-конца приблизительно на то же расстояние, что и последовательность CCUCC

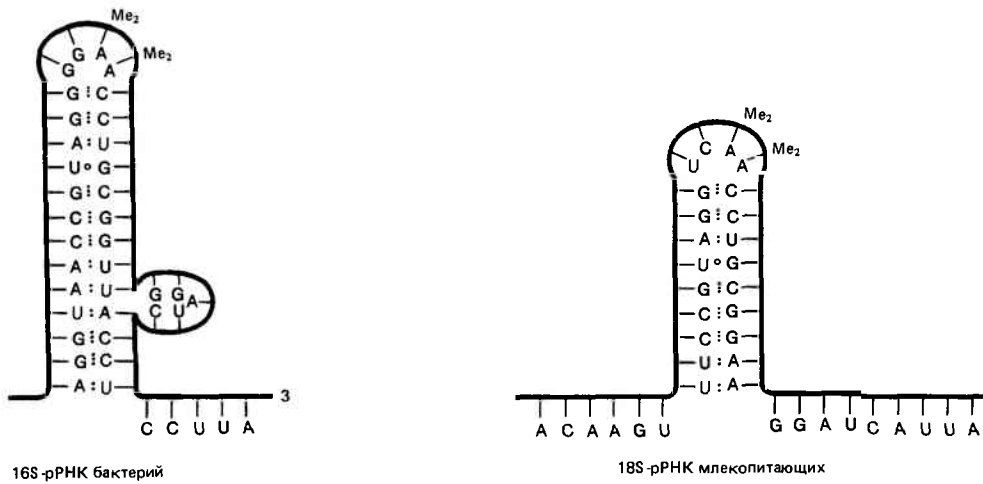


Рис. 9.9. На 3'-конце рРНК может быть образована структура типа «шпильки» в результате спаривания комплементарных оснований.

у бактерий. С другой стороны, эта последовательность может принимать участие в образовании шпильки, как показано на рис. 9.9. Некоторые эукариотические мРНК содержат четырех- или пятичленные последовательности нуклеотидов, теоретически способные образовывать пары с этой частью 3'-концевой последовательности рРНК (при условии, что шпилька в рРНК разрушена). Однако удаленность этого возможного сайта взаимодействия от кодона AUG варьирует. Часто этот сайт располагается так далеко от кодона AUG, что они не могут быть одновременно покрыты рибосомой. Нам не известно, принимает ли эта последовательность участие в иницировании, но если даже она и присутствует, то ее использование является необязательным, так как некоторые эукариотические мРНК полностью лишены ее.

Известен один случай, для которого совершенно точно установлено, что механизм инициации не включает в себя комплементарное взаимодействие между мРНК и 3'-концевой областью рРНК. В митохондриях млекопитающих 3'-концевая область 12–13S-рРНК имеет сходство с аналогичными участками рРНК из бактерий и цитоплазмы эукариот, но характеризуется меньшей степенью консервативности. Способность образовывать концевую шпильку сохраняется, однако последовательности, комплементарной какой-либо области, расположенной перед кодоном AUG в мРНК, не обнаружено.

Малые субъединицы могут перемещаться в сайты инициации эукариотических мРНК

Средние размеры мРНК в цитоплазме эукариотических клеток составляют от 1000 до 2000 оснований; у нее имеется метилированный «кэп» на 5'-конце и poly(A)-последовательность длиной 100–200 оснований – на 3'-конце. При сравнении размеров индивидуальных мРНК с размерами тех белков, которые они кодируют, оказывается, что последовательность мРНК всегда длиннее, чем это необходимо для кодирования соответствующего белка. Однако эта избыточная длина предназначается отнюдь не для кодирования какого-то второго белка, поскольку каждая мРНК моноцистронна. Она состоит из короткой (обычно не более 100 нуклеотидов) лидерной последовательности

Она содержит участок, который может играть роль в распознавании мРНК (показан черной линией).

одного кодирующего участка и концевой некодирующего участка (который во многих случаях имеет значительные размеры и иногда может достигать 1000 оснований). В силу своей 5'-концевой локализации лидерная последовательность оказывается вовлеченной в инициацию. Что же касается еще 3'-концевого участка, то пока нам не известно ни одной функции, которую бы он выполнял в трансляции.

Рибосомы в цитоплазме эукариотических клеток не связываются непосредственно с сайтом инициации перед началом кодирующей области. Вместо этого первой опознаваемой структурой является метилированный кэп на 5'-конце. Некоторые молекулы мРНК не имеют кэпов – либо потому, что образование этих структур было заблокировано, либо в результате их ферментативного удаления из мРНК. Такие лишенные кэпов мРНК не транслируются с достаточной степенью эффективности в системах *in vitro*. За исключением всего лишь нескольких вирусных мРНК, не имеющих кэпов, все остальные мРНК в цитоплазме эукариотических клеток (но не в органеллах) содержат на 5'-конце такую модификацию. Только эти особые вирусные мРНК могут транслироваться *in vitro*, не имея кэпов. Следовательно, можно думать, что большей части мРНК кэп необходим для трансляции и что только вирусные мРНК, вероятно, имеют какие-то специфические особенности, позволяющие им обходиться без этих структур.

Кэп узнается рибосомной 40S-субчастицей. Какие его структуры обеспечивают это узнавание? Существенное значение имеет введение первой метильной группы в положение 7 концевого G. Последующие акты метилирования (переводящие кэп 0 в кэп 2), возможно, улучшают эффективность связывания с рибосомой, но не имеют первостепенного значения. 40S-субчастица может связываться с poly(U) или другими синтетическими полинуклеотидами, к 5'-концу которых присоединен кэп. Это, очевидно, свидетельствует о том, что именно кэп (а не следующая за ним последовательность) играет основную роль в процессе распознавания; однако для того, чтобы 60S-субчастица могла присоединиться к иницирующему комплексу, необходима информация, закодированная в последовательности самой мРНК.

В некоторых случаях иницирующий кодон AUG расположен всего лишь на расстоянии 40 нуклеотидов от

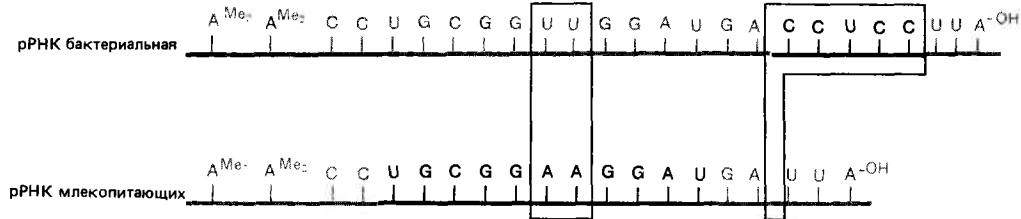


Рис. 9.10. 3'-конец малой рРНК бактерий и млекопитающих консервативен.

Различающиеся области обведены рамкой, а области, которые могут участвовать в распознавании мРНК, изображены красным цветом.

5'-конца мРНК; таким образом, и кэп, и сайт AUG находятся в пределах участка связывания рибосомы. Однако известно, что в некоторых мРНК кэп и кодон AUG локализованы довольно далеко друг от друга — на расстоянии 200–300 оснований. Все же даже в этих случаях, для того чтобы у иницирующего кодона мог образоваться стабильный комплекс, необходимо присутствие кэпа. Однако совершенно не ясно, каким образом рибосома может одновременно взаимодействовать с этими двумя участками, так далеко расположенными друг от друга?

Одно из возможных объяснений состоит в том, что мРНК в результате комплементарных взаимодействий между основаниями приобретает такую вторичную структуру, при которой кэп и кодон AUG могут располагаться рядом и одновременно контактировать с субчастицей рибосомы. Однако данных, подтверждающих эту точку зрения, нет.

Согласно другой модели, 40S-субчастица сначала узнает 5'-концевой кэп и затем «передвигается» вдоль мРНК, пока не встретит иницирующий кодон AUG (см. рис. 6.7). Обычно, хотя и не всегда, это будет первый триплет AUG, который она встретит на своем пути. После этого связь стабилизируется и к 40S-субчастице присоединяется 60S-субчастица в том участке, который выявляется как защищенный рибосомой. В случаях длинной лидерной последовательности можно предположить, что вторая 40S-субчастица узнает 5'-конец еще до того, как первая покинула инициаторный участок; в результате на этом отрезке могла бы выстроиться целая очередь субъединиц.

Идея о том, что рибосомы должны начинать свое движение по матрице с 5'-конца, согласуется с моноцистронной природой эукариотических мРНК. Существуют примеры особых вирусных мРНК, содержащих более одной кодирующей области. Однако в этих случаях из всей молекулы мРНК транслируется только та кодирующая область, которая расположена ближе всех к 5'-концу. Остальные же могут быть прочитаны только после того, как в мРНК произошел разрыв, который привел к образованию нового 5'-конца, находящегося вблизи следующего инициаторного кодона. Это свидетельствует в пользу тех представлений, согласно которым внутренние сайты инициации не узнаются непосредственно в составе полицистронной мРНК.

Связь белкового синтеза с внутриклеточной локализацией

Эукариотическая клетка — высокоупорядоченная структура, все функции которой имеют специфическую локализацию. Это общее правило, в частности, относится и к белковому синтезу. Полирибосомы можно подразде-

лить на два класса — свободные и связанные с мембраной. Они участвуют в синтезе разных групп белков.

Название «свободные полирибосомы» не совсем точно отражает истинную ситуацию, так как на деле они не могут свободно диффундировать в цитоплазме, поскольку ассоциированы с клеточным **цитоскелетом**. Если клетки обработать неионным детергентом (тритоном) в гипертоническом буфере, то большая часть липидов и растворимых белков удаляется, оставляя «цитоскелетный каркас», связанный с остатками ядра. Этот каркас представляет собой сложное сплетение фибрилл. Все «свободные» полирибосомы ассоциированы с цитоскелетом, как это можно увидеть на электронной микрофотографии, приведенной на **рис. 9.11**, или определить при помощи биохимического фракционирования.

Полисомы имеют тенденцию располагаться цепочкой поблизости от ядра, в тех местах, где мРНК входит в цитоплазму. Большая часть продуктов трансляции представляет собой растворимые белки, которые сразу же после освобождения из полирибосом быстро диффундируют от места синтеза. Те белки, которые являются компонентами клеточного цитоскелета, имеют обыкновение включаться в его состав в участках, расположенных неподалеку от места своего синтеза.

В противоположность полирибосомам мономерные рибосомы находятся в клетке в свободном состоянии. Связывание с клеточным цитоскелетом, по-видимому, является функциональным свойством мРНК, которая остается прикрепленной к нему даже после обработки

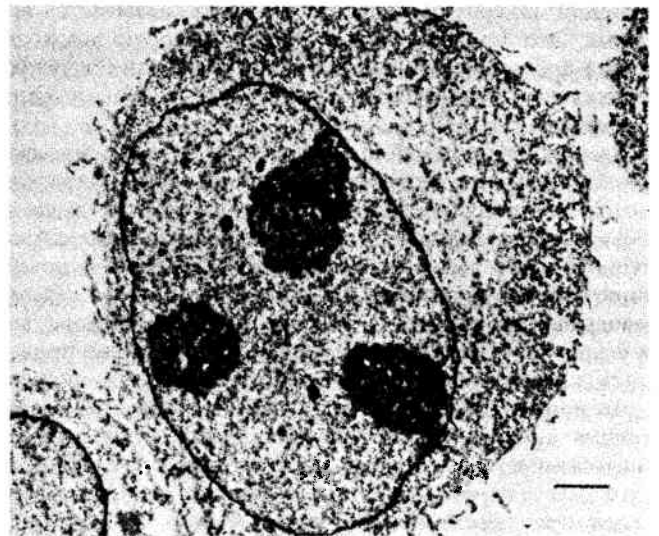


Рис. 9.11. Цитоскелет представляет собой сплетение фибрилл, ассоциированных с полирибосомами.

агентами, дезагрегирующими полирибосомы. Трансляция мРНК, возможно, зависит от ее ассоциации с цитоскелетом, что наиболее отчетливо проявляется при вирусной инфекции. Поливирусная инфекция ингибирует трансляцию в клетке-хозяине, причем ингибирование трансляции сопровождается организационной перестройкой цитоскелета и освобождением клеточной мРНК. В случае инфекции вирусом везикулярного стоматита ассоциация вирусных мРНК с клеточным цитоскелетом происходит вскоре после их синтеза; они транслируются только в течение этого периода, и освобождаются из цитоскелета.

Существование систем трансляции мРНК *in vitro* показывает, что процесс белкового синтеза может происходить независимо от остальных клеточных структур. Однако это отнюдь не умаляет значения данных о том, что при естественном ходе событий мРНК должна быть во время трансляции ассоциирована с цитоскелетом. Это, возможно, имеет важное значение как для метаболизма мРНК *in vivo*, так и для правильной локализации новосинтезированных цитоскелетных белков.

Класс «свободных» полирибосом в основном включает все те полирибосомы, которые не связаны с мембранами. Они ответственны за синтез тех белков, которые не взаимодействуют с мембранами. Другая группа белков имеет специфическую локализацию внутри клетки или на ее наружной поверхности — в зависимости от того, способны ли они взаимодействовать с мембраной или беспрепятственно проходить через нее. Такие белки синтезируются полирибосомами, связанными с мембранами.

Некоторые белки находятся в обособленных клеточных органеллах (компартаментах), таких, как митохондрии или лизосомы. Некоторые являются компонентами мембран. Среди них наибольший интерес представляют те, которые расположены в плазматической мембране (окружающей цитоплазму). Другие белки **секретируются** из клетки в окружающую среду. Каким же образом эти белки находят конечный пункт своего назначения?

Для многих мембранных белков первичная последовательность зрелого полипептида сама по себе недостаточна для встраивания белка в мембрану. Для этого необходима дополнительная информация, которая в большинстве случаев представлена в форме **лидерной последовательности**, локализованной на N-конце белка. Белок, имеющий лидерную последовательность, называется **пре-белком**. Это промежуточная стадия, которую проходит белок в процессе созревания, так как лидерная последовательность отщепляется в процессе встраивания белка в мембрану.

Пре-последовательность отличается от так называемой **про-последовательности**, под которой подразумеваются те дополнительные участки, которые имеются в белках, существующих в виде *стабильных* предшественников. В некоторых белках могут иметься и те и другие. Например, инсулин первоначально синтезируется как **пре-проинсулин**; пре-последовательность отщепляется во время секреции, образуя **проинсулин**, который далее подвергается процессингу с превращением в зрелый инсулин.

Лидерная последовательность выполняет разные функции при различных условиях. Белки, образующиеся в цитоплазме, но предназначенные для функционирования в хлоропластах или митохондриях, синтезируются в виде предшественников с мол. массой ~ 5000 дальтон (около 45 аминокислот), превышающей молекулярную массу зрелого продукта. В таком виде предшественник выходит из полисом. Будучи добавлен *in vitro*, он мо-

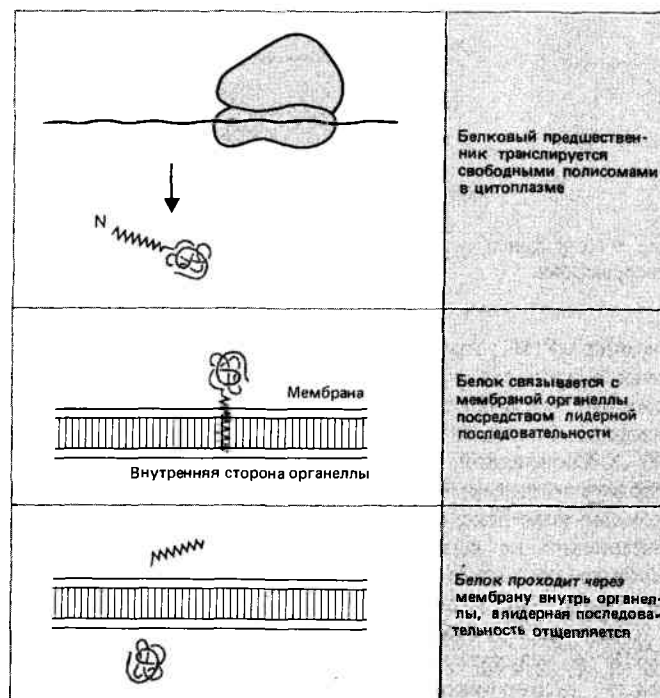


Рис. 9.12. Лидерные последовательности используются белками для узнавания поверхностей митохондрий и хлоропластов.

жет включаться в интактные органеллы. При этом он должен пройти через мембрану органеллы (рис. 9.12) (процесс, во время которого лидерная последовательность белка отщепляется, по всей видимости, под действием протеиназ, локализованных на наружной стороне мембраны). Лидерная последовательность обеспечивает информацию, необходимую для узнавания белка мембранной органеллы, и служит чем-то вроде проводника **после трансляции**. Следует заметить, что отщепляемый лидер не является единственной приемлемой формой информации такого рода; некоторые митохондриальные белки распознаются как таковые уже в зрелой форме и могут иметь внутреннюю последовательность, обеспечивающую прохождение через мембрану без какого-либо расщепления.

Для тех белков, которые секретируются или встраиваются в другие клеточные мембраны, процесс связывания с мембраной в большинстве случаев начинается еще во время трансляции. Полирибосомы, синтезирующие такие белки, ассоциированы с мембраной эндоплазматического ретикулума. Пребелки не освобождаются в цитоплазму и не образуют там пула предшественников, а направляются из рибосомы непосредственно в мембрану. Из мембраны они поступают в аппарат Гольджи, а затем в конечные пункты назначения, например в лизосому или плазматическую мембрану.

Модель, объясняющая механизм встраивания в мембрану, была предложена на основе работ с эукариотическими системами микросом (содержащими рибосомы и эндоплазматический ретикулум). Эти системы способны упаковывать новосинтезированные белки в мембраны, но не функционируют в случае добавления уже выделенного препротейна. В свое время была выдвинута **гипотеза сигнальной последовательности**, согласно которой наличие лидерной последовательности почти во всех секретрируемых белках служит своего рода **сигналом**, присутствие

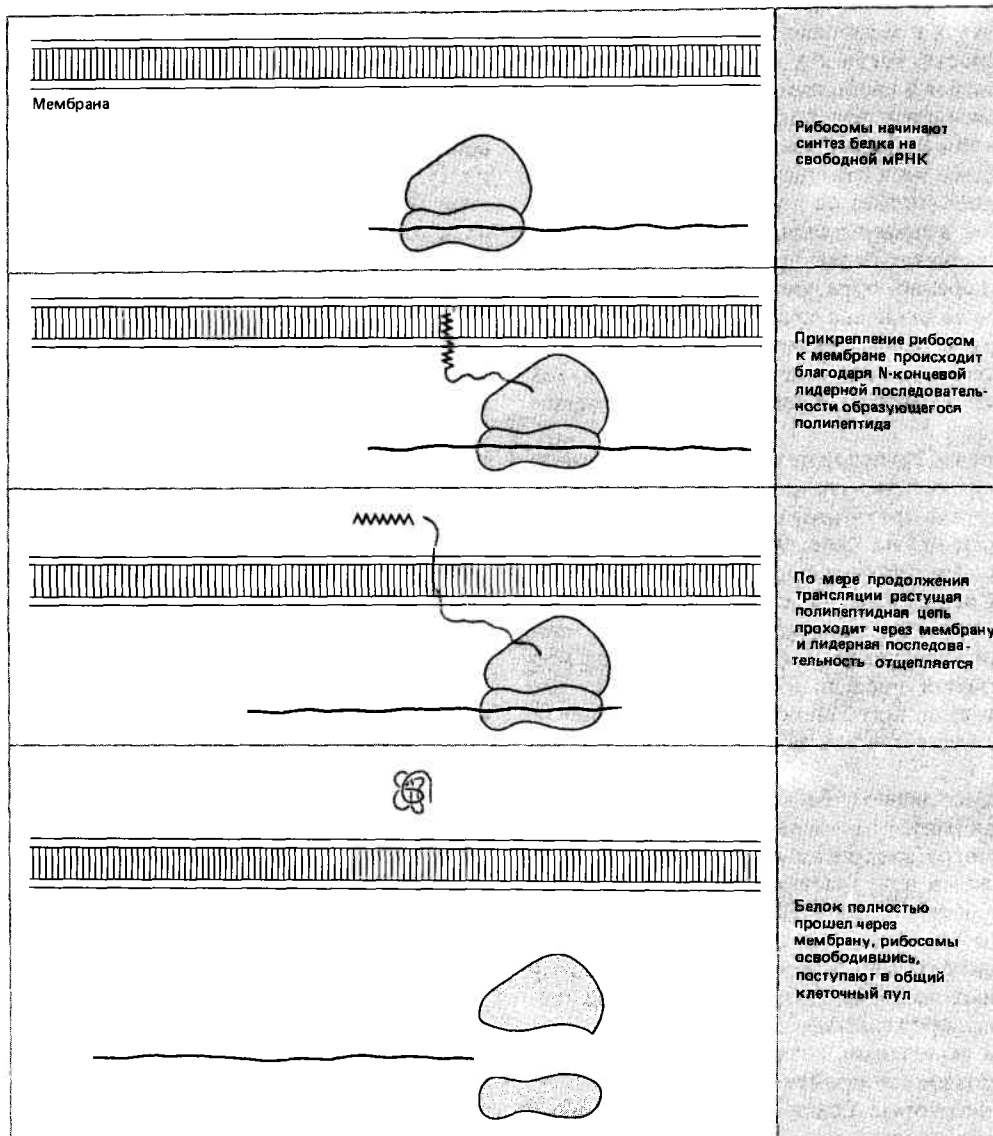


Рис. 9.13. Согласно гипотезе сигнальной последовательности, рибосомы, синтезирующие секреторные белки, прикрепляются

к мембране с помощью лидерной последовательности образующегося полипептида.

которого отличает их от других белков. За некоторыми редкими исключениями, N-концы секретруемых белков состоят из отщепляемой лидерной последовательности, имеющей от 16 до 29 аминокислот. Первые два-три аминокислотных остатка лидерной последовательности являются полярными, хотя в остальной своей части она отличается высоким содержанием гидрофобных аминокислот. Кроме упомянутых особенностей, в первичной структуре лидерной последовательности не обнаружено какой-либо консервативности.

Сигнальная последовательность обеспечивает прикрепление комплекса, образованного из рибосом, мРНК и синтезирующейся полипептидной цепи к мембране. Предполагается, что мембранный рецептор узнает сигнальную последовательность по ее гидрофобности и встраивает белковый предшественник непосредственно в мембрану, — возможно, сразу же после того, как синтезировался фрагмент, состоящий из этой последовательности и еще нескольких аминокислот. Исследования рецепторных белков начались после того, как было обнаруже-

но, что промытые солевым раствором мембраны не способны удерживать рибосомы. Однако это свойство можно восстановить, добавив к ним использованный для промывки солевой раствор. Очищенный активный компонент представляет собой комплекс, состоящий из шести белков.

По мере того как продолжается синтез полипептидной цепи, наступает момент, когда белок уже достаточно хорошо встроен в мембрану и сигнальная последовательность может быть удалена (рис. 9.13). К тому времени, когда рибосома заканчивает трансляцию, белок уже в значительной степени транспортирован через мембрану. Белковые предшественники никогда не находятся в свободном состоянии в клетке.

Только ли сигнальная последовательность отвечает за прикрепление рибосом к мембране? В этом случае синтез всех белков должен был бы начинаться одинаково, а именно путем связывания рибосом со свободной мРНК. Затем после начала синтеза те рибосомы, которые транслируют мРНК, кодирующие секретруемые белки,

должны прикрепляться к мембране при помощи сигнальной последовательности, тогда как другие полирибосомы останутся в цитоплазме в свободном состоянии. По-видимому, этим и объясняется тот факт, что всегда трудно выделить две фракции рибосом—свободную и связанную с мембранами. Однако ни в отношении рибосом, ни в отношении мРНК окончательно не доказано, что они абсолютно не участвуют в прикреплении полисом к мембране. Во всех случаях синтез белка происходит поблизости от мембраны, и перенос образующейся молекулы начинается, когда еще не окончена трансляция—явление, называемое **транспортом, сопряженным с трансляцией**. Оно характерно для различных секретируемых белков, в том числе для некоторых иммуноглобулинов и многих гормонов.

Каков же механизм транспорта белков через мембрану? Первоначально считали, что за этот процесс ответствен N-конец, который протягивает белок через мембрану, как это показано на рис. 9.13. Однако вполне возможно, что N-конец, проникая в мембрану, не проходит ее насквозь. Сигнальная последовательность могла бы, например, «заякориться» в месте внедрения, а белок при этом постепенно втягивался бы в мембрану. Во всех случаях для сигнальной последовательности характерно то, что она может выполнять свою функцию только во время трансляции, а не как часть образовавшегося белка-предшественника.

Какие свойства отличают белки, секретируемые из клетки через мембрану, от белков, *встраивающихся* в мембрану? Механизм внедрения в мембрану, по-видимому, один и тот же—в нем участвует сигнальная последовательность. Но белки, которые не секретируются наружу, обладают еще одним внутренним сигналом, **останавливающим их транспорт**. Он может представлять собой кластер гидрофобных аминокислот, расположенных рядом с несколькими заряженными остатками. Этот кластер служит как бы «крючком», который зацепляет белок в мембране и не позволяет пройти через нее.

Известно, что некоторые белки не имеют лидерных последовательностей, но тем не менее встраиваются в мембраны. Вероятно, какая-то другая последовательность обеспечивает этот перенос, происходящий одновременно с трансляцией, хотя в данном случае отщепления сигнальной последовательности не происходит.

Гипотеза сигнальной последовательности, первоначально предложенная для эукариотических клеток, применима и к бактериям. Мутации в N-концевой лидерной последовательности могут нарушить секрецию. Однако их можно супрессировать мутациями в других генах, в число которых входит по крайней мере один из генов рибосомных белков. Таким образом, сама рибосома каким-то образом участвует в механизме прикрепления к мембране. Это означает, что сигнальная последовательность взаимодействует не только с мембраной и что при этом должны существовать другие белок-белковые контакты с рибосомой. Как и в случае эукариот, у бактерий могут существовать дополнительные сигналы, необходимые для правильной локализации белка после его проникновения в мембрану. Например, C-концевая область β-лактамазы *E. coli* нужна для того, чтобы белок вышел из мембраны и поступил в переплазматическое пространство.

В ряде случаев показано, что секреция у бактерий может происходить посттрансляционно. Наиболее хоро-

шо изученным примером такого рода является белок оболочки фага M13. Он синтезируется в форме предшественника, способного встраиваться в мембрану. В состав предшественника входит отщепляемая лидерная последовательность. Возможная роль этой последовательности постулируется **мембранно-триггерной гипотезой**. Согласно этой модели, лидерная последовательность должна изменять характер упаковки белка в третичную структуру. В присутствии лидерной последовательности полипептид образует конформацию, характерную для водных растворов, которая благоприятствует его существованию в цитоплазме, где он синтезируется. Удаление лидерной последовательности сопряжено с переходом белка в водонерастворимую конформацию, что способствует его проникновению в мембрану.

По крайней мере в ряде случаев информация, предназначенная для локализации белка в клетке, не обладает ни видовой, ни тканевой специфичностью. Например, овциты *X. laevis* транслируют инъецированную глобиновую мРНК кролика, образуя при этом свободный глобин, тогда как при трансляции альбуминовой мРНК крысы появляется белок, который встраивается в мембранные пузырьки. Сигнал, который имеется у препронинсулина млекопитающих, обеспечивает секрецию этого белка даже тогда, когда его мРНК транслируется в клетках *E. coli*. Хотя еще многое предстоит узнать о том, что лежит в основе различий между мембранами (эукариотические клетки содержат, кроме всего прочего, ядерную, цитоплазматические, мембраны аппарата Гольджи, митохондриальные и другие мембраны), можно думать, что использование сигнальной последовательности является универсальным механизмом, обеспечивающим процесс секреции или встраивание белка в мембрану.

Рекомендуемая литература

Бактериальные мРНК были описаны главным образом на примерах некоторых индивидуальных генов, с которых они транскрибируются; ссылки можно будет найти в гл. 14, 15 и 23 данной книги. По структуре и функции эукариотической мРНК основным источником информации является книга Льюина (*Lewin, Gene Expression, 2, Eucaryotic Chromosomes, Wiley, New York, 1980*), особенно стр. 653–693. Обзор по кэпированию написан Банерджи (*Banerjee, Microbiol. Rev., 44, 175–205, 1980*). Вопрос о локализации белков в клетке является спорным, и все ведущие специалисты по этому вопросу высказали свою точку зрения относительно недавно. Основная статья о роли цитоскелета написана Церверой, Дрейфусом и Пенменом (*Gervera Dreyfuss, Penman, Cell, 23, 113–120, 1981*). Обзоры по гипотезе сигнальной последовательности написаны Блобелом и др. (*Blobel et al., Symp. Soc. Exp. Biol., 33, 9–36, 1979*) и Блобелом (*Blobel, Proc. Nat. Acad. Sci. USA, 77, 1496–1500, 1980*), а также Дэвисом и Тай (*Davis, Tai Nature, 283, 433–438, 1980*). Гипотеза сигнальной последовательности в применении к бактериальным полипептидам описана у Эмра, Холла и Силвахи (*Emr, Hall, Silvahy, J. Cell. Biol., 86, 701–711, 1980*). Другие гипотезы были рассмотрены Викнером (*Wickner, Anp. Rev. Biochem., 48, 23–45, 1979* и *Science, 210, 861–868, 1980*).

Часть III

СИНТЕЗ РНК

Кодовое письмо само по себе должно быть действенным фактором, обуславливающим развитие организма. Но... теперь, когда известна молекулярная структура гена, перестает казаться невероятным, что миниатюрный код может точно соответствовать в высшей степени сложному и специфическому плану развития и каким-то образом содержать в себе инструменты для реализации этого плана.

Эрвин Шрёдингер, 1945

РНК-ПОЛИМЕРАЗЫ – ОСНОВНОЙ ТРАНСКРИПЦИОННЫЙ АППАРАТ КЛЕТКИ

Каким образом РНК оказалась в процессе эволюции посредником между ДНК и белком, остается предметом умозрительных гипотез. К сожалению, современным исследователям невозможно взглянуть на компоненты прототипов живых клеток. Возможно, что в первичных примитивных клетках не существовало различий между типами нуклеиновых кислот; иными словами, то, что считалось геномом, участвовало непосредственно как в репликации, так и в трансляции. В какой-то момент отделение трансляционного аппарата от генома, по-видимому, стало полезным для клетки, и белки начали синтезироваться на молекулах-посредниках, отличавшихся от самого генома. Невозможно установить, как это коррелировало во времени с появлением других видов рибонуклеиновых кислот, участвующих в трансляции, но поразительно, что РНК, с одной стороны, содержится в рибосомах, а с другой – образует адапторную тРНК. (Не удивительно, если выяснилось бы, что некогда РНК выполняла в рибосомах более важную роль, чем та, которую она выполняет теперь.)

Все это говорит о том, что РНК играет центральную роль в генной экспрессии – и не только в качестве матрицы, используемой при синтезе белка, но и как инструмент, необходимый для трансляции матрицы. В некотором смысле эти функции РНК являются отражением различных свойств молекул РНК, в целом связанных с генной экспрессией, но выполняющих несколько различающихся и явно независимых функций. Все типы РНК имеют общее происхождение, а именно транскрибируются с ДНК. В случае мРНК продукт транскрипции является посредником, предназначенным для трансляции. тРНК-и рРНК-транскрипты представляют собой активные молекулы. У прокариот (а, возможно, также и у эукариот) транскрипция – это тот уровень, на котором в основном осуществляется контроль экспрессии генов. Первый (и иногда единственный) этап этого контроля состоит в «принятии решения» о транскрипции данного гена. Рассматривая разные этапы транскрипции, надо всегда учитывать, каким образом они могут влиять на регуляцию активности генов.

Транскрипция – это не единственный способ синтеза РНК. Вирусы, геномы которых состоят из РНК, детерминируют специфические ферменты, осуществляющие различные реакции, но обладающие одним общим свойством: способностью синтезировать РНК на матрице, которая сама представлена молекулами РНК. Этот процесс служит как для образования мРНК, необходимой для поддержания инфекционного процесса (транскрипция РНК), так и для воспроизведения вирусного генома (репликация РНК). Это происходит и в прокариотических, и в эукариотических клетках. Однако и в случае (эукариотических) ретровирусов возможна дальнейшая реакция, в которой вирусная РНК используется в качестве матрицы для обратной транскрипции, приводящей к образованию ДНК.

Что представляет собой РНК-полимераза?

Транскрипция происходит в три различные стадии: **инициации, элонгации и терминации**. На первой и последней стадиях фермент взаимодействует с ДНК, чтобы начать синтез РНК, или отделяется от нее по завершении синтеза. Между этими стадиями осуществляется основной процесс синтеза РНК (транскрипция). Фермент передвигается вдоль молекулы ДНК; в результате образуется цепь РНК, последовательность которой детерминирована матрицей.

РНК-полимераза, определяемая ранее по способности включать рибонуклеозидтрифосфаты в рибонуклеиновые кислоты в присутствии ДНК-матрицы, сейчас может рассматриваться как часть более сложного комплекса, участвующего в транскрипции. Исходная ферментативная активность – это лишь *минимальный* компонент того, что можно назвать РНК-полимеразой. Она лишь контролирует правильное спаривание рибонуклеотидов с ДНК и катализирует образование фосфодиэфирных связей между ними.

И для инициации, и для терминации синтеза РНК необходимы дополнительные активности. Тут очевидна аналогия «разделения труда» между рибосомой и белковыми факторами трансляции. Правда, иногда бывает довольно трудно решить, можно ли считать тот или иной полипептид, участвующий на определенных стадиях транскрипции, составной частью РНК-полимеразы или же дополнительным компонентом.

Продукты транскрипции представляют собой дискретные молекулы, имеющие предопределенные 5'- и 3'-концы. Из этого следует, что иницирование и терминирование транскрипции должны происходить в определенных сайтах ДНК. Инициация предполагает образование комплекса между РНК-полимеразой и ДНК в участке, расположенном вблизи нуклеотида, с которого начинается транскрипция. Последовательность ДНК, необходимая для образования этого комплекса, называется **промотором**. Промотор может включать в себя как последовательности, непосредственно связывающиеся с иницирующим комплексом, так и те, которые прилегают к этому участку. (Распознавание этих последовательностей необходимо для инициации, но они не являются частью стабильного сайта связывания.) То место, с которого начинается включение первого нуклеотида в синтезирующийся транскрипт, называется **стартовой точкой**.

Условия, необходимые для инициации, обычно оказываются более сложными, чем те, которые требуются для элонгации. Все компоненты, участвующие в элонгации, также необходимы и для инициации, однако в последнем случае подключаются еще и дополнительные полипептиды. Вспомогательные факторы не всегда могут рассматриваться как субъединицы РНК-полимеразы. В основе принципа, по которому проводятся эти различия, лежит

универсальность их включения в реакцию транскрипции. Если они, являясь частью обычного ферментативного препарата, используются только на начальных стадиях, но необходимы для распознавания всех промоторов, то, вероятнее всего, эти факторы должны быть классифицированы как составные части фермента. Если же они требуются только для транскрипции определенных генов, то их скорее следует отнести к регуляторным факторам.

Терминация означает распознавание такой точки, после которой не происходит дальнейшего роста цепи РНК. Соответствующий участок ДНК называется **терминатором**. От соответствующего компонента РНК-полимеразы требуется выполнение двух функций. Во-первых, его ферментативная активность должна прекратить дальнейшее образование фосфодиэфирных связей. Во-вторых, он должен диссоциировать от ДНК, освобождая при этом вновь синтезированную РНК.

Терминация не так хорошо исследована, как инициация. У бактерий существуют разные классы терминаторных участков. По крайней мере некоторые из них не узнаются той формой РНК-полимеразы, которая осуществляет элонгацию, и требуют дополнительных белковых факторов. В данном случае опять оказывается нелегко определить, что считать компонентом фермента и что – нет. Процесс терминации у эукариот изучен еще хуже, чем у бактерий, и что следует подразумевать под термином «РНК-полимераза» – в какой-то степени дело вкуса.

В случае бактериальной РНК-полимеразы можно попытаться определить роль индивидуальных полипептидов, функционирующих на разных стадиях транскрипции. Эукариотический фермент очищен значительно хуже, и основная трудность заключается в выделении истинной РНК-полимеразной активности из неочищенного экстракта. В последующих главах мы подробно разберем все факторы, участвующие в реакции транскрипции РНК с ДНК-матрицы.

В данной главе мы рассмотрим механизм транскрипции и свойства препаратов РНК-полимеразы, очищенных обычным путем. В гл. 11 будет обсуждаться природа промоторов. Затем в гл. 12 и 13 мы подойдем к вопросу о контроле процесса транскрипции РНК-полимеразой бактерий. В клетках бактерио-хозяина и в клетках, инфицированных фагом, процессы инициации и терминации регулируются путем использования дополнительных белковых факторов, которые модифицируют или коренным образом изменяют важнейшие свойства РНК-полимеразы.

Субъединичная структура бактериальной РНК-полимеразы

У бактерий синтез мРНК, рРНК и тРНК осуществляется одной и той же РНК-полимеразой. Общее количество молекул РНК-полимераз, присутствующих в клетках *E. coli*, может составлять 7000. Значительная часть из них, по-видимому, участвует в транскрипции. В зависимости от условий роста в синтезе РНК может быть занято одновременно от 2000 до 5000 молекул фермента. Кроме основной РНК-полимеразы в клетке существует еще один белок (продукт гена *dnaG*), осуществляющий синтез РНК-загавок, необходимых для репликации ДНК. Однако функции этого белка в синтезе РНК ограничиваются только этим процессом (гл. 33).

Лучше всего охарактеризована РНК-полимераза из *E.*

coli; ее субъединичная структура аналогична структуре всех других изученных видов бактерий. **Полный фермент**, или **голофермент**, имеет мол. массу около 480 000 дальтон. Он состоит из следующих субъединиц:

2	α	по 40 000 каждая
1	β	155 000
1	β'	160 000
1	σ	85 000

α -, β - и β' -субъединицы разных видов бактерий имеют довольно близкую молекулярную массу: молекулярная масса σ -субъединицы варьирует в значительных пределах – от 44 000 до 92 000 дальтон. Фермент имеет несколько вытянутую форму, и его максимальные размеры составляют 15 нм.

Голофермент ($\alpha_2\beta\beta'\sigma$) можно разделить биохимическими методами на два компонента: **минимальный фермент** ($\alpha_2\beta\beta'$) и **сигма-фактор** (σ -полипептид). В названии компонентов отражен тот факт, что только голофермент может иницировать транскрипцию, а далее сигма-фактор освобождается из комплекса и собственно элонгация осуществляется минимальным ферментом. Таким образом, минимальный фермент способен синтезировать фосфодиэфирные связи на ДНК-матрице, но он не может иницировать транскрипцию в нужном участке.

Сигма-фактор контролирует связывание РНК-полимеразы с ДНК

Основная функция сигма-фактора состоит в том, чтобы обеспечить стабильное связывание РНК-полимеразы с промоторами, а не с другими участками ДНК.

Минимальный фермент сам по себе имеет сродство к ДНК; в основе этого сродства лежит электростатическое взаимодействие между положительно заряженным белком и отрицательно заряженной нуклеиновой кислотой. Возможно, что способность к связыванию любой ДНК независимо от ее нуклеотидной последовательности – характерная особенность всех белков, имеющих специфические участки узнавания на ДНК (гл. 14). Случайная последовательность ДНК, которая связывается РНК-полимеразой, называется **слабым участком связывания**, а комплекс ДНК-фермент закрытым, поскольку ДНК в этом комплексе находится в двухцепочечной форме. Константа связывания при формировании закрытого комплекса в участке слабого связывания составляет $2 \cdot 10^{11} \text{ M}^{-1}$, а полупериод диссоциации комплекса на фермент и ДНК – около 60 мин.

Сигма-фактор принципиально изменяет характер взаимодействия РНК-полимеразы с ДНК. У голофермента сродство к слабым участкам связывания, т.е. любым последовательностям ДНК, резко *снижено*. Константа связывания для этой реакции составляет 10^7 M^{-1} , а ее период полужизни – менее 1 с. Таким образом, сигма-фактор нарушает способность фермента связываться со случайным участком ДНК примерно в 10 000 раз, и образующиеся комплексы ДНК-фермент становятся *очень короткоживущими*.

Но сигма-фактор сообщает ферменту способность узнавать *специфические* участки связывания. Голофермент может очень прочно связываться с промоторами; константа связывания при этом составляет 10^{14} M^{-1} , а полупериод жизни образующегося комплекса длится несколько часов. Эта константа ассоциации является

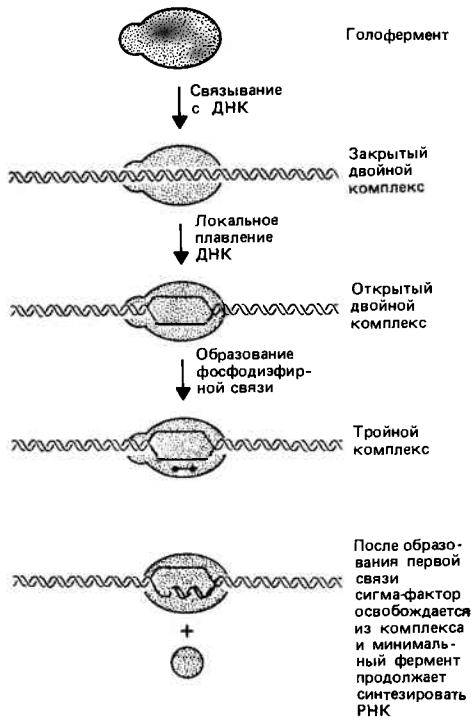


Рис. 10.1. Иницирование транскрипции происходит в несколько последовательных этапов, в процессе которых закрытый двойной комплекс переходит в открытый, а затем в тройной комплекс.

усредненной величиной. На самом деле существуют примерно стократные вариации в скорости связывания фермента с различными промоторами (в зависимости от конкретной нуклеотидной последовательности ДНК), что имеет принципиальное значение для частоты инициации с соответствующего промотора.

Как показано на рис. 10.1, распознавание промоторов голоферментом существенным образом отличается от реакции минимального фермента со слабыми участками связывания. Взаимодействие голофермента с промотором начинается с формирования аналогичного закрытого (бинарного) комплекса. Но затем он переходит в **открытый комплекс**. При этом происходит плавление небольшого участка ДНК в пределах той последовательности, которая связана с ферментом. Обычно фермент экранирует примерно 60 нуклеотидных пар молекулы ДНК, внутри которых 12–17 пар нуклеотидов (согласно разным определениям, см. далее) находятся в неспаренном состоянии и образуют участок с одноцепочечной структурой. В результате ряда последовательных событий, приводящих к формированию открытого комплекса, происходит **прочное связывание**.

Поскольку РНК-полимеразе необходимо нарушить структуру ДНК, образование открытых комплексов и транскрипция быстрее протекают на отрицательно суперспирализованной кольцевой ДНК, чем на линейной. То напряжение, которое возникает в двойной спирали в результате суперспирализации, облегчает расплетание двух цепей. Это может иметь важное значение для иницирования транскрипции по крайней мере на некоторых промоторах (гл. 11). Поскольку образование промоторов предполагает разделение цепей ДНК, то открытый комплекс между голоферментом и ДНК менее стабилен в условиях низкой температуры и высокой ионной силы, тог-

да как для закрытого комплекса между минимальным ферментом и ДНК характерна противоположная зависимость от этих параметров.

Особенности этих типов связывания указывают на то, что РНК-полимераза может находить промоторы методом проб и ошибок, как это показано на рис. 10.2. Любой не работающий в данный момент в клетке минимальный фермент скорее всего существует в форме закрытых слабых комплексов, поскольку образование таких комплексов происходит быстро, а распад — медленно. (К сожалению, точно не известно, какая часть молекул свободных РНК-полимераз клетки существует в форме минимального фермента и какая представлена голоферментом.) Голофермент очень быстро ассоциирует со слабыми участками связывания и также быстро отделяется от них. Таким образом, продвигаясь вдоль молекулы ДНК, голофермент образует и разрушает ряд закрытых комплексов до тех пор, пока в процессе поиска (случайно) не натолкнется на промотор. Тогда в результате узнавания специфической последовательности он может прочно связаться с ДНК в нужном участке и образовать открытый комплекс.

Константа скорости связывания с промоторами очень близка к тому пределу, который определяется скоростью диффузии. Поскольку связывание происходит очень быстро, то скорость данного процесса ограничивается в основном скоростью диссоциации голофермента от слабых

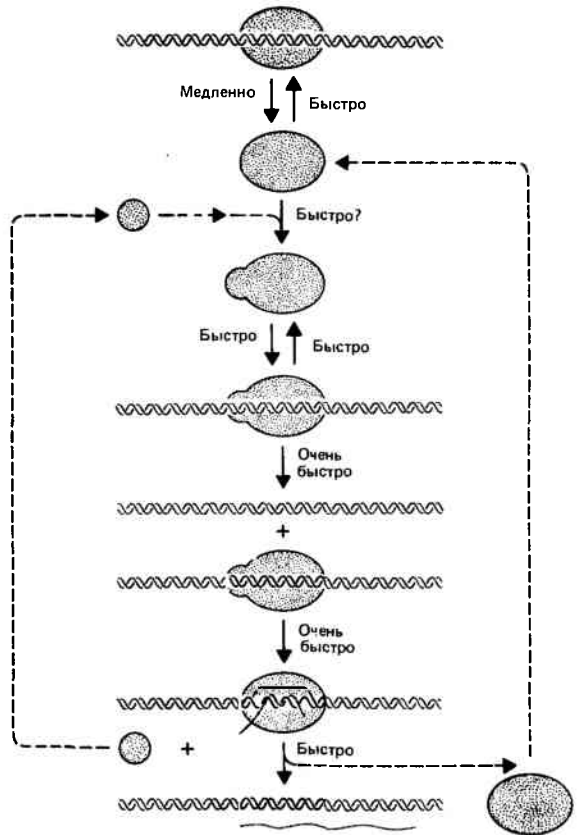


Рис. 10.2. Цикл сигма-фактора и минимального фермента при транскрипции.

Освобождение минимального фермента происходит при терминеции, после чего он либо связывается со свободными участками ДНК, либо с сигма-фактором, образуя голофермент, который стабильно взаимодействует только с промоторами. Сигма-фактор отделяется сразу же, как только образуется тройной комплекс.

участков связывания. Однако это происходит слишком медленно, примерно 1 диссоциация в секунду. Следовательно, РНК-полимераза может использовать другое средство для поиска участков связывания, возможно путем прямой замены одной связывающей последовательности на другую. Эта реакция может происходить со скоростью около 1 замены в миллисекунду. Иными словами, вместо того чтобы двигаться вдоль цепи ДНК, оставляя один участок связывания и диффундируя к другому, фермент, вероятно, хватается за одну последовательность ДНК, меняет ее быстро на другую и продолжает в той же беспорядочной манере обменивать последовательности, пока не доберется до промотора. После этого фермент образует стабильный открытый комплекс, и начинается инициация.

Каким образом сигма-фактор влияет на способность минимального фермента ассоциировать с ДНК? Проведенные ранее эксперименты указывали на то, что сигма-субъединица непосредственно не связывается с дуплексом ДНК. Однако возможно, что она способна взаимодействовать с ДНК, находящейся в суперспирализованном состоянии, хотя мы не имеем данных о том, может ли она сама узнавать специфические нуклеотидные последовательности. В то же время известно, что когда голофермент образует прочно связанный комплекс с промотором, сигма-фактор контактирует с ДНК в области начального плавления, в точке, расположенной непосредственно перед стартовой точкой транскрипции. Добавление сигма-фактора может изменить конформацию минимального фермента таким образом, что он менее эффективно распознает слабый участок связывания и уже в *форме голофермента* может специфически контактировать с участками прочного связывания.

Рабочий цикл сигма-фактора

Сигма-фактор нужен только для инициации. Он сразу же освобождается из комплекса с минимальным ферментом, как только начинается синтез РНК.

После того как сформируется открытый (бинарный) комплекс, следующим шагом является включение двух первых нуклеотидов и образование фосфодиэфирной связи между ними. Это приводит к образованию **тройного комплекса** между минимальным ферментом ДНК и синтезирующейся РНК (рис. 10.1). Данная реакция протекает настолько быстро, что период полужизни бинарного комплекса в участке прочного связывания составляет всего лишь 0,2 с. Инициация заканчивается формированием тройного комплекса, и сигма-фактор уже не является его составной частью.

Прямая взаимосвязь между освобождением сигма-фактора и синтезом фосфодиэфирной связи неизвестна. Возможно, сигма-фактор непосредственно замещается динуклеотидом образующейся РНК; возможно также, что его удаление из комплекса может быть индуцировано изменением свойств минимального фермента, происшедшим под воздействием синтезирующейся РНК. В любом случае, начиная с момента образования первой фосфодиэфирной связи, транскрипция осуществляется тройным комплексом, содержащим минимальный фермент.

Минимальный фермент в составе тройного комплекса очень прочно связан с ДНК. Он как бы заперт в нем до тех пор, пока не закончится элонгация. Когда же транскрипция завершается, минимальный фермент отделяется

от ДНК в виде свободного белкового тетрамера. Как показано на рис. 10.2, освободившийся минимальный фермент затем должен найти другой сигма-фактор и включиться в следующий цикл транскрипции.

Наличие цикличности во временном объединении сигма-фактора с минимальным ферментом решает дилемму, стоящую перед РНК-полимеразой: привести в соответствие взаимодействие фермента с матрицей при инициации и элонгации. Это в самом деле дилемма, поскольку для инициации требуется прочное взаимодействие *только* с определенными последовательностями (промоторами), тогда как при элонгации необходимо прочное связывание со *всеми* последовательностями, вдоль которых происходит движение фермента. Минимальному ферменту присуще высокое сродство к ДНК, которое увеличивается в присутствии новосинтезированной РНК. Однако его сродство к слабым участкам связывания слишком велико, чтобы позволить ферменту эффективно находить промоторы. При этом поиск участков прочного связывания методом проб и ошибок путем ассоциации и диссоциации может длиться много часов. Сигма-фактор значительно ускоряет этот процесс, уменьшая стабильность слабых комплексов. В то же время, стабилизируя ассоциацию в участках прочного связывания, сигма-фактор необратимо сдвигает реакцию в сторону образования открытых комплексов. Но затем действия голофермента «парализуются» его же собственным специфическим сродством к промоторам. Поэтому, освобождаясь от сигма-фактора, фермент снова способен связываться с любой последовательностью ДНК, что позволяет ему продолжать транскрипцию.

Минимальный фермент синтезирует РНК

РНК транскрибируется только с одной цепи ДНК-матрицы. Для осуществления этого процесса **цепи ДНК** должны быть локально расплетены в данном участке. Минимальный фермент начинает транскрипцию на расплетенных цепях ДНК открытого промоторного комплекса. По мере передвижения фермента вдоль матрицы и роста цепи РНК область локального расплетения цепей движется вместе с ним.

Участок матрицы, ассоциированный с ферментом, составляет около 60 нуклеотидов, однако область локального расплетения цепей будет короче, и точно оценить ее размер довольно трудно. Судя по чувствительности ДНК в открытом комплексе к нуклеазе, специфичной к одноцепочечным участкам, длина расплетенного участка равна 12 парам нуклеотидов (гл. 11). Если же РНК-полимеразная реакция протекает на кольцевых двуспиральных молекулах ДНК, то протяженность неспаренного участка ДНК составляет 17 нуклеотидных пар.

По мере того как происходит расплетение цепей ДНК, каждая из отдельных цепей взаимодействует с определенным участком фермента. Как показано на **рис. 10.3**, одноцепочечный участок матрицы, расположенный перед точкой присоединения первых рибонуклеотидов цепи РНК, будет находиться в свободном состоянии, а в том месте, где только что прошел синтез РНК, он будет существовать в форме гибрида ДНК—РНК. Таким образом, протяженность гибридной области может быть несколько меньше, чем расплетенный участок ДНК. Видимо, гибридный комплекс ДНК—РНК имеет в длину 12 пар нуклеотидов.

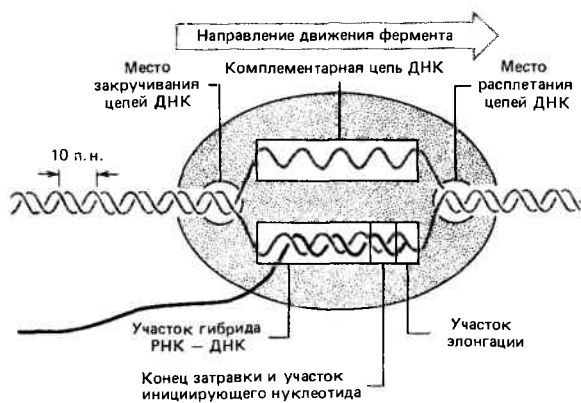


Рис. 10.3. РНК-полимераза экранирует участок ДНК примерно в 60 пар нуклеотидов; она имеет несколько активных центров.

После того как РНК-полимераза продвинулась далее по матрице, в оставленном позади нее расплетенном участке восстанавливается двухцепочечная структура ДНК, а РНК вытесняется в виде свободной полинуклеотидной цепи. В каждый данный момент транскрипции около 50 последних рибонуклеотидов растущей цепи находится в комплексе с ДНК и ферментом. Связанный с ферментом участок матрицы, в котором расплетаются цепи ДНК и происходит синтез РНК, иногда называют **точкой роста** (хотя этот термин чаще используют при описании процесса репликации).

К сожалению, мы пока еще не понимаем топологии процесса расплетания цепей ДНК и последующего восстановления структуры двойной спирали, однако, исходя из способности очищенной РНК-полимеразы транскрибировать двуспиральную ДНК *in vitro*, можно думать, что этот процесс контролируется самим ферментом. Мало вероятно, чтобы при этом фермент вращался вокруг оси ДНК, так как это повлекло бы за собой движение всей цепи РНК и ассоциированных с ней рибосом. Возможно, что вращательные движения совершает матрица; при этом участок ДНК, находящийся перед ферментом, вращается по направлению раскручивания и одновременно происходит закручивание цепей ДНК позади фермента – в обратном направлении. Для осуществления этого процесса *in vivo*, по-видимому, необходимо участие других ферментативных активностей, которые соответствующим образом изменяют пространственную структуру ДНК.

Все нуклеиновые кислоты синтезируются из предшественников – 5'-нуклеозидтрифосфатов. На рис. 10.4 показана реакция полимеризации между 5'-концевой трифосфатной группой присоединяемого нуклеотида и 3'-ОН-группой последнего нуклеотида цепи.

Минимальный фермент должен удерживать две реагирующие группы в таком положении, которое благоприятно для образования фосфодиэфирной связи. Как только ковалентная связь возникла, фермент продвигается далее по матрице на один нуклеотид, с тем чтобы повторить реакцию. Скорость реакции при 37°C высокая – около 40 присоединенных нуклеотидов в 1 с (гл. 9). Как показано на рис. 10.3, последний нуклеотид растущей цепи определяет **участок связывания затравки**. Область, занимаемая присоединяемым нуклеозидтрифосфатом, образует **участок элонгации**.

Различие в природе обоих реагирующих групп стано-

вится очевидным при инициации, когда фермент должен непосредственно связать два нуклеозидтрифосфата. Первое основание в цепи РНК чаще всего (но не всегда) является пурином, и фермент может *in vitro* присоединять АТР или GTP к участку, который отличается от участка элонгации. Он является **участком иницирующего нуклеотида** и должен в целом перекрываться с участком связывания затравки. Первый нуклеотид цепи содержит все три 5'-фосфатных остатка.

Все четыре нуклеозидтрифосфата могут поступать в участок элонгации нуклеотидов. По-видимому, сама РНК-полимераза отбирает нужные предшественники путем спаривания новопривывшего нуклеотида с цепью ДНК-матрицы. Возможно, структура участка элонгации такова, что фосфодиэфирная связь образуется лишь в том случае, если нуклеотид формирует правильную пару с соответствующим основанием ДНК. Если же оказалось, что нуклеотид не способен образовать совершенную пару, то он скорее всего удаляется и вместо него испытывается другой. (Вероятность случайного правильного выбора нуклеотида равна примерно 25%, что намного превосходит вероятность вхождения правильной аминокислоты в А-участок рибосомы, составляющую всего лишь 3%.)

Для функционирования РНК-полимеразы важное значение имеют двухвалентные катионы. Рибонуклеотиды поступают в участок элонгации нуклеотидов в форме хелатных соединений $Mg^{2+} - NTP$. Минимальный фермент содержит два атома цинка. Возможно, что они нужны для поддержания нуклеотидов в участке связывания затравки и участке иницирующего нуклеотида в конформации, благоприятной для реакции с очередным нуклеотидом.

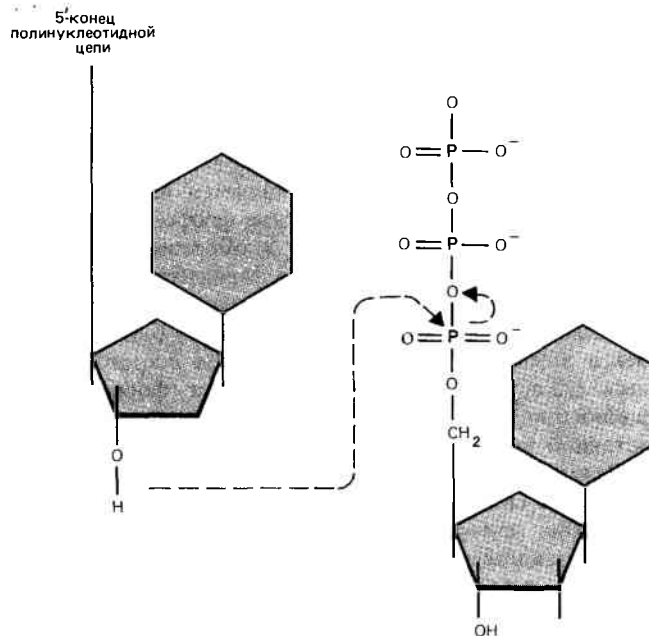


Рис. 10.4. Образование фосфодиэфирной связи предполагает гидрофильную атаку 3'-ОН-группой последнего нуклеотида цепи 5'-трифосфата нового нуклеотида, сопровождающуюся освобождением пиррофосфата.

Считается, что нуклеотидная цепь растет в направлении от 5'-конца к 3'-концу, поскольку 5'-конец цепи остается неизменным и каждый новый нуклеотид увеличивает ее длину на один остаток со стороны 3'-конца.

Функции субъединиц минимального фермента

Мы располагаем весьма ограниченными данными о пространственной организации минимального фермента, и самое большее, что мы можем сделать, – это построить схематическую диаграмму участков, определяющих различные ферментативные функции (как это показано на рис. 10.3). Ни один из этих участков пока еще не локализован ни на одной из полипептидных субъединиц. Однако некоторая общая информация о роли отдельных субъединиц уже имеется.

На β -субъединицу действуют два типа антибиотиков. Это было выяснено путем локализации мутаций, определяющих устойчивость к ним. **Рифамицины** (из которых наиболее широко используется рифампицин) блокирует инициацию, действуя на стадии, предшествующей формированию первой фосфодиэфирной связи. Если же эта стадия уже завершилась, то образовавшийся тройной комплекс устойчив к ингибированию. **Стрептолидигины** ингибируют элонгацию. β -Субъединица является мишенью и для этой группы антибиотиков. Об этом свидетельствует тот факт, что в опытах по реконструкции именно источник добавленной β -субъединицы определяет устойчивость к стрептолидигинам. β -Субъединица также мезится некоторыми аффинными аналогами нуклеозидтрифосфатов. Все эти данные дают основание думать, что β -субъединица участвует в связывании нуклеотидных субстратов.

Гепарин является полианионом, который взаимодействует с β' -субъединицей и ингибирует транскрипцию *in vitro*. Гепарин конкурирует с ДНК при связывании с матрицей до образования стабильного иницирующего комплекса. β' -Субъединица является наиболее основной, что могло бы играть определенную роль в связывании фермента с матрицей.

Выделенная в чистом виде α -субъединица существует в форме димера, способного связываться с β -субъединицей; затем комплекс $\alpha_2\beta$ присоединяет β' , образуя минимальный фермент. При инфицировании *E. coli* фагом T4 α -субъединица модифицируется в результате ADP-рибозилирования аргинина. Это сопровождается уменьшением сродства к обычным промоторам, распознаваемым голоферментом. Следовательно, вполне возможно, что α -субъединица играет какую-то роль в распознавании промотора.

Ни одну из субъединиц нельзя считать участком связывания для σ . Все они, по-видимому, прямо или косвенно участвуют в превращении минимального фермента в голофермент.

РНК-полимеразы фагов, возможно, являются «минимальными» ферментами

Почему для осуществления своих функций бактериальной РНК-полимеразе необходима огромная мультимерная структура? Некоторые фаги кодируют РНК-полимеразы значительно меньшего размера, состоящие всего лишь из одной полипептидной цепи. Этот факт указывает, что транскрипционный аппарат, необходимый для синтеза РНК, может быть намного меньше, чем тот, который имеется у клетки-хозяина.

Существование таких ферментов подводит к мысли о том, что транскрипция может производиться «мини-

мальным» аппаратом. Эти ферменты узнают всего лишь несколько промоторов в фаговой ДНК и не могут узнавать другие. Таким образом, свойства, присущие этим ферментам, накладывают на них определенные ограничения, способствуя связыванию только с некоторыми специфическими участками ДНК и синтезированию РНК. Как же устроены эти ферменты?

РНК-полимеразы, кодируемые родственными фагами T3 и T7, представляют собой полипептиды, состоящие всего лишь из одной цепи, примерно по 110 000 дальтон каждая. Они очень быстро синтезируют РНК (при 37°C скорость синтеза составляет около 200 нуклеотидов в 1 с). Полимераза фага T3 начинает синтез всех РНК с образования динуклеотида pppGpG. Эта характерная особенность указывает на то, что процесс инициации у фага, вероятно, проще, чем в случае полимераз *E. coli*.

В отличие от этого фермент клетки-хозяина способен транскрибировать любую из многих (около 1000) транскрипционных единиц. Некоторые из них могут транскрибироваться только полимеразой без участия каких-либо других факторов, хотя эффективность транскрипции будет зависеть от степени сродства фермента к определенному промотору. Однако многие гены транскрибируются только в присутствии дополнительных белковых факторов. В некоторых случаях эти факторы специфичны в отношении какой-либо одной транскрипционной единицы; иногда они участвуют в координированной транскрипции многих единиц. Заражение некоторыми фагами вызывает резкие изменения в сродстве РНК-полимеразы клетки-хозяина к промоторам, выражающиеся в том, что она перестает узнавать специфические участки своего генома, а вместо этого начинает транскрипцию на фаговых промоторах. Таким образом, ферменту клетки-хозяина приходится действовать в соответствии с различными клеточными и фаговыми функциями, что изменяет его транскрипционные свойства. Следовательно, сложность фермента может, по крайней мере частично, отражать наличие множества сигналов контроля, на которые он должен отвечать.

Сложные эукариотические РНК-полимеразы

Транскрипционный аппарат эукариотических клеток устроен значительно сложнее и изучен хуже, чем бактериальный. У эукариот обнаружены три ядерные РНК-полимеразы, характеризующиеся разной локализацией и состоящие из большого числа полипептидных компонентов. Кроме того, в митохондриях и хлоропластах обнаружены другие РНК-полимеразные активности.

Основная часть полимеразной активности приходится на долю РНК-полимеразы I, которая обнаруживается в ядрышках и ответственна за транскрипцию генов рРНК. Этот фермент обеспечивает 50–70% клеточного синтеза РНК. РНК-полимераза примерно одинаково стимулируется ионами Mn^{2+} и Mg^{2+} .

Другим важным ферментом является РНК-полимераза II. Она локализована в нуклеоплазме (ядро, за исключением ядрышка); на долю этого фермента приходится 20–40% клеточной РНК-синтезирующей активности. Он ответствен за синтез гетерогенной ядерной РНК (гя-РНК), предшественника мРНК. Фермент намного эффективнее стимулируется ионами Mn^{2+} , чем ионами Mg^{2+} .

РНК-полимераза III представляет собой минорную ферментативную активность, обеспечивающую до 10% клеточного синтеза РНК. Фермент находится в нуклеоплазме и осуществляет синтез многих малых ядерных РНК и тРНК. Он стимулируется ионами Mg^{2+} несколько эффективнее, чем ионами Mg^{2+} .

Различная зависимость активности РНК-полимераз от бивалентных катионов в свое время считалась основной отличительной особенностью этих ферментов (вот почему мы упоминаем о ней). Сейчас же различные полимеразные активности дифференцируют, исходя из их чувствительности к бициклическому октапептиду — α -аманитину. В клетках животных, растений и насекомых РНК-полимераза II быстро инактивируется низкими концентрациями α -аманитина (около 0,03 мкг/мл). Эта особенность сохраняется у довольно широкого спектра организмов. Активность РНК-полимеразы III в животных клетках ингибируется при значительно более высокой концентрации α -аманитина — около 20 мкг/мл; однако в дрожжевых клетках и в клетках насекомых этот фермент устойчив и к этой концентрации. Активность РНК-полимеразы I не ингибируется ни в каких клетках. (Это означает, что для проявления ингибирующего действия α -аманитина его концентрация должна быть выше 500 мкг/мл.)

В неочищенных экстрактах все полимеразные активности связаны с крупными белковыми агрегатами с мол. массой от 500 000 дальтон и более и со сложным субъединичным составом. Каждый фермент имеет две большие субъединицы — одна из них с мол. массой примерно 200 000 дальтон, другая — около 140 000 дальтон, — и до 10 маленьких субъединиц, варьирующих по молекулярной массе от 10 000 до 90 000 дальтон. Являются ли какие-либо из этих субъединиц одинаковыми или сходными у разных полимераз, пока не известно.

Поскольку ни одну из этих РНК-полимераз до сих пор не удалось реконструировать из субъединиц, остается неясным, все ли белковые субъединицы входят в состав каждого фермента. Не известно также, какие из них отвечают за каталитическую активность и могут ли остальные выполнять регуляторные функции. Возможно, что ферментные препараты представляют собой основные транскрипционные комплексы — в общем сходные для всех клеток, но регулируемые дополнительными белковыми факторами. Или, скажем, эти факторы могут входить в состав ферментного препарата в такой же степени, как и основной каталитический комплекс. Иными словами, остался нерешенным следующий важный вопрос, связанный с устройством транскрипционных аппаратов: способны ли компоненты выделяемого транскрипционного комплекса самостоятельно решать, какие гены следует транскрибировать, или же для этого необходимы вспомогательные белковые факторы, которые еще надлежит идентифицировать?

Выделенные ферментные препараты можно использовать для синтеза РНК *in vitro*. Были разработаны некоторые гетерологичные системы, в которых инициация происходит в соответствующих промоторных участках. Так, например, препарат РНК-полимеразы II из культуры

человеческих клеток KB может транскрибировать глобиновый ген мышцы. Это означает, что для узнавания промотора РНК-полимеразой не имеют значения ни тканевые, ни видоспецифические особенности. Следовательно, существует некая универсальная и консервативная структура, которая узнается как промотор. Конечно, это не исключает возможности того, что в естественных условиях дополнительные белковые факторы и другие последовательности ДНК участвуют в модулировании РНК-полимеразной реакции.

В другой модельной системе РНК-полимераза III из *Xenopus laevis* может специфически транскрибировать гены 5S-РНК только в том случае, если в систему добавлен дополнительный белковый фактор, полипептид с мол. массой 37 000 дальтон, находящийся в ооцитах в комплексе с 5S-РНК. По-видимому, этот вспомогательный белок необходим для основного транскрипционного комплекса, работающего с генами 5S-РНК, но не нужен для транскрипции других генов, поскольку гены тРНК могут транскрибироваться РНК-полимеразой III без какого-либо фактора (гл. 11).

Поскольку общая сложность транскрипционного аппарата эукариот выявляется только в различных конкретных системах, в настоящее время можно с уверенностью говорить о существовании трех мультимерных ферментов, необходимых для синтеза РНК. Все ли их компоненты имеют одинаково важное значение и сколько при этом требуется других белковых факторов — пока не известно.

РНК-полимеразы митохондрий и хлоропластов имеют, по-видимому, меньшие размеры и отличаются от ядерных ферментов. Безусловно, геном этих органелл намного меньше клеточного генома, и собственным полимеразам нужно транскрибировать только ограниченное число генов; следовательно, и транскрипционный контроль у них может быть более простым. Эти ферменты могут быть аналогичны фаговым полимеразам, отвечающим за транскрипцию только определенных единичных генов и не способных отвечать на более сложные сигналы. Ни один из этих ферментов до сих пор не очищен в количествах, достаточных для тестирования в системе *in vitro*.

Рекомендуемая литература

В качестве источника при написании многих обзоров и оригинальных исследовательских статей широко используется книга RNA polymerase, Cold Spring Laboratory, New York, 1976, вышедшая под редакцией Лозика и Чамберлина (*Losick, Chamberlin*). Две главы, освещающие основные моменты излагаемого здесь материала, написаны Чамберлином (*Chamberlin*). В них дается вначале общее представление о проблеме (стр. 17–68), а затем подробно рассматриваются вопросы взаимодействия бактериального фермента с матрицей (стр. 159–192).

Глава 11

ПРОМОТОРЫ: САЙТЫ ИНИЦИАЦИИ
ТРАНСКРИПЦИИ

Основной вопрос, возникающий при исследовании взаимодействия между РНК-полимеразой и ее промотором, состоит в следующем: каким образом белок узнает специфические последовательности в ДНК? Имеется ли в молекуле фермента активный центр, способный различать химическую структуру, образованную определенными основаниями в двуспиральной молекуле ДНК? Насколько фермент специфичен? Промоторы различаются своим сродством к РНК-полимеразе, что, возможно, имеет большое значение для контроля частоты инициации, а следовательно, и уровня генной экспрессии. Каким образом изменения в последовательности ДНК сказываются на способности взаимодействовать с ферментом?

Иницирование транскрипции – важнейший этап в осуществлении контроля генной экспрессии. При решении вопроса о необходимости экспрессии определенного гена часто этот этап является первым, а иногда единственным. Как контролируется способность РНК-полимеразы иницировать транскрипцию на конкретном промоторе? На некоторых промоторах РНК-полимераза в принципе способна эффективно иницировать транскрипцию, хотя регуляторные белковые факторы могут *препятствовать* этому процессу. В других случаях одной полимеразы для инициации недостаточно и необходимо присутствие дополнительных белков. Эти белковые факторы узнают последовательности ДНК, расположенные рядом с промотором или перекрывающиеся с ним. Таким образом, несмотря на то что промотор определяется как последовательность, с которой взаимодействует РНК-полимераза, прилегающие области также могут оказывать влияние на способность фермента иницировать транскрипцию.

Промотор и прилегающие к нему участки контроля являются последовательностями ДНК, чьи функции – *узнаваться* белками, и этим они принципиально отличаются от последовательностей, предназначенных для транскрибирования и последующего транслирования. Информация, необходимая для функционирования промотора, заложена в самой последовательности ДНК, т.е. сама структура ДНК служит сигналом. Экспрессируемые же области приобретают смысл только после того, как закодированная в них информация преобразуется в форму другой нуклеиновой кислоты или белка.

Промотор необходим для транскрипции определенной области ДНК, непосредственно к нему прилегающей. Согласно принятой в генетике терминологии, последовательности ДНК, не экспрессирующиеся в виде диффузионных молекул РНК или белка, а выполняющие какие-либо другие функции, называются *цис-доминантными* или *цис-активными* участками. Это означает, что данные участки оказывают влияние только на последовательности, расположенные в той же молекуле ДНК, в которой находятся они сами. Они не могут влиять на гомологичные последовательности, расположенные в других молекулах ДНК (гл. 14).

После того как на стартовой точке началась транскрипция, РНК-полимераза продолжает продвигаться

вдоль матрицы до тех пор, пока не достигнет терминатора. Последовательность, которая находится между промотором и терминатором, называется **единицей транскрипции**. Как это видно на **рис. 11.1**, транскрипционная единица является участком ДНК, кодирующим образование одной молекулы РНК. В зависимости от конкретной системы в состав транскрипционной единицы могут входить один или несколько генов. Иногда область, находящаяся ближе к промотору, обозначается как **проксимальная** в отличие от **дистальной**, расположенной вдали от него.

Определение стартовой точки
in vivo и in vitro

Стартовой точкой служит пара оснований в ДНК, соответствующая нуклеотиду, который первым включается в транскрипт, синтезируемый РНК-полимеразой (другими словами, это нуклеотид, определяющий включение в мРНК иницирующего нуклеотида). Как правило, он занимает строго определенное положение, но иногда может быть представлен двумя соседними основаниями, а в некоторых случаях любым из нескольких близлежащих оснований. Для определения стартовой точки необходимо исследовать 5'-конец первичного транскрипта, непосредственного продукта РНК-полимеразы, раньше, чем с ним произойдут какие-нибудь изменения. Наибольшее количество информации о стартовых точках можно получить в опытах *in vitro*, так как *in vivo* очень трудно выделить первичный транскрипт, прежде чем он начнет деградировать (в случае прокариот) или подвергнется модификации (у эукариот).

На **рис. 11.2** приведена схема универсального метода идентификации стартовой точки с использованием гибридизации транскрипта с его ДНК-матрицей. Суть метода

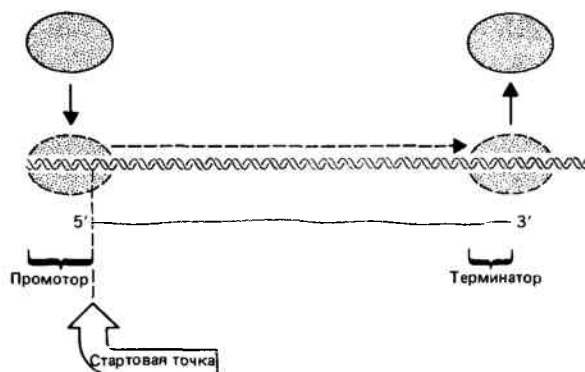


Рис. 11.1. Транскрипционной единицей является последовательность ДНК, считываемая с образованием единой молекулы РНК, начинающейся на промоторе и заканчивающейся на терминаторе.

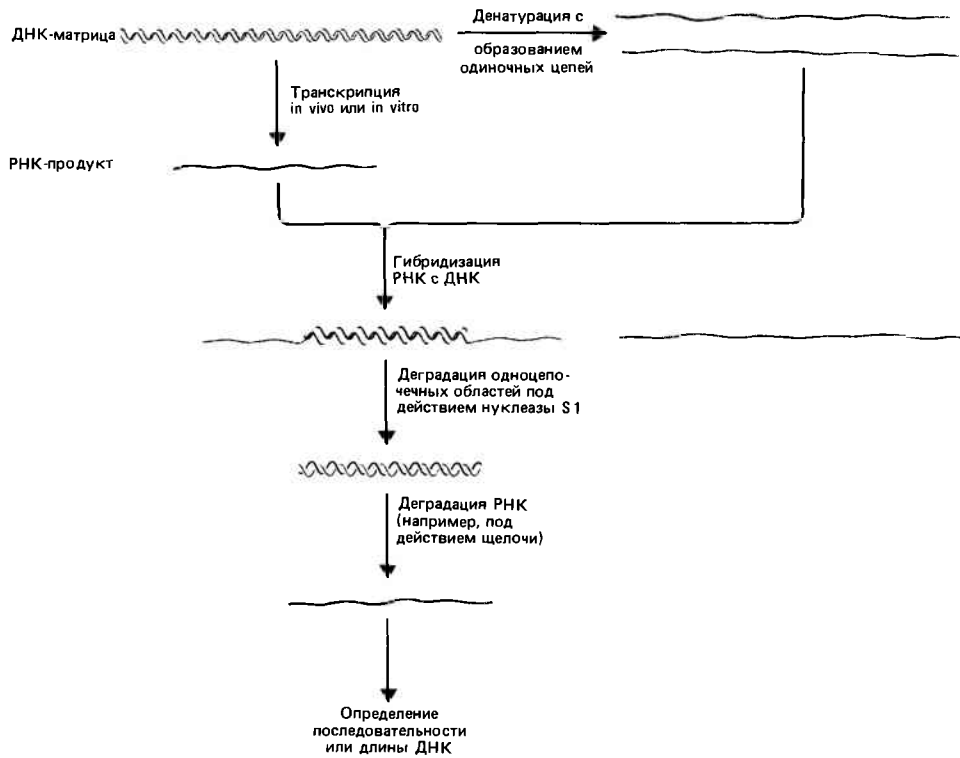


Рис. 11.2 Местоположение стартовой точки можно определить, сравнивая образовавшуюся РНК с ДНК-матрицей. При гибридизации РНК с денатурированной ДНК образуется гибрид РНК—ДНК, в состав которого входит матричная цепь. Другая цепь ДНК остается неспаренной. Обработка нуклеазой S1, специфически рас-

щепляющей одноцепочечную ДНК, гидролизует как неспаренную цепь, так и области матричной цепи за пределами транскрипционной единицы. РНК может быть удалена из гибрида РНК—ДНК. Далее можно определить последовательность ДНК или, зная ее длину, локализовать местоположение РНК на матрице.

состоит в том, что подвергают деградации всю ДНК, не гибридизовавшуюся с РНК, а затем определяют последовательность сохранившегося участка ДНК.

У бактерий 5'-конец РНК идентифицируется однозначно благодаря наличию трифосфата. Выделение мРНК *in vivo* затрудняется ее нестабильностью, но в случае рРНК и тРНК часто оказывается возможным блокировать процессинг первичного транскрипта (являющегося предшественником зрелой РНК) и таким образом идентифицировать 5'-конец. Клонированные гены можно использовать в качестве матрицы при транскрипции *in vitro*, что в результате дает возможность изолировать образующуюся РНК и с ее помощью точно локализовать соответствующий участок в ДНК. Некоторые гены были исследованы как *in vivo*, так и *in vitro*, и при этом определение стартовой точки дало одинаковые результаты. Это придает уверенность в правильности получаемых результатов в тех случаях, когда информация *in vivo* недоступна и для идентификации стартовой точки используется только система *in vitro*.

У эукариот предшественник мРНК копируется вскоре после инициирования транскрипции, и поэтому *in vivo* невозможно выделить молекулы, сохранившие первичный 5'-трифосфатный конец. В реакции копирования в качестве субстрата используется 5'-три- или 5'-дифосфат. Из этого можно сделать вывод, что сайт копирования соответствует истинной стартовой точке. Этот вывод нашел подтверждение в опытах по транскрипции *in vitro*. Оказалось, что во всех исследованных случаях первое основание транскрипта соответствует нуклеотиду, к которому *in vivo* присоединяется кэп.

Исходя из этого, стартовые точки для РНК-полимеразы II локализовали, идентифицируя те последовательности ДНК, которые соответствовали стартовой точке мРНК, образующейся *in vivo*.

Поскольку транскрипты, синтезируемые РНК-полимеразой I (рРНК) и РНК-полимеразой III (малые ядерные РНК и тРНК) не копируются, можно выделить молекулы, содержащие 5'-трифосфатный конец. Это позволяет однозначно идентифицировать стартовую точку, которая используется *in vivo* (некоторые маленькие ядерные РНК копируются и поэтому не могут быть исследованы аналогичным образом).

В органеллах процесс копирования не обнаруживается, и, следовательно, в принципе здесь нет препятствий при идентификации стартовых точек для всех видов митохондриальных и хлоропластных РНК.

Сайт связывания РНК-полимеразы *E. coli*

Прочные участки связывания, на которых РНК-полимераза (голофермент) образует стабильные иницирующие комплексы, располагаются внутри промоторов. Эти последовательности ДНК могут быть получены с помощью методики, приведенной на рис. 11.3. Как показано на рисунке, РНК-полимераза *in vitro* взаимодействует с матрицей, содержащей определенную транскрипционную единицу. Затем с помощью фермента ДНКазы гидролизуют все участки ДНК, не защищенные РНК-полимеразой.

Образующиеся фрагменты варьируют по длине от 41 до 44 пар оснований. Фактически полимеразы могут ини-

цировать транскрипцию непосредственно на экранируемом фрагменте. В результате синтезируется короткая РНК размером 17–20 оснований. Синтез этой РНК оканчивается, когда полимераза достигает конца фрагмента. Это показывает, что защищенный фрагмент является последовательностью, на которой иницируется транскрипция, а стартовая точка располагается примерно в середине этого фрагмента.

ДНК, предшествующая стартовой точке, называется последовательностью, расположенной **против хода транскрипции**; ДНК после стартовой точки (которая образует транскрибируемую последовательность) называется последовательностью, расположенной **по ходу транскрипции**. Последовательность ДНК принято записывать так, что транскрипция происходит слева направо. Это соответствует обычному написанию мРНК в направлении от 5'-конца к 3'-концу. Часто последовательность ДНК записывается только для той цепи, с которой транскрибируется РНК. Положения оснований нумеруются в обоих направлениях, начиная от стартовой точки, которую обозначают как +1. По ходу транскрипции происходит нарастание порядковых номеров оснований. Позиции оснований, расположенных против хода транскрипции, обозначаются со знаком минус, и значения номеров возрастают по мере удаления от стартовой точки. Используя эти обозначения, можно написать, что экранируемый РНК-полимеразой фрагмент расположен в пределах от -20 до +20.

Участки связывания обладают той характерной особенностью, что, будучи выделенными в чистом виде, они не могут повторно связываться с РНК-полимеразой. Это означает, что хотя данный фрагмент и является последовательностью, прочно связывающей РНК-полимеразу и, кроме того, способной иницировать транскрипцию, его одного недостаточно для изначального связывания фермента. Для осуществления процесса узнавания необходимы другие последовательности, расположенные вне этого фрагмента. Следовательно, в состав промотора входят как прочный участок связывания, так и эти добавочные последовательности. Исходя из приведенных данных, можно предположить существование двух возможных типов структур промоторов.

Возможно, весь промотор является участком связывания фермента, но дополнительные последовательности в меньшей степени ассоциированы с ферментом или менее защищены и, следовательно, чувствительны к нуклеазе, хотя прочно связанные последовательности экранируются ферментом. В этом случае разногласия между строением экранируемого фрагмента и его неспособностью к повторному связыванию фермента отражают геометрию взаимодействия РНК-полимеразы и ДНК.

Можно допустить также, что промотор состоит из двух видов последовательностей. Взаимодействие фермента с первой последовательностью, возможно, является необходимым условием для связывания со второй. При переходе от узнавания к прочному связыванию могут происходить конформационные изменения — например, укорачивание ДНК, в результате которого фермент закрывает только экранируемую область. С другой стороны, эти изменения могут быть связаны с перемещением фермента от места первичного узнавания к прочному участку связывания.

Способность РНК-полимеразы взаимодействовать с определенными участками ДНК можно исследовать с помощью **метода отпечатков**. Для этого РНК-полиме-

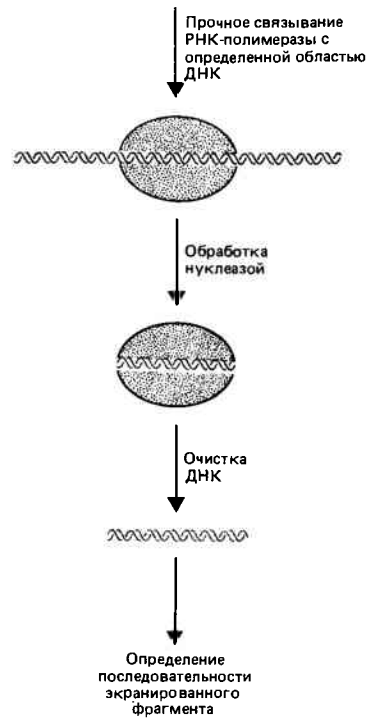


Рис. 11.3. Фрагменты ДНК, связывающие РНК-полимеразу, могут быть получены благодаря тому, что они защищены ферментом от деградации нуклеазой.

разе или другому белку дают возможность связаться с ДНК, которую затем частично гидролизуют **эндоуклеазой**.

В результате независимо гидролизуются различные фосфодиэфирные связи нуклеиновой кислоты. При определенных условиях каждая фосфодиэфирная связь подвергается действию нуклеазы, но при этом гидролиз данной связи происходит только в ограниченном числе молекул ДНК. Если РНК-полимераза блокирует доступ нуклеазы к ДНК, то определенные связи вообще не будут разорваны.

Местоположение подвергшихся гидролизу связей определяют, вводя метку в одну из цепей ДНК, причем только в один ее конец. Как показано на **рис. 11.4**, фрагменты ДНК, полученные после нуклеазной обработки, разделяют по размеру при помощи электрофореза в геле. (Принцип, лежащий в основе *частичного* расщепления каждого гидролизующего участка с последующим выделением меченных по концу фрагментов, тот же самый, который был использован при определении последовательности ДНК; см. рис. 3.4.) Каждой гидролизованной связи на геле соответствует полоса, подвижность которой определяется расстоянием от меченого конца до места разрыва. В участке, защищенном от гидролиза РНК-полимеразой, разрывов не происходит и соответствующие полосы на геле отсутствуют. Каждая из двух цепей ДНК, меченных радиоактивным изотопом, может быть проанализирована по отдельности. Сравнивая полученные отпечатки с результатами определения последовательности ДНК, устанавливают нуклеотидную последовательность участка связывания и его местоположение.

В результате такого анализа одного из промоторов (для *lac*-генов) было показано, что участок связывания РНК-полимеразы более обширный, чем тот, который обнаруживается при непосредственном выделении экрани-

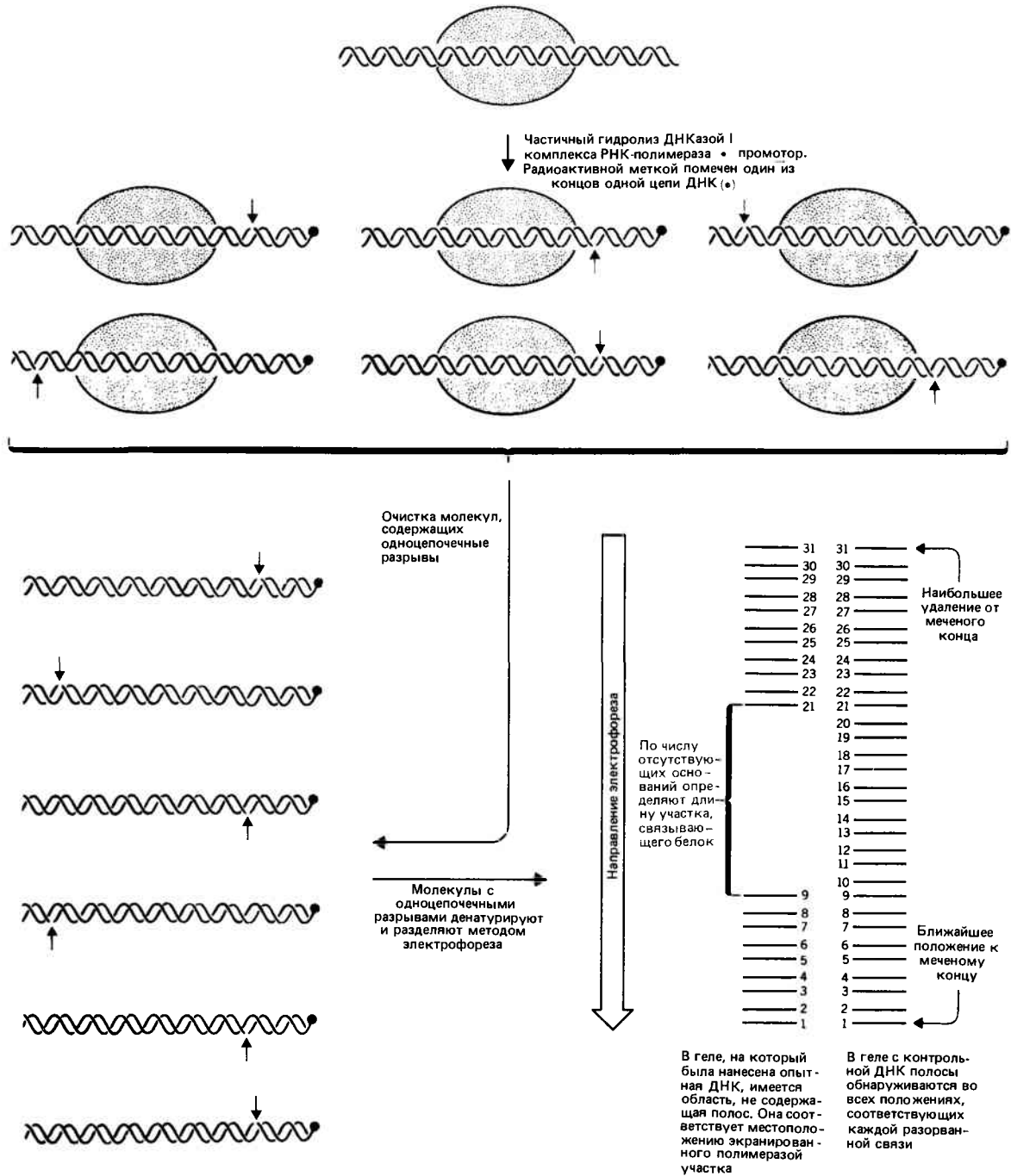


Рис. 11.4. Исследование участков ДНК, связывающих белки, с помощью метода «отпечатков».

Принцип метода состоит в том, что в условиях частичного гидролиза каждая фосфодиэфирная связь в свободной ДНК разрушается в какой-либо из молекул. Участок ДНК, экранированный связавшимся белком, защищен от нуклеазы во всех молекулах. Параллельно проводятся две реакции с контрольной, очищенной ДНК и с опытной, взаимодействовавшей с белком.

Затем, если цепи ДНК разделить и подвергнуть электрофоретическому анализу, то образуются радиоактивные полосы, соответствующие фрагментам ДНК с мечеными концами. Местоположение фрагмента соответствует числу содержащихся в нем оснований. Более короткие фрагменты обладают большей подвижностью, поэтому последовательность читается в направлении от основания геля вверх (см. рис. 3.4). В контроле все

связи гидролизуются, образуя серию полос, соответствующих каждому основанию. На рисунке можно насчитать 31 полосу. В экранированном фрагменте связи не могут быть гидролизованы, поэтому полосы определенного размера, соответствующие фрагменту, связанному с белком, не образуются. На рисунке полосы от 10 до 20 отсутствуют. Следовательно, сайт, связывающий блок, располагается от 10-го до 20-го основания, считая от меченого конца ДНК.

Можно изменить условия опыта и разделять фрагменты ДНК не только по длине, но и одновременно определять их нуклеотидную последовательность. Для этого контрольную и опытную ДНК можно обработать, как показано на рис. 3.5, и далее приготовить по четыре геля для определения нуклеотидной последовательности ДНК. Сравнение двух наборов гелей позволяет прочитать последовательность ДНК непосредственно в том участке, который связывает белок. Коричневыми стрелками показано действие ДНКазы. Стрелками показан гидролиз ДНК-азой.

руемых фрагментов. Размеры этого участка составили около 50 п.н., если считать от стартовой точки против хода транскрипции, и около 20 п.н.—по ходу транскрипции. Обе цепи ДНК не одинаково хорошо защищены—в особенности на концах связывающего участка. Из этого следует, что РНК-полимераза, связываясь с ДНК, находится в асимметричной конформации. Приведенные факты согласуются с данными о том, что при транскрипции считывается только одна цепь ДНК.

Похожие результаты могут быть получены в экспериментах с использованием экзонуклеазы. Этот фермент действует на концевой участок ДНК и последовательно его деградирует. Если с ДНК связан белок, то при столкновении с ним действие нуклеазы прекращается. В данном случае также можно определить границы промотора, т.е. точки, расположенные по обе стороны от него, далее которых экзонуклеаза не может продвинуться. Таким образом, еще раз подтвердилось, что последовательность, связывающаяся с РНК-полимеразой, расположена примерно на 44 п.н. левее стартовой точки транскрипции.

В совокупности эти результаты показывают, что РНК-полимераза связывается, причем асимметрично, с участком ДНК размером около 44–50 п.н., если считать от стартовой точки против хода транскрипции, и около 20 п.н.—по ходу транскрипции. Последние 25–30 п.н. в связывающем участке, расположенные против хода транскрипции, менее прочно взаимодействуют с ферментом. Эта область связывается с ферментом достаточно хорошо, чтобы быть экранированной при мягкой обработке эндо- или экзонуклеазами, но не может выдерживать более жесткие условия обработки, используемые для получения целых фрагментов связанного ферментом участка. Приведенные данные хорошо согласуются с моделью, согласно которой существует единый связывающий сайт для РНК-полимеразы, в котором (как мы увидим) некоторые точки контакта имеют более важное значение, чем другие.

Консервативная последовательность в промоторах *E. coli*

Одним из путей конструирования промотора является, очевидно, сборка уникальной последовательности ДНК, способной создать сигнал для связывания РНК-полимеразы. Исходя из размеров бактериального генома, можно думать, что минимальная длина участка, достаточная для создания необходимого сигнала, составляет 12 п.н. Более короткие последовательности, вероятно, встречаются довольно часто (просто в результате случайных событий) и могут создавать ложные сигналы. Последовательность из 12 п.н. не обязательно должна быть непрерывной. Так, если определенное число пар оснований разделяет две короткие постоянные последовательности, их совместная длина может составлять менее чем 12 п.н., так как само разделяющее расстояние создает часть сигнала (даже если разделяющая последовательность не обладает никакой специфичностью).

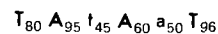
Были предприняты попытки определить характерные особенности ДНК, необходимые для связывания РНК-полимеразы. Для этого сравнивали последовательности различных промоторов. Любая необходимая для связывания нуклеотидная последовательность должна присутствовать во всех промоторах. Такую последовательность принято называть **консервативной**. Однако в результате

анализа большого количества промоторов *E. coli* (более 50) была обнаружена удивительная особенность: среди 60 п.н., взаимодействующих с РНК-полимеразой, не обнаружено сколько-нибудь протяженной консервативной последовательности, иными словами, большая часть последовательности участка связывания может быть необязательной. Но некоторые короткие участки внутри промотора, очевидно, являются консервативными. Они-то и могут быть теми сигналами, которые узнаются РНК-полимеразой.

Обычно, но не всегда, стартовой точкой является пуриновое основание (А встречается в два раза чаще, чем Г). По-видимому, в более редких случаях, когда обнаруживается пиримидин, РНК-полимераза способна преодолеть трудность, связанную с включением этого основания в качестве иницирующего нуклеотида. Для стартовой точки характерно то, что она находится в середине последовательности САТ, но консервативность этого триплета не настолько выражена, чтобы считать его обязательным сигналом.

Против хода транскрипции от стартовой точки располагается область из 6 п.н., которая может быть обнаружена почти во всех промоторах. Эта последовательность не является строго постоянной, но всегда имеет большое сходство с последовательностью ТАТААТ, которая часто называется **блоком Прибнова**. Фактически как таковая последовательность ТАТААТ встречается довольно редко, но она является **среднестатистической**, или **канонической последовательностью**, составленной из оснований, наиболее часто встречающихся в каждой позиции.

Каноническая последовательность может быть представлена в следующем виде

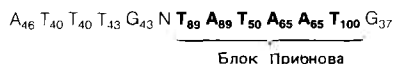


где цифрами показана частота (процент) встречаемости оснований, наиболее обычная для данного положения. Заглавные буквы используются для обозначения оснований, присутствующих в данном положении более чем в 54% случаев, а маленькими буквами—основания, не столь консервативные, но встречающиеся чаще, чем при случайном распределении. То положение, в котором не дается преимущества какому-либо основанию, обозначено как N.

Обнаружение этой последовательности позволило сделать важное обобщение. Стало понятно, что полезно рассматривать возможные регуляторные сайты в ДНК, имея в виду идеализированные последовательности, составленные из наиболее часто встречающихся оснований. Среднестатистическая последовательность выводится на основе уравнивания всех известных образцов с получением максимальной гомологии. Для последовательности, которая выведена как среднестатистическая, каждое конкретное основание должно наиболее часто встречаться в своем положении (например, оно может присутствовать в более чем 60% случаев), причем большинство реальных последовательностей должно отличаться от среднестатистической на небольшое число замен—например, на 1 или 2.

Степень консервативности каждого основания в определенном положении в некоторой степени варьирует. Если в состав рассматриваемой области включать основания, встречающиеся в небольшом числе промоторов, например в 40% случаев, то ее можно продолжить за пределы блока Прибнова. Имеющиеся данные можно сум-

мировать следующим образом:



Цифрами показан процент встречаемости наиболее часто обнаруженного основания; N означает, что в данном положении может находиться любое основание.

Между блоком Прибнова и окружающими последовательностями существует вполне четкое разграничение. Если частота встречаемости оснований указывает на их важное значение для взаимодействий с РНК-полимеразой, то мы вправе ожидать, что высококонсервативный димер ТА в начале блока Прибнова и полностью консервативный Т в конце являются наиболее важными основаниями. Прочие основания в блоке Прибнова, обладающие промежуточной степенью консервативности, вероятно, также вовлечены в процесс взаимодействия с ферментом. Следующие несколько позиций, находящихся против хода транскрипции, имеют склонность быть обогащенными А–Т-парами, которые, возможно, способствуют взаимодействию, но не являются абсолютно необходимыми.

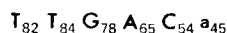
Среднестатистическая последовательность блока Прибнова состоит исключительно из пар А–Т. Это наводит на мысль о том, что ее функционирование может быть связано с плавлением и образованием одноцепочечных участков. Для разрыва двух водородных связей в каждой паре А–Т требуется меньше энергии, чем для разрыва трех водородных связей в паре G–C. Из этого следует, что для расхождения цепей в участке, состоящем из пар А–Т, требуется минимальное количество энергии. В действительности же большинство блоков Прибнова, идентифицированных в бактериальных промоторах, содержат одну пару G–C, поэтому в реакции плавления обычно расходуется больше энергии, чем то минимальное количество, которое возможно теоретически.

Логично предположить, что среднестатистическая последовательность должна наилучшим образом выполнять функции узнающего участка. Однако окружающие основания также могут влиять на это узнавание. Например, две трансверсии от А–Т к Т–А могут компенсировать одна другую, создавая другую последовательность, в такой же мере эффективную, как и среднестатистическая. Поэтому бывает часто затруднительно предсказать эффективность конкретной последовательности по сравнению со среднестатистической.

Центр блока Прибнова обычно располагается примерно на расстоянии 10 п.н. от стартовой точки против хода транскрипции. Поэтому иногда этот блок обозначают просто как **положение –10**. В действительности положение гексануклеотидного участка варьирует от –11 до –5 или от –14 до –8, причем последний интервал встречается чаще.

Сходная последовательность обнаруживается в другом положении с центром около нуклеотидного остатка –35.

Этот участок называется **последовательностью –35**. Поскольку эта последовательность представляет собой часть той последовательности, которая должна узнаваться РНК-полимеразой, но собственно не входит в область прочного связывания, то иногда ее называют *областью узнавания*. Эта последовательность –TTGACA; более детально она описывается так:



В исследованных промоторах расстояние между положениями –35 и –10 варьирует от 16 до 19 п.н. Хотя реально, может быть, и неважно, какая последовательность находится между этими областями, ее размеры могут иметь критическое значение для того, чтобы взаиморасположение рассматриваемых участков соответствовало геометрии РНК-полимеразы. Делеции, уменьшающие промежуток до 15 п.н., или вставки, увеличивающие расстояние до 20 п.н., снижают активность промоторов, в которых они возникают. Имеется ряд промоторов, не содержащих специальной последовательности в положении –35 (это определяют по отсутствию гомологии). Однако обычно такие промоторы содержат расположенный рядом участок, который узнается регуляторным белком, способствующим инициации. Поэтому возможно, что регуляторный белок заменяет функцию положения –35. Но поскольку другие промоторы, в которых наблюдается такая же регуляция транскрипции, сохраняют положение –35, наблюдаемая корреляция не абсолютна. Вероятно, положение –35 в какой-то степени определяет эффективность, с которой РНК-полимераза узнает определенный промотор.

Обнаружено также очень небольшое количество промоторов, лишенных положения –10. Таким образом, ни одна из консервативных последовательностей не является абсолютно необходимой для функционирования промотора. Нам пока еще не известны все детали реакции узнавания, и сейчас мы с уверенностью можем сказать только одно: «типичный» промотор использует положения –35 и –10 для взаимодействия с РНК-полимеразой. Поскольку у ряда промоторов рассматриваемые последовательности отсутствуют, можно предполагать существование других способов узнавания, однако, сколько существует таких механизмов и как они восполняют отсутствие среднестатистических последовательностей, пока не известно.

Промоторные мутации, усиливающие и ослабляющие экспрессию генов

Мутации в промоторах, влияя на уровень экспрессии контролируемого гена (генов), который находится под их контролем, не изменяют кодируемых генами продуктов. Большинство таких мутаций было обнаружено в результате выделения бактериальных мутантов, у которых способность к транскрипции смежных с промотором генов была либо совсем утрачена, либо находилась на очень низком уровне. Такие мутации называют **ослабляющими**. Реже обнаруживаются мутации, способствующие транскрипции, иницируемой с промотора. Это так называемые **усиливающие** мутации.

Удивительно, что оба типа мутаций могут возникнуть в результате изменений, затрагивающих только одну пару оснований. Следовательно, взаимодействие РНК-полимеразы с этими участками в промоторе должно быть высокоспецифичным. Рассматриваемое основание должно либо узнаваться ферментом, либо непосредственно участвовать в процессе инициации. (Очевидно, что мутации, ослабляющие экспрессию, могут возникать также в результате делеции более протяженных участков промотора.)

Информация, полученная в результате экспериментов *in vivo*, лишь указывает в целом на природу изменения, обусловленного мутацией. Чтобы проанализировать ее действие на молекулярном уровне, необходимо охарактер-

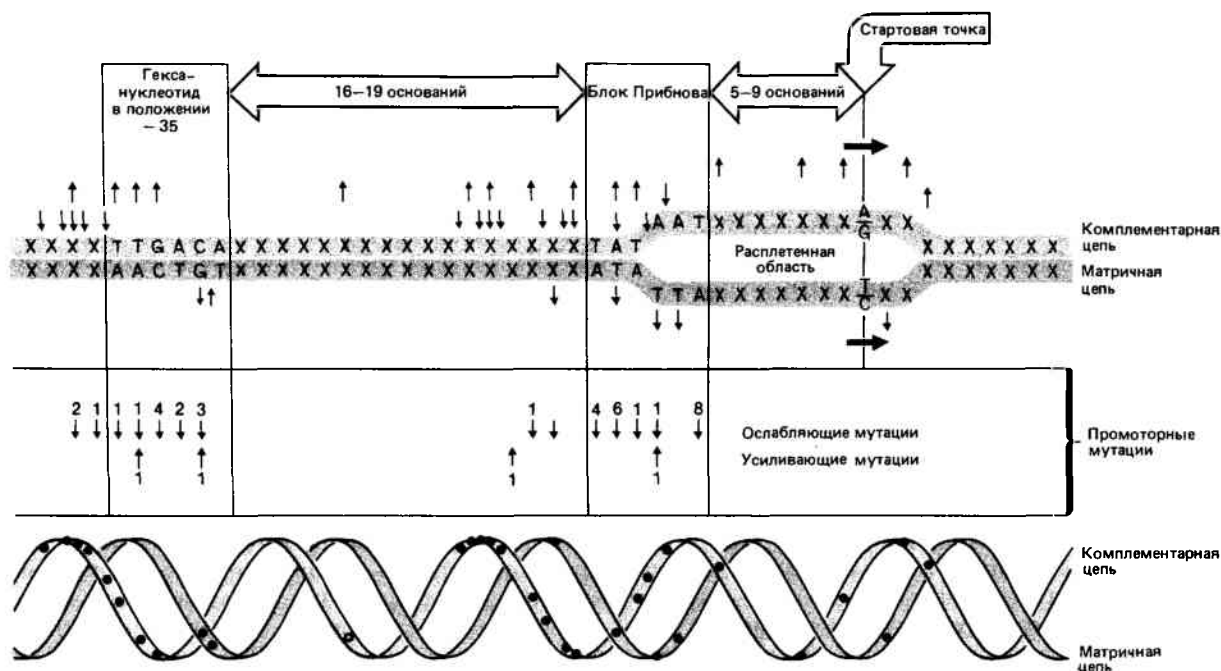


Рис. 11.5. Точки контакта для РНК-полимеразы находятся на одной поверхности ДНК, образующей промотор.

Последовательность ДНК представляет типичный промотор с консервативными последовательностями в положениях -35 и -10 . Область первоначального разделения цепей начинается от блока Прибнова и заканчивается сразу за стартовой точкой.

В верхней части рисунка стрелками, направленными к двойной спирали, указаны сайты, модификация которых нарушает связывание РНК-полимеразы с ДНК. Сайты, экранируемые ферментом от модификаций, обозначаются стрелками, направленными от двойной спирали. Стрелки, направленные на основание, указывают само модифицированное основание. Стрелки между основаниями обозначают модификацию фосфодиэфирной связи.

ризовать взаимодействие между РНК-полимеразой и мутантным промотором *in vitro*. При этом можно показать, на какое свойство промотора влияет мутация. Может ли она изменять сродство промотора к РНК-полимеразе? Может ли она, не нарушив способности фермента, связываться с ДНК, помешать инициации? Изменяется ли под влиянием мутаций действие регуляторных факторов?

На рис. 11.5 обобщены данные по локализации точковых мутаций, влияющих на функцию промотора. Почти все они попадают в две области, соответствующие положениям -35 и -10 . Основания, расположенные в других положениях, очевидно, имеют менее важное значение или даже безразличны в случае подавляющего большинства (но не всех) промоторов.

Распределение мутаций внутри блока Прибнова с центром в положении -10 подтверждает важную роль трех консервативных оснований. Большинство (18 из 20) мутаций, ослабляющих транскрипцию, возникает в каком-либо из этих положений. Любая замена в димере ТА, расположенном в начале блока Прибнова или в последнем Т, приводит к возникновению таких мутаций. Почти все эти мутации представляют собой замены пар А-Т на пары Г-С. Это служит еще одним доводом в пользу вывода о важном значении высокой концентрации пар А-Т. Однако незначительное число мутаций, возникших в результате трансверсий от А-Т к Т-А или наоборот, показывает, что отсутствие пар Г-С не единственное условие, определяющее функцию промотора. Например, один

В средней части рисунка маленькими стрелками показаны сайты, влияющие на функционирование промотора. Стрелки, направленные вниз, обозначают ослабляющие мутации, а стрелки, направленные вверх, — мутации, усиливающие иницирование транскрипции. Показано суммарное число мутаций, обнаруженных для каждого положения в случае всех изученных промоторов.

В нижней части рисунка представлена диаграмма, изображающая двойную спираль ДНК, спроектированную на плоскость. Как видно из рисунка, все точки контакта расположены на одной поверхности. Большинство находится на комплементарной цепи (той, которая не является матричной).

промотор в начале блока Прибнова вместо обычного ТА содержит АА, но в другом промоторе, имеющем ТА, мутационная замена, приводящая к появлению АА, проявляется в виде мутации, ослабляющей транскрипцию. Таким образом, основания должны действовать совместно; окружающие последовательности также играют важную роль в определении значения каждой позиции.

Мутация, локализуемая в области блока Прибнова и усиливающая транскрипцию, обнаружена в *lac*-промоторе. Промотор дикого типа имеет структуру ТАТГТТ. Мутация, усиливающая транскрипцию, приводит к возникновению последовательности ТАТАТТ. Таким образом, ее эффект состоит в увеличении степени гомологии со среднестатистической последовательностью. Должна существовать какая-то причина, в силу которой это увеличение эффективности не произошло спонтанно и не было закреплено отбором у бактерий. (Очень небольшое число других мутаций, усиливающих транскрипцию, было обнаружено в участке -10 . Все они возникают в промоторах, у которых гомология с блоком Прибнова слишком незначительна, чтобы можно было понять механизм действия этих мутаций. Однако вполне возможно, что они повышают степень гомологии со среднестатистической последовательностью.)

Непосредственно слева от блока Прибнова (но еще в частично консервативной области -10) возникают другие мутации, способные как усилить, так и ослабить экспрессию гена. Место возникновения таких мутаций свиде-

тельствует о том, что рассматриваемая часть промотора взаимодействует с РНК-полимеразой таким образом, что индивидуальное основание играет важную роль в данном промоторе, хотя они могут отличаться в других промоторах.

Большинство мутаций в положении — 35 представляет собой замены, в результате которых в промотор вводится основание, неприемлемое в данном промоторе, хотя и вполне обычное для других промоторов. Это еще раз подчеркивает важную роль эффекта, создаваемого окружающими последовательностями: РНК-полимеразе требуется не присутствие конкретного основания в каждом положении, а наличие определенного сочетания оснований, локализованного в определенных участках. Некоторые основания никогда не обнаруживаются в определенных положениях, и возможно, что их отсутствие имеет более важное значение для процесса узнавания, чем наличие любых других оснований.

Влияние мутаций в положении — 35 трудно интерпретировать, поскольку большинство из них возникает в промоторах, для связывания с которыми необходим дополнительный регуляторный фактор. Поэтому вполне возможно, что данная мутация влияет на взаимодействие этого фактора с ДНК, а не на саму РНК-полимеразу (см. далее). Более равномерное распределение мутаций в этой области говорит о том, что существует меньшая концентрация определенных позиций внутри положения — 35 или в ближайшем его окружении, любая из которых более или менее одинаково встречается в других промоторах.

При анализе всех этих вопросов необходимо помнить, что усиливающие и ослабляющие мутации определяются относительно *обычной* эффективности, с которой работает промотор. Эффективность различных промоторов варьирует в широких пределах. Поэтому изменения, выявляемые как ослабляющие транскрипцию в одном промоторе, в другом могут быть не обнаружены, причем такой промотор даже в отсутствие мутаций может быть менее эффективен, чем мутантная форма промотора первого типа. Поэтому для определения истинного эффекта этих мутаций необходимо проводить исследования *in vitro*, что дает возможность изучать сродство РНК-полимеразы к различным промоторам дикого и мутантного типов в одинаковых условиях. *In vitro* наблюдаются приблизительно стократные различия в степени сродства РНК-полимеразы к различным промоторам. Этот разброс коррелирует с частотой транскрипции многочисленных генов *in vivo*.

Основные точки контакта в промоторе

Точки, с которыми РНК-полимераза контактирует, находясь на промоторе, можно определять, обрабатывая комплекс фермент-промотор реагентами, модифицирующими определенные основания. Присутствие фермента может увеличивать или уменьшать доступность определенного основания по сравнению с его доступностью в составе ДНК, не контактирующей с РНК-полимеразой. Изменение чувствительности оснований отражает геометрию комплекса и может быть использовано для описания его формы.

Реагент диметилсульфат (ДМС) специфически взаимодействует с пуринами, метилируя остатки G в большой

бороздке и остатки A в малой бороздке ДНК. В результате метилирования пурины становятся чувствительными к нагреванию. Поэтому после прогрева ДНК лишается основания, на месте которого остается брешь; в этом положении фосфодиэфирная цепь может быть разорвана в присутствии щелочи. (Данная реакция используется при определении последовательности ДНК, как это было описано в гл. 3.)

Если тимин в ДНК заменяется на бромурацил (BUdR), то в результате ультрафиолетового облучения высвобождается бром. Это приводит к появлению свободной радикала, разрывающего цепь ДНК.

Этилнитрозомочевина непосредственно атакует каркас ДНК. В результате фосфодиэфирная связь, соединяющая два нуклеотида, превращается в фосфотриэфирную (в результате алкилирования). Эта связь чувствительна к щелочи, и поэтому в данном месте ДНК может быть разорвана.

Общее свойство всех этих модификаций состоит в том, что они дают возможность разорвать соответствующую связь в полинуклеотидной цепи. Такой сайт можно идентифицировать с помощью тех же подходов, которые использовались в экспериментах по определению участков связывания РНК-полимеразы (рис. 11.4). Одну из цепей ДНК метят по концу; тогда в результате каждого разрыва образуется фрагмент, который обнаруживается при электрофоретическом анализе как полоса в геле, соответствующая определенному размеру. Используя такой подход, сравнивают чувствительность комплекса РНК-полимераза-промотор с чувствительностью свободной ДНК. В результате оказывается, что ряд полос пропадает. Таким путем выявляются участки промотора, экранированные ферментом от модификаций. Интенсивность ряда других полос может усиливаться, выявляя тем самым участки, в которых ДНК должна находиться в более доступной конформации.

Эксперимент с использованием метилирующего агента ДМС может быть проведен в обратном порядке. Сначала можно прометилировать ДНК, а затем воздействовать на нее РНК-полимеразой. При этом образуется набор молекул ДНК, метилированных в некоторых, но не во всех возможных положениях. Те молекулы ДНК, которые не смогли провзаимодействовать с РНК-полимеразой, подвергаются обычной обработке. В результате образуются фрагменты, размеры которых при электрофоретическом разделении соответствуют местам разрывов. Это дает возможность локализовать те точки, в которых предшествующая модификация *предотвращает* связывание РНК-полимеразы с ДНК. Аналогичные эксперименты можно осуществить, проводя алкилирование фосфодиэфирного каркаса ДНК.

В результате этих экспериментов было показано, что большинство точек, по которым происходит взаимодействие фермента с ДНК, содержится в областях — 35 и — 10. В пределах этих областей одни и те же группы оснований, будучи способными предотвращать связывание, если они предварительно были модифицированы, проявляют повышенную или пониженную чувствительность к модификации после связывания фермента. На рис. 11.5 точки контакта сравниваются с сайтами, затронутыми мутациями. Хотя они не совпадают полностью, но присутствуют в одних и тех же небольших областях. В этих двух участках общая длина контактирующей области составляет 12–15 п.н., что немного больше максимальных размеров консервативных участков.

Особенно интересно, что тождественные *позиции* в различных промоторах могут являться точками контакта, несмотря на то что содержат различные основания. Из этого, очевидно, можно сделать вывод о существовании общего механизма связывания РНК-полимеразы, хотя реакция не зависит от наличия определенных оснований в какой-либо из точек контакта. Эта модель может объяснить, почему некоторые точки контакта не совпадают с сайтами, подверженными мутированию. Не каждая мутация также затрагивает точку контакта; может быть, в данном случае сказывается некоторое влияние соседних оснований, которые фактически не соприкасаются с ферментом?

Узнавание промоторов и расплетание двойной спирали ДНК

Особенно важно, что эксперименты с предварительной модификацией и с экранированием в присутствии фермента выявили в одной и той же области *одинаковые сайты*. Эти два экспериментальных подхода дают возможность исследовать образование разных контактов. В первом случае выявляются все те сайты, которые фермент должен узнать в процессе связывания с ДНК. Во втором определяются все те сайты, которые действительно образуют контакты в бинарном комплексе. Поэтому если имеются какие-либо контакты, участвующие в первичном узнавании, но в дальнейшем не существенные для поддержания устойчивого связывания, то они были бы обнаружены как сайты, в которых предварительная модификация предотвращает связывание с ферментом, но которые не экранируются в составе бинарного комплекса. Поскольку сайты с такими свойствами не были обнаружены, а экранированные сайты включают все те, которые участвуют в узнавании, и еще ряд дополнительных позиций, то, по-видимому, фермент сначала узнает некий набор сайтов, необходимый для своей «посадки», и только после этого образует более обширные контакты.

Другой тип экспериментов по модификации позволяет непосредственно обнаружить расплетенную область ДНК, входящую в состав бинарного открытого комплекса. При разделении цепей ДНК неспаренные аденины становятся чувствительными к дополнительному метилированию в присутствии ДМС. Если затем фрагмент ДНК освободить от фермента, расплетенная область ренатурирует, за исключением метилированных остатков аденина, которые теперь содержат метильные группы в положении N 1, принимающем обычно участие в комплементарном взаимодействии оснований. Метильные группы мешают основаниям А узнавать своих партнеров Т. Это несовершенство в спаривании можно обнаружить при помощи фермента нуклеазы S1, которая специфически расщепляет ДНК в любом одноцепочечном участке. Как обычно, результаты такого расщепления анализируют электрофоретически, определяя размеры образующегося фрагмента, особенно по концу.

В результате такого эксперимента образуется набор полос с размерами, соответствующими положениям от -9 до $+3$. Эти данные показывают, что область, расплетаемая в реакции инициирования, включает участок, начинающийся от правого конца блока Прибнова и заканчивающийся как раз за стартовой точкой. Возможно, само расплетение начинается непосредственно внутри блока Прибнова, но здесь его не удается зафиксировать.

(Оценка размеров расплетенного участка по степени раскручивания показала, что он составляет менее 17 п.н., см. также гл. 10.)

При рассмотрении в трех измерениях все точки контакта, находящиеся против хода транскрипции от блока Прибнова, располагаются на одной и той же поверхности ДНК (рис. 11.5). Эти основания, вероятно, узнаются при первоначальном образовании закрытого бинарного комплекса. Это создает возможность для реакции узнавания между РНК-полимеразой и ДНК путем образования контакта между ферментом и одной из поверхностей ДНК, что является простейшей из возможных моделей. После того как расплетание ДНК началось, возможно узнавание и связывание со следующими сайтами, расположенными на другой стороне ДНК.

Мало что известно о точках контакта, имеющих в составе фермента, но РНК-полимераза, находящаяся на промоторе, может быть ковалентно пришта к ДНК, в которой тимин замещен на бромурацил. При этом было показано, что σ -субъединица (сигма-субъединица) взаимодействует с нематричной цепью в положении -3 , тогда как β -субъединица контактирует с положением $+3$. Такое расположение σ -субъединицы, видимо, дает ей возможность принимать участие в поиске промотора и (или) в стабилизации первоначально расплетенной области. Но нам не известно, участвует ли σ -субъединица в этом непосредственно, или ее функция опосредована взаимодействием с коферментом.

О важном значении расплетания цепей при инициировании транскрипции свидетельствует эффект, оказываемый суперспирализацией на матричные свойства ДНК. Как мы упоминали в гл. 2, введение отрицательных сверхвитков в ковалентнозамкнутую молекулу ДНК (т.е. в молекулу, у которой нет свободных концов) сопряжено с расплетанием двойной спирали. В последнее время стало известно, что как прокариотические, так и эукариотические РНК-полимеразы *in vitro* более эффективно иницируют транскрипцию на суперспирализованных матрицах, чем на релаксированных ДНК. Возможно, это объясняется тем, что суперспирализованная структура требует меньше энергетических затрат для первоначального плавления ДНК в иницирующем комплексе. Значение этого эффекта для контроля активности промоторов у бактерий обнаруживается при нарушениях в работе ферментов, влияющих на степень суперспирализации. К таким ферментам относятся ДНК-гираза, *вносящая* отрицательные сверхвитки, и топоизомераза I, релаксирующая (устраняющая) отрицательные сверхвитки (гл. 32). Добавление ингибиторов ДНК-гиразы подавляет транскрипцию. Обратный эффект, усиливающий транскрипцию, наблюдается в присутствии мутаций, инактивирующих топоизомеразу I. Оба этих эффекта проявляются только в отношении некоторых промоторов.

Все это свидетельствует о том, что степень отрицательной суперспирализации ДНК оказывает влияние на эффективность работы промоторов. Ингибирование ДНК-гиразы нарушает процесс суперспирализации ДНК, подавляя в результате экспрессию определенных транскрипционных единиц. Сначала думали, что мутация в топоизомеразе I повышает экспрессию генов, поскольку предотвращает снижение степени суперспирализации под действием фермента. Однако вполне возможно, что это действие обусловлено наличием дополнительных мутаций. В целом взаимосвязь между работой топоизомеразы и транскрипцией еще недостаточна ясна.

Почему одни промоторы в отличие от других чувствительны к суперспирализации? Одна из возможностей состоит в том, что зависимость каждого промотора от степени суперспирализации определяется его нуклеотидной последовательностью. Это означало бы, что некоторые промоторы имеют последовательности, которые плавятся более легко (а поэтому в меньшей степени зависят от суперспирализации), тогда как другие обладают более тугоплавкими последовательностями (и поэтому они в большей степени зависят от суперспирализации). Однако вполне возможно, что важную роль играет местоположение промотора, если разные области бактериальной хромосомы характеризуются различной степенью суперспирализации.

Позитивная регуляция работы промоторов

Итак, мы определили промотор как последовательность ДНК, способную связывать РНК-полимеразу и затем инициировать транскрипцию. Но существует ряд промоторов, на которых РНК-полимераза не способна инициировать транскрипцию в отсутствие вспомогательных регуляторных белков. Такие белки принято называть **позитивными регуляторами**, так как их присутствие необходимо для активирования единицы транскрипции. Известно несколько позитивных регуляторов. Некоторые из них — это фагоспецифические белки, другие присутствуют в клетке-хозяине. Необходимо сразу сказать, что мы, в сущности, не понимаем природы различий между промоторами, способными функционировать, *per se* и теми, которым необходимы позитивные регуляторы.

Ряд локусов у *E. coli* (а также у других бактерий) подвержен позитивному контролю, при котором для их экспрессии необходима высокая концентрация циклического АМР. Действие этого нуклеотида состоит в активации белка БАК (САР) или, как его еще называют, СРР^{*1}-фактора. (Окончательный выбор названия еще не сделан.) Этот белок, имеющий важное значение для транскрипции тех промоторов, которые он контролирует, представляет собою димер с мол. массой 45 000 дальтон. В гл. 15 мы более подробно обсудим роль этого белка в координации экспрессии генов.

БАК непосредственно связывается с ДНК, и комплекс сАМР·БАК·ДНК можно выделить для любого промотора, работающего в присутствии этого белка. Получены мутации по *lac*-оперону, локализующиеся в сайте связывания, которые делают транскрипцию *in vivo* независимой от БАК. Эти мутации также предотвращают связывание БАК с ДНК *in vitro*. В каждом промоторе сайт связывания БАК находится слева от РНК-полимеразного сайта. Во всех БАК-связывающих сайтах можно обнаружить предполагаемую 11-членную среднестатистическую последовательность, содержащую различные отклонения. Но вызывает удивление тот факт, что в ряде промоторов БАК-связывающие сайты располагаются на различном расстоянии от стартовой точки. Это затрудняет создание единой модели действия БАК.

В *lac*-опероне область ДНК, экранируемая БАК, простирается приблизительно от положения — 72 до — 52.

Вероятно, что два димера белка БАК связываются с ДНК. Результаты проведенных экспериментов показывают, что БАК преимущественно располагается с одной стороны двойной спирали ДНК, причем с той самой, которая связывает РНК-полимеразу. Это позволяет расположить два белка в непосредственной близости друг от друга.

Однако в *gal*-опероне БАК-связывающая зона находится между положениями — 50 и — 23. Последовательность, расположенная слева от этой зоны, не нужна для связывания. По всей видимости, этот участок связывает только один БАК-димер, и он, вероятно, непосредственно контактирует с РНК-полимеразой, так как БАК-связывающий сайт перекрывается с участком, экранируемым РНК-полимеразой.

В *ara*-опероне БАК-связывающая зона для единственного БАК-димера находится еще дальше от стартовой точки и располагается между положениями — 107 и — 78. В данном случае БАК не может контактировать с РНК-полимеразой потому, что *другой* регуляторный белок связывается с областью, расположенной между сайтами связывания БАК и РНК-полимеразы.

Одна из моделей, описывающая действие БАК, утверждает, что оно состоит в белок-белковом взаимодействии, опосредованном связыванием с ДНК. Незначительного взаимодействия между двумя белками будет достаточно для достижения существенного увеличения сродства РНК-полимеразы к промотору. Геометрия взаимодействия БАК и РНК-полимеразы в случае *lac*- и *gal*-оперонов, по-видимому, должна быть различной, а в случае *ara*-оперона необходимо постулировать участие трех белков.

Другая возможность состоит в следующем: эффект, оказываемый БАК, полностью проявляется в результате его связывания с ДНК. Возможно, что в БАК-зависимых промоторах РНК-полимераза не может осуществлять первоначальное плавление ДНК. Это мог бы сделать БАК таким образом, чтобы расхождение цепей началось от места его связывания и продолжилось бы на соседние участки. С этой моделью согласуются данные о том, что мутации в топоизомеразе I, описанные в последнем разделе, устраняют зависимость этих оперонов от БАК, позволяя им функционировать без его помощи. Возможно, что увеличение суперспирализации (и связанное с этим расхождение цепей) представляет собой альтернативу действию БАК.

С помощью рентгеноструктурного анализа кристаллов БАК удалось обнаружить, что он состоит из двух доменов. Возможно, что малый С-концевой домен (~ 65 аминокислот) связывается с ДНК, поскольку известно, что большой N-концевой домен (~ 135 аминокислот) связывает сАМР. В одной из моделей строения С-концевого домена постулируется, что он способен образовывать надлежащие контакты только с ДНК, находящейся в *левозакрученной* В-конформации. Если в результате связывания БАК с ДНК возникает левозакрученный участок рядом с промотором, это вызывает образование расплетенной области, внутри которой происходит переход в обычную правозакрученную форму. Этому (предполагаемому) эффекту содействует суперспирализация.

Различные опероны, контролируемые БАК, различаются по степени зависимости от этого белка. Как это связано со структурой конкретного промотора, не ясно, но одно из предположений состоит в том, что эта зависимость коррелирует с наличием среднестатистической последовательности — 35. В ее отсутствие степень зависи-

¹ БАК (САР) — белковый активатор катаболизма, а БРЦ (СРР) — белковый рецептор цАМР.

мости от БАК возрастает. Возможно, взаимодействие РНК-полимеразы с положением -35 и БАК с ДНК выполняют родственные функции при иницировании транскрипции.

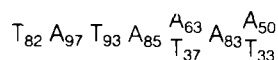
Возможные консервативные последовательности для РНК-полимеразы II

Первые попытки обнаружить сходство в эукариотических промоторах были предприняты при сравнении нуклеотидных последовательностей, расположенных против хода транскрипции от стартовых точек у различных генов. Поскольку РНК-полимераза I транскрибирует только гены, кодирующие идентичные рРНК, то в данном случае нет возможности вычленить последовательность промотора. В случае РНК-полимеразы III число изученных генов, выделенных из одного источника, было недостаточным (правда, все равно оказалось, как мы увидим далее, что промотору не обязательно располагаться непосредственно слева от кодирующей области). В случае РНК-полимеразы II была определена последовательность множества генов (нескольких сотен) из разных источников и охарактеризованы их стартовые точки. Исходили из предположения, что первое основание мРНК соответствует сайту инициации, а не образуется в результате процессинга большего предшественника. В нескольких случаях правильность этого предположения была строго доказана по совпадению сайтов инициации *in vitro* со стартовой точкой для мРНК.

Так же как и в случае бактериальных промоторов, гомология вблизи стартовой точки неполная. Существуют три участка, в которых была обнаружена значимая гомология: стартовая точка, участок, расположенный в районе положения -25 , и участок в районе положения -75 .

У стартовых точек не наблюдается значительной степени гомологии, но обнаруживается тенденция, что первым основанием в мРНК является А, граничащий с обеих сторон с пиримидинами.

Почти повсеместно встречается определенная последовательность из 7 п.н., отделенная от стартовой точки расстоянием от 19 до 27 п.н. Эта семичленная последовательность была обнаружена у млекопитающих, птиц, амфибий и насекомых. Во всех изученных случаях (независимо от источника) среднестатистическая последовательность имеет следующий вид:



Часто эту последовательность называют ТАТА или **блоком Хогнесса**. Эта каноническая последовательность всецело состоит из А-Т-пар (в двух положениях ориентация этих пар может изменяться). Только в незначительном числе описанных случаев присутствует пара Г-С. Блок, как правило, окружен последовательностями, богатыми Г-С, которые, возможно, принимают участие в его функционировании. Он почти идентичен с блоком Прибнова, обнаруженным в бактериальных промоторах. Действительно, единственное существенное отличие состоит в локализации этих участков. Блок Хогнесса располагается предпочтительно в положении -25 , тогда как блок Прибнова — в положении -10 . Видимо, эта структура не абсолютно необходима для транскрипции, так как извест-

на пара случаев, когда данная последовательность вообще отсутствует.

Далее против хода транскрипции имеется другая последовательность, которая у некоторых промоторов консервативна, хотя в ряде других случаев она отсутствует. Эта последовательность располагается между положениями -70 и -80 , и ее канонический состав имеет вид



Иногда она обозначается как **СААТ-блок**. Мы еще не знаем, является ли она характерным свойством всех промоторов или же присуща только определенному классу.

Системы транскрипции *in vitro* и *in vivo*

При попытках выявить промоторы для эукариотических РНК-полимераз были использованы те же подходы, с помощью которых ранее исследовали бактериальные РНК-полимеразы. Эукариотическим системам свойственны два ограничения. Первое: *in vivo* фактически не было получено мутаций, затрагивающих промотор. Поэтому мы не располагаем какой-либо предварительной информацией о локализации эукариотических промоторов. Второе: пока не было возможности непосредственно охарактеризовать участки, связывающиеся с какой-либо из РНК-полимераз. Это объясняется сложностью выделения ферментного препарата и отсутствием информации о том, что именно образует активную структуру фермента, хотя, конечно, создание удобной системы, с помощью которой можно будет извлекать ДНК-связывающий сайт из состава иницирующегося комплекса, — дело времени.

Менее прямой, но в некоторых случаях более информативный подход состоит в определении промотора как структуры, способной иницировать транскрипцию в подходящей тест-системе. Для этого используются три типа систем.

В системе *in vitro* используется классический подход: проводят очистку всех компонентов и подбирают условия, при которых наблюдается правильная инициация. «Правильная» инициация определяется как процесс образования РНК, начинающийся в сайте, соответствующем 5'-концу мРНК. В последнее время появилась возможность осуществить это в отношении каждой из трех эукариотических РНК-полимераз. Именуемые системы характеризуются различной степенью очистки. В состав некоторых систем входят неочищенные клеточные экстракты, содержащие РНК-полимеразу, в других — фермент добавляют к клеточному экстракту. В дальнейшем эти системы должны быть заменены препаратами, содержащими все охарактеризованные компоненты, и тогда *in vitro* можно будет сравнивать активности РНК-полимераз из различных тканей и объектов.

Система, с помощью которой исследуют транскрипцию в ооцитах, в принципе напоминает систему, использованную для исследования трансляции. В ее основе лежит введение подходящей ДНК-матрицы в ядра ооцитов *X. laevis*. Синтезирующийся РНК можно выделить и проанализировать. Основное ограничение этой системы состоит в том, что она полностью лимитирована условиями, существующими в ооците.

В случае систем *in vivo* не обязательно (как можно было бы предположить из их названия) исследовать способность используемых клеток транскрибировать свои

обычные гены. В данном случае исследуется только способность клеточной культуры транскрибировать матрицу, которая вводится в систему посредством трансфекции (механизм, с помощью которого экзогенная ДНК может проникнуть в клетку и существовать в ней, более детально описывается в гл. 38). Данная система действительно является нативной системой *in vivo*, так как транскрипция осуществляется тем же самым ферментным аппаратом, который обеспечивает экспрессию собственного клеточного генома. Но система оказывается в непривычной ситуации в том отношении, что в используемой матрице могут содержаться гены, которые в обычных условиях в этих клетках не транскрибируются.

Во всех трех системах подходы, используемые для характеристики промотора, одинаковы. Исследователь стремится *in vitro* внести в матрицу определенные изменения и только после этого помещает ее в систему транскрипции. Если оказывается, что определенный фрагмент ДНК способен инициировать транскрипцию, то делается вывод, что он способен выполнять функцию промотора. Затем определяют границы последовательности, образующей данный промотор. Для этого уменьшают размеры фрагмента с каждой из сторон до тех пор, пока в некоторой точке не пропадает инициирующая активность. Обычная схема такого эксперимента приведена на рис. 11.6. Левую границу промотора, находящуюся против хода транскрипции, можно легко определить, прослеживая постепенное удаление оснований с соответствующего конца фрагмента до тех пор, пока промотор не перестанет функционировать. Для определения границы промотора, находящейся по ходу транскрипции, необходимо сконструировать ДНК, в которой укороченный промотор повторно пришивается к транскрибируемой последовательности (поскольку иначе отсутствует продукт, необходимый для тестирования).

При выполнении этих экспериментов необходимо соблюдать ряд предосторожностей, позволяющих избежать нежелательных эффектов. Вероятно, для узнавания РНК-полимеразой ДНК необходимо выполнение определенных топологических требований; поэтому в таких экспериментах последовательность промотор-ген обычно помещают вслед за стандартной последовательностью ДНК. Это позволяет избежать влияния, обусловленного близостью к концу фрагмента, когда функционирование промотора прекращается просто потому, что он расположен слишком близко к концу фрагмента ДНК. Кроме того, это гарантирует стандартное окружение фрагмента при исследовании его промоторной функции. В системах *in vitro*, в которых терминирование транскрипции, возможно, происходит недостаточно эффективно, матрица может быть разрезана на некотором расстоянии от промотора (обычно около 500 п.н. по ходу транскрипции); это дает гарантию того, что все полимеразы пройдут одинаковое расстояние, образовав идентичные транскрипты.

После того как границы промотора уже определены, можно исследовать значение конкретных оснований, вводя точечные мутации или используя другие способы изменения последовательности ДНК. Как и в случае бактериальной РНК-полимеразы, такие мутации можно охарактеризовать как *усиливающие* или *ослабляющие* транскрипцию. Некоторые из этих изменений могут сказываться только на скорости инициирования транскрипции, другие могут влиять на ее специфичность, так как приводят к изменению местоположения стартовой точки. Для уверенности в том, что мы имеем дело с молекулами, ко-

торые можно сравнивать, в каждом случае необходимо охарактеризовывать 5'-конец РНК так, как это было показано на рис. 11.2.

В системе *in vitro* РНК-полимераза II функционирует правильно

Результаты, получаемые в системе *in vitro*, в большинстве случаев можно интерпретировать непосредственно. Последовательность ДНК, расположенная левее стартовой точки, можно удалить без сколько-нибудь заметного влияния на транскрипцию вплоть до точки, локализованной где-то между положениями -45 и -30 . Точные координаты варьируют в зависимости от конкретной системы, но всегда определяются левее блока ТАТА. Если же продвинуться далее этой области, то транскрипция подавляется в двадцать раз и более. Рассмотренные данные соответствуют нашим представлениям о бактериальном промоторе, который представляет собой короткую и четко ограниченную последовательность, расположенную непосредственно слева от стартовой точки.

Правая граница промотора варьирует. В двух хорошо охарактеризованных случаях она располагается приблизительно между положениями -10 и -12 ; видимо, последовательность, расположенная непосредственно слева от стартовой точки, имеет менее важное значение. В этих случаях при делеции стартовой точки инициация происходит в сайте, который расположен на *той же* *расстоянии* от промотора, что и первоначальная стартовая точка. Наблюдаемые при этом отклонения, которые могут составлять одно или два основания, связаны с поиском пурина (обычно остатка А), с которого инициируется транскрипция. Однако в отсутствие первоначальной стартовой точки эффективность инициации может быть несколько снижена. В другом случае нужная для инициации последовательность простирается до положения $+6$ и, следовательно, включает в себя стартовую точку.

Мы точно не знаем, какая особенность последовательности ДНК в пределах этой границы необходима для узнавания РНК-полимеразой. Поскольку стартовая точка далеко не всегда попадает в эту область, можно заключить, что от геометрии комплекса зависит, в каком месте происходит инициация. Возможно, когда фермент связывается слева от стартовой точки, комплекс способен вытягиваться по направлению хода транскрипции. Блок ТАТА всегда расположен внутри промоторной области; его особое значение для транскрипции подтвердилось в опытах с введением двух мутаций в ген, кодирующий кональбумин у цыпленка. Замена Т в третьей позиции блока Хогнесса на G приводит к возникновению мутации, сильно ослабляющей транскрипцию гена. Эффект этот обусловлен не просто изменением состава пар оснований, так как замена, приводящая к появлению А, оказывает такое же действие (при этом только изменяется ориентация пары Т—А). Следовательно, важное значение имеет, видимо, точная последовательность блока ТАТА. Как было показано, этой области достаточно для инициирования транскрипции *in vitro*. Об этом свидетельствует тот факт, что последовательность положения от -32 до -12 , относящаяся к области поздних генов транскрипционной единицы аденовируса, встроившись в различные участки ДНК бактериальной плазмиды, способна инициировать транскрипцию на расстоянии 30 п.н. от места своего включения.

Все эти результаты были получены *in vitro* с одной

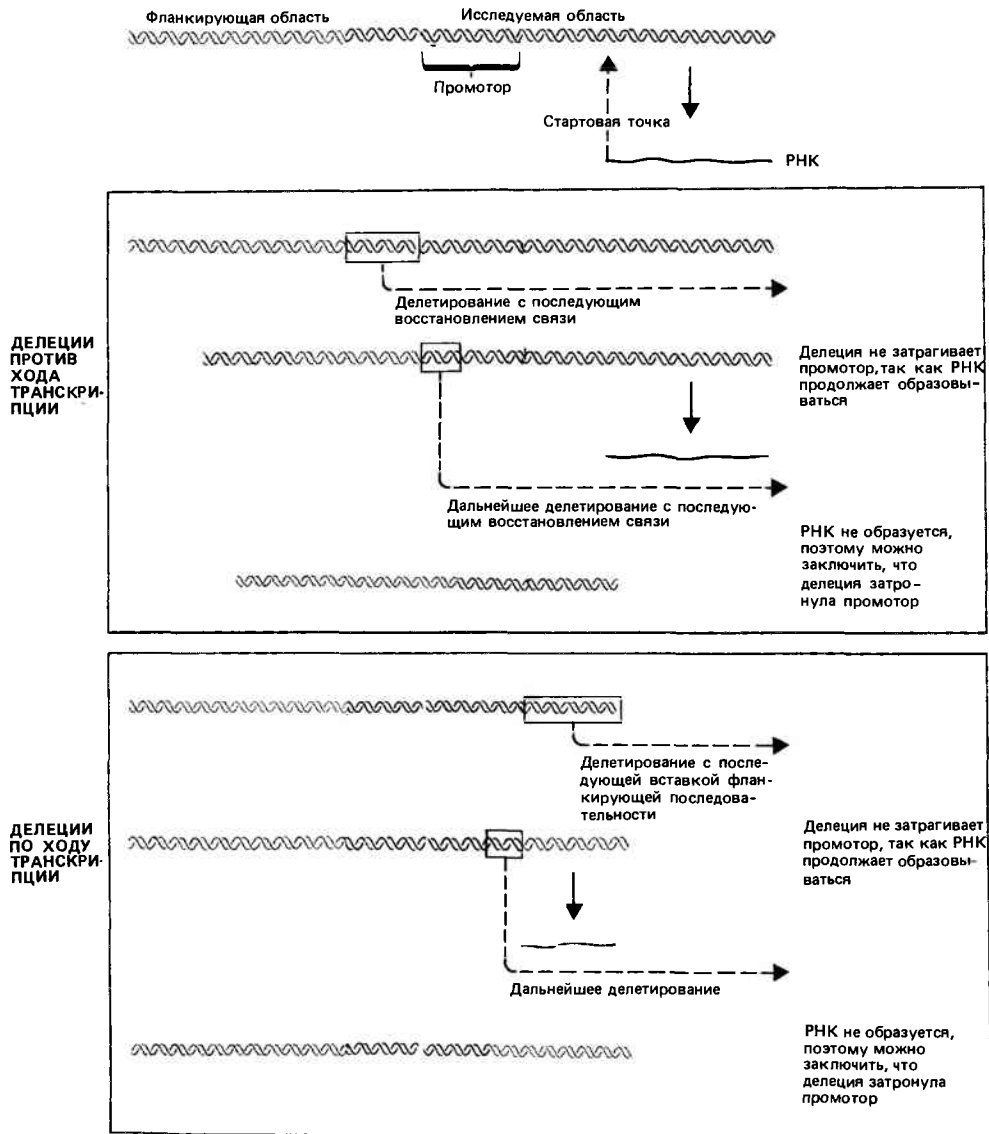


Рис. 11.6. Границы промотора могут быть определены с помощью делеций.

С каждого конца исследуемой области создаются делеции. Если одна делеция не в состоянии повлиять на синтез РНК, а следующая уже оста-

вливает его, то граница промотора находится между ними. (В случае мутаций, расположенных по ходу транскрипции, последовательность, входящая в состав транскрипционной единицы, изменяется.)

и той же системой из культуры клеток млекопитающих. Является ли свидетельством эволюционной консервативности способность данной системы функционировать в присутствии не только генов млекопитающих, но также и генов птиц и насекомых? Для гена, кодирующего фиброин у насекомых, была получена аналогичная система из шелкоотделительной железы шелкопряда. И в этом случае у промотора обнаружены точно те же самые пределы — положения от -29 до $+6$. Однако бесклеточный экстракт из шелкоотделительной железы шелкопряда транскрибирует гены фиброина приблизительно в три раза более эффективно, чем аденовирусную ДНК, тогда как экстракт клеток млекопитающих функционирует в три раза лучше с аденовирусной ДНК, для которой он является гомологичной системой. Система из шелкоотделительной железы шелкопряда также специфически повышает эффективность транскрипции гомологичного гена фиброина, если сохраняется последовательность, расположенная левее от положения -73 . Обнаружение этого

эффекта только в гомологичной системе наряду с тенденцией каждой системы *in vitro* работать более эффективно в присутствии гомологичной матрицы говорит о том, что *in vitro* мы изучаем только те свойства промотора, которые могут преодолеть видоспецифические барьеры. Но фактическая функция промотора является безусловно более обширной.

Промоторы РНК-полимеразы II МНОГОКОМПОНЕНТНЫ

При анализе промоторных участков *in vitro* обнаружен некоторый элемент, отвечающий по своим свойствам концепции бактериального промотора. Это короткая и хорошо охарактеризованная последовательность, располагающаяся сразу левее стартовой точки (против хода транскрипции).

В системе ооцитов или *in vivo* мы обнаруживаем силь-

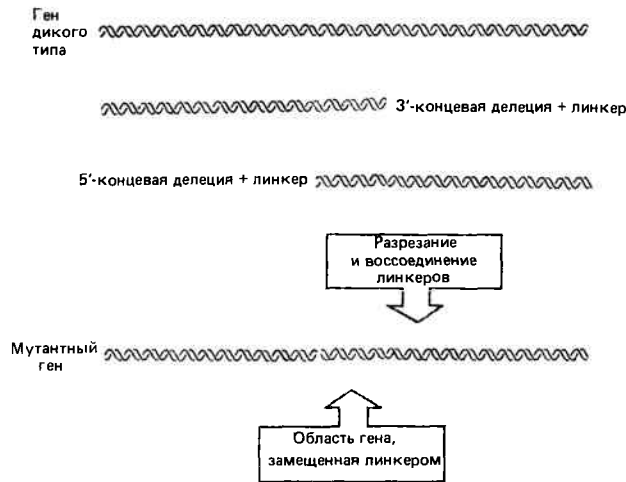


Рис. 11.7. Метод сканирования линкером позволяет заменить короткую последовательность гена дикого типа последовательностью линкера, точно соответствующего по размеру вводимым делециям.

ную зависимость в функционировании промотора от последовательностей, расположенных против хода транскрипции от блока ТАТА. Координаты этих последовательностей варьируют, но в большинстве случаев они располагаются на расстоянии 40 п.н. от блока ТАТА.

При удалении этой области наблюдается снижение частоты инициации со стартовой точки примерно до 2% по отношению к начальному уровню. В противоположность этому, если удаляется ТАТА-блок, то инициация продолжается, однако стартовая точка не имеет точной локализации.

Более детальный анализ последовательностей, необходимых для функционирования промотора, производится с помощью метода сканирования линкером. Этот подход позволяет вводить наборы мутантных последовательностей в определенные участки гена. Постановка опыта изображена на рис. 11.7. В область исследуемого гена вводят делеционные мутации. В одном случае создается набор фрагментов, имеющих делеции с 5'-конца гена, в другом случае — с 3'-конца. Делетируемые области заменяются последовательностью линкера, который представляет собой короткий синтетический олигонуклеотид, содержащий участок узнавания определенной рестриктазой. Фрагменты обоих типов расщепляют рестриктирующей эндонуклеазой и затем соединяют друг с другом, таким образом, чтобы произошло объединение линкеров (смотри реакцию расщепления рестриктирующими ферментами в гл. 17).

В результате описанных манипуляций происходит замена первоначальной области гена, которая делетирована, последовательностью линкера. Используя наборы 5'- и 3'-концевых делеций, захватывающих разные участки гена дикого типа, таким образом можно «просканировать» всю исследуемую область, определяя ее чувствительность к мутациям.

С помощью такого подхода было установлено, что в промоторной части гена тимидин-киназы вируса герпеса имеется три различных элемента. Область, окружающая ТАТА-блок, необходима для точной инициации. Две другие обособленные области, локализованные между положениями — 47 и — 61, а также между — 80 и — 105, нужны для эффективного протекания этого процесса.

Оказалось, что при одновременном введении мутаций в оба участка наблюдается тот же самый эффект, что и при введении мутации в один из них. Из этого следует, что эти области выполняют сходную функцию.

Такая многокомпонентная организация означает, что нельзя исследовать внутреннюю структуру промотора, просто делетируя его отдельные участки. Наблюдаемое при этом исчезновение функции может быть вызвано изменением критических расстояний между компонентами системы, даже если делетируемая последовательность сама по себе не несет какой-либо информации. Следовательно, правильным было бы создание маленьких делеций, которые заменяются последовательностями, равными по длине.

Области, расположенные между тремя указанными элементами, не выполняют промоторной функции. Оказалось, что расстояние между этими тремя сайтами можно изменять в определенных пределах. Для двух более удаленных (влево от ТАТА-блока) областей оно может быть увеличено более чем на 15 пар оснований, не оказав при этом влияния на их функцию. А их удаленность от ТАТА-блока может быть увеличена более, чем на 30 пар нуклеотидов, прежде чем будет затронута функция промотора.

В β -глобиновом гене мыши также были обнаружены три типа участков, расположенных аналогичным образом. Однако существенные различия между двумя промоторами состоят в том, что центральная последовательность в глобиновом гене содержит консервативный СААТ-блок, необходимый для работы этого промотора. В случае же тимидин-киназного гена этот блок расположен между двумя более удаленными влево (против хода транскрипции) компонентами и его присутствие не имеет столь существенного значения.

Эксперименты с делеционными мутантами гена тимидин-киназы (ТК) вируса герпеса показали, что область между положениями — 100 и — 60 контролирует частоту инициации. В отсутствие этого участка частота инициации в обычной стартовой точке падает до 2% от первоначального уровня. Если делетируется блок ТАТА, инициация продолжает происходить, но при этом снижается точность узнавания первоначальной стартовой точки. Аналогично в случае глобиновых генов млекопитающих делеция области размером 20–30 п.н. с центром примерно около положения — 70 вызывает значительное снижение транскрипции. Для некоторых генов дрожжей также показано, что последовательность, расположенная влево от стартовой точки, играет важную роль. В случае гистоновых генов морского ежа рассматриваемая последовательность расположена еще дальше от стартовой точки — между положениями — 139 и — 111. Данные последовательности не активны, если область, предшествующая стартовой точке, делетирована, но, как правило, они значительно увеличивают частоту инициации.

Поэтому *in vivo* в составе промотора можно выделить две области, расположенные, как это изображено на рис. 11.8.

Обычно слева, недалеко от положения — 60, находится область, оказывающая сильное влияние на частоту инициации. Ее делеция приводит к значительному (в 10–15 раз) понижению уровня транскрипции. Рассматриваемая область содержит консервативную последовательность, обозначаемую как блок СААТ. Возможно, что она отвечает основному действию на связывание РНК-полимеразы. Однако остаточная транскрипция, происходящая в отсут-

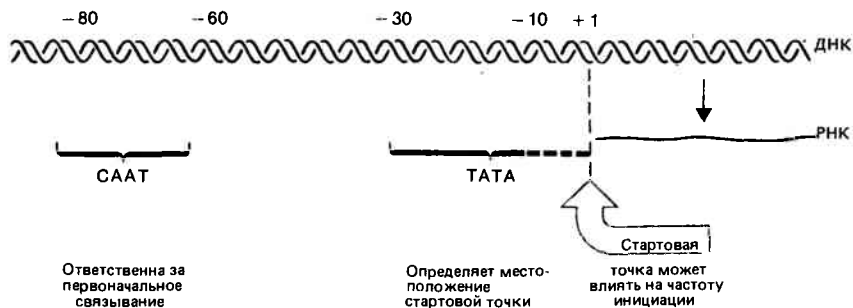


Рис. 11.8. Промоторы для РНК-полимеразы II содержат обособленные участки, выполняющие, по-видимому, различные функции. Промежуточные последовательности скорее всего не несут определенных функций.

стве этой последовательности, иницируется в надлежащей стартовой точке.

Рядом со стартовой точкой расположена область, содержащая блок ТАТА. В результате делеции этой области процесс выбора стартовой точки приобретает более неустойчивый характер, хотя наблюдаемое при этом подавление транскрипции относительно невелико. Поэтому значение данной последовательности, вероятно, состоит в такой фиксации РНК-полимеразы, при которой она может иницировать транскрипцию в надлежащем сайте. Это согласуется с имеющимися данными о существовании промоторов, исходно лишенных последовательности ТАТА. В результате синтезируемые мРНК могут начинаться более чем в одной точке вместо обычной ситуации, когда используется единственная стартовая точка.

Каким же образом в состав промотора могут входить области, расположенные на расстоянии, превышающем то, с которым может взаимодействовать РНК-полимераза, связавшаяся с ДНК? Две возможные модели изображены на рис. 11.9. Согласно одной из них, первоначально РНК-полимераза взаимодействует с сайтом, наиболее удаленным влево от стартовой точки (считая против хода транскрипции). Затем фермент, продвигаясь, приближается к сайту инициации. В основе другой точки зрения лежит то обстоятельство, что *in vivo* вероятность существования ДНК в линейном виде мала. Компактная организация молекулы может привести к сближению сайтов, которые разделены в двухцепочечной ДНК (рис. 29.9). Поэтому связывающим участком для РНК-полимеразы могут быть последовательности, удаленные друг от друга в составе линейной ДНК, но пространственно сближенные ДНК-связывающими белками. Из этого, очевидно, следует, что для работы промотора важное значение имеет взаимное расположение областей ДНК, входящих в его состав (а не только их первичная структура), что несколько противоречит данным, полученным при анализе тимидин-киназного промотора.

Как мы видим, в любой из приведенных моделей область, расположенная намного левее стартовой точки, вероятно, имеет очень важное значение для связывания РНК-полимеразы, но при этом может не быть необходимой для придания ферменту правильной конфигурации, позволяющей узнавать истинную стартовую точку. Эту функцию может осуществлять область, расположенная ближе к стартовой точке. Когда блок ТАТА делетирован, РНК-полимераза сохраняет способность к связыванию с промотором благодаря области, находящейся левее, но при этом фермент образует менее специфичные контакты вблизи стартовой точки. В результате инициация происходит более чем в одной точке. С другой стороны, когда отсутствует наиболее далеко отстоящая область промотора, способность РНК-полимеразы к связыванию

с ДНК сильно снижена, хотя наличие блока ТАТА гарантирует правильность инициации.

Как объяснить тот противоречивый факт, что для транскрипции, происходящей *in vivo*, необходима последовательность, расположенная слева от положения -60 , при том, что *in vitro* эта последовательность явно не представляет большого значения. Основная причина этого кроется, очевидно, в различной эффективности процесса транскрипции *in vivo* и *in vitro*. Система *in vitro* относительно малоэффективна. В ней реально транскрибируется менее 1% матриц. Нам точно не известно, какая часть матриц используется при транскрипции *in vivo* или *in vitro*, но, видимо, она намного больше, чем *in vitro*. Уровень транскрипции *in vitro* может соответствовать остаточной экспрессии, которую мы наблюдаем в отсутствие промоторной последовательности, расположенной слева от положения -60 . Во всех случаях *in vitro* матрица обладает более простым строением, так как молекулы ДНК не организованы в обычную компактную структуру. Это может привести к другому способу узнавания матрицы РНК-полимеразой, который скорее всего менее эффективен, возможно потому, что она использует только некоторые участки промотора.

Даже принимая во внимание, что в системе *in vivo* существуют более сложные требования, мы пока рассматриваем промотор как обособленную область, ответственную за связывание РНК-полимеразы. Ряд дальнейших результатов показал, что ситуация, возможно, не столь проста. ДНК вируса SV40 содержит две одинаковые последовательности размером 72 п.н., расположенные тандемно на расстоянии 200 п.н. левее стартовой точки одной из транскрипционных единиц. Эти последовательности находятся в области ДНК, характеризующейся необычной структурой нуклеопротеина (гл. 30). Эксперименты по делеционному картированию показали, что удаление обеих 72-нуклеотидных последовательностей-повторов значительно подавляет транскрипцию *in vivo*. В присутствии хотя бы одной из этих последовательностей транскрипция происходит нормально. На основе этих данных мы можем утверждать, что рассматриваемая последовательность образует наиболее удаленную от стартовой точки область промотора.

Опыты, в которых 72-нуклеотидную последовательность вырезали из ДНК и затем снова вводили в какой-либо участок молекулы, показали, что в ее присутствии может восстанавливаться нормальный уровень транскрипции, причем вне зависимости от места ее введения. Более того, если β -глобиновый ген встроить в молекулу ДНК, содержащую эту последовательность, уровень его транскрипции увеличивается в 200 раз. Этот эффект наблюдается, даже когда 72-нуклеотидная последовательность находится от точки старта на расстоянии, превы-

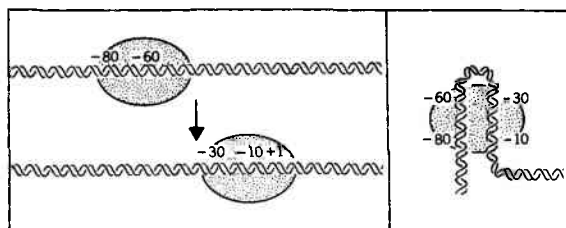


Рис. 11.9. Согласно одной модели, РНК-полимераза может передвигаться вдоль матрицы от участка первичного узнавания к стартовой точке, что устраняет противоречие между размерами промотора и фермента.

Другое предположение основано на том, что *in vivo* ДНК организована более компактно.

шающем 1400 п.н. против хода или 3300 п.н. по ходу транскрипции. Что происходит за пределами этих границ, пока не исследовано. Нам еще предстоит определить, на каком расстоянии 72-нуклеотидная последовательность перестает функционировать.

Рассмотренный элемент называется усилителем транскрипции, или **энхансером**. Он не входит в состав промотора, но способен увеличивать частоту иницирования транскрипции. Как он это делает? Как он может осуществлять свое действие на таких больших расстояниях? Одна из возможностей состоит в том, что энхансер изменяет всю структуру матрицы, например влияя на организацию хроматина или изменяя плотность суперспирализации ДНК. Но возможно также, что энхансер обеспечивает расположение матрицы в определенном месте в клетке, например прикрепляя ДНК к ядерному матриксу. Существует еще одна возможность, хотя и менее вероятная, — энхансер непосредственно участвует в связывании РНК-полимеразы (после чего фермент начинает двигаться собственно к промотору).

В геноме ряда вирусов имеются элементы, несущие ту же функцию, что и 72-нуклеотидная последовательность вируса SV40, хотя они не гомологичны друг другу. В пользу предположения о том, что эти элементы могут выполнять сходные функции, свидетельствует тот факт, что аналогичные компоненты генома ретровирусов могут заменять энхансеры в ДНК вируса SV40. Возможно, что эта часть ретровирусной ДНК вовлечена в активацию клеточных генов.

К настоящему моменту энхансеры не были обнаружены в составе природных клеточных единиц транскрипции¹. Характер взаимодействия вирусных энхансеров с клеточными промоторами может быть различным, так как не все промоторы чувствительны к их действию. Например, транскрипция с промотора α -глобиновых генов не усиливается в присутствии 72-нуклеотидного повтора вируса SV40.

Значение этих результатов очевидно. Нам необходимо соблюдать осторожность при определении элементов, составляющих промотор. Недостаточно только показать, что делеция определенной последовательности нарушает процесс транскрипции. Мы должны также установить, насколько важна *локализация* делетированной последовательности. Итак, что же такое промотор? Если мы ис-

пользуем рабочее определение, согласно которому это последовательность или последовательности ДНК, которые должны определенным образом располагаться относительно стартовой точки, то под промотором надо понимать последовательность вокруг блоков СААТ и ТАТА.

Из этого определения становится ясно, что мы не вправе исследовать структуру промотора, просто делетируя отдельные его участки. При этом промотор может инактивироваться в результате изменения расстояний между его определенными областями, даже если сами делетированные последовательности и не играют никакой роли. Поэтому необходимо создавать маленькие делеции, на место которых помещались бы другие последовательности таких же размеров. С помощью такого подхода можно определить те последовательности ДНК вокруг стартовой точки, которые необходимы для инициации транскрипции.

Промотор РНК-полимеразы III расположен в самой транскрипционной единице

Во всех случаях, когда предпринимались попытки нарушить промоторную последовательность РНК-полимеразы, предполагалось, что промотор окажется расположенным против хода транскрипции (считая от стартовой точки). Так продолжалось до тех пор, пока не был охарактеризован промотор для генов, кодирующих 5S-РНК у *X. laevis*. У этих генов, транскрибируемых РНК-полимеразой III, промотор расположен в пределах транскрипционной единицы на расстоянии более чем 50 оснований правее стартовой точки по ходу транскрипции. Такое его расположение объясняет отсутствие скольких-нибудь очевидных консервативных последовательностей в участках, расположенных слева от стартовых точек генов, транскрибируемых РНК-полимеразой III.

Транскрипцию гена 5S-РНК, взятого в качестве матрицы в составе плазмидной ДНК, можно осуществить в ядерном экстракте из ооцитов *X. laevis*, в которых этот ген экспрессируется в норме. В рассматриваемом случае местоположение промотора было определено при использовании рекомбинантных плазмид, у которых делеции распространялись внутрь гена, с обоих его концов. Оказалось, что синтез 5S-РНК продолжается даже в том случае, если делетируется вся последовательность, расположенная перед геном. Это означает, что промотор располагается не за пределами гена.

Когда делеции захватывают сам ген, то образующийся продукт очень схож с обычной 5S-РНК. При этом синтез продолжается до тех пор, пока делеция не достигнет положения +55. Образующаяся РНК состоит из двух частей. Проксимальная часть соответствует плазмидной ДНК, тогда как дистальная часть представляет транскрипт с оставшегося участка, кодирующего 5S-РНК. Но когда делеция захватывает позиции, находящиеся за положением +55, это приводит к полному прекращению транскрипции. Следовательно, промотор расположен справа от положения +55, но при этом дает возможность РНК-полимеразе иницировать транскрипцию на более или менее определенном расстоянии от него. В гене дикого типа инициация происходит в строго фиксированной точке. При ее отсутствии инициация начинается с пуринового основания, ближайшего к положению —55 п.н. от промотора.

¹ В последнее время установлено существование энхансеров в целом ряде клеточных генов (генов иммуноглобулинов, инсулина и др.). —Прим. ред.

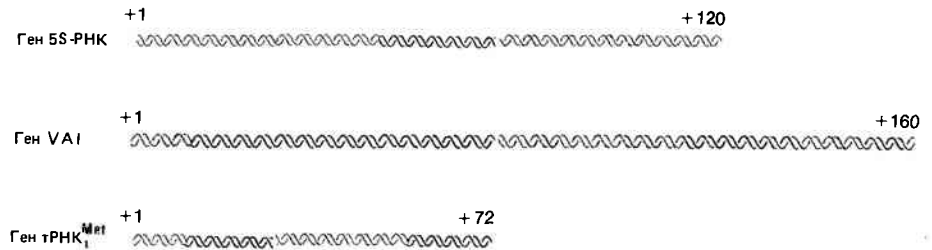


Рис. 11.10. Все три гена для РНК-полимеразы III содержат внутренние промоторы (участки, изображенные красным).

Цифрами обозначены первые и последние основания в гене.

Когда делеция захватывает дистальную часть гена, транскрипция не нарушается до тех пор, пока проксимальные 80 п.н. остаются интактными. Транскрипция прекращается, если делеция затрагивает эту область. Таким образом, правая граница промотора, расположенная по ходу транскрипции, находится около положения +80.

Поэтому промотор для транскрипции 5S-РНК находится внутри самого гена между положениями +55 и +80. Фрагмент, содержащий эту область, способен осуществлять иницирование транскрипции любой ДНК на расстоянии 55 п.н. перед сайтом своей интеграции. Каким же образом РНК-полимераза иницирует синтез РНК в области, предшествующей расположению своего промотора? Наиболее предпочтительное объяснение состоит в том, что размеры фермента позволяют ему одновременно связываться с промотором и взаимодействовать с областью, отстоящей от него на расстояние 55 п.н. Так же как и в случае с РНК-полимеразой II, геометрия связывания фермента с промотором должна определять местоположение стартовой точки, но пиримидины при этом не могут быть использованы при инициации. Поэтому фермент обладает некоторой свободой при выборе ближайшего пурина, участвующего в инициации. Конечно, между ферментами существуют различия: так, РНК-полимераза II устанавливает контакт с точкой старта, расположенной от промотора по ходу транскрипции, тогда как РНК-полимераза III связывается в противоположной ориентации (при условии, что промотор является местом связывания).

Все ли промоторы для РНК-полимеразы III находятся внутри транскрипционных единиц? Границы некоторых из них указаны на рис. 11.10. У гена VA I-аденовирусов промотор расположен между положениями +9 и +72. Делеции, затрагивающие эту область, приводят к потере активности гена. В случае гена, кодирующего синтез тРНК₁^{Met} у *X. laevis*, промотор располагается в двух отдельных областях, лежащих внутри гена. С помощью делеционного картирования были определены координаты этих областей, находящихся на участках от +8 до +30 и от +51 до +72. Любая делеция, уменьшающая расстояние между ними, нарушает инициацию, но изменения в участке между ними или вставки, увеличивающие расстояние между этими двумя областями (вплоть до 30 п.н.), не оказывают никакого действия. Таким образом, промотор состоит из двух отдельных областей размером около 20 п.н. каждая, которые должны быть разделены расстоянием минимум в 20 п.н. У ряда других генов для тРНК обнаруживается похожее строение промоторов. В эукариотических тРНК последовательности этих областей высококонсервативны. Ранее данный факт интерпретировался исключительно с точки зрения функции молекулы тРНК. Но сейчас установлено, что такая консервативность может быть объяснена необходимостью сохранения промоторных последовательностей.

В связи с локализацией промотора во внутренней части гена возникает принципиальный вопрос. Когда промотор располагается с внешней стороны транскрипционной единицы, то он способен эволюционировать независимо, удовлетворяя лишь потребностям ферментов. Но последовательности ДНК, необходимые для иницирования транскрипции, осуществляемой РНК-полимеразой III, должны отвечать условиям, диктуемым структурой довольно различных синтезирующихся молекул, таких, как тРНК, 5S-РНК, VA РНК и малые ядерные РНК. Как же эти последовательности обеспечивают все необходимое для взаимодействия с РНК-полимеразой III?

Одна из возможностей состоит в том, что фермент способен узнавать широкий диапазон промоторных последовательностей. Это не кажется невероятным, если учесть, что в промоторах, узнаваемых РНК-полимеразой других типов, консервативные участки очень малы по размеру. Возможно также, что фермент самостоятельно не способен узнавать промотор, а нуждается для этого во вспомогательном факторе. В этом случае необходимо существование различных факторов для каждого типа промотора. Действительно, оказывается, что РНК-полимераза III способна транскрибировать гены, кодирующие 5S-РНК, только в присутствии вспомогательного фактора. Это белок с мол. массой 37 000 дальтон, взаимодействующий с участком от +45 до +96. Данный белок (несколько копий которого должно связаться с ДНК, чтобы покрыть столь обширную область) одновременно служит двум целям, поскольку в ооцитах он также связывается с образовавшейся 5S-РНК. Рассматриваемый регуляторный белок не связывается с ДНК, кодирующей тРНК₁^{Met}. Мы можем предположить, что в дальнейшем будет охарактеризован ряд аналогичных факторов, играющих важную роль в селекции промоторов.

Одна молекула белка связывается с регуляторной областью, образуя «преинициаторный комплекс» (в состав этого комплекса входит также белковый фактор, необходимый для транскрипции всех генов класса III). Наличие такого комплекса служит сигналом, сообщающим о том, что ген находится в «активном» состоянии и может быть транскрибирован. РНК-полимераза способна связаться с геном только при условии, что предварительно с ним связался белковый фактор. Сформировавшийся комплекс стабилен и может функционировать в течение многих циклов репликации.

Вполне возможно, что способность образовывать преинициаторный комплекс – широко используемый регуляторный механизм. Связывание белкового фактора с промотором приводит к тому, что промотор становится способным к взаимодействию с РНК-полимеразой, так что по существу регуляторный фактор включает ген.

На рис. 11.11 приведена модель, объясняющая, каким образом одна молекула фактора способна стимулировать несколько последовательных циклов транскрипции, иницируемых с определенного промотора. При инициации

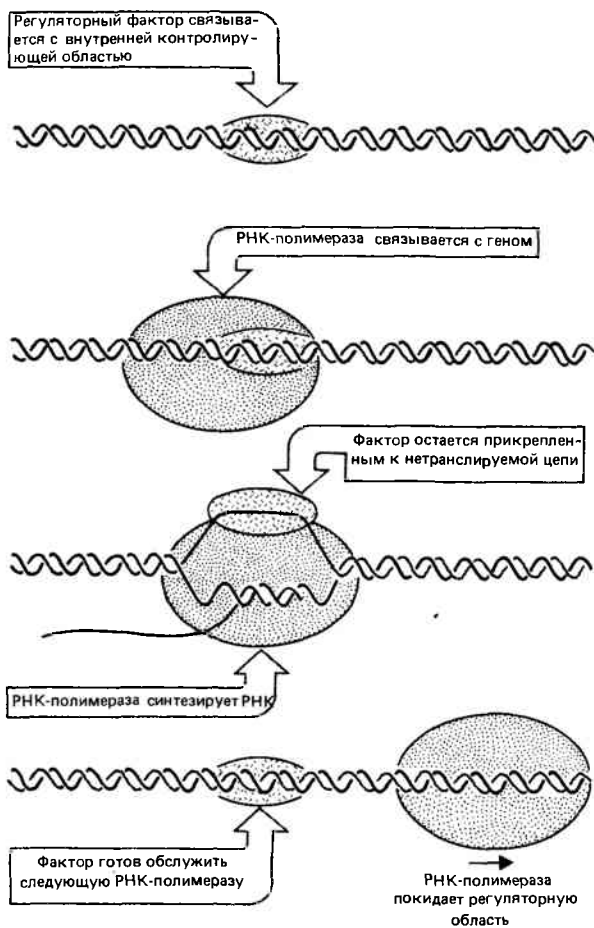


Рис. 11.11. Вспомогательный регуляторный фактор для гена 5S-РНК связывается с некодирующей цепью и, следовательно, может оставаться прикрепленным к ДНК в течение многочисленных циклов транскрипции.

фактор остается связанным с некодирующей цепью и благодаря этому не вытесняется. Мы думаем, что будут охарактеризованы и другие аналогичные факторы, играющие важную роль в выборе промотора.

Хотя способность рассматриваемых генов транскрибироваться определяется промотором, расположенным внутри кодирующей последовательности, стартовая точка также способна оказывать некоторое влияние на этот

процесс. Изменения в области, находящейся непосредственно левее стартовой точки, могут влиять на эффективность транскрипции. Действительно, описан один яркий пример, когда два гена для тРНК имеют идентичные последовательности, а области, находящиеся перед генами, различны и соответственно эти гены транскрибируются с различной эффективностью. Таким образом, за первичное узнавание гена ответственен внутренний промотор, возможно совместно с регуляторными факторами, но частота инициации определяется областью, прилегающей к стартовой точке, и возможно, принимающей непосредственное участие во взаимодействии с РНК-полимеразой III.

Рекомендуемая литература

Взаимодействие РНК-полимеразы с ее промоторами подробно описали Гилберт (Gilbert, RNA Polymerase Losick and Chamberlin, Eds, Cold Spring Harbor Laboratory, New York, 1976, pp. 193–206), Сибенлист, Симпсон и Гилберт (Siebenlist, Simpson, Gilbert, Cell, **20**, 269–281, 1980). Анализ канонических последовательностей промоторов был проведен Розенбергом и Куртом (Rosenberg, Court, **13**, 319–353, 1979). Влияние суперспирализованного состояния ДНК на функцию промотора было описано Смитом (Smith, Cell, **24**, 599–600, 1981). Некоторые методические приемы, используемые для характеристики эукариотических промоторов РНК-полимеразы II, и полученные результаты рассматриваются Корденом и др. (Corden et al., Science, **209**, 1406–1414, 1981); ряд особенностей канонических последовательностей коротко описаны Брезнатчем и Шамбоном (Breathnach, Chambon, Ann. Rev. Biochem. **50**, 349–383, 1981). Вопрос о различиях в функционировании промотора in vivo и in vitro освещен Макнайтом и др. (McKnight, Cell, **25**, 385–398, 1981). Обзор по транскрипции генов 5S написан Корном (Korn, Nature, **295**, 101–105, 1982). Природа двухкомпонентного промотора для генов тРНК исследована Галли, Хофстеттером и Бирнстилом (Galli, Hofstetter, Birnstiel, Nature, **294**, 626–631, 1981) и описана в обзоре Холла и др. (Holl et al., Cell, **29**, 3–5, 1982). Всесторонний анализ влияния энхансеров вируса SV40 на транскрипцию генов осуществили Банерджи, Рускони и Шаффнер (Banerji, Rusconi, Schaffner, Cell, **27**, 299–308, 1981); о существовании ретровирусных энхансеров сообщили Левинсон и др. (Levinson et al., **295**, 568–572, 1981).

Глава 12

СИСТЕМНЫЕ ПЕРЕКЛЮЧЕНИЯ ИНИЦИИРОВАНИЯ ТРАНСКРИПЦИИ

При обычном делении клетки бактериальная РНК-полимераза остается неизменной. Один и тот же фермент транскрибирует все гены и поэтому способен узнавать широкий круг промоторов. Под контролем этих промоторов находятся гены, экспрессирующиеся с различной эффективностью и в различных условиях. Способность иницировать транскрипцию на определенном промоторе

контролируется, по-видимому, путем многочисленных модификаций, придающих ферменту различную специфичность или же подавляющих его активность. При этом изменения никогда не затрагивают самих субъединиц фермента. По-видимому, такой механизм не является совершенным, так как модификации, позволяющие РНК-полимеразе специфически узнавать один вид промоторов,

могли бы препятствовать взаимодействию с другим. Это привело бы к ограничению возможностей фермента.

Сразу же, как только стало ясно, что функции РНК-полимеразы подразделяются между минимальным ферментом, ответственным за элонгацию синтезирующейся РНК, и сигма-фактором (σ -фактором), участвующим в выборе промотора, возник вопрос о возможности существования нескольких типов сигма-факторов, специфичных для разных классов промоторов. Как правило, такой механизм сам по себе, по-видимому, не используется для контроля транскрипции у бактерий. Но при определенных обстоятельствах в жизненном цикле бактериальной клетки происходят коренные изменения. При этом наблюдается выключение транскрипции ранее экспрессируемых генов и включение новых транскрипционных единиц. В этих случаях, возможно, происходит введение долговременных изменений непосредственно в РНК-полимеразу.

Известно два случая, когда выключение экспрессии одних генов и включение других связано с заменой сигма-фактора. Одно из этих явлений — **спорообразование**, или **споруляция** — состоит в резких морфологических изменениях, переводящих бактерии в покоящуюся форму (спору), способную пережить неблагоприятные условия. Другое явление обнаруживается при **литической инфекции** клетки бактериофагом. Когда инфекция развивается по этому пути, то в конце концов в результате размножения фага клетка погибает. Во всех наиболее простых случаях при развитии фага происходит переключение транскрипции. Однако известен только один хорошо изученный случай, когда изменения транскрипционной специфичности обусловлены заменой клеточного сигма-фактора на фаговый. (Это обнаружено в бактериях, способных образовывать споры.) Чаше изменения происходят под действием других механизмов — обычно с использованием дополнительных факторов транскрипции. Создается впечатление, что регуляторный механизм, основанный на возникновении изменений в самой РНК-полимеразе, неохотно используется клеткой, и только в качестве последней возможности. Вероятно, что способность использовать заменяемые друг друга сигма-факторы эволюционно возникла только у очень ограниченного круга бактерий.

Спорообразование

Спорообразование характерно для ряда бактерий, таких, как *Bacillus subtilis*. Это процесс, альтернативный обычному жизненному циклу бактерии. В конце **вегетативной стадии** логарифмической фазы роста количество питательных веществ начинает истощаться и одновременно у бактерий обнаруживается ряд морфологических изменений. При этом происходят репликация ДНК и сегрегация генома, который локализуется на одном конце клетки и покрывается прочной споровой оболочкой. Этот процесс занимает около 8 ч.

При споруляции обнаруживаются значительные изменения в биосинтетических процессах; в этом участвует множество бактериальных генов. Преимущественно контроль осуществляется на уровне транскрипции. Синтез РНК продолжается на всем протяжении споруляции; при ингибировании синтеза РНК процесс спорообразования прекращается. Транскрипция ряда генов, функционировавших при вегетативной фазе роста, выключается, но большинство продолжает экспрессироваться. Кроме того, начинается экспрессия большого числа генов, специ-

фичных только для споруляции. В конце споруляции около 40% бактериальной мРНК специфично именно для этой стадии, тогда как 60% остаются такими же, что и при вегетативной фазе роста.

Гены, участвующие в споруляции, можно идентифицировать с помощью *spo*⁻ мутаций, не оказывающих влияния на обычный вегетативный рост бактерий, но блокирующих споруляцию. Эти не способные к споруляции мутанты могут быть классифицированы на основе той стадии процесса, которую они блокируют. Соответственно мутанты обозначаются как *spo0* (споруляция вообще не может начаться), *spo1*, *spoII* и так далее для последующих стадий. Некоторые из этих мутаций возникают в генах, кодирующих ферменты или структурные белки, необходимые для возникновения спор, но другие картируются в генах, контролирующих переключение различных фаз жизненного цикла.

Первыми указаниями на то, что в спорулирующих клетках происходят изменения транскрипционной специфичности, послужили данные о том, что некоторые бактериофаги способны инфицировать вегетативные, но не спорулирующие клетки. При этом развитие бактериофагов нарушается из-за неспособности транскрибировать свою ДНК. Мы не можем описать все изменения, происходящие при споруляции; не знаем мы и возможных механизмов, принимающих участие в этом процессе, но ряд превращений, происходящих на ранних стадиях (как сейчас стало ясно), связан с изменениями в специфичности инициирования транскрипции РНК-полимеразой.

Фермент, обнаруживаемый у *B. subtilis* и участвующий в обычном вегетативном развитии, имеет такую же субъединичную структуру, как и фермент *E. coli*, $\alpha_2\beta\beta'\sigma$. Сигма-фактор в случае РНК-полимеразы *B. subtilis* имеет мол. массу 55 000 дальтон (σ^{55}) и выполняет те же функции, что и σ -фактор из *E. coli*. В спорулирующих клетках можно обнаружить по крайней мере еще две формы РНК-полимеразы. Они содержат такой же минимальный фермент, как и РНК-полимераза вегетативных клеток, но σ^{55} заменена в этом случае на другие белки. Изменения, происходящие в транскрипционной специфичности, суммированы на **рис. 12.1**.

Форма РНК-полимеразы, способная инициировать споруляцию, была обнаружена благодаря использованию гена, который активируется в начале этого процесса. Этот ген невозможно транскрибировать *in vitro* той формой РНК-полимеразы, которая активна в вегетативной фазе развития. Однако его удается протранскрибировать ферментом, содержащим вместо σ^{55} белок с мол. массой 37 000 дальтон. Этот белок называют σ^{37} , исходя из предположения, что он способен действовать как новый сигма-фактор, придающий минимальному ферменту способность транскрибировать новые гены.

Что вызывает замену одного сигма-фактора другим? В вегетативных клетках уже присутствует σ^{37} , хотя он и не связан с минимальным ферментом. Поэтому, вероятно, какой-то другой белок, возможно продукт гена *spo0*, непосредственно участвует в замене σ^{55} или же модифицирует минимальный фермент, повышая его сродство к σ^{37} . При этом мутация в любом из пяти генов *spo0* может блокировать экспрессию генов, специфичных для ранних стадий процесса споруляции, которые обычно транскрибируются $\alpha_2\beta\beta'\sigma^{37}$ -ферментом. Поэтому замена σ^{55} на σ^{37} может представлять собою весьма сложный процесс. Не исключено также, что эта замена происходит только у части молекул РНК-полимераз, так как некото-

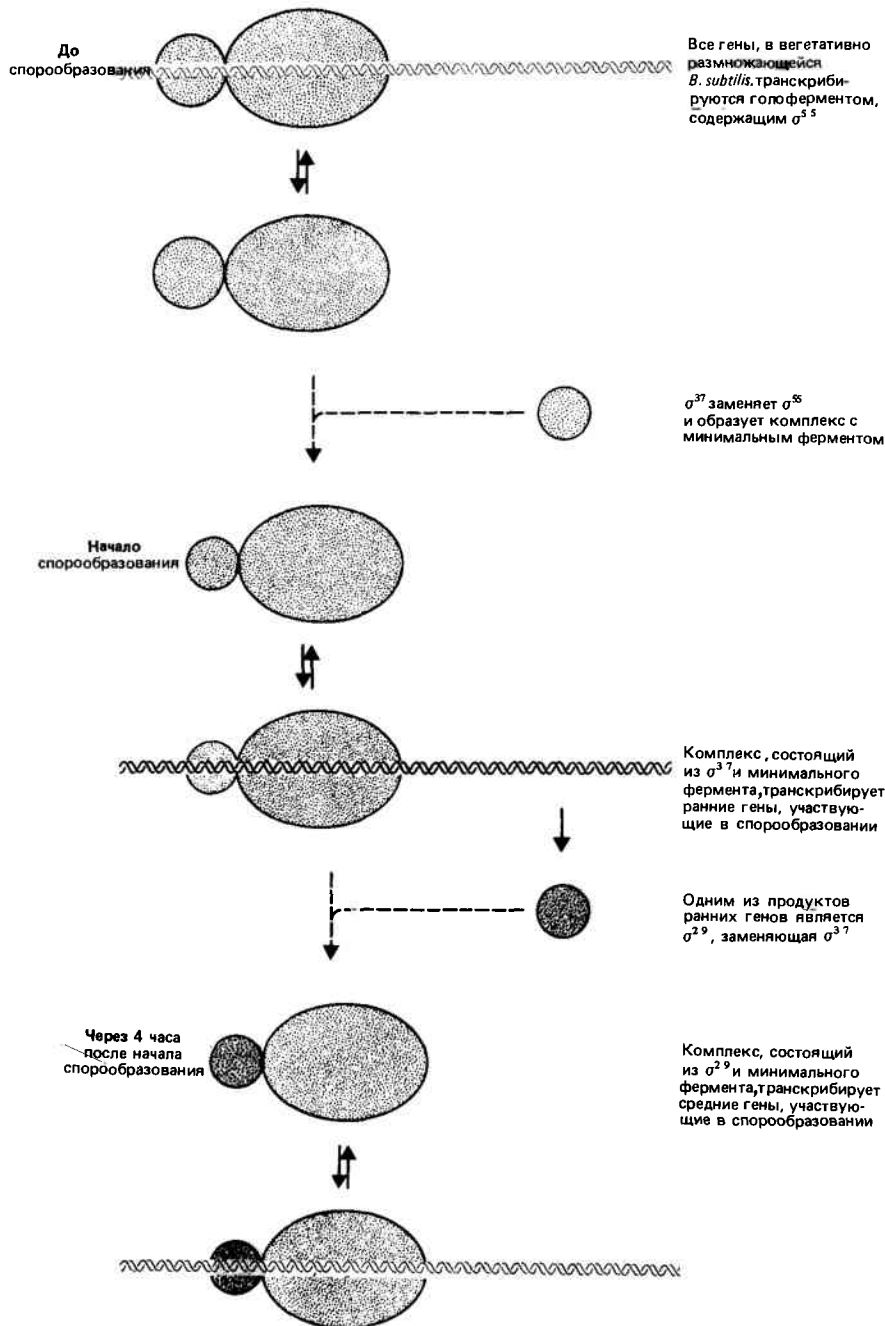


Рис. 12.1. При споруляции происходят последовательные изменения в специфичности иницирования транскрипции РНК-полимеразой *B. subtilis*.

рое количество вегетативного фермента обнаруживается на стадии споруляции. Вытесненный фактор σ^{55} не деградирует и может быть очищен из экстракта спорулирующих клеток.

Через 4 ч после начала споруляции в клетках впервые появляется другая форма фермента. Она содержит белок с мол. массой 29 000 дальтон, ассоциированный с обычным минимальным ферментом. Новый белок, обозначаемый σ^{29} , придает способность минимальному ферменту транскрибировать *in vitro* новые гены. По крайней мере в одном из *spod*-мутантов σ^{29} не синтезируется. Поэтому можно думать, что рассматриваемый белок является продуктом гена, участвующего в споруляции на ранней стадии и необходимого для включения следую-

щей стадии процесса. В этом случае опять-таки не ясно, каким образом σ^{29} способен заменять σ^{37} или σ^{55} в составе фермента.

РНК-полимераза, содержащая σ^{37} или σ^{29} , в действительности сохраняет способность транскрибировать некоторые из вегетативных генов (а в последнем случае по крайней мере ряд генов, включающихся на ранних стадиях споруляции). Таким образом, все три формы РНК-полимеразы транскрибируют перекрывающиеся наборы генов; каждый фермент способен транскрибировать ряд новых генов, но теряет способность узнавать некоторые старые. До тех пор пока не будут охарактеризованы промоторы, узнаваемые каждой формой фермента, мы не сумеем разобраться, что лежит в основе всех этих процессов.

Сигма-факторы, специфические для различных стадий фаговой инфекции

Фаг SPO1— крупный вирулентный вирус, заражающий *B. subtilis*. В зависимости от того, какие фаговые гены экспрессируются, литический цикл можно подразделить на три стадии. Сразу же после начала инфекции транскрибируются **ранние** гены фага. Через 4–5 мин транскрипция ранних генов прекращается и начинается транскрипция средних генов. Затем на 8–12-й минутах транскрипция средних генов заменяется транскрипцией **поздних** генов. Каковы механизмы, обуславливающие переключение транскрипции?

Ранние гены фага транскрибируются голоферментом бактерии-хозяина. Они по существу не отличаются от клеточных генов, чьи промоторы способна узнавать РНК-полимераза с субъединичной структурой $\alpha_2\beta\beta'\sigma^{55}$. Транскрипционная специфичность этого фермента и голофермента РНК-полимеразы из *E. coli* очень похожи или даже идентичны. Оба фермента *in vitro* узнают одни и те же промоторы. В положениях –35 и –10 ранних генов бактериофага SPO1 находятся шестичленные последовательности, соответствующие каноническим последовательностям промоторов *E. coli* (гл. 11). (Несмотря на то что у фага SPO1 основание Т заменено на гидроксимитилурацил, это, по-видимому, не вызывает никаких различий.)

Для транскрипции средних и поздних генов необходима экспрессия фагового генома. Три регуляторных фаговых гена, обозначаемых 28, 33 и 34, контролируют очередность транскрипции. Функции этих генов рассмотрены на **рис. 12.2**. Регуляция осуществляется последовательно; при этом сначала бактериальный фермент транскрибирует ранний ген, чей продукт необходим для транскрипции средних генов. В свою очередь продукты некоторых средних генов необходимы для транскрипции поздних генов.

Мутанты, дефектные по раннему гену 28, не способны транскрибировать средние гены. Продукт этого гена (обозначаемый gr28) представляет собой белок с мол. массой 26 000 дальтон, который присоединяется к минимальному ферменту и заменяет сигма-фактор. Образующаяся при этом РНК-полимераза специфически транскрибирует средние гены. В тех случаях, когда фермент выделяют из клеток, в которых фаговая инфекция достигла средней стадии, или когда к очищенному минимальному ферменту добавляют gr28, РНК-полимераза становится способной связываться с промоторами средних генов. При этом она утрачивает какую-либо способность взаимодействовать с промоторами ранних генов. Следовательно, единственным событием, необходимым для переключения транскрипции с ранних генов на средние, является замена бактериального сигма-фактора на фаговый gr28 (иногда называемый σ^{gr28}). Каким образом gr28 заменяет σ^{55} и какова судьба бактериального сигма-фактора, неизвестно. Возможно, что gr28 обладает более высоким средством к минимальному ферменту.

Два средних гена принимают участие в следующем переключении транскрипции. Мутация в гене 33 или 34 предотвращает транскрипцию поздних генов. Продукты этих генов представляют собою белки с мол. массой 13 000 и 24 000 дальтон соответственно. Они, присоединяясь к минимальному ферменту, заменяют gr28. И

в этом случае нам также не известно, как gr33 и gr34 вытесняют gr28 (или некоторое остаточное количество бактериального фактора σ^{55}). Но если они уже связались с минимальным ферментом, то он будет способен инициировать транскрипцию только на промоторах поздних генов.

Последовательная замена сигма-фактора имеет двойное значение. При каждой замене этой субъединицы РНК-полимераза становится способной узнавать новый класс генов и фермент уже не транскрибирует предшествующие гены. Таким образом, эти переключения приводят к системным изменениям в активности РНК-полимеразы. Вероятно, весь или фактически весь минимальный фермент клетки связывается с той сигма-субъединицей, которая должна функционировать в данный момент, причем это взаимодействие необратимо.

Для характеристики фермента представляет интерес тот факт, что субъединицы различного размера могут выполнять одну и ту же функцию. Они придают способность минимальному ферменту распознавать определенную группу промоторов. Совокупность данных о процессах, протекающих при спорообразовании и фаговой инфекции, открыла нам существование (по крайней мере) пяти различных сигма-факторов, с которыми может взаимодействовать минимальный фермент. Это σ^{55} , σ^{37} , σ^{29} , σ^{gr28} и $\sigma^{gr33-34}$. Другой возможный фактор σ^{28} был обнаружен в вегетативно делящихся клетках. В них обнаружен также полипептид δ , необходимый, вероятно, для увеличения точности выбора промоторов в присутствии различных сигма-факторов. Таким образом, в ходе эволюции минимальный фермент приобрел важное свойство воспринимать сигналы от любого из сигма-факторов.

Для каждого сигма-фактора может существовать своя собственная консервативная последовательность –35 и –10

Чем отличаются различные классы промоторов? Гены, принимающие участие в спорообразовании, не содержат обычных консервативных последовательностей, обнаруживаемых в положениях –35 и –10 промоторов вегетативных генов. Вероятно, каждый класс промоторов, контролирующих гены спорообразования, обладает характерными только для него консервативными последовательностями. Этот вывод, возможно, справедлив также и в случае промоторов, контролируемых вторым вегетативным фактором σ^{28} . Похожая ситуация обнаруживается при исследовании средних генов бактериофага SPO1, среди которых было изучено достаточное количество последовательностей промоторов, чтобы определить среднестатистическую.

Участки ДНК, с которыми связываются РНК-полимераза, содержащая белок gr28, и бактериальный фермент, имеют одинаковые размеры и одинаковую локализацию относительно структурной части гена. РНК-полимераза, работающая на средних стадиях процесса споруляции, экранирует область ДНК с координатами от –40/–50 до +20.

В положении –10 расположена консервативная последовательность, которая по своему А–Т-богатому составу аналогична блоку Прибнова, хотя все основания, образующие пары, находятся в одной ориентации

TTTNTTT

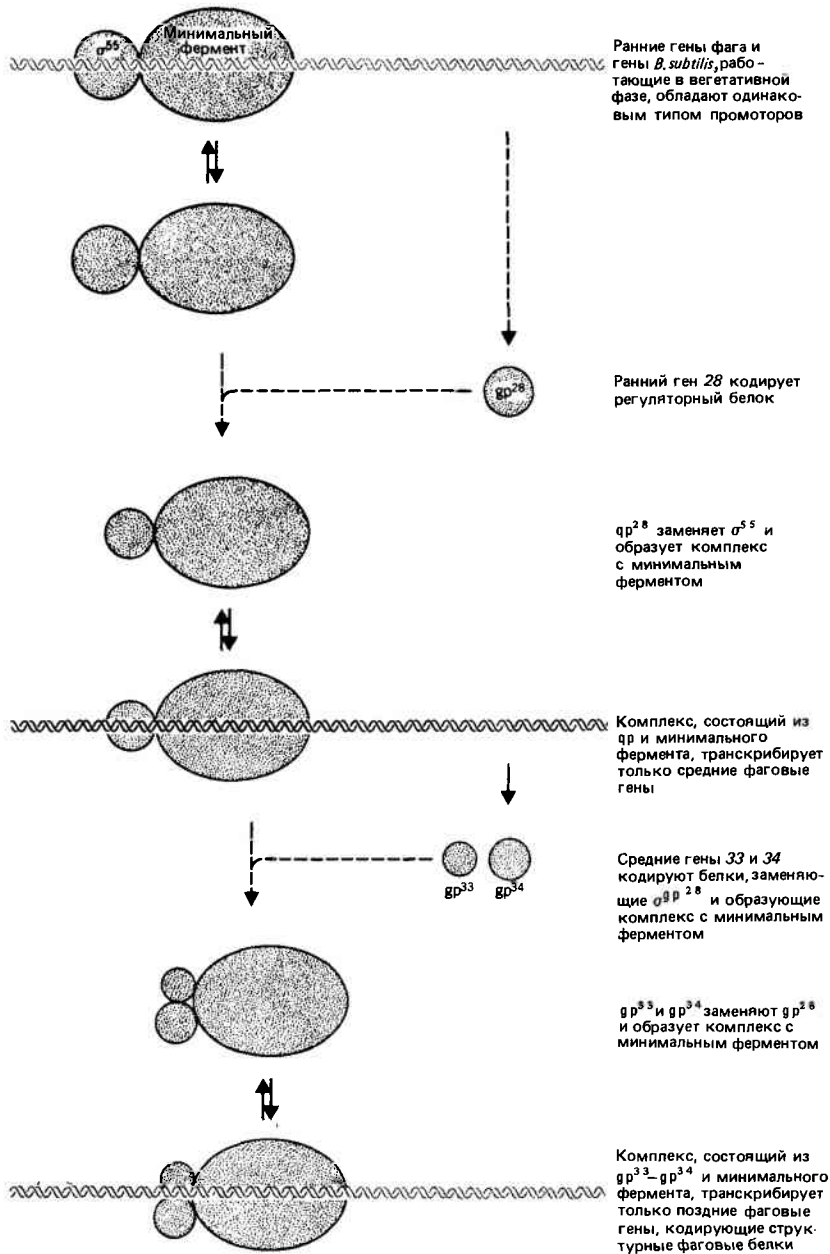


Рис. 12.2. Транскрипция генов фага SPO1 контролируется двумя последовательными заменами сигма-фактора, приводящими к изменению специфичности транскрипции.

в отличие от консервативной последовательности
ТАТААТ,

характерной для генов вегетативной стадии развития. Возможно, эти две последовательности, локализованные в положении -10 , выполняют одинаковые функции.

Но в области -35 обнаруживаются различия. Как и в случае промоторов *E. coli*, только в этой области (кроме блока Прибнова) промоторов средних генов можно выявить сколько-нибудь консервативную последовательность. Эта последовательность представляет собою гексануклеотид

AGGAGA,

отличающийся от соответствующей последовательности
TTGACA,

характерной для генов вегетативной стадии роста.

Либо области -35 или -10 , либо обе вместе обуславливают, возможно, принадлежность промоторов к опре-

деленным классам, которые узнаются РНК-полимеразой. У *B. subtilis* минимальный фермент может проявлять (как минимум) шесть различных видов транскрипционной специфичности. Поэтому мало вероятно, что каждый сигма-фактор индуцирует в минимальном ферменте образование определенной конформации, ответственной за узнавание конкретного класса промоторов. Скорее всего, что сигма-факторы непосредственно сами узнают характерные особенности в промоторах. Однако это предположение возвращает нас обратно к вопросу о том, как относительно небольшой полипептид может взаимодействовать с участком ДНК длиной более 20 п.н.

В связи с вопросом о специфичности узнавания промотора следует вспомнить, что уже довольно давно известно об изменениях в транскрипционной специфичности при инфицировании *E. coli* фагом T4. Это обусловлено изменениями, происходящими как в сигма-факторе, так и в субъединицах минимального фермента. Их влияние трудно разделить, но, вероятно, оба компонента воз-

рос о промоторах подробно рассматривается у Роуза (Roza, J. Mol. Biol., 147, 199–204, 1981). С большим упором на генную экспрессию представлен материал по споруляции и индуцируемым ею изменениям в транскрипции в обзорах Лозика и Перо (Losick, Pero,

in RNA-Polymerase (Losick, Chamberlin, Eds.), Cold Spring Harbor Laboratory, New York, 1976, pp. 227–246 и Cell, 25, 582–584, 1981). Вопрос о РНК-полимеразах, индуцированных фагом, рассматривается Баутом (Bautz, in RNA Polymerase – в упомянутом выше сборнике, pp. 273–284).

Глава 13

ТЕРМИНАЦИЯ И АНТИТЕРМИНАЦИЯ

Как только РНК-полимераза инициирует транскрипцию, она начинает продвигаться вдоль матрицы, синтезируя РНК до тех пор, пока не достигнет терминирующей последовательности. Здесь фермент перестает включать нуклеотиды в растущую цепь РНК, освобождает образовавшийся продукт и отделяется от ДНК-матрицы. (Точно не известно, в какой последовательности происходят два последних события.) При терминации все водородные связи, удерживающие вместе гибрид РНК—ДНК, должны быть разорваны; только после этого происходит восстановление двухцепочечной структуры ДНК.

Последовательность ДНК, необходимая для остановки транскрипции, называется **терминатором**. О терминаторах у эукариот известно относительно немного, но терминаторы у бактерий и фагов исследованы очень хорошо. Терминаторы значительно различаются по своей эффективности, а также по зависимости от регуляторных белков, по крайней мере при их исследовании *in vitro*.

На некоторых терминаторах терминационные события могут быть *предотвращены* под действием специфических регуляторных белков, взаимодействующих с РНК-полимеразой. В результате **антитерминации** фермент продолжает синтезировать РНК за пределами терминатора. Это явление называется **прочтыванием терминатора** (этим же термином обозначают проскакивание рибосомой терминирующего кодона при супрессии).

Антитерминация представляет собой особый регуляторный механизм. Различные белки (**факторы антитерминации**) позволяют РНК-полимеразе проскакивать определенные терминирующие последовательности. Таким образом, контролируется способность фермента транскрибировать гены, расположенные за терминатором. Данный механизм характерен для регуляторных систем фагов.

Описывая события, происходящие при терминации, этот процесс необходимо рассматривать не просто как способ, обуславливающий образование 3'-конца молекулы РНК, а как удобный механизм, контролирующий генную экспрессию. Таким образом, и прикрепление РНК-полимеразы к ДНК (инициация), и отделение фермента от матрицы специфически контролируются. Обнаруживаются интересные аналогии между системами, участвующими в инициации и терминации. В обоих случаях должен обязательно происходить разрыв водородных связей (первоначальное плавление ДНК при инициации и диссоциация гибрида РНК—ДНК при терминации), и в обоих случаях требуются дополнительные белковые факторы, взаимодействующие с минимальным ферментом. Не исключено, что эти процессы осуществляются двумя различными формами ферментов.

Обнаружение терминаторов в системе *in vitro*

Для молекулы РНК, синтезированной в клетке, трудно определить точку терминации. Как в случае прокариот, так и в случае эукариот возможно, что 3'-конец молекулы образовался в результате расщепления первичного транскрипта и, следовательно, не соответствует реальному сайту, в котором терминировала РНК-полимераза. Для 3'-конца не существует маркера, подобного трифосфату, присутствующему исходно на 5'-конце.

Наилучшим образом сайт терминации можно охарактеризовать, используя систему *in vitro*, в которой может происходить терминация. Способность фермента останавливать синтез РНК во многом зависит от условий реакции, например от ионной силы. Поэтому терминация, наблюдаемая *in vitro* в какой-либо точке, еще не означает, что данный сайт является истинным терминатором. Но сейчас известно несколько примеров, когда у определенного вида РНК (прокариотической) имеется одинаковый 3'-конец независимо от того, синтезирована ли эта РНК *in vivo* или *in vitro*. Это убедительно показывает, что удалось обнаружить истинную терминирующую последовательность.

При сравнении последовательности ДНК с 3'-концевым участком синтезированной РНК можно локализовать сайт терминации в ДНК. Таким образом были охарактеризованы многие прокариотические терминаторы. При этом не было обнаружено какого-либо сходства между последовательностями, находящимися за точками, в которых синтез РНК прекращается. Это свидетельствует о том, что участок, ответственный за терминацию, расположен в пределах последовательности, транскрибируемой РНК-полимеразой. Таким образом, решение полимеразы о целесообразности транскрипции последующих участков ДНК зависит от свойств той области, которую полимеразы транскрибирует в данный момент, а не от свойств области, лежащей впереди.

Существуют р-зависимые и р-независимые терминаторы

Когда в системе *in vitro* определенные транскрипционные единицы используются в качестве матриц для синтеза РНК, правильной терминации не происходит. Минимальный фермент может на некоторое время остановиться на терминаторе, но затем продолжит транскрипцию, наращивая цепь РНК до тех пор, пока какие-либо случайные события не заставят его отделиться от

матрицы. В результате образуется гетерогенный набор молекул РНК, отличающихся различными 3'-концами.

Фактор ρ был описан как белок, добавление которого в систему транскрипции, полученную из *E. coli*, приводит к образованию дискретных молекул РНК с уникальными концами. Из этого, очевидно, следует, что соответствующие терминаторы на ДНК-матрице могут функционировать только в присутствии фактора ρ . Их называют **ρ -зависимыми терминаторами**. Фактор ρ *in vitro* оказывает свое действие только в среде с относительно низкой ионной силой.

Степень зависимости от фактора ρ варьирует. Для некоторых терминаторов необходима относительно высокая концентрация фактора ρ , тогда как другие хорошо функционируют при более низких его концентрациях. Эти вариации отражают истинные различия в эффективности терминации, обнаруживаемые *in vivo*.

Существуют ли терминаторы полностью независимые от фактора ρ ? Некоторые терминаторы *in vitro* при встрече с минимальным ферментом, по видимому, способны в одиночку осуществлять весь комплекс терминационных реакций. Это так называемые **ρ -независимые терминаторы**. Однако мы должны с осторожностью трактовать эти данные, так как возможно, что в действительности в бактериальной клетке фактор ρ также необходим для узнавания таких сигналов терминации. *In vitro* в отсутствие фактора ρ эти терминаторы проявляют свое действие только тогда, когда транскрипция происходит в условиях относительно высокой ионной силы. При этом, возможно, истинная (при низкой ионной силе) зависимость от фактора ρ не тестируется.

Результаты, получаемые *in vitro*, не вносят в этот вопрос окончательной ясности, поскольку ρ -зависимые и ρ -независимые терминаторы необходимо исследовать в различных условиях. Какую информацию можно получить *in vivo*? Существуют бактериальные мутации, нарушающие активность фактора ρ . Хотя такие мутации по-разному влияют на различные ρ -зависимые терминаторы, как правило, они не сказываются на процессе терминации в сайтах, которые *in vitro* были охарактеризованы как ρ -независимые.

Указанием на то, что между двумя типами сайтов терминации действительно существуют определенные различия, явились данные по изучению их последовательностей. Ни в одном случае в области, предшествующей сайту терминации, не обнаружено сколько-нибудь выраженной консервативной последовательности. Однако в случае ρ -независимых терминаторов на конце РНК всегда имеется небольшой отрезок из остатков U. Этого не обнаруживается у РНК, терминирующих на ρ -зависимых сайтах, которые полностью лишены какой-нибудь гомологии. Но у обоих типов терминаторов наблюдается поразительное совпадение их вторичной структуры, образующейся благодаря наличию **палиндрома**, расположенного непосредственно перед терминационным сайтом.

Немного об инвертированных повторах

Палиндромные последовательности играют важную роль не только при терминации, но также и в других процессах, контролирующих функционирование мРНК (а возможно, и ДНК). Поэтому мы сначала очень коротко рассмотрим их свойства, а потом вернемся к терминации.

Палиндром представляет собою двухцепочную последовательность ДНК, которая одинаково читается в обеих цепях, если каждую цепь читать в соответствии с ее ориентацией. Например, последовательность



является палиндромом, так как при ее чтении от 5' к 3' она идентична (GGTACC) в обеих цепях (при этом верхняя строчка прочитывается слева направо, а нижняя справа налево).

Чисто формально такая последовательность носит название «область с **двойной симметрией**». **Ось симметрии** является центральной точкой, относительно которой последовательность остается одинаковой при чтении в противоположных направлениях. Проведя линию через ось симметрии



мы можем видеть, что с каждой стороны от нее находится одна и та же последовательность $\begin{array}{l} \text{GGT} \\ \text{CCA} \end{array}$. Но ориентация данных последовательностей противоположная. По этой причине последовательность $\begin{array}{l} \text{GGT} \\ \text{CCA} \end{array}$ называется **инвертированным повтором**.

Две копии инвертированного повтора не обязательно должны быть непрерывными. Рассмотрим тот же повтор в других условиях:



Мы по-прежнему имеем инвертированный повтор, состоящий из тех же трех пар оснований, хотя ось симметрии переместилась в центр последовательности из трех неповторяющихся пар оснований, разделяющих две копии повтора.

Инвертированные повторы обладают важным свойством, влияющим на вторичную структуру ДНК или РНК. Копии повтора, находящиеся в одной и той же цепи ДНК, **взаимно комплементарны**, если их прочитывать в противоположных направлениях. В вышерассмотренном случае при чтении верхней строчки от оси симметрии влево мы получим последовательность TGG, а при чтении вправо — комплементарную ей последовательность ACC. При образовании комплементарных пар возможно возникновение **шпильки** в РНК или **структуры креста** в ДНК, как это показано на **рис. 13.1**.

Таким образом, именно присутствие в ДНК палиндромной последовательности обуславливает способность соответствующей РНК формировать двухцепочечную вторичную структуру в результате взаимодействия между двумя копиями инвертированного повтора. Например, как мы уже рассматривали в гл. 7, благодаря такой шпильчатой структуре образуются участки, формирующие структуру клеверного листа молекулы тРНК.

Возможность возникновения крестообразных структур *in vivo* остается под вопросом. Для их формирования необходимо, чтобы обычная двуспиральная ДНК была способна образовывать одноцепочечные петли. Мало вероятно, чтобы такое могло происходить спонтанно. Но этому

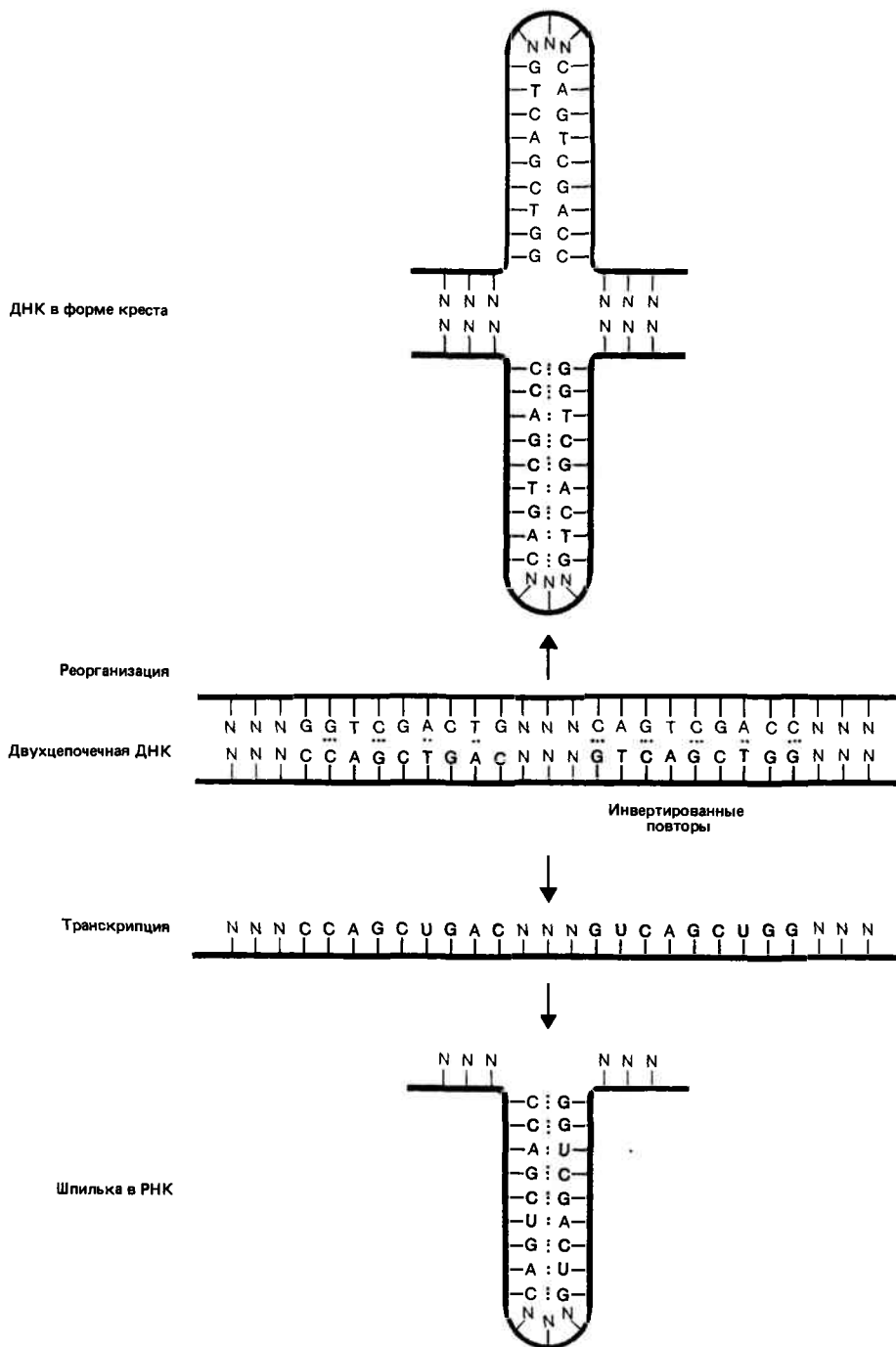


Рис. 13.1. Инvertированные повторы могут принимать участие в образовании вторичной структуры ДНК или РНК.

В центре изображен двухцепочечный участок ДНК, содержащий инvertированные повторы. Структура креста способна образовываться при условии, что ДНК реорганизуется таким образом, при котором возможно комплементарное спаривание между последовательностями, находящимися в составе *одной цепи*. Если транскрибируется верхняя цепь ДНК, то

последовательность синтезированной РНК будет такой же, как у нижней цепи матрицы. Поэтому комплементарные участки смогут спариваться, образуя шпильку. Получившаяся шпилька в РНК точно такая же, как нижняя часть крестообразной структуры ДНК. Таким образом, инvertированные повторы формируют двухцепочечный стебель шпильки. Три основания, разделяющие повторы, образуют одноцепочечную петлю на конце стебля.

может способствовать суперспирализация, а также наличие белков, связывающихся с ДНК.

Достигнув палиндрома, минимальный фермент приостанавливается

У прокариот палиндромы обнаруживаются в каждом терминаторе. Ряд примеров приведен на рис. 13.2, где

показаны два r -независимых терминатора наибольшей длины. В первом случае шпилька состоит из 54 оснований (включая петлю и двухцепочечный стебель). Во втором шпильку образуют только 28 оснований. Как и во всех r -независимых терминаторах, шпилька в большей степени обогащена парами G—C, а за ней непосредственно следует участок из шести остатков U. В случае r -зависимого терминатора шпилька находится приблизительно на таком же расстоянии от фактического сайта термина-

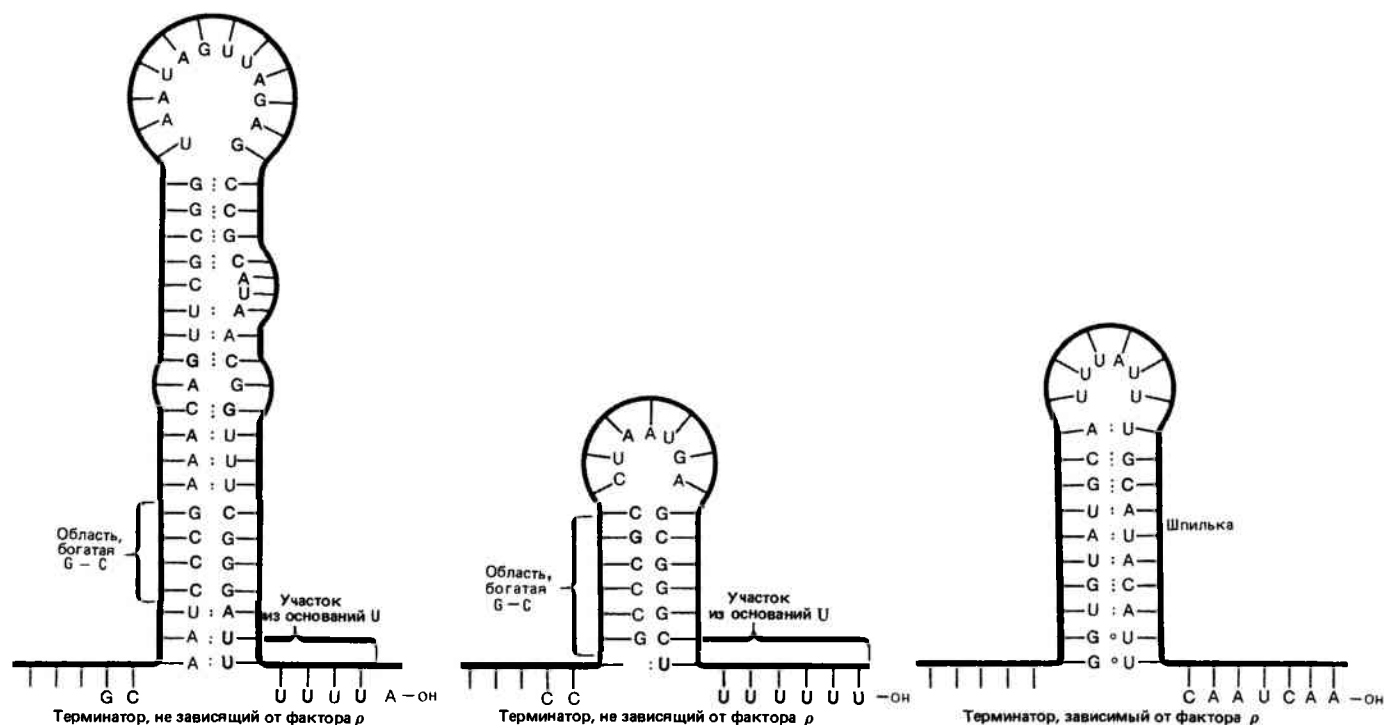


Рис. 13.2. В состав терминаторных последовательностей входят G—С-богатые области, которые образуют шпильки различ-

ной длины. В ρ-независимых терминаторах в отличие от ρ-зависимых за шпилькой следует участок из оснований U.

ции, но она содержит меньше пар G—С, и, кроме того, в этом случае отсутствует участок, состоящий из U.

Последовательности, образующие палиндромы, не обнаруживают гомологии между собой, но сходство между ними может существовать на уровне вторичной структуры. Какую же форму должны принимать такие характерные сигналы? Как мы уже упоминали, по-видимому, мало вероятно, чтобы в ДНК образовывались крестообразные структуры, хотя вполне возможно, что такие структуры могут возникать во время транскрипции, когда цепи молекулы расплетены. Более вероятно, что шпильки, образующиеся в синтезированной молекуле РНК, могут узнаваться РНК-полимеразой. Известно, что существенное значение для процесса терминации имеют свойства самой РНК. Так (в качестве примера), замена GTP на ITP подавляет терминацию, возможно, потому, что I образует менее стабильные связи с С, чем с G.

С помощью точковых мутаций, нарушающих процесс терминации, было показано, что палиндромные последовательности непосредственно участвуют в этом процессе. Такие мутации возникают в области, находящейся приблизительно на расстоянии 35 п.н. перед фактической точкой терминации. В данном случае обнаруживается некоторое сходство в размерах с последовательностью, участвующей в связывании РНК-полимеразы с промотором. Все исследованные мутации затрагивают область, образующую стебель шпильки. Большинство из них нарушает комплементарное связывание, но некоторые не оказывают такого действия. Следовательно, прочность стебля, очевидно, не единственный фактор, необходимый для процесса терминации.

Как сказывается образование шпильки на работе РНК-полимеразы? Видимо, все шпильки, возникающие в синтезирующейся РНК, заставляют полимеразу дви-

гаться медленнее или останавливаться. На обычном терминаторе полимеразы останавливается приблизительно на 60 с. (Если бы транскрипция продолжалась, то за 1 мин фермент преодолел бы расстояние в 2000 п.н.) Однако продолжительность реальной паузы варьирует от случая к случаю. Известен случай, когда РНК-полимераза может фактически остановиться на терминаторе, и хотя она не продолжает более транскрипции, но и не терминирует. События, происходящие после паузы, зависят от типа сайта терминации.

На ρ-независимых терминаторах полимеразы доводит реакцию до конца, отделяясь от матрицы и высвобождая РНК. Последовательность из остатков U, расположенная сразу же за шпилькой, вероятно, принимает участие в этом процессе. РНК—ДНК-гибрид rU—dA характеризуется необычайно слабыми комплементарными взаимодействиями. Поэтому достаточно минимального количества энергии, чтобы разрушить связи, удерживающие вместе две цепи. Таким образом, остановка РНК-полимеразы в области шпильки в ρ-независимом терминаторе создает благоприятные условия для высвобождения РНК из комплекса rU—dA, образующегося в области терминатора. Реально терминация часто происходит на любом из нескольких остатков U, расположенных ближе к концу U-участка. Можно предположить, что, терминируя, фермент начинает как бы «спотыкаться».

Значение для терминации U-участка подтверждается опытами, в которых было показано, что после его частичной делеции терминация не происходит, хотя РНК-полимераза продолжает делать паузу в районе шпильки. Последовательность из оснований U соответствует А—Т-богатой области в ДНК. Таким образом, мы можем видеть, что А—Т-богатые участки играют важную роль как при терминации, так и при инициации.

Как работает фактор ρ ?

Фактор ρ функционирует только в процессе терминации. Это белок с мол. массой 55 000 дальтон. Активной формой, по-видимому, является тетрамер. *In vitro* фактор проявляет две активности. Он функционирует в качестве фактора терминации, причем он требуется в каталитических количествах. Обычно максимальная активность проявляется, когда концентрация фактора составляет 10% от количества РНК-полимеразы. Фактор обладает также РНК-зависимой НТРазной активностью, необходимой для терминации.

Взаимодействует ли фактор ρ в процессе терминации с ДНК, РНК или РНК-полимеразой? Для проявления своей НТРазной активности, фактору ρ необходимо присутствие полирибонуклеотида длиной в 50 или более остатков. Это свидетельствует о том, что фактор ρ взаимодействует с РНК. Согласно существующей модели, фактор ρ связывается с 5'-концом синтезирующейся РНК, а затем продвигается вдоль нее, используя для этого энергию гидролизующихся НТР.

На рис. 13.3 показано, что остановка РНК-полимеразы на терминаторе благоприятствует тому, чтобы фактор ρ смог догнать ее. Затем фактор ρ взаимодействует с РНК-полимеразой, вызывая освобождение РНК и отделение полимеразы и самого фактора от матрицы. Эта модель предполагает, что основное свойство терминатора — это его способность заставлять РНК-полимеразу делать паузу. Исходя из этого, различная эффективность

многочисленных терминаторов может объясняться тем, что они отличаются по длительности паузы в передвижении РНК-полимеразы по матрице. На это может влиять стабильность шпильки, зависящая от ее протяженности и количества пар G—C. Нам не известно, способен ли фактор ρ осуществлять терминацию в *любой* точке, в которой РНК-полимераза делает остановку, или же последовательность терминатора должна содержать некоторую дополнительную информацию.

Идея о том, что фактор ρ передвигается вдоль РНК, позволила высказать важное предположение о взаимосвязи процессов транскрипции и трансляции. Для того чтобы фактор ρ мог осуществлять свое действие, ему необходимо, во-первых, связаться с 5'-концом мРНК и, во-вторых, передвигаться вдоль молекулы. Рибосомы, транслирующие мРНК, препятствуют реализации каждого из этих условий. Поэтому достигнет ли фактор ρ терминатора, на котором сделала паузу РНК-полимераза, во многом зависит от событий, происходящих во время трансляции.

Исходя из этих предпосылок, можно объяснить наблюдение, которое длительное время вызывало недоумение. В ряде случаев нонсенс-мутация в одном из генов нарушает экспрессию расположенных за ним генов, входящих в состав той же транскрипционной единицы. Данное явление было названо **полярностью** (гл. 9). Причина этого явления — исчезновение мРНК, соответствующей дистальной части транскрипционной единицы. Длительное время продолжались споры о природе этого явления. Вызвано

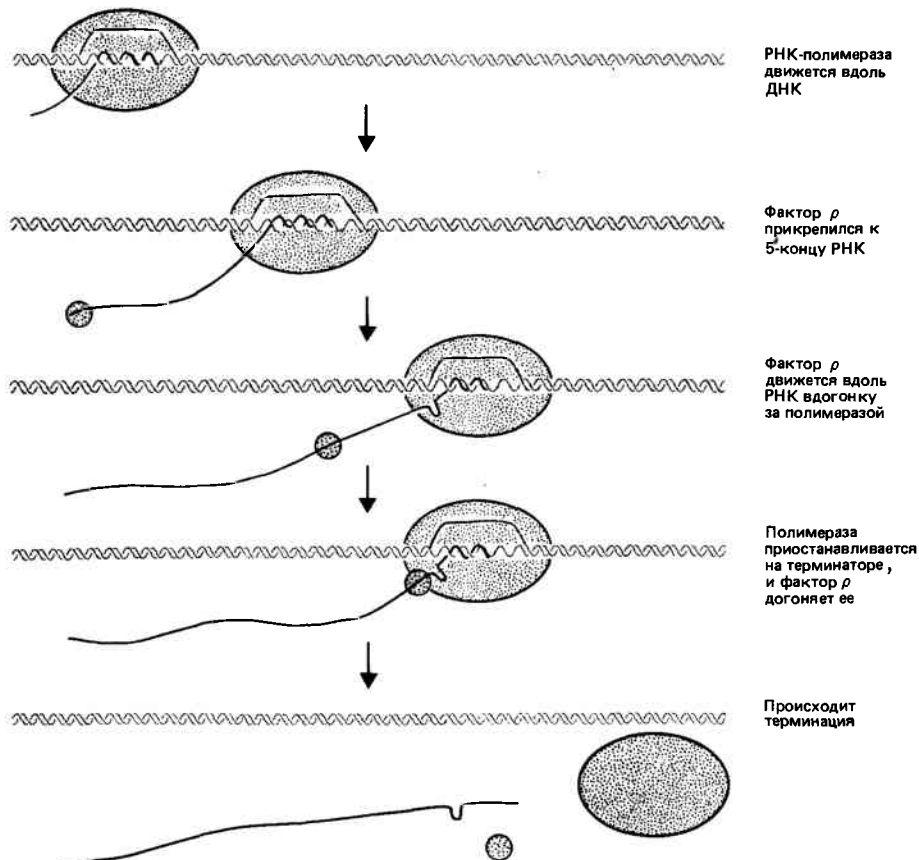


Рис. 13.3. Фактор ρ движется за РНК-полимеразой вдоль РНК. Когда на терминаторной шпильке происходит остановка поли-

меразы, у фактора появляется возможность догнать ее, вызвав тем самым терминацию.

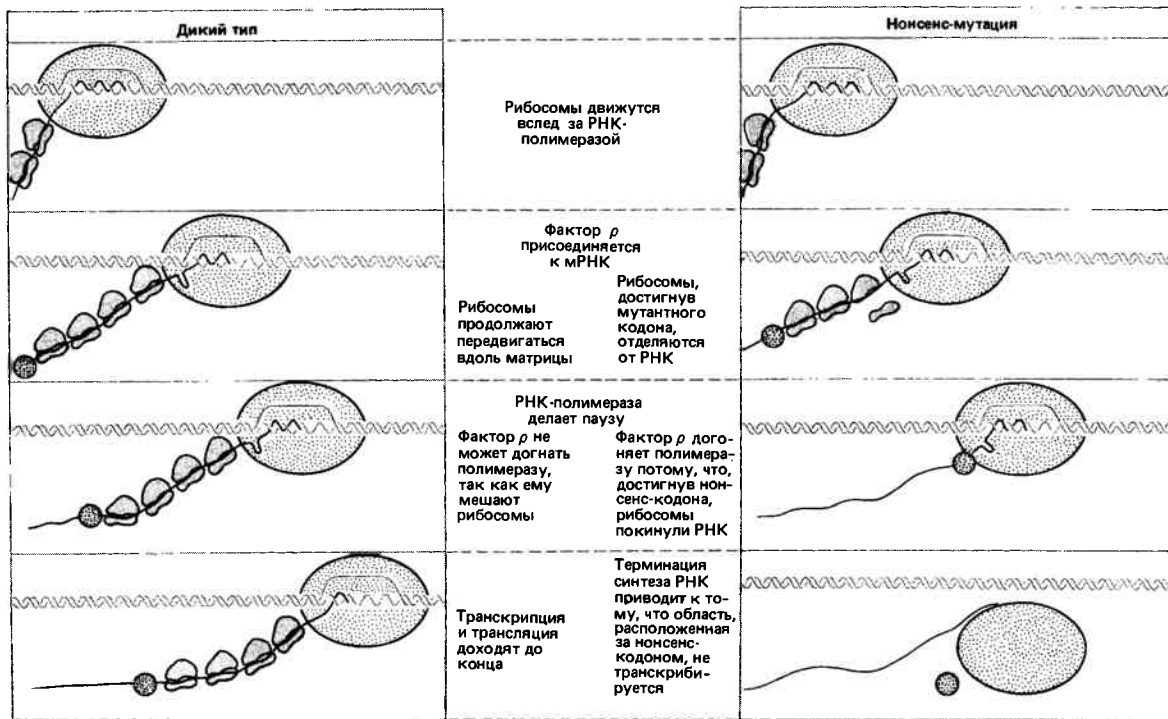


Рис. 13.4. Благодаря своему механизму действия фактор ρ способен осуществлять связь между процессами транскрипции

и трансляции при условии, что ρ -зависимый терминатор находится поблизости нонсенс-мутации.

ли это тем, что мРНК синтезируется и затем быстро деградирует, или же преждевременное завершение трансляции приводит к остановке синтеза РНК?

Предположим, что внутри транскрипционной единицы могут существовать ρ -зависимые терминаторы, располагающиеся перед основным терминатором. Тогда явление полярности можно объяснить как на рис. 13.4. Обычно терминаторы, находящиеся внутри транскрипционной единицы, не работают, так как из-за присутствия рибосом фактор ρ на этих терминаторах никогда не взаимодействует с РНК-полимеразой, хотя она делает здесь паузу. Нонсенс-мутации приводят к освобождению мРНК от рибосом, и поэтому у фактора ρ появляется возможность свободно передвигаться вдоль мРНК. В результате он может провзаимодействовать с полимеразой, остановившейся на терминаторе. Это приведет к терминации и освобождению РНК-полимеразы. Следовательно, дистальная часть транскрипционной единицы не сможет экспрессироваться. (Для чего существуют внутренние терминаторы? Возможно, что они являются просто шпильками, похожими на те, которые возникают на обычных терминаторах. Терминация с участием этих структур может происходить только в необычных условиях полярности).

Мутации по гену фактора ρ

Взаимодействие фактора ρ с другими белками, а также некоторые особенности его действия были обнаружены в результате выделения мутантов *E. coli* с измененным фактором ρ . Ген, кодирующий фактор ρ , получил название *rho*. Эти мутации вызывают многочисленные эффекты. Поэтому первоначально их выделяли, используя различные фенотипические проявления, а мутации получали различные названия; среди них наиболее хо-

рошо изученной является *suA*. Свое название *suA* получила за способность супрессировать полярность.

Мутации в гене *rho* в разной степени сказываются на процессе терминации. Общий эффект этих мутаций заключается в нарушении терминации. Но степень повреждения, если ее оценивать по проценту случаев прочитывания терминатора *in vivo*, зависит от конкретного терминатора. По-видимому, это отражает тот факт, что большинство мутаций, влияющих на активность фактора ρ , являются *leaky*, поэтому у них сохраняется вполне определенная терминирующая активность. Если для работы различных терминаторов необходимо разное количество фактора ρ , то оставшаяся в мутантных клетках активность может обеспечить работу некоторых терминаторов, для функционирования которых требуется низкая концентрация ρ , тогда как другие не смогут функционировать в этих условиях и РНК-полимераза сможет преодолеть их.

Способность мутаций *suA* супрессировать полярность объясняется тем, что они понижают вероятность взаимодействия фактора ρ с внутренними терминаторами, следовательно, за нонсенс-кодонами. Благодаря этому остановка трансляции не вызовет завершения транскрипции и область мРНК, расположенная за мутантным кодоном, сможет транслироваться вновь прикрепившимися рибосомами.

Некоторые мутации в гене фактора ρ можно супрессировать мутациями, возникающими в других генах. В результате появляется отличная возможность для выявления белков, контактирующих с фактором ρ . Участие β -субъединицы РНК-полимеразы во взаимодействии с фактором ρ обнаруживается благодаря существованию мутаций двух типов. Во-первых, мутации в гене *rhoB*, кодирующем эту субъединицу, могут ослаблять терминацию

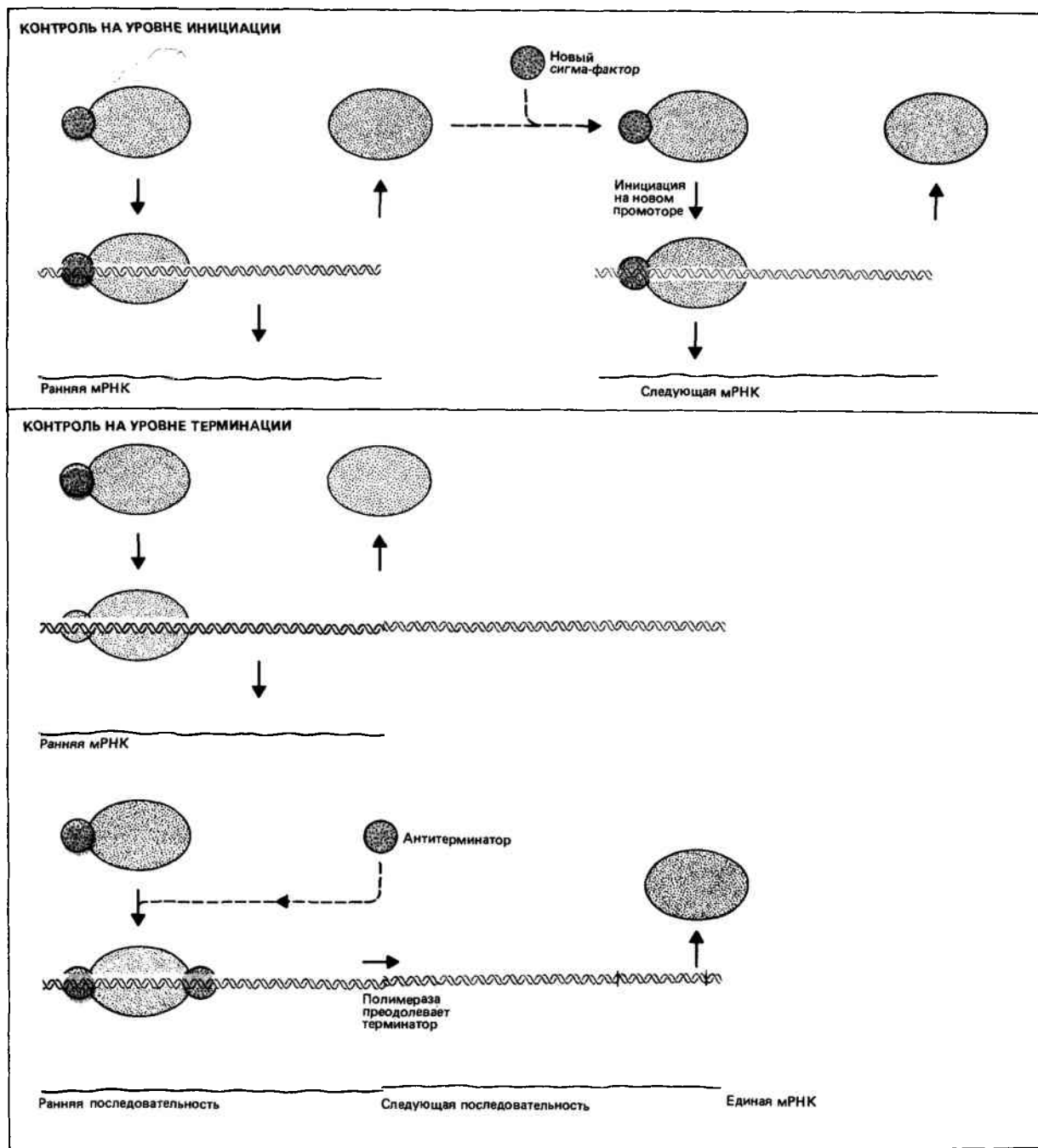


Рис. 13.5. Изменения транскрипционной специфичности могут контролироваться на уровне инициации или терминации.

в ρ -зависимых сайтах. Во-вторых, введение мутаций в ген *groV* может восстановить способность терминировать транскрипцию в ρ -зависимых сайтах у бактериальных мутантов, дефектных по этому фактору.

Механизм антитерминации, контролируемый фаговым геномом

У определенных фагов терминация используется в качестве удобного механизма, позволяющего контролировать переключение транскрипции с ранних генов на по-

следующую стадию экспрессии. Как можно применить антитерминацию для этих целей, показано на рис. 13.5. Для сравнения изображен случай, когда для регуляции генной экспрессии используется новый тип промоторов.

В гл. 12 мы видели, что одним из механизмов, позволяющим осуществлять переключение транскрипции с ранних генов на следующую стадию развития, является замена сигма-фактора в бактериальной РНК-полимеразе на другой фактор, изменяющий специфичность инициации фермента. Другая возможность состоит в синтезе новой фагоспецифической РНК-полимеразы. В любом из этих случаев основной особенностью, позволяющей

РНК-полимеразе отличать разнообразные классы генов, является наличие у них *различных промоторов*. Таким образом, как только синтезировались необходимый сигма-фактор или новая РНК-полимераза, экспрессия последующего класса генов уже не зависит от экспрессии ранних генов. Эти классы генов становятся независимыми друг от друга. Часто экспрессия ранних генов прекращается после переключения транскрипции на следующую стадию.

С другой стороны, возможность использования анти-терминационного механизма в качестве регуляторного зависит от конкретного расположения генов. Ранние гены должны непосредственно примыкать к генам, которые будут транскрибироваться вслед за ними. При этом необходимо, чтобы два класса генов были разделены терминатором. Если по какой-либо причине терминации не произойдет, то РНК-полимераза сможет транскрибировать гены, расположенные по другую сторону от сайта терминации. Таким образом, в результате антитерминации РНК-полимераза продолжает узнавать те же самые промоторы. Поэтому новые гены экспрессируются путем образования более длинных молекул РНК, содержащих на 5'-конце последовательности, соответствующие ранним генам, а на 3'-конце – последовательности, соответствующие новым генам. Поскольку оба типа последовательностей остаются объединенными, экспрессия ранних генов продолжается.

Наиболее хорошо изученным примером антитерминации является регуляция транскрипции при развитии фага лямбда. Характер контролируемых процессов у всех фагов, использующих бактериальную РНК-полимеразу для транскрипции ранних генов, имеет большое сходство. Во всех случаях один из ранних фаговых белков необходим для транскрипции последующего набора генов. У фага лямбда гены, транскрибируемые бактериальной РНК-полимеразой сразу же после начала инфекции, называются *предранними* или *немедленно ранними*, а гены, экспрессирующиеся на следующей стадии, – *задержанно ранними*.

Регуляторный ген, продукт которого контролирует переключение транскрипции с предранних генов на задержанно ранние, был идентифицирован благодаря мутациям, нарушающим переключение. Мутанты фага лямбда по гену *N* способны транскрибировать только предранние гены, и поэтому инфекционный процесс останавливается на этой стадии. Наблюдаемый эффект во многом похож на то, что происходило с фагом SPO1, несущим в гене 28 мутацию, которая нарушает образование σ_{SP28} . С генетической точки зрения безразлично, обусловлено ли включение транскрипции новых классов генов изменением специфичности инициации или антитерминацией. Оба процесса находятся под позитивным контролем со стороны раннего фагового гена, кодирующего белок, необходимый для включения следующего класса генов.

На **рис. 13.6** изображена генетическая карта ранней области фага лямбда. Два предранних гена *N* и *cro* транскрибируются РНК-полимеразой *E. coli* соответственно влево и вправо от промоторов, обозначенных P_L и P_R . Это означает, что при *левосторонней* и *правосторонней* транскрипции в качестве матрицы используются разные цепи ДНК. На **рис. 13.7** показано, что транскрипция, осуществляемая РНК-полимеразой *E. coli*, останавливается в конце генов *N* и *cro* соответственно на терминаторах t_{L1} и t_{R1} . Оба этих терминатора являются ρ -зависимыми. (Первоначально фактор ρ был обнаружен благодаря функционированию именно этих терминаторов.)

Ситуация изменяется в результате экспрессии гена *N*. Продуктом этого гена является белок антитерминатор, позволяющий РНК-полимеразе преодолевать t_{L1} и t_{R1} . В результате транскрибируются задержанно ранние гены, расположенные за этими терминаторами. Поскольку рN является очень нестабильным белком с периодом полужизни около 5 мин, для транскрипции задержанно ранних генов необходима постоянная экспрессия гена *N*. Тут не возникает никаких проблем, поскольку ген *N* является частью транскрипционной единицы, содержащей задержанно ранние гены, и его транскрипция предшествует транскрипции задержанно ранних генов.

Как и у других фагов, в случае фага лямбда для экспрессии поздних генов, кодирующих компоненты фаговой частицы, необходимы дополнительные регуляторные механизмы. Такой контроль осуществляется геном *Q*, относящимся к задержанно ранним генам. Продукт этого гена рQ является еще одним белком-антитерминатором, позволяющим РНК-полимеразе, иницирующей транскрипцию на позднем промоторе P_R , преодолевать терминатор, располагающийся между этим промотором и поздними генами. Так, наличие белков-антитерминаторов с различной специфичностью позволяет осуществить последовательную экспрессию фаговых генов.

Антитерминация зависит от определенных сайтов в ДНК

Антитерминирующее действие белка рN высокоспецифично. Он не подавляет терминацию на всех ρ -зависимых сайтах. Например, после синтеза рN терминация по-прежнему происходит в конце бактериальных генов. В то же время антитерминационные события не определяются терминаторами t_{L1} и t_{R1} , так как если бактериальный ген, содержащий ρ -зависимый терминатор, встраивается в раннюю область фага лямбда, то белок рN способен обеспечивать антитерминацию на его терминаторе. Следовательно, антитерминация происходит на *любом* терминаторе, который встречается на пути РНК-полимеразы,

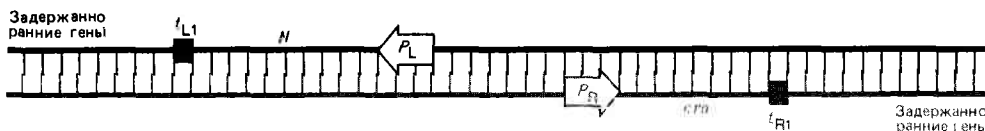


Рис. 13.6. У фага лямбда имеются две ранние транскрипционные единицы.

Карта бактериофага лямбда изображена в виде параллельных линий, соответствующих двум цепям ДНК. Верхняя цепь транскрибируется в ле-

вом направлении, тогда как нижняя – в правом. Промоторы показаны стрелками. Закрашенными квадратиками обозначены терминаторы. Гены *N* и *cro* являются предранними; они отделяются от задержанно ранних генов терминаторами.

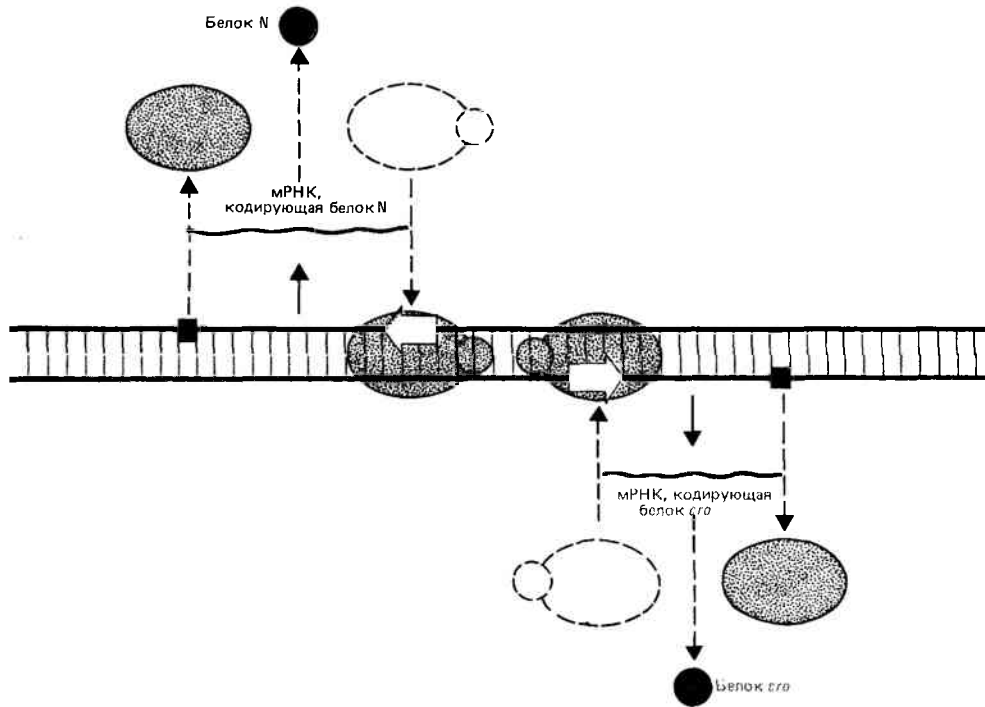


Рис. 13.7. Предранние гены фага лямбда транскрибирует бактериальная РНК-полимераза.

Фермент связывается с двумя промоторами P_L и P_R и транскрибирует гены N и cro , останавливаясь на соответствующих терминаторах t_{L1} и

t_{R1} . В обоих случаях для терминации необходимо присутствие фактора ρ (это не отражено на рисунке). В результате транскрипции образуются две предранние мРНК, кодирующие белки pN и cro .

транскрибирующей раннюю область бактериофага лямбда. Сведения по этому вопросу обобщены в табл. 13.1.

Таким образом, в каждой из этих транскрипционных единиц должен существовать некий сайт, который узнается белком pN . Этот сайт является сигналом, указывающим на то, что антитерминация должна произойти на первом же встретившемся терминаторе. Другими словами, сайт, необходимый для антитерминации и узнаваемый белком pN , и сайт терминации, в котором при определенных обстоятельствах данный белок обеспечивает антитерминацию, находится в разных местах. На основе этих данных можно сделать более общее заключение.

Если нам известен участок ДНК, взаимодействуя с которым белок проявляет свою активность, это еще не дает оснований утверждать, что данный участок является последовательностью, изначально узнаваемой рассматриваемым белком. Эти участки могут находиться далеко друг от друга.

Таблица 13.1

Процесс антитерминации может затрагивать все РНК-полимеразы, инициирующие синтез на ранних промоторах фага лямбда, независимо от вида терминатора

Вид локусов промотор → терминатор		События, происходящие в клетках, инфицированных фагом лямбда
Бактериальный Ранний фага лямбда	Бактериальный Ранний фага лямбда	Терминация происходит Белок pN предотвращает терминацию
Ранний фага лямбда	Бактериальный	Белок pN предотвращает терминацию

Участок ДНК, узнаваемый белком pN , называется *nut* (N utilization). Участки, ответственные за антитерминацию при левосторонней и правосторонней транскрипции, обозначаются соответственно как *nutL* и *nutR*. Нам известно, что эти участки должны находиться между P_L и t_{L1} в одной транскрипционной единице и между P_R и t_{R1} — в другой. Где же точно они расположены? Их можно картировать с помощью гибридов фага лямбда с другими фагами, у которых имеются замены в определенных участках генома, либо с помощью делеций в ДНК фага лямбда, предотвращающих антитерминацию, а также путем выделения точковых мутаций *nut⁻*, блокирующих антитерминацию.

С помощью таких экспериментов удалось локализовать *nutL* между стартовой точкой промотора P_L и началом области, кодирующей ген N ; *nutR* находится между концом гена *cro* и терминатором t_{R1} . Как видно из рис. 13.8, это означает, что оба сайта располагаются в совершенно различных областях соответствующих транскрипционных единиц. В то время как *nutL* находится в начале, про *nutR* можно сказать, что он расположен близко к терминатору.

Как же осуществляется антитерминация? Когда белок pN узнает сайт *nut*, он должен провзаимодействовать с РНК-полимеразой в этом сайте, модифицировав ее таким образом, что она в дальнейшем, находясь на терминаторе, уже не отвечает на фактор ρ . Различная локализация сайтов *nut* указывает на то, что это событие не связано ни с инициацией, ни с терминацией, а осуществляется под действием РНК-полимеразы в момент элонгации РНК в сайте *nut*.

Затем РНК-полимераза становится в такой степени

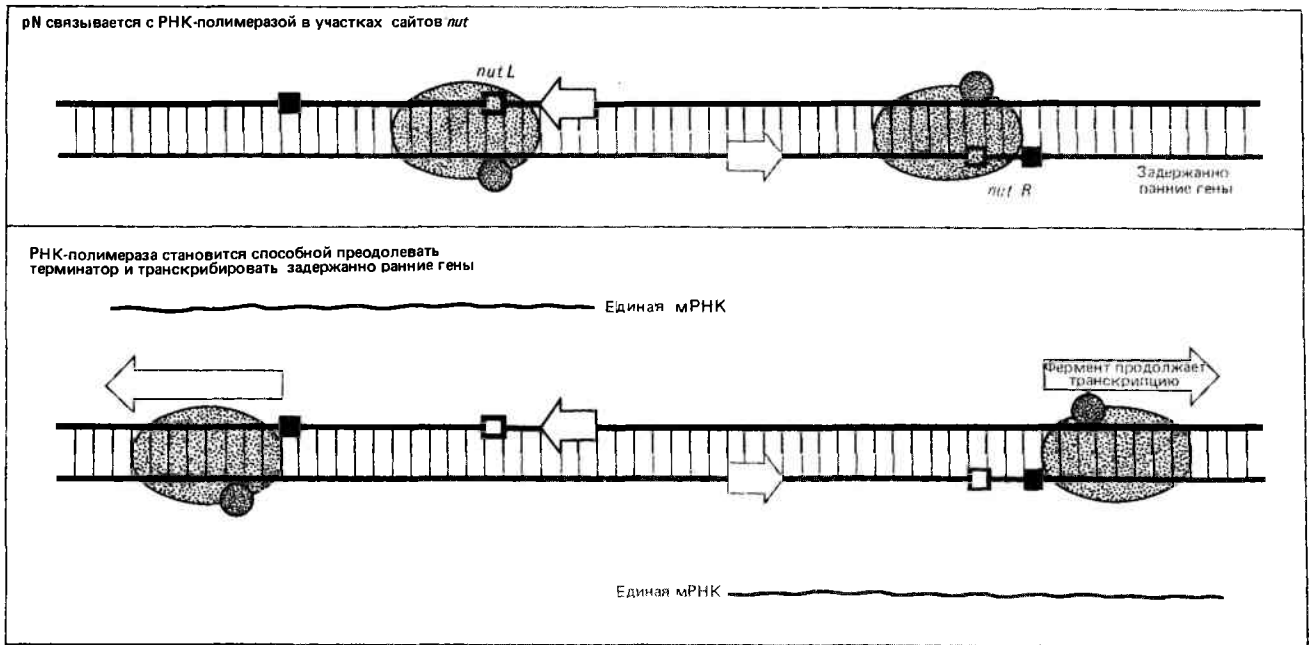


Рис. 13.8. Белок рN связывается с РНК-полимеразой, когда она проходит через сайт *nut*. В результате фермент становится способным преодолеть терминаторы t_{L1} и t_{R1} , образуя при этом

единую мРНК, содержащую последовательности, соответствующие как задержанно ранним, так и предранним генам.

неконтролируемой, что преодолевает терминатор, не обращая внимания на его сигнал.

Анализ нуклеотидной последовательности ДНК выявил два участка из 17 пар оснований, идентичных, за исключением одной нуклеотидной замены, и одинаково ориентированных по отношению к направлению транскрипции. Последовательность *nut* включает инвертированный повтор из 5 пар оснований, две копии которого разделены несимметричной областью следующим образом:

A G C C C T G A A P_u A A G G G C A
T C G G G A C T T P_y T T C C C G T

где окрашенными прямоугольниками обозначены инвертированные повторы, а пара оснований P_u-P_y указывает на единственное различие между *nutL* и *nutR*.

Является ли способность белка рN узнавать короткую последовательность в транскрипционной единице примером более широко используемого механизма антитерминации? Существуют другие фаги, родственные лямбда, которые имеют различные гены *N* и различную антитерминаторную специфичность. Область фагового генома, в которой находятся сайты *nut*, имеет различную последовательность у разных фагов, и, вероятно, каждый фаг обладает только ему свойственными сайтами *nut*, которые специфически узнаются собственным белком рN. Каждый из этих белков рN должен обладать одинаковой способностью взаимодействовать с транскрипционным аппаратом, осуществляя антитерминаторную функцию, но при этом иметь различную специфичность по отношению к последовательности ДНК, активирующей этот процесс.

Существуют ли дополнительные субъединицы у РНК-полимеразы?

Открытие антитерминации как фагоспецифического механизма помогло обнаружить другие бактериальные белки, являющиеся компонентами транскрипционного аппарата. Бактериальные белки, с которыми взаимодействует белок рN в процессе функционирования, могут быть выявлены с помощью мутантов *E. coli*, в которых белок рN не может обеспечить антитерминацию. Эти мутанты резистентны к инфицированию фагом лямбда, так как в этих клетках фаг способен экспрессировать только свои предранние гены. Некоторые из этих мутаций картируются в гене *rpoB*. Это свидетельствует в пользу того, что белок рN взаимодействует с β -субъединицей минимального фермента.

Другие мутации *E. coli*, блокирующие функцию белка рN, расположены в генах, не кодирующих компонентов РНК-полимеразы. Они обозначаются как локусы *nus*, которых насчитывается по крайней мере три: *nusA*, *nusB* и *nusE* (от англ. N utilization substance - N утилизирующая субстанция). Локус *nus* должен кодировать белки, которые являясь частью транскрипционного аппарата, не выделяются в составе обычной формы РНК-полимеразы.

Функции, контролируемые *nusA* и *nusB*, по-видимому, имеют отношение исключительно к терминированию транскрипции, в то время как *nusE* оказывается идентичен локусу *rpsJ*, кодирующему рибосомный белок S10. Вероятно, функция белка S10 при транскрипции отличается от той функции, которую он выполняет в рибосоме, хотя вполне возможно, что этот белок принимает определенное участие в сопряжении процессов транскрипции и трансляции.

Функция, контролируемая *nusA*, в значительной степени связана с действием белка рN. Можно получить мутации по гену *N*, которые преодолевают блок антитерминации, обусловленный мутациями *nusA*. Предположение о том, что *nusA* необходим для функционирования белка рN, подтверждается тем фактом, что *in vitro* два этих белка связываются между собой. (*nusA* является одним из двух белков, прочно связывающихся *in vitro* с белком рN. Второй белок — это полипептид с мол. массой 25000 дальтон с неизвестной генетической природой и функцией.) рN представляет собою маленький (13500 дальтон), сильно щелочной и довольно асимметричный по форме белок. Мы пока еще не располагаем удовлетворительной системой *in vitro* для изучения его функции.

Продукт гена *nusA* является кислым белком с мол. массой 69000 дальтон. *In vitro* он функционирует как фактор терминации на сайтах t_{L2} и t_{R2} фага лямбда, которые классифицированы как р-независимые сайты; они располагаются дистально по отношению к р-зависимым сайтам t_{L1} и t_{R1} . Фактически суть действия белка NusA состоит в том, что он вызывает остановку РНК-полимеразы на этих терминаторах на длительное время — около 15 мин. Для освобождения новосинтезированной цепи РНК требуется присутствие фактора р, но в очень низких концентрациях, которые сами по себе не эффективны в отсутствие NusA-белка. Это свидетельствует в пользу того, что действие белка NusA нетождественно действию фактора р, а дополняет его. Генетические эксперименты показывают, что мутации *nusA* и *nusB* обладают сходными свойствами (обе супрессируют полярность, как можно было бы предполагать из-за отсутствия фактора терминации); поэтому, по-видимому, они функционируют аналогичным образом.

Белок NusA связывается с минимальным ферментом, но не взаимодействует с голоферментом. Однако при добавлении сигма-фактора к комплексу $\alpha_2\beta\beta'\text{NusA}$ он замещает белок NusA, восстанавливая, таким образом, структуру голофермента $\alpha_2\beta\beta'\sigma$. Можно полагать, что РНК-полимераза совершает цикл, изображенный на рис. 13.9, в котором она существует поочередно либо в форме фермента, готового к инициации ($\alpha_2\beta\beta'\sigma$), либо в форме фермента, готового к терминации ($\alpha_2\beta\beta'\text{NusA}$).

В гл. 10 мы уже описывали, что инициация происходит поэтапно. Голофермент $\alpha_2\beta\beta'\sigma$ связывается с промотором и освобождает сигма-фактор. Это приводит к образованию минимального фермента $\alpha_2\beta\beta'$, способного осуществлять элонгацию. Затем белок NusA узнает минимальный фермент и связывается с ним, образуя комплекс $\alpha_2\beta\beta'\text{NusA}$. Мы не знаем, на каком этапе это случается; это могло бы произойти как в любой момент элонгации, так и просто в сайте терминации. В любом случае, пока комплекс $\alpha_2\beta\beta'\text{NusA}$ связан с ДНК, компонент NusA не может быть вытеснен. Но когда терминация произошла, фермент переходит в состояние, в котором NusA или автоматически отделяется от минимального фермента, или может быть замещен сигма-фактором.

С точки зрения минимального фермента жизнь состоит из чередования процессов ассоциации с сигма-фактором для инициации или с белком NusA для терминации. Одновременное взаимодействие обоих факторов с минимальным ферментом невозможно. Нет основания считать один фактор более присущим ферменту, чем другой, поскольку оба они, по-видимому, его альтернативные компоненты. Таким образом, выделение мини-

мального фермента представляет собой получение формы РНК-полимеразы, способной осуществлять основную функцию при синтезе РНК, но лишенной субъединиц, необходимых для других функций. Как мы уже отмечали, до конца не ясно, где проходит граница между собственно минимальным ферментом и полным транскрипционным комплексом.

По-видимому, фактор NusA обеспечивает связь между белком рN и минимальным ферментом. Это приобретает смысл при условии, что белок NusA является терминационной субъединицей, чье присоединение к ферменту обеспечивает специфичность по отношению к определенным терминаторам. Таким образом, белок рN может действовать как антитерминатор, мешая NusA проявлять функцию, необходимую для терминации. Отсюда напрашивается вывод, что бактерия-хозяин может, вероятно, содержать другие белки, аналогичные NusA (или рN), чьи взаимодействия с минимальным ферментом контролируют узнавание сайтов терминации.

Это заключение объяснило бы любопытный факт, касающийся NusA. Он идентичен белку, ранее описанному как фактор L, который необходим для транскрипции *in vitro* некоторых генов. Это свойство не имеет ничего общего с тем, что можно было бы ожидать от фактора терминации. Но если именно присутствие белка NusA позволяет этим гипотетическим бактериальным функциям подавлять терминацию в определенных сайтах, то NusA мог бы быть необходимым для транскрипции любого гена, расположенного за таким сайтом. Возможно также, что сама субъединица NusA могла бы придавать минимальному ферменту способность прочитывать некоторые *t*-сайты и в то же время распознавать другие.

Пока еще мы не располагаем точными сведениями о том, как взаимодействуют белок NusA, минимальный фермент и фактор р при терминации, но идентификация NusA как терминаторной субъединицы, а также существование фаговых белков рN, различных по специфичности, но, несмотря на это, по-видимому, способных взаимодействовать с *nusA*, предполагают возможность существования целого ряда других (неизвестных) факторов, влияющих на минимальный фермент или модифицирующих его свойства. Это означает также, что минимальный фермент, возможно, сам по себе не способен осуществлять терминацию, и поэтому эксперименты *in vitro* по терминации с минимальным ферментом могут неадекватно отражать его свойства. Таким образом, природу р-независимых сайтов необходимо исследовать заново в системе *in vitro*, содержащей другие возможные терминационные компоненты.

Трудности в изучении терминации у эукариот

В отношении большинства эукариотических РНК-полимераз мало что известно как о сигналах терминации, так и о самом процессе терминации. Наибольшая трудность при анализе транскриптов — это отсутствие уверенности в истинности сайта терминации. Хотя молекулы РНК, синтезированные *in vivo*, имеют определенные 3'-концы, можем ли мы узнать, были ли они образованы в результате терминации или в результате разрыва в более длинном первоначальном транскрипте?

В случае транскриптов РНК-полимеразы II проблема осложняется процессингом, происходящим у 3'-конца,

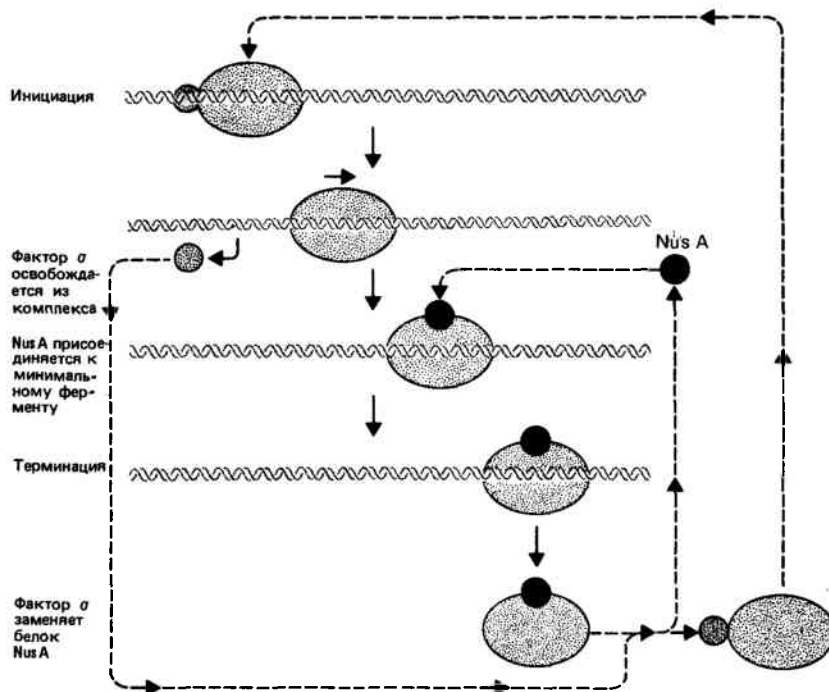


Рис. 13.9. РНК-полимераза может осуществлять либо инициацию, либо терминацию в зависимости от того, какой из двух белков, σ или NusA, присоединится к минимальному ферменту.

к которому присоединена длинная полиадениловая кислота. По крайней мере в некоторых случаях доказано, что 3'-концы для полиаденилирования в действительности образованы в результате разрыва, а не терминации. Следовательно, мы располагаем некоторой информацией о сигналах, необходимых для образования конца, к которому присоединяется poly(A), но 3'-область для подлинно первоначального транскрипта в целом остается нехарактеризованной. В ДНК вируса SV40 имеется сайт, вызывающий терминацию; его последовательность похожа на р-независимый бактериальный терминатор, имеющий структуру шпильки, за которой следует последовательность оснований U.

Для РНК-полимеразы I единственным транскриптом является большой предшественник, содержащий последовательности для высокомолекулярных рРНК. Предшественник подвергается интенсивному процессингу, и еще не доказано, действительно ли 3'-конец образуется в результате терминации. Тем не менее конец соответствует участку из оснований U.

При транскрипции *in vitro* РНК-полимераза III образует молекулы, имеющие такие же 5'- и 3'-концы, как синтезированные *in vivo*. Из этого можно сделать вывод, что произошли истинные инициация и терминация. Данная система может быть изучена благодаря введению изменений в последовательность матрицы вокруг терминационной области.

Когда гены для 5S-РНК *X. laevis* транскрибируются *in vitro* гомологичным ферментом, терминация происходит внутри участка из четырех оснований U. Среди образовавшихся молекул наблюдается некоторая гетерогенность за счет того, что часть их оканчивается на третьем или даже на четвертом основании U, тогда как обычно терминация происходит на втором основании U. Это, по-видимому, истинное свойство терминационного процесса, так как такая же гетерогенность наблюдается среди молекул, синтезированных *in vivo*.

Точно так же, как и в прокариотических терминаторах, участок U расположен в G—C-богатом районе. Не-

смотря на то что в эукариотических терминаторах присутствуют последовательности со вторичной симметрией, они не нужны для терминации. Так, мутации в матрице, нарушающие симметрию, не препятствуют нормальному окончанию синтеза РНК. Последовательность, расположенная за участком U, не существенна, поскольку все нуклеотиды справа от него могут быть заменены на другие последовательности без какого-нибудь ущерба для терминации.

Самого по себе U-участка недостаточно для терминации, так как область из четырех последовательных остатков U, встречающаяся внутри транскрипционных единиц, прочитывается РНК полимеразой III. (Однако не существует внутренних участков из пяти U—возможно, в связи с тем, что они являются более эффективными терминаторами. Это следует из сравнения терминаторной способности участков, состоящих из пяти и четырех оснований U). Следовательно, решающим моментом в терминации должно быть распознавание последовательности U₄, находящейся в области, богатой G—C-парами.

Как же происходит реакция терминаций? Нельзя полагаться на слабое взаимодействие рU-dA в РНК—ДНК гибридной области, которая расположена на конце транскрипта, потому что часто только два первых остатка UU транскрибируются перед терминацией. Возможно, G—C-богатая область играет определенную роль в остановке фермента, но, по-видимому, не существует чего-либо подобного шпильке, которая участвует в терминации у прокариот. Мы не знаем, каким образом фермент способен так специфично отвечать на столь короткий сигнал.

Рекомендуемая литература

Терминация у бактерий и ее использование в качестве контролирующего механизма описана Робертсом [Roberts, in RNA Polymerase (Lasick, Chamberlin, Eds.), Cold Spring Harbor Laboratory, New York, 1976, 247–271],

а также в обзоре Адхьяса и Готтесмана (*Adhya, Gottesman, Ann. Rev. Biochem.*, **47**, 967–996, 1978). Природа сигналов терминации обсуждается в обзоре Розенберга и Курта в связи с рассмотрением вопроса о канонических последовательностях (*Rosenberg, Court, Ann. Rev. Genet.*, **13**, 319–353, 1979). Описание более сложных бактериальных транскрипционных комплексов дано Гринблаттом (*Grinblatt, Cell*, **24**, 8–9, 1981). Поскольку о терминации у эукариот известно очень мало, публикация обзоров по этому вопросу – пока еще дело будущего.

Часть IV

КОНТРОЛЬ ГЕННОЙ ЭКСПРЕССИИ У ПРОКАРИОТ

В строго структурной концепции геном рассматривается как мозаика, составленная из независимых молекулярных программ, планирующих создание отдельных клеточных компонентов. Однако для выполнения этих программ абсолютно жизненно важное значение имеет координация событий. Открытие генов-регуляторов и генов-операторов, а также регуляции активности структурных генов посредством репрессии показало, что геном содержит не только серию программ, но также координированную систему белкового синтеза и средства, позволяющие контролировать их выполнение.

Франсуа Жакоб и Жак Моно, 1961

ОПЕРОН НА ПРИМЕРЕ ОРГАНИЗАЦИИ ЛАКТОЗНЫХ ГЕНОВ

Бактериям необходимо уметь быстро отвечать на изменения в окружающей среде. Их выживаемость зависит от способности переключать метаболизм с одного субстрата на другой, поскольку в условиях их обитания поступление питательных веществ может постоянно меняться. Одноклеточные эукариоты способны проявлять такую же зависимость от непрерывно меняющихся условий; более же сложно организованные многоклеточные организмы имеют более постоянный набор метаболических путей и могут не отвечать на изменение внешних условий перестройкой своего метаболизма.

Легкая приспособляемость – большое преимущество бактерий. Однако при этом им необходима экономичность, поскольку бактерия, которая в ответ на изменение условий окружающей среды использует пути, требующие больших энергетических затрат, может оказаться в невыгодном положении. Совершенно очевидно, что было бы неэкономично производить все ферменты метаболического пути, который не может быть использован из-за отсутствия субстрата. Бактерия идет по компромиссному пути: она не синтезирует ферментов того или иного пути в отсутствие соответствующего субстрата, но способна в любое время начать их синтез при появлении такового.

Этим объясняются главные особенности организации бактериальных генов – их объединение в кластеры таким образом, что все ферменты, необходимые для осуществления определенного пути биосинтеза, детерминируются генами, сцепленными друг с другом. Вся группа генов может транскрибироваться в одну полицистронную мРНК, которая последовательно транслируется рибосомами с образованием каждого из белков.

В предыдущих главах мы упоминали, что лишь несколько генов *E. coli* транскрибируются индивидуально, тогда как большинство является частью больших единиц транскрипции. Такая форма организации позволяет координированно регулировать выражение всех генов одной единицы транскрипции посредством контроля ее транскрипции. Последнее достигается взаимодействием регуляторного белка с регуляторным сайтом, расположенным рядом с промотором.

Индукция и репрессия контролируется малыми молекулами

Уже к началу этого века стало известно, что определенные ферменты дрожжей образуются в клетках только при выращивании их на определенных субстратах, – эффект, получивший название «индукция ферментов». Индукция была исследована у бактерий, и в частности у *E. coli*, у которой контролирующийся механизм такого типа представлен организацией генов лактозной системы.

При выращивании клеток в отсутствие β-галактозида они содержат очень незначительное число молекул фермента β-галактозидазы – скажем, меньше пяти. Функция фермента сводится к разложению молекулы β-галактозида на составляющие ее сахара. Например, лактоза разла-

гается на глюкозу и галактозу, которые в свою очередь используются в дальнейшем метаболизме. В отсутствие субстрата в ферменте нет необходимости.

При добавлении соответствующего субстрата в бактериальных клетках очень быстро проявляется активность фермента в результате синтеза его новых молекул. Они появляются уже через 2–3 мин, и вскоре их количество достигает свыше 5000 молекул фермента в расчете на одну бактериальную клетку. (При определенных условиях на долю β-галактозидазы может приходиться 5–10% от общего растворимого белка бактерии.) При удалении из среды субстрата синтез фермента прекращается так же быстро, как был начат.

Такая реакция на изменения в составе питательной среды наблюдается не только при необходимости использовать новые субстраты; она используется также для исключения синтезов эндогенных соединений при их внешнем появлении в среде. Например, триптофан, одна из существенных аминокислот, синтезируется при участии фермента триптофан-синтазы. Однако если в среде, на которой выращиваются бактерии, присутствует триптофан, синтез фермента немедленно прекращается. Это явление получило название **репрессия**. Она позволяет бактериальной клетке избежать перевыпуска своих ресурсов на ненужную в данный момент синтетическую активность.

Индукция и репрессия представляют собой разные стороны одного и того же явления. В одном случае бактериальная клетка регулирует свою способность к использованию данного субстрата, в другом она регулирует способность к синтезу определенного промежуточного компонента метаболического пути. Функционирование каждого типа регуляции определяется малыми молекулами, которые являются либо субстратом для фермента, либо продуктом ферментативной деятельности соответственно. Малые молекулы, индуцирующие образование ферментов, способных метаболизировать их, названы **индукторами**. Те же, которые предотвращают образование ферментов, способных синтезировать их, названы **корепрессорами**.

Способность соединений действовать в качестве индуктора или корепрессора высокоспецифична. Их функцию могут осуществлять только субстраты/продукты или близкородственные молекулы. В обоих случаях активность малых молекул не зависит от их взаимодействия с ферментом. Так, существуют молекулы, которые достаточно сходны с природными индукторами β-галактозидазы, но не могут быть метаболизированы ферментом. Наиболее показательным примером может служить изопронилиогалактозид (ИПТГ), один из нескольких триогалактозидов с таким свойством. ИПТГ представляет собой очень эффективный индуктор, хотя он и не узнается β-галактозидазой.

Молекулы, индуцирующие синтез ферментов, но не являющиеся субстратами, получили название **добровольных индукторов**. Они оказались крайне полезными благодаря их способности сохраняться в клетке в исходном состоянии. (Истинный индуктор был бы использован

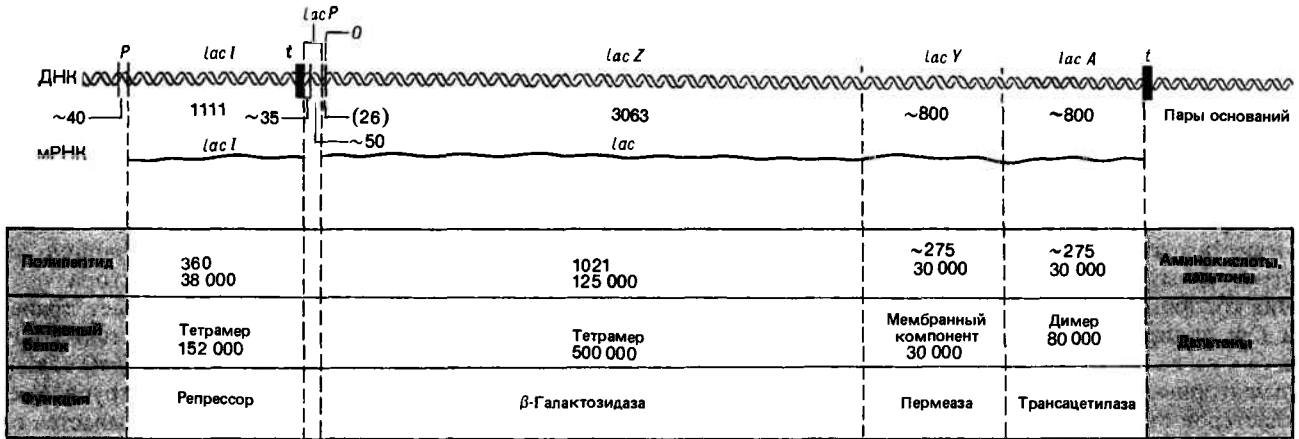


Рис. 14.1. На *lac*-оперон приходится около 6000 пар нуклеотидов ДНК.

В левой части расположен ген *lacI*, имеющий свой собственный промотор и, по-видимому, терминатор. Конец кодирующей области гена *lacI* непосредственно прилегает к *lacP*-промотору, который разделяется на

левую половину, связывающуюся с БАК, и правую половину, связывающуюся с РНК-полимеразой. Оператору соответствуют первые 26 пар оснований гена *lacZ*, который отличается чрезвычайно большой протяженностью. За геном *lacZ* следуют гены *lacY*, *lacA* и, по-видимому, терминатор.

в качестве субстрата фермента, а это препятствовало бы изучению системы.) Существование добровольных индукторов позволило установить важный факт. Помимо ферментативной активности система должна обладать еще какой-то функцией, способной узнавать соответствующий субстрат; эта способность узнавать родственные потенциальные субстраты отличается от таковой у фермента.

Кластеры генов регулируются координированно

Процессы индукции и репрессии обычно связаны не с одним ферментом. Мы уже указывали, что часто все ферменты одного пути биосинтеза регулируются совместно. Наряду с ферментами, действительно участвующими в определенном процессе биосинтеза, в координированно контролируемую единицу могут включаться другие родственные активности, например белок, ответственный за транспорт малых молекул в клетку. Примером такого типа регуляторной единицы служат *лактозные (lac)* гены.

На рис. 14.1 изображена группа из трех генов. Один из них, *lacZ*, кодирует фермент β -галактозидазу, который в активной форме представляет собой тетрамер с мол. массой около 500 000 дальтон. Второй, *lacY*, кодирует β -галактозидпермеазу, белок с мол. массой 30 000 дальтон, связывающийся с мембраной и являющийся компонентом системы транспорта. Третий, *lacA*, кодирует β -галактозид-трансацетилазу, фермент, который переносит ацетильную группу от ацетил-СоА к β -галактозидам.

Мутации в генах *lacZ* или *lacY* могут приводить к *lac⁻*-генотипу, при котором клетки не могут использовать лактозу. Мутации *lacZ⁻* ведут к утрате ферментативной активности, что препятствует сбраживанию лактозы. Мутанты *lacY⁻* не способны использовать лактозу среды. У клеток *lacA⁻* не обнаружено идентифицируемого дефекта, и функция этого гена остается невыясненной. Значение реакции ацетилирования непонятно, однако вполне возможно, что она обеспечивает бактериям преимущество при росте в присутствии определенных неме-

таболизируемых аналогов β -галактозидов, поскольку эта реакция ведет к их детоксикации и выделению.

Кластер трех генов *lacZYA* транскрибируется в виде одной мРНК с промотора, прилегающего к гену *lacZ*. Их индукция контролируется на уровне транскрипции. В отсутствие индуктора группа генов не транскрибируется. При добавлении индуктора транскрипция инициируется в промоторе *lacP* и продолжается, проходя через гены, в направлении терминатора, расположенного где-то за пределами гена *lacA*. Этим достигается **координированная регуляция**, при которой *все* гены выражаются (или не выражаются) одновременно.

Трансляция мРНК осуществляется последовательно с 5'-конца. Это делает понятным, почему при индукции сначала появляется β -галактозидаза, затем β -галактозидпермеаза и, наконец, β -галактозид-трансацетилаза. Выражение трех генов с помощью единой мРНК объясняет также, почему относительные количества трех ферментов остаются одинаковыми при варьирующих условиях индукции. Индукция, по существу, служит сигналом, вызывающим выражение генов. Индукторы могут различаться по своей эффективности. Кроме того, другие факторы могут влиять на абсолютный уровень транскрипции и трансляции, однако соотношение между продуктами трех генов предопределено организацией генов.

мРНК отличается крайней нестабильностью: период ее полураспада составляет около 3 мин. Именно это свойство мРНК позволяет очень быстро выключать этот процесс. Транскрипция прекращается, как только удаляется индуктор. В очень короткий промежуток времени вся лактозная мРНК распадается, и клетка прекращает образование ферментов.

Регуляторный ген контролирует структурные гены

Гены, кодирующие необходимые для клетки белки с ферментативными или структурными функциями, называются **структурными генами**. Подавляющее большинство

бактериальных генов попадает в эту категорию и обеспечивает огромное разнообразие структуры и функций белков. В группу структурных генов входят также гены, кодирующие рРНК и тРНК. Все типы структурных генов имеют тенденцию быть организованными в группы, которые могут координированно контролироваться таким же способом, как в случае трех структурных генов *lacZYA*.

Что служит мишенью для малых молекул индуктора? Нам известно, что такая мишень отличается от продуктов структурных генов. Это другой белок, единственная функция которого – контроль за выражением структурных генов. Ген, кодирующий этот белок, назван **регуляторным геном**.

Регуляторные гены ответственны за контроль выражения кластеров структурных генов, осуществляя его обычно путем синтеза белков, контролирующих транскрипцию. Регуляторные белки выполняют эту функцию, связываясь с определенными сайтами в ДНК.

Мы можем отличить структурные гены от регуляторных по эффекту мутаций. Мутация в структурном гене ведет к отсутствию в клетке определенного белка, кодируемого этим геном. Мутация же в регуляторном гене влияет на выражение всех структурных генов, которые он контролирует. Природа такого влияния зависит от типа регуляции.

Выражение *lac*-генов контролируется по типу **негативной регуляции**. Из этого следует, что гены транскрибируются *при условии*, что они *не выключены* регуляторным белком. Следовательно, при мутации, инактивирующей репрессор, гены остаются в активном состоянии. Поскольку функция регулятора сводится к *предотвращению* выражения структурных генов, он был назван **белком-репрессором**.

Регуляторные белки обладают и другими свойствами. Мы уже видели, что существуют **позитивные регуляторы**, которые названы так потому, что в их присутствии выражение структурных генов *включается*. В отсутствие регулятора гены не могут выражаться. Примером регуляции такого типа является инициирование транскрипции путем образования новых сигма-факторов (гл. 12) или специфическая антитерминация транскрипции (гл. 13).

Контролирующая система оперона

Концепцию, согласно которой два разных класса генов различаются по своим функциям, сформулировали Жакоб и Моно в 1961 г. при создании классической модели **оперона**. Они определили оперон как полную единицу выражения генов, включающую структурные гены, регуляторный ген (или гены) и контролирующие элементы (сайты действия регуляторных белков). Каждый отдельный оперон может быть охарактеризован системой взаимоотношений между регуляторными белками и их сайтами-мишенями. Это можно рассматривать как регуляторную систему. Первая модель была создана для *lac*-оперона, который до настоящего времени остается наиболее полно охарактеризованным опероном, однако подобные принципы были использованы для создания аналогичных систем в случае других оперонов.

На **рис. 14.2** приведена основная система контроля лактозного оперона. Группа из трех структурных генов *lacZYA*, условно представленная в ориентации слева на-

право, образует единицу транскрипции. Транскрипция иницируется в промоторе *lacP*, расположенном сразу же слева (против хода транскрипции) от первого гена *lacZ*. Регуляторный ген, *lacI*, лежит немного дальше влево и образует независимую единицу транскрипции. Он транскрибируется со своего собственного промотора в моностронную РНК, которая кодирует белок-репрессор. Основные свойства контролирующей системы определяются двойственной природой этого репрессора: с одной стороны, он способен предотвращать транскрипцию, с другой – узнавать небольшие молекулы индуктора.

Каким образом репрессор предотвращает транскрипцию? Он связывается с последовательностью ДНК, названной **оператором** и обозначаемой как *lacO*. Оператор расположен между промотором *lacP* и группой структурных генов *lacZYA*. Если репрессор связывается с этим сайтом, его присутствие предотвращает инициирование транскрипции в промоторе, которая осуществляется РНК-полимеразой.

Как влияют малые молекулы индуктора на такого рода взаимодействие? Белок-репрессор имеет очень высокое сродство к оператору. В отсутствие индуктора он связывается с ним таким образом, что прилегающие структурные гены не могут быть транскрибированы. Однако индуктор способен присоединяться к репрессору, образуя комплекс репрессор-индуктор, который более не связывается с оператором. Основной характер такого взаимодействия определяется наличием в белке-репрессоре двух участков связывания, один из которых предназначен для индуктора, а другой – для оператора. При связывании индуктора с определенным участком конформация белка изменяется, что оказывает влияние на активность другого участка. Такой тип взаимоотношений получил название **аллостерического контроля**. В результате при добавлении индуктора репрессор превращается в форму, которая отделяется от оператора. Затем РНК-полимераза может инициировать транскрипцию структурных генов.

Белок-репрессор лактозного оперона представляет собой тетрамер, построенный из идентичных субъединиц с мол. массой 38 000 дальтон каждая. На одну клетку приходится около 10 таких тетрамеров (это значит, что в одном клеточном цикле должно произойти примерно 40 трансляционных событий, осуществляемых при участии *lacI*-мРНК). Регуляторный ген транскрибируется со скоростью, которая, по-видимому, определяется степенью сродства его промотора к РНК-полимеразе. (Мутации в промоторе могут в значительной степени увеличивать степень выражения гена. Такой способ был скомбинирован с амплификацией локуса для получения большего количества репрессора, чем то, которое обычно присутствует в клетке.)

Конститутивные мутации определяют действие репрессора

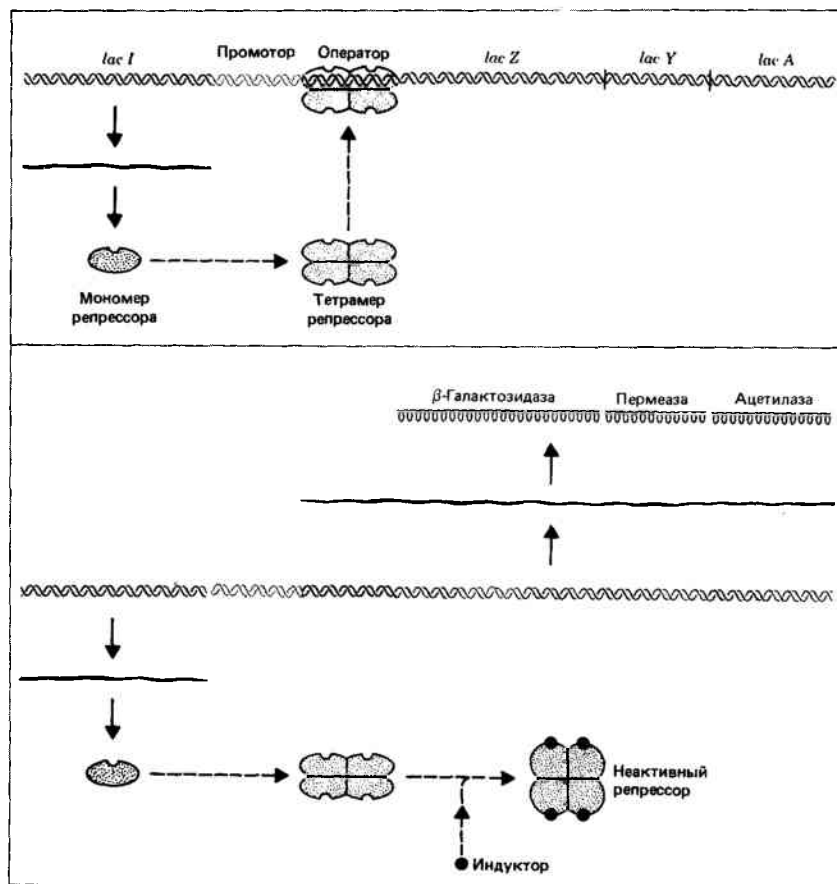
Регуляторный ген *lacI* был первоначально идентифицирован при выделении мутаций, влияющих на выражение *всех трех* структурных генов, но картирующихся вне этих генов. Поскольку мутации *lacI* комплементировали мутации в структурных генах, они были использованы для идентификации другого гена, кодирующего диффундирующий продукт.

Генотип *lacI⁻* может быть обусловлен как точковыми мутациями, так и делециями. Последнее свидетельствует

Рис. 14.2. Включение *lac*-оперона с помощью индуктора.

Вверху. Ген *lacI* синтезирует репрессор, который в форме тетрамера связывается с оператором и предотвращает транскрипцию структурных генов.

Внизу. Добавление индуктора превращает репрессор в неактивную форму, которая не способна более связываться с оператором. В результате в промоторе иницируется транскрипция и осуществляется синтез трех ферментов.



об утрате опероном обычного способа регуляции. У *lacI*⁻-мутантов одновременно выражаются все структурные гены независимо от того, присутствует индуктор или нет. Такое поведение называют **конститутивным** способом выражения генов.

Способность *lacI*⁻-мутантов к конститутивному выражению генов согласуется с поведением системы негативной регуляции. Ген *lacI*⁺ кодирует белок-репрессор, способный выключать транскрипцию группы генов *lacZYA*. Мутация гена, приводящая к *lacI*⁻, позволяет генам экспрессироваться конститутивно, поскольку репрессор становится неактивным.

В пользу этого вывода свидетельствуют результаты, полученные при изучении взаимоотношений между конститутивным мутантным *lacI*⁻-геном и *lacI*⁺-геном дикого типа, присутствующими в одной и той же клетке. Образующийся **частичный диплоид** содержит одну копию оперона в составе бактериального генома, а другую — в независимой самореплицирующейся молекуле ДНК, несущей только несколько генов, в том числе копию оперона или часть его. Такая дополнительная молекула ДНК называется **плазмидой**.

В клетках, содержащих один регуляторный ген *lacI*⁺, а другой *lacI*⁻, восстанавливается нормальная регуляция. Структурные гены и в этом случае подчиняются нормальной регуляции, оказываясь репрессированными в отсутствие индуктора. Этот процесс проиллюстрирован на схеме, представленной на **рис. 14.3**. Конститутивные мутации в репрессоре ведут к утрате им способности связываться с оператором. В результате в промоторе может свободно иницироваться транскрипция. Однако введе-

ние второго регуляторного гена дикого типа восстанавливает присутствие нормального репрессора. Это ведет к тому, что оперон снова выключается в отсутствие индуктора. Говоря языком генетики, индуцибельность дикого типа доминирует над мутантной конститутивностью. Это отличительная особенность негативного контроля.

Регуляторный ген *lacI* расположен сразу же слева от группы структурных генов. Однако благодаря тому, что он детерминирует *транс*-диффундирующий продукт, нет необходимости в его близкой локализации к структурным генам. В самом деле, как мы уже видели, ген *lacI*, находясь в составе независимой молекулы ДНК, способен контролировать группу генов *lacZYA*, расположенных в бактериальной хромосоме (ситуация может быть и обратной). У других оперонов *E. coli* регуляторный ген действительно локализован на некотором расстоянии от структурных генов. Можно, однако, предположить, что имеется преимущество в локализации гена-регулятора вблизи структурных генов, выражающееся в том, что он функционирует только вместе с ними. Поэтому его отделение не имеет смысла.

Функция оператора *цис*-доминантна

Исходно оператор был идентифицирован благодаря выделению конститутивных мутаций, обозначенных *o*^c, характерные свойства которых свидетельствовали о том, что область, повреждаемая мутациями, не детерминирует диффундирующего продукта. Структурные гены,

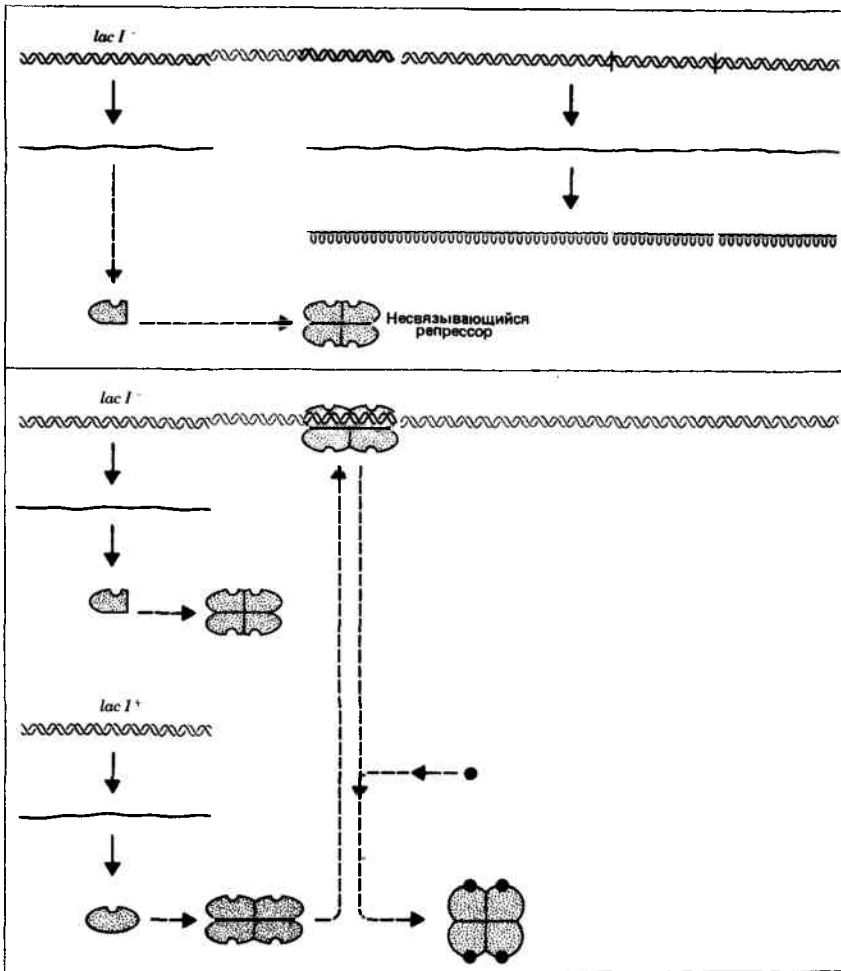


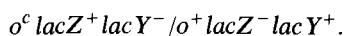
Рис. 14.3. Конститутивные мутации в гене *lacI* являются рецессивными.

Вверху. Мутация *lacI*⁻ вызывает образование неактивного репрессора. Он не способен связываться с оператором. В результате оперон выражается. Внизу. Введение гена *lacI*⁺ ведет к синтезу нормального репрессора. Он отличается активностью в *транс*-положении, способен связываться с оператором и отделяться от него при добавлении индуктора.

смежные с местом мутации *o*^c, выражались конститутивно. Объяснялось это мутационным изменением оператора, предотвращающим связывание с ним репрессора. В результате репрессор оказывался не способным препятствовать иницированию транскрипции РНК-полимеразой. Оперон, таким образом, выражается конститутивно, как это показано на рис. 14.4.

Оператор способен контролировать только расположенные рядом с ним гены *lac*. Если в бактериальную клетку ввести второй *lac*-оперон в независимой молекуле ДНК, он будет содержать свой собственный оператор. Два оператора не будут влиять друг на друга. Так, в том случае, когда оперон имеет оператор дикого типа, он будет репрессироваться при обычных условиях, тогда как второй оперон с мутацией *o*^c будет выражаться характерным для него путем.

Способ проверки таких мутаций на практике сводится к определению способности двух оперонов, отличающихся как своими структурными генами, так и операторами, комплементировать между собой. Предположим, что мы сконструировали диплоидную комбинацию.



Хотя в клетке присутствует по одному дикому аллелю каждого локуса, мы не видим выражения генов дикого типа. Конститутивный оператор позволяет выражаться аллелям *lacZ*⁺ и *lacY*⁻. Однако ген *lacY*⁺ другого опе-

рона не выражается в условиях репрессии, так как его оператор сохраняет дикый тип.

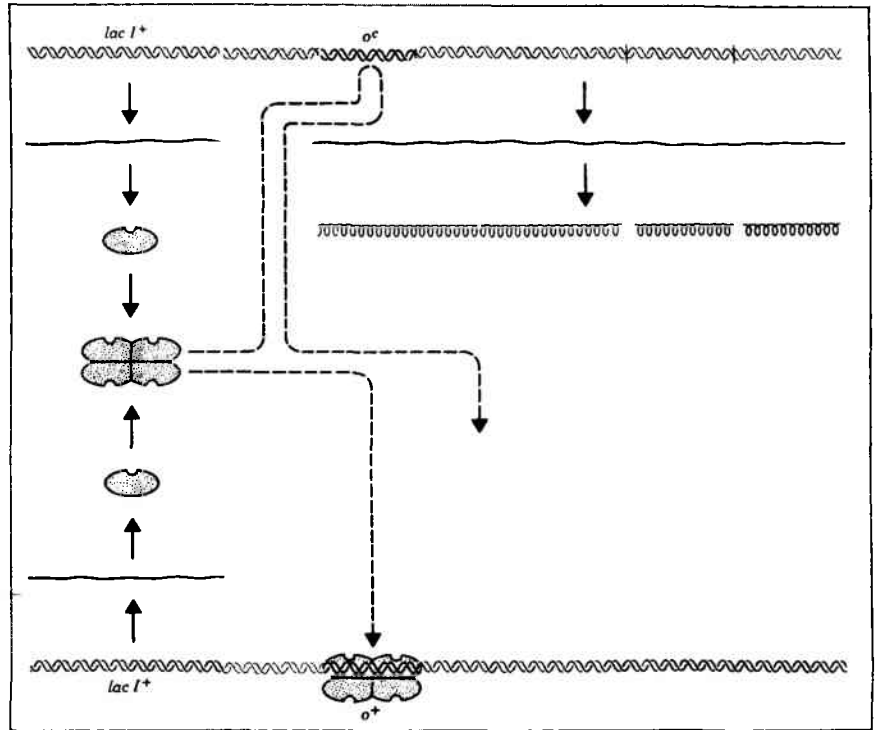
Способность какого-либо сайта контролировать смежные гены независимо от присутствия в клетке других аллелей данного сайта называется **цис-доминантностью**. Она указывает, что мутантный сайт не связан с синтезом какого-либо продукта (обычно белка, но в принципе возможно также РНК), способного диффундировать в клетке, проявляя свой эффект и в отношении другого аллеля. Иногда **цис-доминантный** сайт называют **цис-активным** в отличие от **транс-активной** функции, которая благодаря образованию диффундирующего продукта способна действовать на все соответствующие сайты в клетке независимо от того, присутствуют ли они в одной и той же или в разных молекулах ДНК (гл. 1).

Концепция о **цис-доминантности** применима к любой последовательности ДНК, функция которой состоит в способности быть узнаваемой, а не в синтезе диффундирующего продукта. Следовательно, мутации в промоторе или терминаторе также относятся к **цис-активным**.

В результате **цис-доминантности** оказывается невозможным определить принадлежность мутации к группе комплементации. (Способность к комплементации характерна только для генов, выражение которых связано с образованием диффундирующих продуктов.) Это значит, что в тех случаях, когда два **цис-активных** сайта расположены вблизи друг от друга — например, промотор

Рис. 14.4. Конститутивные мутации оператора (o^c) являются *цис*-доминантными.

Верхний оперон содержит мутантный оператор, не способный связываться с репрессором. Поэтому прилегающие к нему структурные гены выражаются конститутивно. Нижний оперон содержит оператор дикого типа, способный связываться с репрессором и поэтому выражаемый только при индукции. Поведение каждого ряда структурных генов контролируется исключительно смежным оператором. Другой оператор не влияет на их свойства.



и оператор, — мы не можем классифицировать мутации с помощью комплементационного теста. Мы можем различить их только по их влиянию на фенотип.

Наконец, следует указать, что *цис*-доминантность характерна для любого сайта, физически связанного с контролируемой последовательностью. Так, если контролируемый сайт функционирует как часть полицистронной мРНК, мутации в нем будут проявлять точно такой же тип *цис*-доминантности, который они проявляли бы, если бы он функционировал в ДНК. Критическое значение имеет тот факт, что контролируемый сайт невозможно отделить физически от генов. С генетической точки зрения не имеет значения, находятся ли сайт и гены вместе в ДНК или в РНК.

В промоторе или гене репрессора могут встречаться неиндуцибельные мутации

Неиндуцибельные мутантные опероны не могут вообще выражаться. Они относятся к тем же двум генетическим классам, что и конститутивные мутанты. Для определения природы мутации могут быть использованы те же приемы изучения взаимоотношений с целью определения доминирования для каждого типа мутаций.

Промоторные мутации, как мы уже указывали, являются *цис*-активными. Те мутации, которые предотвращают связывание РНК-полимеразы с локусом *lacP*, превращают оперон в нефункционирующий из-за того, что он не может быть транскрибирован.

Мутации в гене *lacI*, влияющие на способность репрессора связываться с индуктором, вызывают тот же фенотип. Репрессор сохраняется в активной форме, которая способна узнавать оператор и предотвращать транскрипцию. Добавление индуктора оказывается бесполезным. Такие мутации описаны как *lacI^s*. На рис. 14.5

показано, что они проявляют *транс*-активность и доминируют над диким типом. Это обусловлено тем, что мутантный репрессор связывается со всеми операторами в клетке, препятствуя их транскрипции, и не может быть отделен от них независимо от свойств любого другого белка-репрессора, который может присутствовать в клетке.

Каким путем репрессор блокирует транскрипцию?

Впервые репрессор был получен при выделении из экстрактов клеток *E. coli* компонента, способного связываться с «добровольным» индуктором ИПТГ. (Поскольку количество репрессора в клетке слишком незначительно, для получения достаточного количества материала необходимо было использовать мутацию промотора, повышающую эффективность транскрипции гена *lacI*, и ввести такой мутантный локус *lacI* в молекулу ДНК, присутствующую в клетке в большом числе копий. В результате удалось увеличить общую продукцию репрессора в 100–1000 раз.)

Связывание очищенного белка-репрессора с ДНК впервые было охарактеризовано с помощью метода связывания на фильтрах. Применение этого метода обусловлено способностью нитроцеллюлозных фильтров удерживать белок, но не ДНК. Однако ДНК, комплексирующаяся с белком, также задерживается на фильтре. В результате специфически отбираются последовательности ДНК, способные связываться с репрессором.

Репрессор связывается с ДНК, содержащей последовательность *lac*-оператора дикого типа, при этом ДНК должна иметь двухцепочечную форму. Связывания репрессора не происходит, если ДНК выделена из мутанта o^c . При добавлении ИПТГ репрессор отделяется от

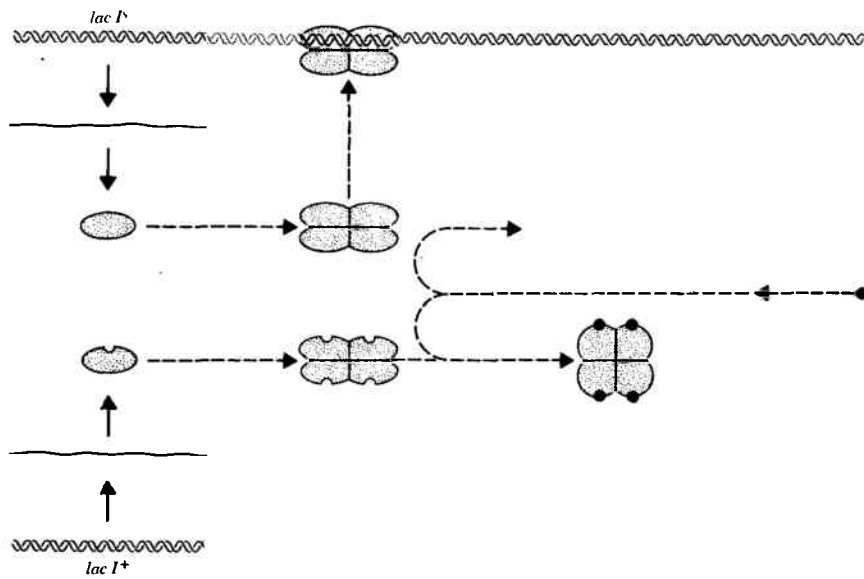


Рис. 14.5. Неиндуцибельные мутации *lacI* являются доминантными.

Ген *lacI^r* продуцирует репрессор, который не может связываться с индуктором, например из-за потери сайта связывания с ним. В результате репрессор остается постоянно связанным с оператором независимо от присутствия индуктора. Такой ген является *транс*-доминантным, так как введение гена *lacI⁺* не приводит к изменению эффекта: неиндуцибельный репрессор остается связанным с оператором.

lacO-ДНК *in vitro*. Таким образом, реакция между белком-репрессором и *lacO*-ДНК *in vitro* контролируется так же, как и *in vivo* и, следовательно, может быть использована для изучения механизма репрессии.

Какова точная локализация оператора по отношению к промотору? Определение размеров *lacO*-локуса по длине ДНК, защищаемой репрессором *in vitro* от действия нуклеаз, показало, что он представляет собой область размером около 26 пар нуклеотидов. Она простирается от положения -5 сразу же слева от стартовой точки мРНК до положения $+21$ в пределах транскрипционной единицы. Следовательно, область оператора может перекрываться с концом промотора, начинающимся от блока Прибнова и кончающимся в пределах транскрипционной единицы.

Препятствует ли такое перекрывание одновременному связыванию репрессора и РНК-полимеразы с *lac*-ДНК? Эксперименты по выяснению конкурентных взаимоотношений белков показали, что присоединение любого из двух белков к ДНК предотвращает связывание другого. Такая ситуация изображена на рис. 14.6. Подобные же результаты были получены для нескольких других оперонов.

Следовательно, в основе обычного механизма репрессии лежит связывание репрессора с оператором, перекрывающего, по крайней мере частично, последовательность, которая должна узнаваться РНК-полимеразой. Перекрывание не всегда происходит в одной и той же области промотора. Например, у фага лямбда оператор лежит в области, расположенной слева от промотора (гл. 16). Однако общий принцип сводится к тому, что связывание репрессора с оператором мешает РНК-полимеразе подойти к своему промотору, что препятствует иницированию транскрипции.

Контакты в операторе

Каким образом репрессор узнает специфическую последовательность ДНК оператора? Взаимодействие между лактозным репрессором и оператором часто используется как образец реакции связывания белка со специфической последовательностью ДНК. Оно наиболее хорошо

охарактеризовано с точки зрения как последовательности ДНК, так и структуры белка.

Оператору присуще свойство, общее для многих сайтов узнавания белками: он имеет ось двойной симметрии. Инвертированные повторы не совсем идентичны. Как видно на рис. 14.7, между ними имеется три отличия. (Явное исключение составляет отсутствие симметрии у промотора, который должен содержать информацию о направлении транскрипции.)

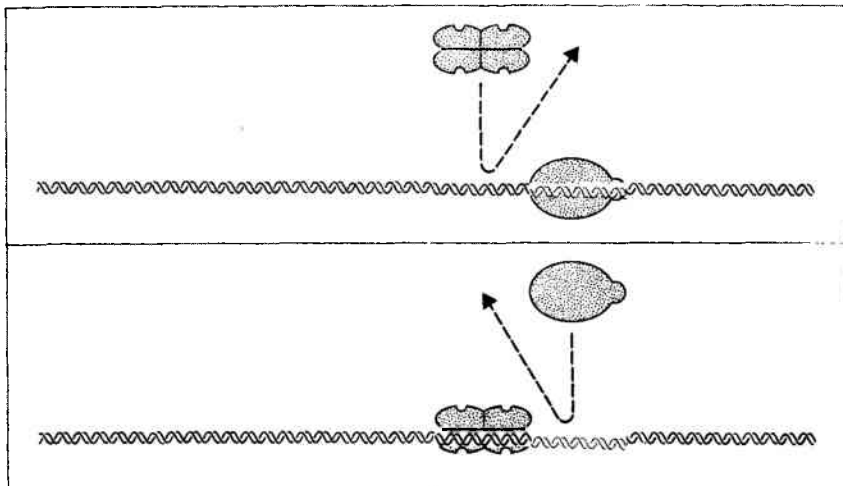
Какую роль играет симметрия в узнавании оператора? Возможность перестроек ДНК с образованием структуры креста была исключена. Присоединение репрессора к ДНК ведет к изменению в структуре дуплекса, эквивалентному незначительному расплетанию (но не настолько, чтобы образовались свободные пары оснований или структура креста).

Отражает ли симметрия в последовательности ДНК наличие симметрии в белке? Такая возможность не исключается, поскольку репрессор — это тетрамер, состоящий из идентичных субъединиц, каждая из которых, следовательно, должна иметь один и тот же участок связывания с ДНК.

Для определения точек, с которыми репрессор контактирует в операторе, используются те же подходы, которые мы уже обсуждали в случае взаимоотношений полимеразы и промотора (гл. 11). Конститутивные точечные мутации дают возможность идентифицировать отдельные пары оснований, имеющие критическое значение; делеции материала с любой стороны позволяют определить концевые точки области связывания. Были проведены эксперименты, в которых связанную и не связанную с репрессором ДНК сравнивали по чувствительности к метилированию или образованию поперечных сшивок в ДНК под действием УФ-облучения. Цель этих экспериментов — идентификация оснований, которые либо защищены, либо более чувствительны к обработкам, когда они связаны с белком.

В тех случаях, когда связанную с репрессором ДНК выделяли посредством деградации незащитенных последовательностей нуклеазой, она не соответствовала по протяженности всей области симметрии. Ранее уже указывалось, что эта область представлена 26 парами ос-

Рис. 14.6. Связывание репрессора и связывание РНК-полимеразы – взаимоисключающие события в операторно-промоторной области *lac*-оперона.



нований, расположенными между нуклеотидами в положениях -5 и +21. Зона, идентифицируемая конститутивными мутациями, имеет еще меньший размер. В пределах центральной области, представленной 13 парами оснований и расположенной между нуклеотидами +5 и +17, содержится восемь сайтов, в которых, как было установлено, замещение одной пары оснований приводит к конститутивности. В свете этих данных особое значение получают результаты изучения мутаций промотора, суммированные ранее на рис. 11.5. За связывание специфических последовательностей ДНК с белком может быть ответственно незначительное число существенных специфических контактов в пределах большей по размеру области.

Распределение сайтов мутаций показывает, что левая часть оператора проявляет более высокую чувствительность к повреждениям: она несет шесть мутаций по сравнению с двумя в правой части. Две мутации в правой части являются строго симметричными копиями мутаций,

обнаруженных в левой части. (В положении +17 Т—А замещает С—G, тогда как в положении +5 А—Т замещает G—C. Подобным же образом в положении +14 С—G замещает Т—А, а в положении +8 G—C замещает А—Т.) Однако эффект конститутивных мутаций не одинаков для членов каждой пары. Таким образом, симметрия имеет отношение к контактам репрессора с оператором, но тем не менее репрессор более тесно связывается с левой частью оператора.

Участки усиления и защиты оснований проявляют некоторую симметрию в пределах основной области тесных контактов. На рис. 14.7 показано, что все (кроме двух) тиминовые остатки в области от +1 до +22 могут быть перекрестно связаны с репрессором. Эксперименты с метилированием показывают, что большая часть пуринов (А и G) между положениями +3 и +19 весьма эффективно защищена благодаря связыванию с репрессором. Немногие пурины проявляют повышенную чувствительность, по-видимому, в результате образования «ги-

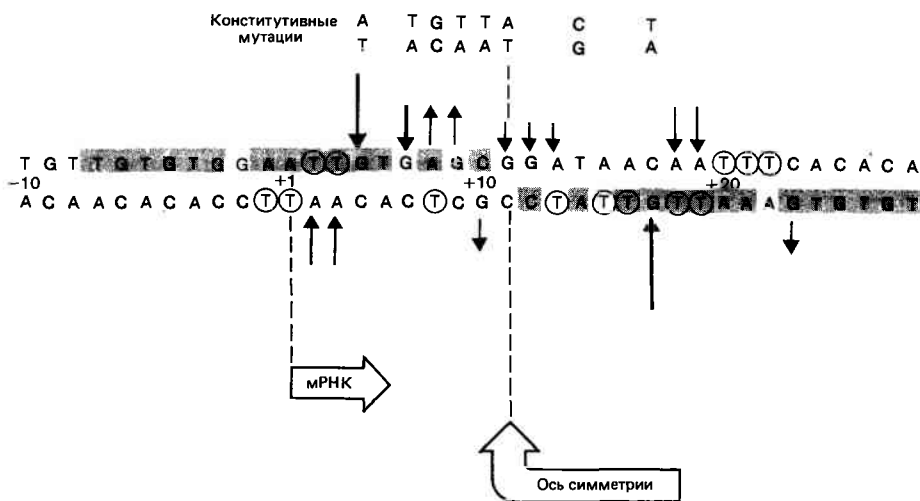


Рис. 14.7. *Lac*-оператор имеет симметричную последовательность нуклеотидов.

Последовательность нуклеотидов пронумерована относительно стартовой точки транскрипции, обозначенной +1. Области двойной симметрии показаны затемненными блоками. Ось симметрии проходит через пару оснований в положении +11. Замены, ведущие к конститутивному выражению генов, показаны парами оснований, расположенными над нуклеотидной последовательностью. Тиминовые основания, находящиеся

в положениях, где замещение бромурцила делает возможным установление перекрестных связей с репрессором, показаны кружками. Стрелки, идущие в направлении последовательности, указывают пуриновые основания, защищаемые репрессором от метилирования. Стрелки, идущие от последовательности, указывают положения, в которых метилирование повышено. Длина стрелок соответствует силе эффекта. Область сильных контактов лежит между основаниями +1 и +23. Вне этой области осуществляются более слабые контакты.

дрофобных карманов» в комплексе репрессор—оператор. Контакты между основаниями +1 и +6 являются по существу симметричными в отношении контактов, образуемых между основаниями +21 и +16; контакты, возникающие ближе к оси симметрии, не симметричны. Таким образом, контакты репрессора с оператором симметричны на концах области связывания; в участке непосредственно вокруг оси симметрии такая симметрия контактов нарушена.

Центральная область, имеющая критическое значение для связывания репрессора, занимает, согласно грубой оценке, первые 20 пар оснований, расположенных по ходу транскрипции от ее стартовой точки. Следует напомнить, однако, что точки контактов РНК-полимеразы в промоторе лежат, как показано на рис. 11.5, вблизи стартовой точки против хода транскрипции. Следовательно, основные области контактов для двух белков являются смежными, а не перекрывающимися. Причина их взаимного исключения может быть обусловлена необходимостью вытягивания каждого белка вдоль ДНК, что и предполагалось на основании способности репрессора защищать ДНК вплоть до положения -5 от расщепления нуклеазой; возможно также, что присутствие репрессора может создавать достаточные стерические помехи, препятствующие стабилизации РНК-полимеразы на промоторе.

Взаимодействие субъединиц репрессора

Белок-репрессор имеет два типа связывающих сайтов, которые (как мы уже видели) взаимодействуют, обеспечивая его способность контролировать выражение генов в ответ на условия окружающей среды. Репрессор непосредственно узнает специфическую нуклеотидную последовательность оператора. Он связывается также с малой молекулой индуктора и вследствие этого теряет способность связываться с операторной ДНК. В субъединице репрессора могут быть идентифицированы два типа участков связывания. Для идентификации используют мутации в гене *lacI*, которые инактивируют репрессор. Их возможная взаимосвязь *in vivo* зависит от мультимерной структуры репрессора.

Мономерные субъединицы репрессора соединяются в клетке случайно, образуя активный тетрамерный белок. В тех случаях, когда в клетке имеется два разных аллеля гена *lacI*, субъединицы двух типов могут ассоциировать с образованием гетеротетрамера, который по своим свойствам может отличаться от гомотетрамера. Такой тип взаимоотношений, называемый **межаллельной комплементацией** (гл. 1), является характерным свойством мультимерных белков.

В случае некоторых мутантных репрессоров имеет место определенный тип взаимодействия, названный **негативной комплементацией**. Она может наблюдаться при комбинировании генов *lacI^{-d}* и *lacI⁺*. Мутация *lacI^{-d}* ведет к образованию репрессора, который не способен связываться с оператором и является поэтому представителем аллелей конститутивного типа. Поскольку мутация типа *lacI⁻* инактивирует репрессор, она является рецессивной по отношению к дикому типу. Однако символ $-d$ указывает, что этот вариант негативного типа проявляет **доминантность** в случае объединения с аллелем дикого типа.

Доминирование обусловлено тем, что аллель *lacI^{-d}* кодирует испорченную субъединицу, которая не только сама по себе не способна связаться с оператором, но, являясь частью тетрамера, препятствует также связыванию с ним любой полноценной субъединицы. Из этого следует, что для достижения репрессии необходим действительно репрессор-тетрамер, а не отдельный мономер. Повреждающее действие может быть также достигнуто *in vitro* при смешивании соответствующих полноценных и испорченных субъединиц.

С помощью мутаций *lacI^{-d}* был идентифицирован сайт связывания с ДНК в субъединице репрессора. Этот факт объясняет способность мутаций препятствовать связыванию смешанных тетрамеров с оператором; уменьшение числа связывающих сайтов должно довольно значительно уменьшать специфическое сродство репрессора к оператору. На рис. 14.8 представлена карта гена *lacI*, на которой показано, что все мутации *lacI^{-d}* сгруппированы в крайнем левом конце гена. Это прямо указывает на то, что в N-концевой области белка находится участок связывания ДНК. Рецессивные мутации типа *lacI⁻* локализируются и в других участках молекулы, но, по-видимому, действуют на связывание белка с ДНК косвенным путем.

Роль N-концевой области в специфическом связывании репрессора с ДНК подтверждается также локализацией в соответствующей части молекулы ДНК мутаций «прочного связывания». Такие мутации повышают сродство репрессора к оператору, причем иногда в такой степени, что он не может отделяться от него даже в присутствии индуктора. Это редкий тип мутаций.

Неиндубибельные мутации типа *lacI^s* делают репрессор нечувствительным к индуктору. Такой эффект может быть обусловлен либо утратой белком участка связывания с индуктором, либо появлением неспособности передавать свой сигнал участку связывания с ДНК. К мутациям этого класса относятся те, которые обуславливают неспособность субъединиц репрессора агрегировать с образованием тетрамера. На рис. 14.8 можно видеть, что мутации *lacI^s* встречаются группами в участках, которые довольно равномерно распределены по длине гена. Промежутки могут соответствовать изгибам в полипептидной цепи. (Систематическое изучение эффектов аминокислотных замен показало, что примерно 6% положений являются нейтральными, не оказывающими влияния на репрессию.)

Репрессор—белок, связывающийся с ДНК

Поведение выделенного белка-репрессора *in vitro* прямо показывает, что в нем имеется участок связывания с ДНК, способность которого оставаться присоединенным к ДНК оператора зависит от структуры остальной части молекулы.

Под действием трипсина репрессор разрезается преимущественно по положению 59. С-концевой фрагмент белка, содержащий аминокислотные остатки 60–360, называют «ядром», **резистентным к трипсину**. Он сохраняет способность собираться в тетрамер и связываться с индуктором; однако он не обладает способностью присоединяться к оператору. N-концевой фрагмент, содержащий аминокислотные остатки 1–59, называют **длинным N-концевым фрагментом**. Он в свою очередь может быть

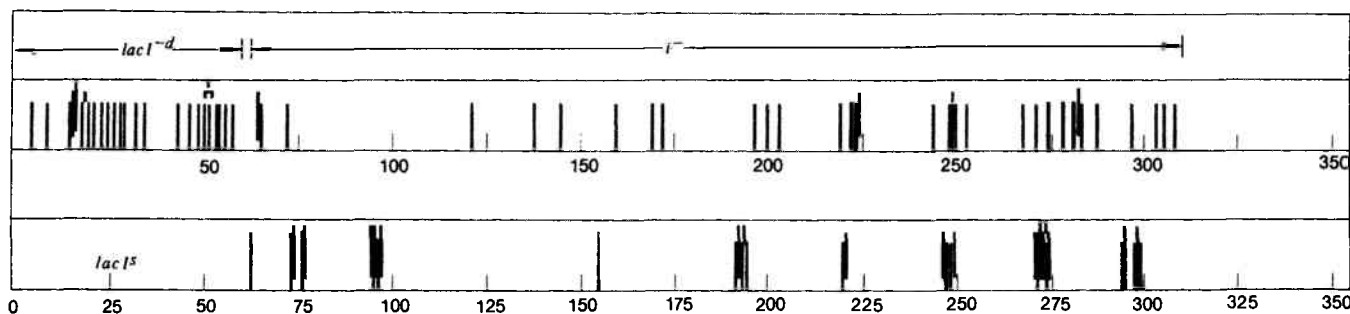


Рис. 14.8. Мутационный анализ гена *lacI* идентифицирует области, ответственные за выполнение различных функций.

Протяженность гена пронумерована числом кодонов. На верхней карте показано распределение мутаций *lacI*⁻ по всей длине гена и наличие мутаций *lacI*⁻ только в N-концевом участке. Мутации *lacI*⁻, нарушающие

образование тетрамеров из мономеров, обозначены красным цветом. Нижняя карта показывает, что мутации *lacI*['] не встречаются в области, предшествующей остатку 62, и группируются между положениями 150 и 300, причем группы обычно отделены друг от друга 26 кодонов или близкими к этому расстояниями.

расщеплен трипсином по 51-му аминокислотному остатку с образованием фрагмента из аминокислотных остатков 1–51, названного **коротким N-концевым фрагментом**. Как короткий, так и длинный N-концевые фрагменты сохраняют способность присоединяться к ДНК, причем их контакт с операторным участком ДНК не отличается от контакта неповрежденного репрессора (хотя они связываются менее прочно по сравнению с интактным белком).

Способность N-концевых фрагментов присоединяться к оператору показывает, что их структура *per se* не зависит от остальной части белка. Этот вывод согласуется с моделями, в которых участок связывания с ДНК представлен как **ответвление** или **выступ**, образуемый 50 аминокислотами N-концевого фрагмента на основной части белка. Ответвление может быть соединено с «ядром» белка областью, напоминающей шарнир, которая образована аминокислотными остатками в положениях 50–80. Остальная часть белка, представленная аминокислотами 81–360, ответственна за сборку в тетрамерную структуру и связывание с индуктором.

Если N-концевой участок является ответвлением, способным связываться с ДНК, какова функция тетрамерной структуры и как она контактирует с ДНК? Согласно одной из моделей, предполагается, что репрессор имеет форму удлинённой гантели или цилиндра длиной примерно 115 Å, ось которой может пересекаться с осью ДНК, образуя небольшой угол, равный 15–20°. Ядра всех четырех субъединиц контактируют своими центральными областями; N-концевые участки связывания находятся по краям молекулы, располагаясь парами. Не исключено, что на самом деле только две из четырех субъединиц контактируют с ДНК, оставляя два других участка связывания с ДНК незанятыми. Хотя «ядро» субъединицы не узнает операторного участка, оно все же способно связываться с ДНК неспецифически, так что, по-видимому, оно также осуществляет контакты (хотя они не являются настолько выраженными, чтобы оказывать влияние на характер защиты ДНК в той же мере, как это делает один N-концевой фрагмент).

Таким образом, согласно этой модели, комплекс репрессор—ДНК представляет собой тетрамерный белок, симметрично расположенный на ДНК, причем только незначительная часть полипептидной цепи (примерно 20%) тесно связана с ДНК. Такое представление принято и другой, совсем отличной моделью, которая предпола-

гает перпендикулярное расположение оси репрессора по отношению к оси ДНК, в результате чего только два N-концевых участка в одном из концов репрессора связаны с оператором.

Отделение репрессора от ДНК

Тетрамер репрессора прочно связывается с оператором. Индуктор вступает в действие, связываясь с репрессором. Каким образом происходит отделение белка от ДНК?

Различные индукторы вызывают характерное уменьшение сродства репрессора к оператору *in vitro*. Эти изменения коррелируют с эффективностью индукторов *in vivo*. Следовательно, механизм индукции основан на уменьшении взаимопротяжения между репрессором и оператором. Чем это обусловлено?

Скорость, с которой происходит отделение репрессора от оператора, довольно низкая. В типичном случае полупериод распада комплекса колеблется в пределах от 10 до 20 мин в зависимости от конкретных условий. Однако при добавлении ИПТГ стабильность комплекса немедленно уменьшается, что проявляется в резком укорачивании полупериода его распада. Этот результат может способствовать выбору между двумя моделями, изображенными на рис. 14.9. Модель равновесия (изображенная слева) предполагает, что репрессор, связанный с ДНК, находится в равновесии со свободным репрессором. Индуктор связывается со свободной формой репрессора, вызывая нарушение равновесия путем предотвращения его повторного связывания с ДНК. Однако скорость диссоциации репрессора и оператора (как показали изменения в отсутствие индуктора) оказалась *значительно* ниже той, которая постулируется данной моделью. Это значит, что индуктор должен связываться *непосредственно* с белком-репрессором, находящимся в комплексе с оператором. Как следует из модели, представленной справа на рис. 14.9, присоединение индуктора должно вызывать изменение в свойствах репрессора, обуславливающее его отделение от оператора.

Связывание комплекса репрессор—ИПТГ с оператором может быть изучено использованием больших концентраций белка при метилировании с определением эффектов защиты/усиления. Большое количество белка

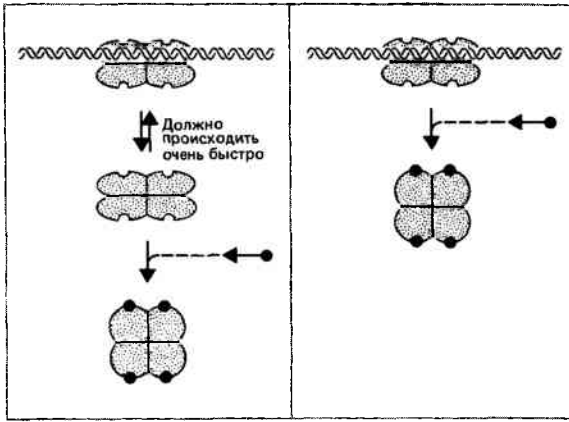


Рис. 14.9. Происходит ли присоединение индуктора к свободно-му репрессору с нарушением состояния равновесия (показано слева) или непосредственно к репрессору, связанному с оператором (показано справа)?

компенсирует низкое сродство комплекса репрессор-ИПТГ к оператору. Комплекс осуществляет такие же контакты с ДНК, которые характерны и для свободного репрессора или N-концевого фрагмента репрессора. Аналогичный результат получен с мутантными репрессорами, сродство которых к операторному участку ДНК повышено; они устанавливают тот же тип контактов.

В целом типы репрессора, у которых различия в сродстве к оператору достигают семи порядков величины, устанавливают одни и те же контакты с ДНК. Изменения в сродстве репрессора к ДНК должны поэтому происходить путем влияния на общую конформацию N-концевого участка, связанного с ДНК, а не путем создания или разрыва одной или нескольких отдельных связей. В результате изменения конформации всей области, связанной с ДНК, многие или все связи с ДНК должны одновременно стать ослабленными. Как это происходит? На основе данных, полученных с помощью разных методов, полагают, что присоединение индуктора вызывает немедленное конформационное изменение белка-репрессора. Присоединение двух молекул индуктора к репрессору-тетрамеру достаточно для снятия репрессии. Однако мы пока не знаем, каким образом общее конформационное изменение влияет на сродство репрессора к оператору.

Наиболее вероятно, что конформационное изменение передается от «ядра» через шарнир в N-концевой участок. На рис. 14.10 представлена такая модель. Конформация участка связывания с ДНК изменяется от состояния, в котором он точно соответствует последовательности ДНК, до состояния, когда N-концевой участок не способен достаточно тесно контактировать с ДНК.

Накопление излишков репрессора

Вероятно, все белки, проявляющие большую специфичность в отношении определенной последовательности, проявляют низкое сродство в отношении каких-то случайных последовательностей ДНК. Большое число участков, отличающихся низким сродством, будут конкурировать с небольшим числом участков с высоким сродством за связывание с репрессором-тетрамером. В геноме

E. coli существует только один сайт, обладающий высоким сродством к репрессору,—оператор. Остальная ДНК может рассматриваться как источник связывающих сайтов с низким сродством к репрессору. Каким образом происходит распределение молекул репрессора между сайтами? Что происходит с репрессором после присоединения к нему индуктора и отделения его от оператора?

Чтобы из всей последовательности генома выбрать последовательность оператора, специфическое сродство связывания у репрессора должно во много раз превышать степень его случайного сродства. Действительно, репрессор присоединяется к операторному участку ДНК в $4 \cdot 10^6$ раз лучше, чем к любой другой случайной последовательности нуклеотидов той же протяженности. (Эта величина представляет соотношение констант ассоциации репрессора с *lacO*-ДНК и другими ДНК.)

Согласно имеющимся данным, оператор представлен 26 парами нуклеотидов в геноме из $4,2 \cdot 10^6$ пар нуклеотидов, так что избыток неоператорной ДНК над операторной составляет $2 \cdot 10^5$ пар оснований. Иными словами, существует $2 \cdot 10^5$ сайтов связывания с низким сродством к репрессору по сравнению с одним сайтом высокого сродства.

Относительная предпочтительность репрессора в отношении операторного участка ДНК в пределах бактериального генома определяется отношением величины, выражающей предпочтительность *per se* в отношении сайта с высоким сродством ($4 \cdot 10^6$), к числу сайтов с низким сродством ($2 \cdot 10^5$). Результат определения показывает, что репрессор связывается с оператором примерно в 20 раз лучше, чем с любым другим сайтом.

В свете этих данных становится понятно, почему мутации, уменьшающие сродство оператора к репрессору в 20–30 раз, приводят к конститутивному выражению генов оперона. В пределах генома мутантные сайты могут утратить преимущество в отношении случайных сайтов. Их специфическое сродство к репрессору по сравнению со случайными сайтами оказывается не настолько выраженным, чтобы они могли сохранять свой статус предпочтительных нуклеотидных последовательностей.

В результате таких взаимоотношений в неиндуцированных клетках одна молекула репрессора-тетрамера присоединяется к оператору. Все или почти все оставшиеся тетрамеры будут присоединяться случайно к другим участкам ДНК, как показано на рис. 14.11. Свободные тетрамеры репрессора, по-видимому, либо не содержатся в клетке, либо их очень мало.

Присоединение индуктора к репрессору приводит почти к тысячекратному уменьшению сродства репрессора к оператору. Сродство же в отношении обычных последовательностей остается неизменным. В результате в индуцированной клетке все или почти все тетрамеры репрессора будут «сохраняться» в случайных сайтах ДНК. Следовательно, эффект индукции ведет не к появлению свободных молекул репрессора, а к изменению его распределения по ДНК.

Исходя из способности репрессора сохраняться на ДНК, можно сделать ряд выводов, имеющих важное биологическое значение. Наиболее прямой из них выражается в следующем: способность репрессора очень быстро связываться с оператором не соответствует времени, которое требовалось бы для множественных диссоциаций и реассоциаций комплекса репрессор—ДНК. Это исключает наличие механизма беспорядочного поиска оператора и дает основание думать, что репрессор способен пере-

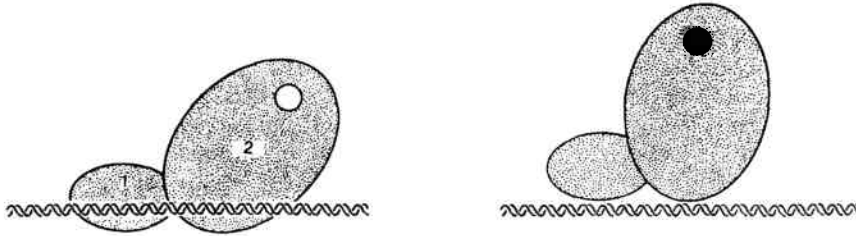


Рис. 14.10. Освобождение репрессора связано с общим изменением его структуры, ослабляющим все его контакты с оператором.

Показано изменение угла наклона, образуемого между N-концевым участком и «ядром», после присоединения индуктора к репрессору.

Слева. N-концевой участок специфически связывается с оператором, располагаясь на ДНК с образованием специфических контактов. Справа. Индуктор изменяет конформацию репрессора таким способом, что N-концевой участок более не лежит непосредственно на ДНК. Для ясности изображена только одна субъединица. 1 – шлем, 2 – ядро.

двигаться непосредственно от случайного сайта в ДНК к оператору. Движение может осуществляться либо очень быстрым скольжением по ДНК (которое, по-видимому, нереально из-за помех, создаваемых другими, связанными с ДНК белками), либо путем непосредственного перемещения из сайта в сайт. Последнее может происходить с помощью того же механизма, который мы уже обсуждали в случае РНК-полимеразы (гл. 10). Реакцию перемещения может облегчать наличие большего числа (четыре-пять) сайтов связывания в расчете на тетрамер, чем на самом деле используется для контактов с ДНК в любое данное время (два из четырех).

Способность большого числа неспецифических сайтов успешно конкурировать с небольшим числом сайтов с высоким сродством свидетельствует о том, что связывание репрессора с оператором должно отличаться высокой чувствительностью как к общей концентрации ДНК, так и к общей концентрации репрессора в клетке. Разница в степени выражения лактозного оперона в его индуцированном и репрессированном состоянии может быть

тысячекратной. Иными словами, даже в отсутствие индуктора существует определенный базальный уровень выражения, составляющий 0,1% от уровня индукции. Он может быть ниже – в присутствии большего количества белка-репрессора – или выше – если количество белка уменьшено. Например, состояние полной репрессии не может быть установлено, если в клетке содержится менее 10 молекул репрессора; точно так же могут возникнуть трудности при индуцировании оперона, если молекул репрессора в клетке слишком много.

Парадокс индукции

В лактозном опероне имеется структурный ген (*lacZ*), детерминирующий β-галактозидазу, необходимую для сбраживания сахара; кроме того, он включает в себя ген, который кодирует белок, необходимый для транспорта субстрата в клетку (*lacY*). Возникает вопрос: каким образом проникает индуктор в клетку, чтобы начать процесс

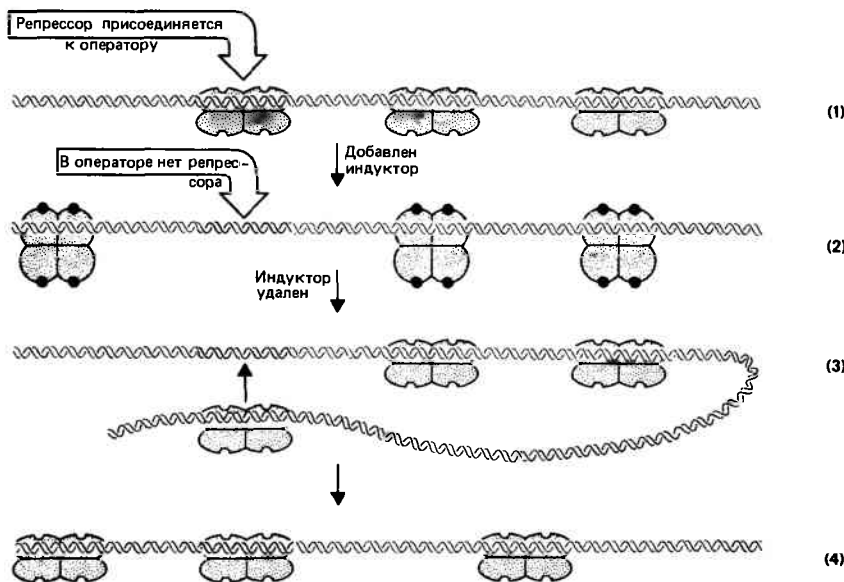


Рис. 14.11. Репрессор тяготеет к ДНК.

Практически весь репрессорный белок находится в связанном с ДНК состоянии. В неиндуцированном состоянии (1) тетрамер находится в операторе. Индукция заставляет его покинуть оператор, но он связывается с другим случайным сайтом ДНК (2). Другие тетрамеры также присоеди-

няют индуктор, но это не влияет на их общее сродство к ДНК. Репрессор может перемещаться из одного сайта в другой путем прямого перемещения (3). При удалении индуктора оператор снова вступает в контакт с репрессором-тетрамером, который находился в случайном сайте ДНК, и перемещается для восстановления неиндуцированного состояния (4).

индукции оперона путем связывания с репрессором, если оперон находится в репрессированном состоянии?

Базальный уровень выражения оперона может составлять тот критический уровень, который обеспечивает наличие в клетке минимального количества белка, достаточного для включения процесса. Это оказывается необходимым не только для введения индуктора (некоторые из индукторов могут вводиться в клетку с помощью каких-либо других систем), но также для того, чтобы был запущен данный метаболический путь.

Хотя лактоза служит субстратом для β -галактозидазы, она не способна индуцировать оперон. На деле она проявляет слабый эффект антииндуктора (связывается с репрессором, повышая его сродство к оператору). Итак, как же происходит индукция?

Природный индуктор оперона был получен при выделении репрессора из индуцированных клеток. Белок оказался связанным с аллолактозой. Основная реакция β -галактозидазы в присутствии лактозы—это процесс разложения лактозы на глюкозу и галактозу. В результате побочной реакции осуществляется перенос галактозы к определенным акцепторным молекулам с образованием (преимущественно) аллолактозы и галактобиозы. Поскольку акцептором, который используется для образования аллолактозы, является на самом деле глюкоза, реакция завершается молекулярным перераспределением.

Следовательно, для индукции пути метаболизма лактозы *in vivo* необходимо введение небольшого количества лактозы, часть которой ферментативно превращается в аллолактозу. Аллолактоза затем индуцирует оперон. Она обладает сильным индуцирующим эффектом, и небольшого количества аллолактозы оказывается достаточно для преодоления слабого антииндуцирующего эффекта лактозного субстрата.

Рекомендуемая литература

Лучшим источником из последних обзоров литературы о функции оперона является книга Миллера и Резникова (Miller, Reznikoff (Eds.), *The Operon*, Cold Spring Harbor Laboratory, New York, 1978). Особый интерес представляют главы (Beckwith, pp. 11–30) об организации системы, (Zabin, Fowler, pp. 89–122) о белковых продуктах, (Miller, pp. 31–88) о гене *lacI*, (Beyreuther, pp. 123–154) и (Weber, Geisler, pp. 155–176) о белке-репрессоре и (Barkley, Bourgeois, pp. 177–200) о связывании репрессора с ДНК и индукторами. Данным о связывании белок–ДНК посвящен обзор Пабо и Зауэра (Pabo, Sayer, *Ann. Rev., Biochem.* 53, pp. 293–321, 1984).

Глава 15

СИСТЕМЫ КОНТРОЛЯ: СРЕДСТВА РЕГУЛЯЦИИ ОПЕРОНОВ

Иногда создается впечатление, что для контроля выражения (экспрессии) гена в той или другой ситуации используется каждый из возможных механизмов. При всем разнообразии механизмов контроля их объединяет то общее, что регуляция во всех случаях осуществляется взаимодействием регуляторного белка с последовательностью нуклеиновой кислоты, часто имеющей двойную симметрию. В этой главе мы рассмотрим некоторые другие примеры взаимодействий такого рода с точки зрения их объединения в системы, контролирующие отдельные опероны. Часто контроль осуществляется при участии сходных механизмов, оперирующих слегка отличным способом, и способ установления связи между ними варьирует от случая к случаю.

Оперон представляет собой структуру, позволяющую координированно регулировать группу структурных генов. Помимо этого возможны другие способы контроля, обеспечивающие различную степень выражения отдельных генов.

Различия между позитивным и негативным контролем

Системы позитивного и негативного контроля определяются ответом оперона в отсутствие регулятор-

ного белка. Характеристики двух типов контролируемых систем являются прямо противоположными.

Гены, находящиеся под негативным контролем, экспрессируются *при условии, что они не* выключены взаимодействием белка-репрессора с соответствующим регуляторным сайтом. Любое действие, создающее помехи для экспрессии генов, может обеспечить негативный контроль, но существует единообразие в этих механизмах: белок-репрессор либо связывается с ДНК, предотвращая иницирование транскрипции ферментом РНК-полимеразой, либо связывается с мРНК, предотвращая иницирование трансляции рибосомами.

Негативный контроль обеспечивает механизм «спасения с потерей функции»: если регуляторный белок инактивирован, система генов функционирует, и клетка, таким образом, не лишается необходимых ферментов. Легко представить, как такая система могла эволюционировать. Исходно система функционировала конститутивно, но затем клетки, способные специфически подавлять ее экспрессию, приобрели преимущество при отборе благодаря их возросшей приспособляемости.

Для генов, которые находятся под позитивным контролем, экспрессия возможна *только* в присутствии активного регуляторного белка. Каким образом это достигается? Существует множество различных систем позитивного контроля. Системы, действующие на иницирование транскрипции определенных оперонов, яв-

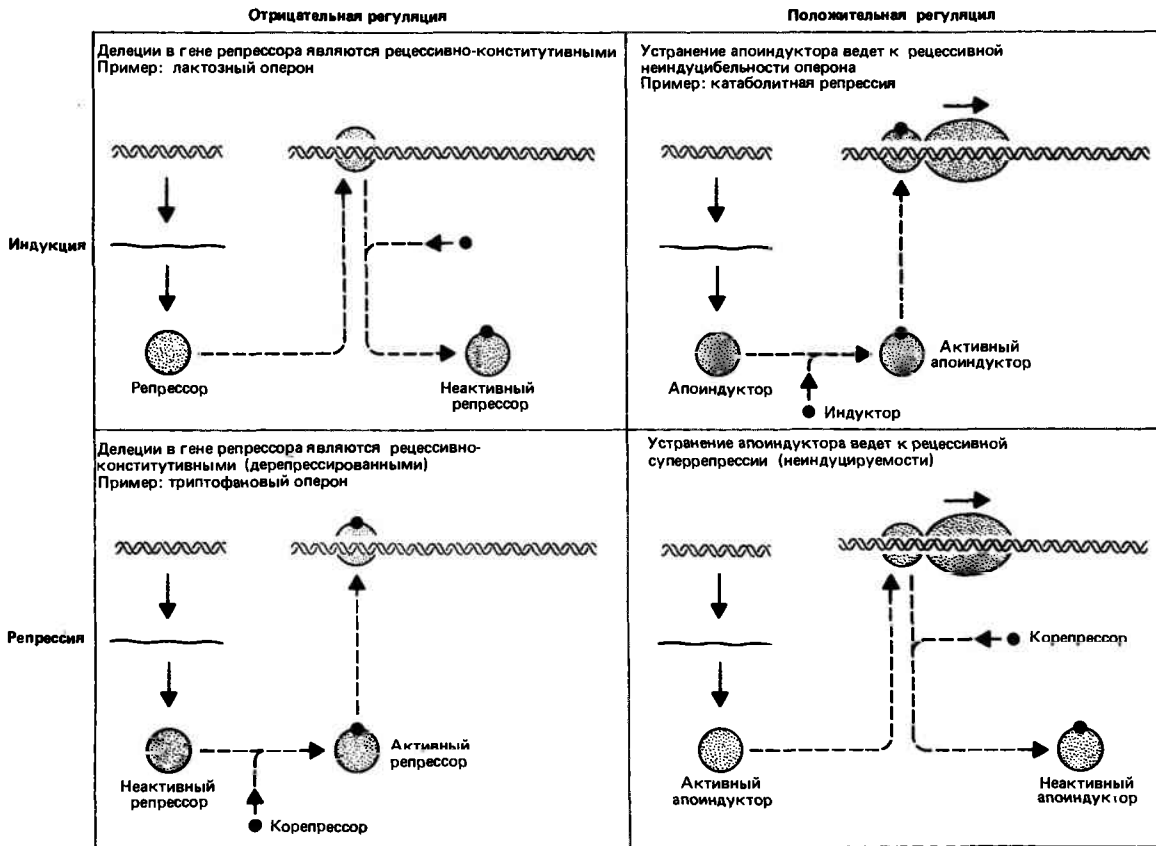


Рис. 15.1. Контролирующие системы отличаются гибкостью и предназначены для осуществления положительного или отрицательного контроля индукции или репрессии.

ляются точными аналогами систем негативного контроля. Однако, вместо того чтобы служить помехой для иницирования транскрипции, регуляторный белок оказывается существенным для нее. Он взаимодействует с ДНК- и РНК-полимеразой, способствуя инициации транскрипции. Такой белок иногда называют **апоиндуктором**. Другие системы позитивного контроля обеспечивают полное замещение РНК-полимеразных субъединиц, участвующих в регуляции специфичности инициации (гл. 12) или действующих как антитерминаторы в транскрипции (гл. 13).

Представить, как эволюционировали системы позитивного контроля, гораздо сложнее, поскольку клетка должна была иметь возможность экспрессировать регулируемые гены еще до того, как появилась какая-либо система контроля. По-видимому, какой-то компонент системы контроля должен был изменить свою функцию. Возможно, первоначально он использовался как постоянная часть аппарата, необходимого для выражения генов; позднее он приобрел способность действовать только при определенных условиях.

Опероны относят к **индуцибельным** или **репрессибельным**, согласно природе их ответа на малые молекулы, регулирующие их выражение. Подобно тому как бактерия получает преимущество благодаря способности индуцировать ряд ферментов только после добавления индуцирующего субстрата, который они метаболизируют, может оказаться полезной и ее способность репрессировать ферменты, способные синтезировать какое-то

соединение, если это соединение содержится в достаточных количествах в среде.

Таким образом, индуцибельные опероны функционируют только в присутствии малых молекул индуктора. Репрессибельные опероны функционируют только в отсутствие малых молекул **корепрессора** (они названы так, чтобы их можно было отличать от белка-репрессора). Согласно терминологии, используемой при описании репрессибельных систем, активное состояние оперона получило название **дерепрессированный**; оно имеет то же значение, что и **индуцированное** состояние. Условия, при которых оперон (мутантный) не может быть дерепрессирован, иногда называют **суперрепрессирующими**. Это название точно соответствует **неиндуцибельному** состоянию.

Как позитивный, так и негативный контроль могут быть использованы для достижения эффекта индукции или репрессии путем использования соответствующих взаимодействий между регуляторным белком и малой молекулой индуктора или корепрессора. На рис. 15.1 показаны четыре простых типа контролируемых систем. Индукция достигается, если индуктор инактивирует репрессорный белок или активирует белок-апоиндуктор. Репрессия осуществляется в том случае, если корепрессор активирует белок-репрессор или инактивирует белок-апоиндуктор.

Генетические последствия инактивации регуляторного белка могут быть использованы для разграничения систем позитивного и негативного контроля. Инактив-

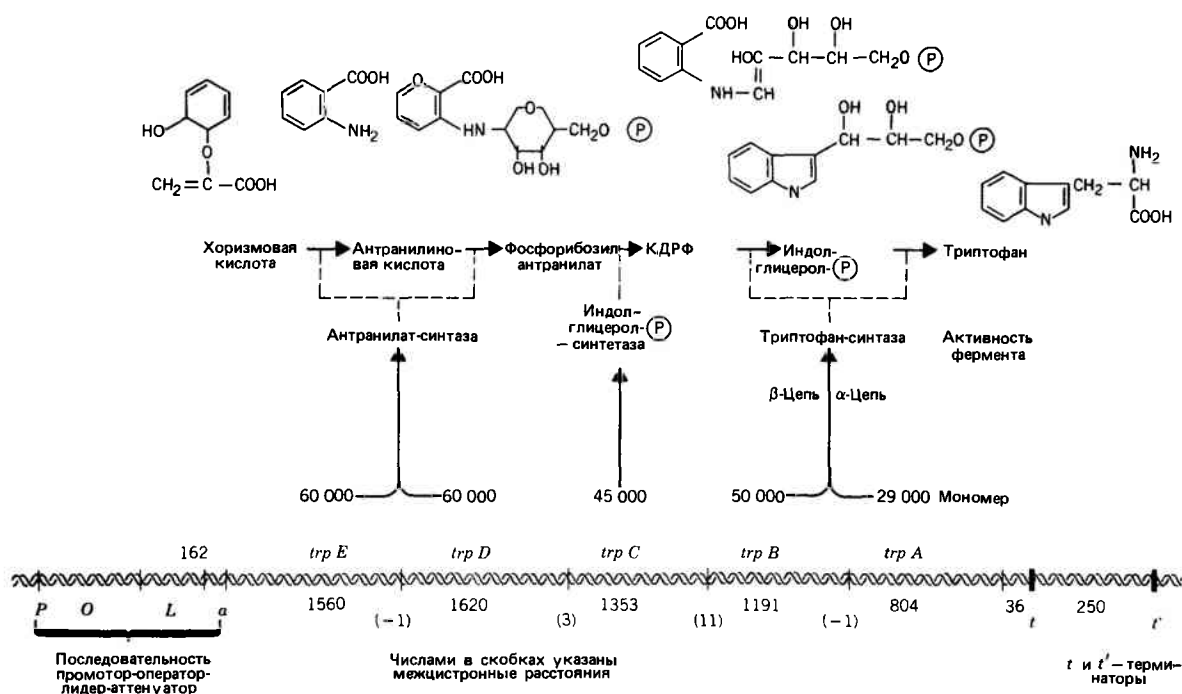


Рис. 15.2. *trp*-оперон содержит пять смежных структурных генов, которым предшествует контролирующая область. КДРФ — карбоксифениламин-1-дезоксирибулозофосфат.

вация белка-репрессора создает рецессивную конституивность (или состояние дерепрессии), типичную для систем негативного контроля. Мутация же, которая инактивирует апоиндуктор, вызывает рецессивную неиндудибельность или суперрепрессию, которая характерна для позитивного контроля.

Триптофановый оперон является репрессибельным

Пять структурных генов триптофанового оперона расположены рядом друг с другом и кодируют три фермента, превращающие хоризмовую кислоту в триптофан. Путь биосинтеза показан на рис. 15.2. Транскрипция инициируется в промоторе, в левом конце кластера генов. Прилегающий к нему оператор связывается с белком-репрессором, кодируемым несцепленным геном *trpR*. Между оператором и кодирующей областью первого гена находится лидерная последовательность. Транскрипция структурных генов терминируется в ρ -независимом сайте *trpL*, локализующемся за 36-й парой оснований от конца последней кодирующей области оперона. Примерно еще через 250 пар оснований расположен другой терминатор, *trpL'*, который проявляет ρ -зависимость. У *E. coli* и у *S. typhimurium* оперон построен, по существу, одинаково.

Как активность, так и синтез триптофановых ферментов контролируются содержанием триптофана в клетке. Классическая обратная связь — **ингибирование конечным продуктом** — используется при синтезе ферментов. Каталитические активности первого фермента пути биосинтеза подавляются триптофаном, конечным продуктом. Это означает, что в тех случаях, когда в клетке имеется достаточное количество триптофана, клетка способна выключить

синтез других аминокислотных молекул путем подавления начального этапа процесса.

Триптофан способен функционировать также как корепрессор, активирующий белок-репрессор. Такой механизм является классическим в случае репрессии. Один из примеров такого механизма приведен на рис. 15.1. Более детально цикл репрессии показан на рис. 15.3. В условиях, когда снабжение триптофаном оказывается достаточным, оперон репрессируется, так как комплекс репрессорный белок-корепрессор связывается с оператором. В тех же случаях, когда триптофана недостаточно, корепрессор не активирует белка-репрессора, что приводит к ослаблению специфичности в отношении оператора, и в результате репрессор оказывается связанным с какими-то другими участками ДНК.

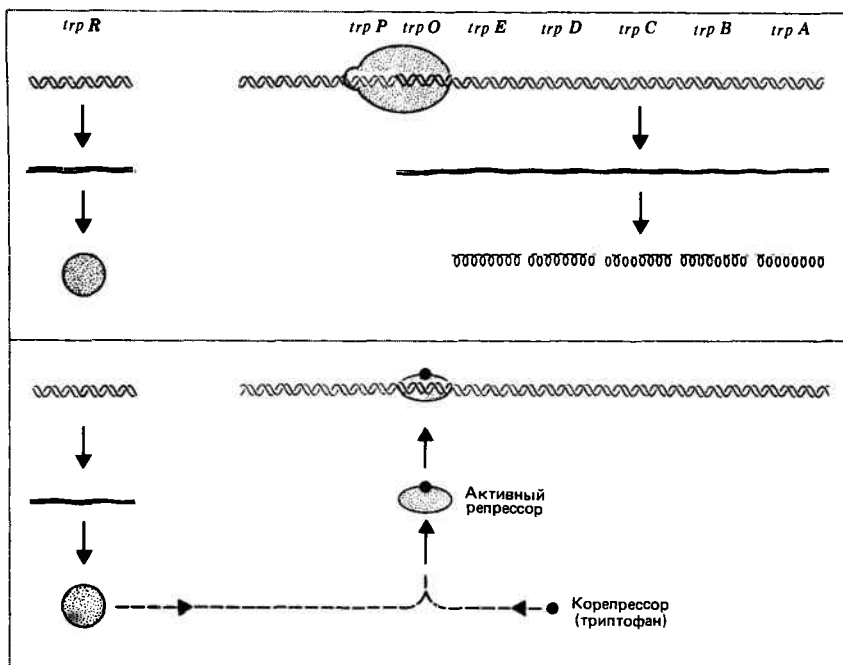
Триптофановый репрессор функционирует в виде тетрамера, состоящего из четырех субъединиц с мол. массой 12000 дальтон каждая. В бактериальной клетке содержится около 20 молекул тетрамера. Репрессор способен связываться с оператором *in vitro* только в присутствии триптофана. Таким образом, единственным существенным отличием в этом случае от взаимодействия в *lac*-опероне является увеличение сродства белка репрессора к оператору в присутствии триптофана.

На рис. 15.4 представлена последовательность нуклеотидов промотор-операторной области триптофанового оперона. РНК-полимераза связывается с областью, распространяющейся примерно от нуклеотида в положении -40 до нуклеотида в положении +18. В пределах этой области она образует контакты, аналогичные описанным в гл. 11. Промотор имеет канонические последовательности в положениях от -35 до -10.

Оператор лежит полностью в пределах промоторной области. Точки контакта с репрессором расположены

Рис. 15.3. Транскрипция *trp*-оперона контролируется репрессором.

Вверху: в отсутствие триптофана репрессор неактивен и РНК-полимераза инициирует транскрипцию. Внизу: добавление триптофана активирует репрессор; в результате он связывается с оператором и предотвращает транскрипцию.



в основном симметрично на двухцепочечной ДНК и занимают область между нуклеотидами -23 и -3 . Оператор содержит область двойной симметрии, которая включает обобщенную последовательность промотора в положении -10 .

Несмотря на то что взаиморасположение операторной и промоторной последовательностей в опероне *trp* отличается от такового в *lac*-опероне, общий механизм репрессии остается тем же: связывание РНК-полимеразы и связывание репрессора являются взаимоисключающими событиями.

Модификация координированной регуляции

Отсутствие репрессора вызывает примерно 70-кратное увеличение частоты актов инициации в *trp*-промоторе. Даже в условиях репрессии структурные гены сохраняют низкий **основной**, или **репрессированный**, уровень экспрессии. Это значит, что эффективность репрессии в операторе значительно слабее, чем та, которую мы видели в *lac*-опероне (базальный уровень которого составляет $\sim 1/1000$ от индуцированного уровня). В условиях репрессии основной уровень трех последних полипептидов оперона примерно в 5 раз выше, чем уровень продуктов генов *trpED*. Некоординированное выражение генов *trpCBA* обуславливается иницированием транскрипции во внутреннем промоторе, обозначенном *trpP2*, который лежит в конце гена *trpD*. *trpP2* — нерегулируемый промотор с низким сродством к РНК-полимеразе, поэтому транскрипция иницируется конститутивно с низкой частотой. В результате транскрипции генов *trpCBA* образуется мРНК, при трансляции которой увеличивается базальный уровень продуктов этих генов.

Дополнительная экспрессия трех последних генов оперона, возможно, обусловлена аминокислотным составом белков. В отличие от других белков продукты генов *trpB*

и *trpC* содержат триптофан. В условиях строгого голодания по триптофану синтез этих белков был бы невозможен, а их отсутствие блокировало бы любое восстановление эндогенного биосинтеза. Однако в тех случаях, когда клетка имеет адекватный основной уровень этих ферментов, их активность может оказаться достаточной для иницирования этого процесса.

Триптофановый оперон контролируется с помощью аттенуации

Почти с самого начала изучения триптофанового оперона было известно, что помимо комплекса промотор—оператор в регуляции участвует еще и другой сайт. Его существование впервые было обнаружено в результате наблюдений, показавших, что делетирование последовательности, расположенной между оператором и кодирующей областью *trpE*, приводит к повышенной экспрессии структурных генов. Этот эффект не зависит от репрессии: возрастают как базальный, так и дерепрессированный уровни транскрипции. Следовательно, происходящие при этом события совершаются *после* того, как РНК-полимераза покинула промотор (независимо от условий, существующих в момент иницирования транскрипции).

Данный регуляторный сайт получил название **аттенуатор**. Он локализован в пределах транскрибируемой лидерной последовательности из 162 нуклеотидов, которая предшествует иницирующему кодону гена *trpE*.

Аттенуатор служит барьером для транскрипции. Он представлен р-независимым сайтом терминации, подобным тем, которые уже обсуждались нами в гл. 13. Короткий G—C-богатый палиндром сменяется восемью следующими друг за другом остатками U. Как *in vivo*, так и *in vitro* РНК-полимераза прекращает транскрипцию в этом участке; в результате образуется транскрипт длиной в 140 оснований.



Рис. 15.4. Оператор *trp*-генов лежит в пределах промотора.

Транскрипция инициируется в положении +1. Обычные гексамеры, родственные каноническим последовательностям, присутствуют в положениях -10 и -35. Существует повтор из 10 пар оснований в пределах оператора (с одним несовпадающим положением). Все мутации σ^c локализируются в пределах области двойной симметрии, но вне канонической последовательности в положении -10, с которой она перекрывается.

Акт терминирования в этом сайте зависит от содержания триптофана (рис. 15.5). В присутствии достаточных количеств триптофана терминирование осуществляется эффективно. Однако в отсутствие триптофана РНК-полимераза способна продолжать транскрибирование структурных генов. Такая регуляция терминирования транскрипции получила название **аттенуация**.

Мы видим, что события, происходящие в аттенуаторе, направлены на осуществление такого же типа контроля, как и события в операторе. В присутствии триптофана оперон репрессирован и большая часть молекул РНК-полимеразы, которые уже переместились из промотора, заканчивают транскрипцию в аттенуаторе. При удалении триптофана РНК-полимераза получает свободный доступ к промотору, и процесс транскрипции, начавшись, не прерывается преждевременно.

Аттенуация может изменять интенсивность транскрипции в 8–10 раз. При эффективной терминировании около 10% молекул РНК-полимеразы, достигающих терминатора, способны проскочить его. В условиях аттенуации, по-видимому, почти все молекулы РНК-полимеразы способны продолжать транскрипцию. Наряду с 70-кратным увеличением эффективности инициации при снятии репрессии аттенуация позволяет варьировать уровень активности оперона в пределах, достигающих 600-кратного различия.

Аттенуация контролируется с помощью альтернативных вторичных структур

Каким образом содержание триптофана в клетке влияет на терминирование транскрипции в аттенуаторе? О механизме этого процесса можно судить на основе данных о нуклеотидной последовательности лидерной области.

На рис. 15.6 показано, что в эту область входит сайт связывания с рибосомами, в котором за кодоном AUG следует область из 13 кодонов. Транслируется ли эта область в **лидерный пептид**? Хотя продукт не был обнаружен *in vivo*, его отсутствие могло объясняться просто его нестабильностью. Известно, что этот сайт связывания с рибосомами является функционально активным, поскольку в тех случаях, когда он слит со структурным геном, поддерживает эффективную трансляцию.

Какова функция лидерного пептида? Он содержит два триптофановых остатка, расположенных непосредственно друг за другом. Триптофан является редкой аминокислотой в составе белков *E. coli*, поэтому трудно себе представить, чтобы такое расположение было обусловлено простой случайностью. О его значении свидетельствуют

данные, показывающие, что мутации в гене триптофан-тРНК—синтазы предотвращают полную репрессию оперона. Эффект этих мутаций выражается в лишении клетки необходимой триптофановой тРНК; в результате клетка оказывается не способной транслировать эти кодоны. По существу, при этом имитируются условия, возникающие в клетке при истощении запасов триптофана.

При таких условиях рибосомы, по-видимому, инициируют трансляцию лидерного пептида, но останавливают ее, когда достигают кодонов триптофана. Исходя из последовательности нуклеотидов в мРНК, можно предположить, что такая **задержка рибосом** в свою очередь может влиять на терминирование в аттенуаторе.

Лидерная последовательность может быть представлена в виде альтернативных структур, образуемых спаренными основаниями. Способность рибосомы проходить через лидерную область может регулироваться переходом из одной структуры в другую. Именно от этой структуры зависит, обеспечит ли мРНК условия, необходимые для терминирования.

На рис. 15.7 изображены эти структуры. В первой имеет место спаривание области 1 с областью 2 и области 3 с областью 4. Спаривание областей 3 и 4 порождает образование структуры шпильки, предшествующей последовательности U_g . В результате появляется существенный сигнал для терминирования трансляции. Вероятно, РНК принимает такую структуру без какого-либо влияния извне.

В тех случаях, когда спаривание областей 1 и 2 невозможно, образуется другая структура. Она возникает при свободном спаривании области 2 с областью 3. Вследствие этого область 4 не находит подходящего партнера для спаривания и остается одноцепочечной. В результате образование шпильчатой структуры терминатора оказывается невозможным.

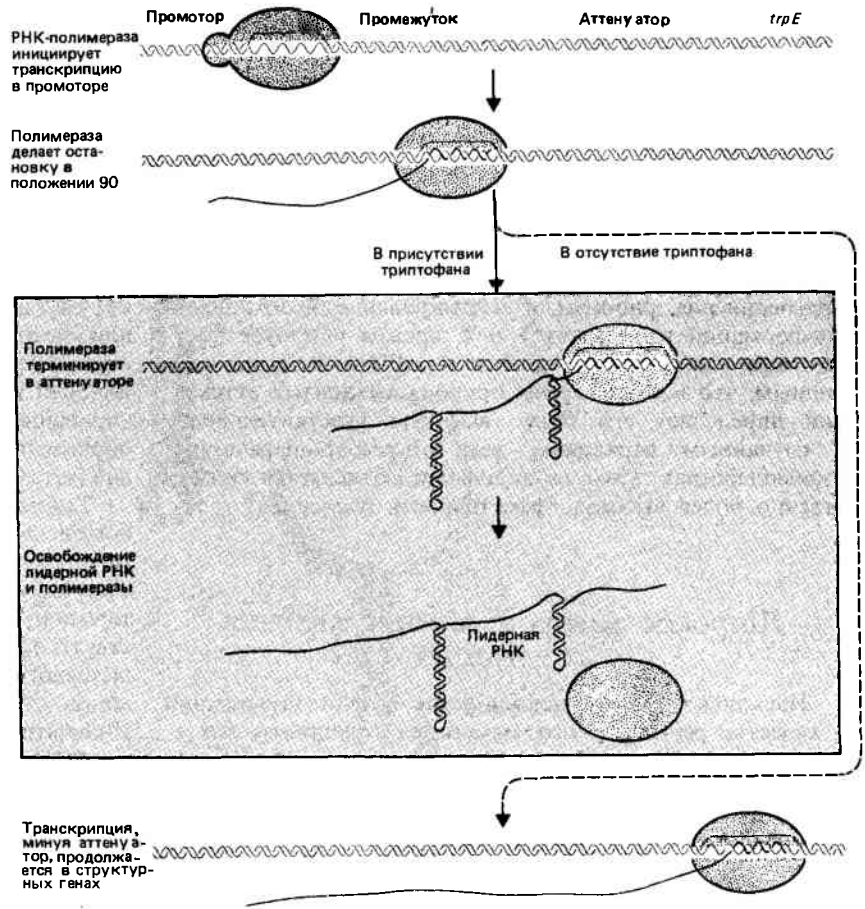
На рис. 15.8 показано, что положение рибосомы может определять, какая из структур будет образована, причем таким способом, что терминирование аттенуируется только в отсутствие триптофана.

В присутствии триптофана рибосомы способны синтезировать лидерный пептид. Они продолжают передвигаться вдоль лидерной области мРНК к кодону UGA, расположенному между областями 1 и 2. Как показано на рисунке, продвигаясь к этой точке, рибосомы переходят в область 2 и предотвращают ее спаривание. Тогда область 3 может спариться с областью 4, образуя шпильчатую структуру терминатора. При таких условиях РНК-полимераза терминирует процесс транскрипции в аттенуаторе.

В отсутствие триптофана рибосомы задерживаются в кодонах триптофана, входящих в состав области 1. В результате область 1 изолируется с помощью рибосом

Рис. 15.5. Аттенуатор контролирует продвижение РНК-полимеразы в область *trpE*-генов.

После инициации в промоторе (независимо от того, имеет ли место базальная экспрессия или депрессия) РНК-полимераза передвигается в положение 90, где она останавливается. При достижении аттенуатора в присутствии триптофана происходит терминация с вероятностью, достигающей примерно 90%, с освобождением лидерной РНК размером 140 пар оснований. В отсутствие триптофана полимераза продвигается в область структурных генов (начало *trpE*-гена соответствует положению +163).



и не может осуществить спаривание с областью 2. Если это происходит несмотря на то, что мРНК еще продолжает синтезироваться, области 2 и 3 должны спариться раньше, чем область 4 начнет транскрибироваться. В результате область 4 остается в одноцепочечной форме. В отсутствие терминаторной структуры шпильки РНК-полимераза продолжает транскрипцию, минуя аттенуатор.

Голодание по другим аминокислотам не приводит к такому результату, так как положения, в которых рибосома останавливается, не препятствуют спариванию областей 1 и 2, что в свою очередь способствует образованию структуры шпильки спарившимися областями 3 и 4. Случай с голоданием по глицину показан на рисунке. Единственное исключение представляет аргинин, кодирующая последовательность которого следует непосредственно за кодонами триптофана. Остановка рибосомы в этом положении также ослабляет терминацию, хотя и менее эффективно, чем в случае ее остановки в кодонах триптофана.

В пользу такой модели свидетельствуют свойства мутаций, влияющих на события в аттенуаторе. Большинство мутаций снижают эффективность терминации, способствуя тем самым повышенному выражению структурных генов. Такие мутации подобны мутациям, обнаруженным в других терминаторах. Один класс мутаций элиминирует пары оснований из двухцепочечной области 3:4, что приводит к ослаблению стабильности шпильчатой структуры терминатора. Мутации другого типа, ло-

кализующиеся в области, соответствующей остаткам U, обладают таким же действием.

Некоторые мутации в лидерной области усиливают эффект терминации в аттенуаторе. Они могут препятствовать снятию терминации при голодании по триптофану. Одна из таких мутаций дестабилизирует спаривание между областями 2 и 3. В результате область 3 приобретает возможность спариться с областью 4 и образовать шпильчатую структуру терминатора даже в тех случаях, когда клетки голодают по триптофану. Другая мутация изменяет иницирующий кодон AUG лидерного пептида, предотвращая в результате трансляцию. Наблюдаемое в результате увеличение эффективности терминации свидетельствует о том, что транскрипция с проскакиванием аттенуатора зависит от способности транслировать лидерную область.

Следовательно, аттенуация обеспечивает механизм, реагирующий на неадекватность снабжения Тгр-тРНК. Иными словами, он отвечает непосредственно на потребность клетки в триптофане для белкового синтеза. На рис. 15.9 показаны два возможных состояния в аттенуаторе — либо происходит терминация, либо — считывание.

Критическое значение для чувствительности механизма имеет положение кодонов триптофана в лидерной области относительно локализации аттенуатора. Детально, необходимой для правильной согласованности всех событий, является наличие сайта, вызывающего остановку РНК-полимеразы в положении, соответствующем 90-му основанию лидерной последовательности. Остановка РНК-полимеразы обеспечивает возможность

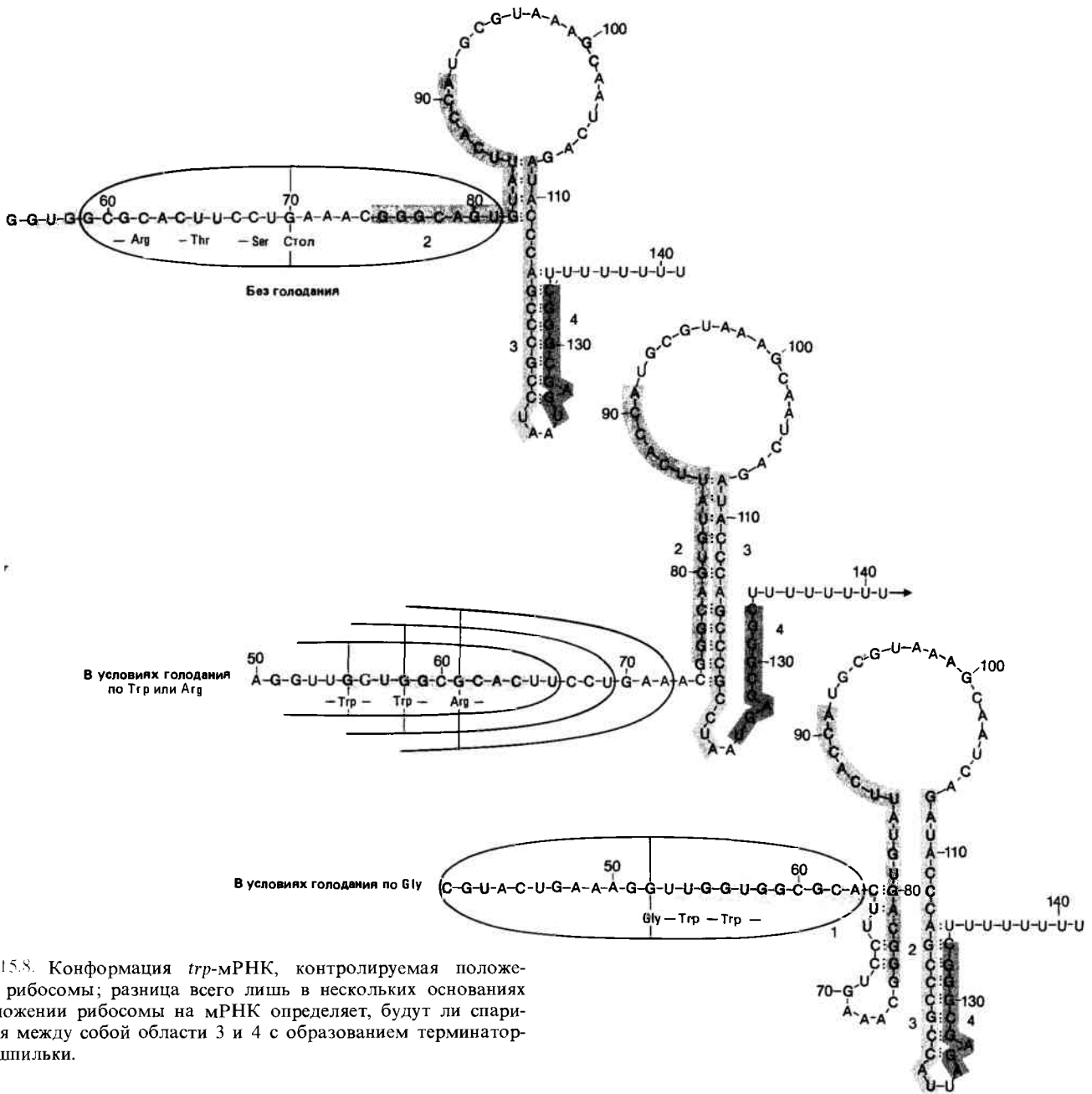


Рис. 15.8. Конформация *trp*-мРНК, контролируемая положением рибосомы; разница всего лишь в нескольких основаниях в положении рибосомы на мРНК определяет, будут ли спариваться между собой области 3 и 4 с образованием терминальной шпильки.

ститутивный характер мутации обусловлен делетированием последовательности аттенуатора. В результате процесс терминации не может осуществляться и РНК-полимераза во всех случаях считывает последовательности структурных генов. Неиндуцибельные мутации могут действовать либо непосредственно на вторичную структуру мРНК, либо опосредованно через лидерный пептид. Мутация, предотвращая образование альтернативной шпилечной структуры, гарантирует обязательное образование терминирующей шпильки; в результате полимераза оказывается не способной транскрибировать структурные гены. Другая мутация превращает кодон САА, обеспечивающий включение глутамина в положение 5 лидерного пептида, в охра-кодон UAA. Терминация трансляции в этой точке эффективно способствует созданию ситуации «отсутствия трансляции» (подобной той, что показана на рис. 15.9), при которой всегда образуется шпилечная структура терминатора.

Сходство между механизмами аттенуации триптофанового и гистидинового оперонов отчетливо выражено. В каждом случае лишение клеток именно аминоксил-тРНК непосредственно предотвращает завершение транскрипции и обеспечивает таким образом транскрибирование структурных генов. Разница лишь в том, что триптофановый оперон подвержен также репрессорному оперонному взаимодействию, тогда как в гистидиновом опероне аттенуация обеспечивает единственный способ контроля. (Мутации в лидерной области *his*-оперона первоначально были обозначены как *hisO*, поскольку считалось, что с их помощью идентифицирован оператор. Этот пример показывает, что природа таких мутаций не может считаться установленной до тех пор, пока не будет исследован биохимически молекулярный механизм контроля.)

Некоторые другие опероны проявляют те же основные свойства. На рис. 15.10 показаны последователь-

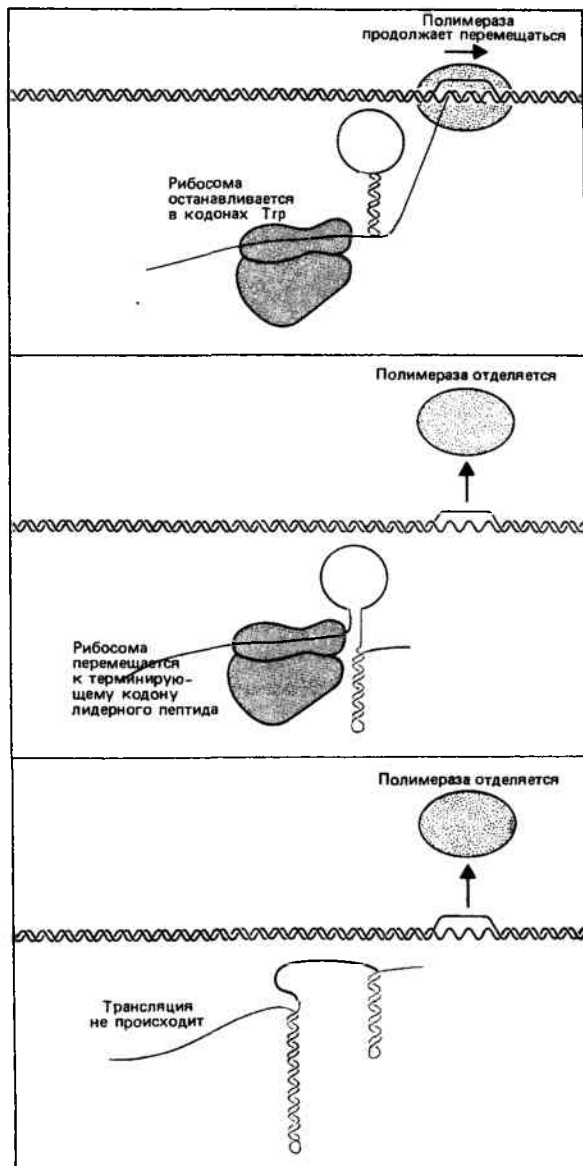


Рис. 15.9. Альтернативные эффекты по отношению к РНК-полимеразе в аттенуаторе зависят от местоположения рибосомы.

ности их лидерных пептидов: в каждом случае остановка рибосомы в кодонах, соответствующих регуляторным аминокислотам, может способствовать изменению вторичной структуры мРНК таким образом, что шпильчатая структура терминатора не будет образовываться. Опероны *trp* и *ilv* подвержены **мультивалентной репрессии**. Каждый из них дерепрессируется при голодании по нескольким аминокислотам: двум – в случае *trp*-оперона и трем в случае *ilv*-оперона. Последовательность лидерного пептида показывает, каким способом это достигается. Кодоны, соответствующие различным аминокислотам, регулирующий оперон, расположены таким образом, что остановка рибосомы в любом из них способна предотвратить образование шпильчатой структуры терминатора.

Отдельные лидерные последовательности значительно длиннее лидера *trp*-оперона, а области, в которых может происходить задержка рибосом, вызывающая дерепрес-

сию, являются более протяженными. Не исключено, что в этих случаях для достижения максимальной дерепрессии может оказаться необходимой остановка более чем одной рибосомы в лидерной области. Точно так же может происходить образование нескольких петель вторичной структуры. Однако в каждом случае сохраняется один и тот же принцип.

Формально аттенуация может быть отнесена к одному из механизмов, в котором функции рибосом эквивалентны функциям белка, действующего по принципу положительной регуляции. Их связывание в сайте (сайтах) остановки – необходимое условие для предотвращения процесса терминации. Рибосомы утрачивают способность связываться в этих сайтах (инактивироваться) при присоединении определенной аминокил-тРНК, которая благодаря этому выступает эквивалентом корепрессора.

Репрессия может иметь место для множества локусов

Белок-репрессор *trp*-оперона действует не только в операторе *trp*-оперона; он репрессирует также транскрипцию двух других локусов.

Ген *aroH* кодирует один из трех ферментов, катализирующих начальную реакцию в общем пути биосинтеза ароматических аминокислот. Его выражение репрессируется триптофаном через активацию *trp*-репрессора. (Другие ферменты, кодируемые генами *aroF* и *aroG*, репрессируются другими регуляторными белками.) Значение таких типов контроля состоит в распространении регуляторной сети не только на пути, связанные с завершением синтеза различных ароматических аминокислот, но и на предшествующие стадии, в процессе которых синтезируются продукты начальных этапов данного пути.

Регуляторный ген *trpR* репрессируется своим собственным продуктом – *trp*-репрессором. Следовательно, белок-репрессор действует, снижая и свой собственный синтез. Такой цикл может служить примером **аутогенного контроля**. Мы увидим, что подобные циклы весьма распространены в регуляторных и других генах и могут осуществлять как негативный, так и позитивный контроль. Наиболее распространен негативный аутогенный контроль; в этом случае белок ингибирует свой собственный синтез, так что его содержание в клетке саморегулируется. В тех случаях, когда содержание данного белка становится слишком высоким, дальнейшее образование репрессора предотвращается, так как белок подавляет транскрипцию своего собственного гена. Если же содержание репрессора падает, белок оказывается не способным подавлять свой собственный синтез; в результате возобновляется транскрипция, содержание белка восстанавливается. (В случае позитивного аутогенного контроля белок способствует своему собственному синтезу; как будет видно в гл. 16, этот тип связи обеспечивает включение/выключение экспрессии генов.)

Родственная последовательность оператора протяженностью свыше 21 пары оснований имеется в каждом из трех локусов, в которых действует *trp*-репрессор. Нуклеотидная последовательность этих операторов приведена на рис. 15.11. Каждый оператор содержит последовательности со значительной (но не идентичной) двойной симметрией. По-видимому, особенности, характерные для всех трех операторов, отражают наличие точек, существенных для контактов с *trp*-репрессором. Это делает

Оперон	Последовательность лидерного пептида	Регуляторные аминокислоты
<i>his</i>	Met- Thr- Arg- Val- Gln- Phe- Lys- His- His- His- His- His- His- His- Pro- Asp	His
<i>pheA</i>	Met- Lys- His- Ile- Pro- Phe- Phe- Phe- Ala- Phe- Phe- Phe- Thr- Phe- Pro	Phe
<i>leu</i>	Met- Ser- His- Ile- Val- Arg- Phe- Thr- Gly- Leu- Leu- Leu- Leu- Asn- Ala- Phe- Ile- Val- Arg- Gly- Arg- Pro- Val- Gly- Gly- Ile- Gln- His	Leu
<i>thr</i>	Met- Lys- Arg- Ile- Ser- Thr- Thr- Ile- Thr- Thr- Thr- Ile- Thr- Ile- Thr- Thr- Gly- Asn- Gly- Ala- Gly	Thr, Ile
<i>iv</i>	Met- Thr- Ala- Leu- Leu- Arg- Val- Ile- Ser- Leu- Val- Val- Ile- Ser- Val- Val- Val- Ile- Ile- Ile- Pro- Pro- Cys- Gly- Ala- Ala- Leu- Gly- Arg- Gly- Lys- Ala	Leu, Val, Ile

Рис. 15.10. Последовательности лидерных пептидов оперонов, контролирующих биосинтез аминокислот, содержат множество кодонов для аминокислот, регулирующих оперон.

понятным, как один белок-репрессор действует в нескольких локусах: каждый локус имеет копию специфической нуклеотидной последовательности, узнаваемой репрессором (точно так же, как каждый промотор содержит специфическую последовательность, характерную для других промоторов).

Отличительной особенностью таких разбросанных операторов является их различная локализация относительно стартовой точки в каждом локусе. На рис. 15.11 показано, что в триптофановом опероне оператор расположен между нуклеотидами -23 и -3 , тогда как в *trpR* лежит между нуклеотидами -12 и $+9$, а в локусе *araH* находится дальше от начала транскрипции, располагаясь между нуклеотидами -49 и -29 . Мы уже видели, что разные репрессоры связываются в сайтах, локализованных в различных участках в пределах промоторов или сцепленных с ними областей. Тот факт, что один и тот же репрессор оказывается эффективным в различно расположенных операторах, подтверждает вывод, что репрессия представляет собой способ блокировать доступ к промотору.

Арабинозный оперон находится под двойным контролем

До сих пор мы имели дело с системами, которые можно формально охарактеризовать как примеры положительного или отрицательного контроля. Однако система контроля арабинозного оперона не укладывается ни в одну из этих категорий. Этот оперон регулируется с помощью белка, взаимодействующего с контролирующим сайтом в промоторе, однако регуляция проявляет черты как отрицательного, так и положительного контроля.

Отдельные гены контролируются как часть этой системы. Организация системы контроля арабинозного оперона показана на рис. 15.12. Основная группа генов кодирует три фермента, превращающие L-арабинозу в D-ксилозу-5-фосфат. Они являются продуктами генов *araBAD*, транскрибируемых в полицистронную мРНК в указанном порядке. Два несцепленных гена *araE* и *araF*, кодирующие белки мембраны и периплазмы соответственно, которые участвуют в поглощении арабинозы, подвержены той же регуляции, что и гены *araBAD*. Иногда систему, объединяющую разобщенные гены общим контролем, называют регулоном.

Ген *araC* является регуляторным. Он локализован перед промотором генов *araBAD* и транскрибируется в направлении, противоположном направлению транскрипции группы генов *araBAD*. Следовательно, область между генами *araBAD* и *araC* должна содержать два промотора в противоположной ориентации. В них иницируется так называемая **дивергентная транскрипция**.

В гене *araC* были обнаружены как неиндуцибельные, так и конститутивные мутации. Делеции или нонсенс-мутации вызывают образование неиндуцибельного *araC⁻*-типа, который является рецессивным по отношению к *araC⁺*. Такая взаимосвязь характерна для системы положительного контроля, в которой *araC* кодирует белок, активный только в присутствии арабинозы и необходимый для того, чтобы РНК-полимераза могла иницировать в промоторе транскрипцию. Однако конститутивные мутации *araC^c* также рецессивны по отношению к *araC⁺*. Такое положение характерно для системы отрицательного контроля, в которой *araC* кодирует белок-репрессор, инактивируемый арабинозой (подобно тому, как в *lac*-опероне). Как разрешается такой генетический конфликт? На рис. 15.13 изображена модель. Ген *araC* кодирует белок-репрессор, исходно названный P1, а впоследствии — C^{rep}. Этот белок связывается с контролирующим сайтом, предотвращая транскрипцию генов *ara*. Функция арабинозы сводится к связыванию с репрессором в качестве малых молекул индуктора. До этого момента имеет место типичный отрицательный контроль. Но связывание с арабинозой не только инактивирует способность репрессора присоединяться к оператору. Вместо этого репрессор превращается в белок-индуктор, который первоначально был назван P2, а теперь известен как C^{ind}. Его связывание с промотором необходимо для включения транскрипции. Это уже характерно для положительного контроля.

Сложная организация регуляторной области *ara*-оперона

Две формы белка C должны иметь различные сайты связывания в контролирующей области *ara*-оперона. Они были идентифицированы с помощью *cis*-мутаций. Оператор *araO* определяется обычным способом, как сайт, в котором происходит связывание C^{rep}. Инициатор, *araI*,

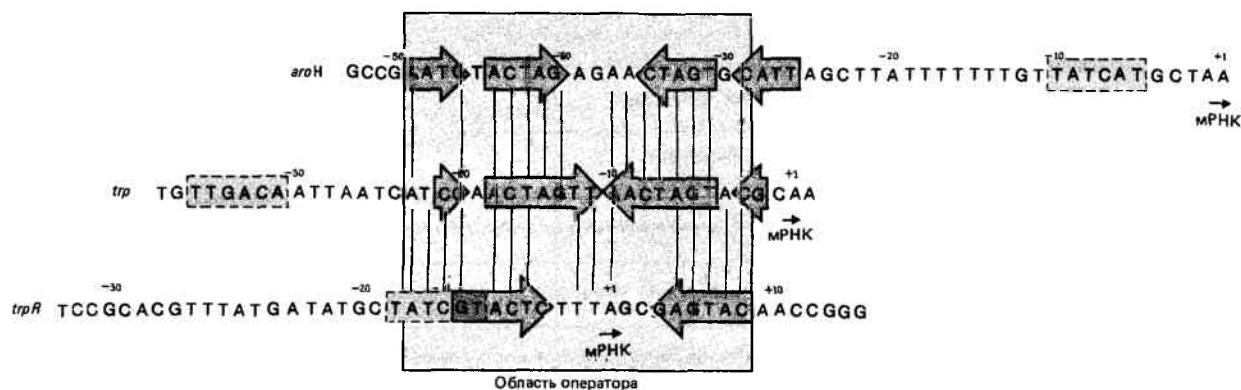


Рис. 15.11. *trp*-Репрессор узнает операторы в трех локусах. Вертикальные линии соответствуют консервативным основаниям в операторах. Стрелками показаны области двойной симметрии. Прямоугольниками обозначены канонические последовательности промотора (с хо-

рошим соответствием в *aroH* и *trp*, менее удачным — в *trpR*). В *aroH* не содержится последовательность —35, а в *trp* нет последовательности —10.

характеризуется как сайт, участвующий в иницировании транскрипции. Он содержит сайт связывания с белком C^{ind} . Примерное расположение этих сайтов показано на рис. 15.13.

Контролирующая область *ara*-оперона расположена между стартовыми точками двух дивергентно транскрибирующихся единиц (*araC* и *araBAD*). Разделяющее их расстояние соответствует только 147 парам оснований. Два промотора занимают значительную часть этой области. Наряду с сайтами, связывающими белки C^{rep} и C^{ind1} , в этой области должен также находиться по крайней мере один сайт связывания с регуляторным белком БАК (гл. 11), так как обе единицы транскрипции являются зависимыми от комплекса БАК-циклический АМР. Схема, представляющая современный взгляд на расположение регуляторных элементов относительно друг друга, приведена на рис. 15.14.

Очистка белка C позволила идентифицировать непосредственно последовательность ДНК, с которой связывается каждая из форм белка. Белок представляет собой димер из идентичных субъединиц с мол. массой 33 000 дальтон. Он может быть получен в C^{rep} -форме при связывании с антииндуктором D-фукозой. Он существует в C^{ind} -форме, если связывается с арабинозой. Эксперименты с использованием метода «отпечатков пальцев» показывают, что в любой форме он перекрывает около 38 пар оснований ДНК.

В форме C^{rep} он связывается только с областью, расположенной между основаниями —44 и —106 относительно стартовой точки оперона *araBAD*. В форме C^{ind} он сохраняет некоторое сродство к этому сайту, но связывается преимущественно в области, расположенной между нуклеотидами в положениях —78 и —40. Таким образом, превращение белка из одной формы в другую сопровождается переключением его связывания из области, представленной 38 парами оснований, центр которой находится на 125-й паре оснований против хода транскрипции, в область, центр которой расположен в 59-й паре оснований, отсчитываемой против хода транскрипции от стартовой точки. По аналогии со свойствами других связывающихся с ДНК белков такое переключение может осуществляться путем вытеснения.

Белок БАК (в случае активации циклическим АМР)

имеет два связывающих сайта в контролирующей области (см. также ниже в данной главе). Основной сайт лежит между нуклеотидами в положении —107 и —78, как показано на рисунке. [Другой сайт с более низким сродством (на рисунке не показан) лежит между нуклеотидами —146 и —121. Не ясно, используется ли он *in vivo*.] Напомним, что белок БАК необходим для активации *обеих* единиц транскрипции. Это может быть осуществлено в том случае, если он действует из одного, расположенного в середине сайта связывания.

Промоторы не были охарактеризованы путем непосредственного связывания их с РНК-полимеразой. Можно, однако, с уверенностью сказать, что каждый имеет протяженность, соответствующую примерно 50 парам оснований, расположенным против хода транскрипции от стартовой точки.

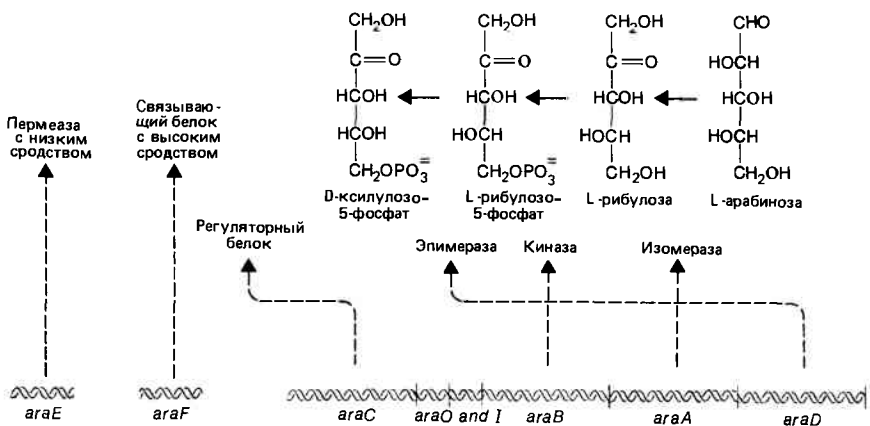
Каким образом взаимодействуют различные связывающие сайты при осуществлении их характерного двойного контроля оперона *araBAD*?

Выражение оперона требует, чтобы три белка связывались с ДНК одновременно, вероятно контактируя друг с другом, как это показано на рис. 15.14. На основе результатов, полученных при изучении свойств двух типов конститутивных мутаций, мы знаем, что как C^{ind} , так и БАК проявляют свое действие через промотор. Мутации *araI^c* позволяют оперону выражаться в отсутствие C^{ind} . Такие мутации локализируются в области —35. Мутации *araX^c* делают возможным выражение оперона в отсутствие как C^{ind} , так и БАК. Они локализованы в области —10. Способность РНК-полимеразы иницировать транскрипцию в промоторе *araBAD* обычно зависит от одновременного присутствия белков C^{ind} и БАК в их сайтах связывания, однако эти требования могут быть сняты в результате мутаций. Непонятным остается характер взаимодействия этих белков. Порядок расположения сайтов связывания дает основание думать, что белок БАК взаимодействует с белком C^{ind} , что является предпосылкой для взаимодействия белка C^{ind} с РНК-полимеразой. Такое поведение отличается от прямого взаимодействия белка БАК с РНК-полимеразой, происходящего, по-видимому, в других локусах.

Пока остается неясным, как осуществляется репрессия оперона *araBAD*. Трудно понять, каким образом белок C^{rep} , связывающийся так далеко в области, расположенной против хода транскрипции, может влиять на промо-

¹ C — от англ. CAP — catabolite activator protein — Прим. ред.

Рис. 15.12. *ara*-Оперон состоит из регуляторного гена *araC*, отделенного от группы структурных генов *araBAD* регуляторной областью (несцепленные гены *araF* и *araE* также контролируются *araC*-геном).



тор. Если это тот сайт, в котором осуществляется отрицательная регуляция, механизм должен быть отличным от действия других репрессоров. Не исключено, что одним из эффектов белка C^{rep} является предотвращение связывания белка БАК с его сайтом. Если это так, то эффект превращения белка C^{rep} в C^{ind} сводится к перемещению белка с одной стороны БАК-сайта (где он подавляет связывание белка БАК) на другую сторону (где он помогает связыванию белка БАК).

Ген *araC* подвержен аутогенной репрессии белком C^{rep} . Сайт связывания белка C^{rep} перекрывается с последовательностью промотора гена *araC* обычным путем. В репрессированном состоянии происходит основная (на низком уровне) транскрипция гена *araC*, приводящая в результате к образованию белка C^{rep} . Создается ауто-

генный цикл, при котором содержание белка C^{rep} остается на уровне, достаточном только для того, чтобы поддерживать репрессированное состояние генов *araC* и *araBAD*. Добавление арабинозы для превращения белка C^{rep} в C^{ind} дает двойной эффект. С одной стороны, снимается репрессия с гена *araC*, что приводит к синтезу больших количеств белка C , с другой – белок C^{ind} активирует группу генов *araBAD*. Следовательно, полный цикл представляет крайне экономичную систему, использующую появление арабинозы для увеличения количества регуляторного белка, а также для перевода генов из выключенного состояния во включенное. Таким образом, в результате индукции образование минимальных адекватных количеств белка C^{rep} сменяется высоким выходом белка C^{ind} .

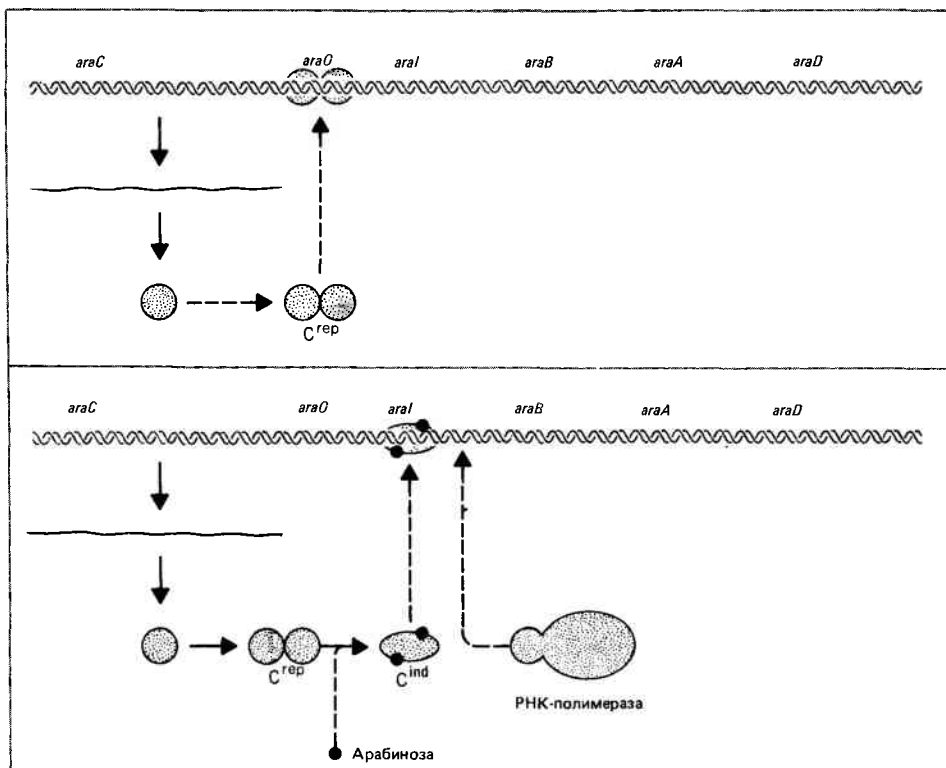


Рис. 15.13. Оперон *ara* подвергается как позитивному, так и негативному контролю белком C . Вверху. В репрессированном состоянии белок C^{rep} связывается с *araO*

и предотвращает транскрипцию. Внизу. В индуцированном состоянии белок C^{ind} связывается с *araI* и способствует иницированию транскрипции.

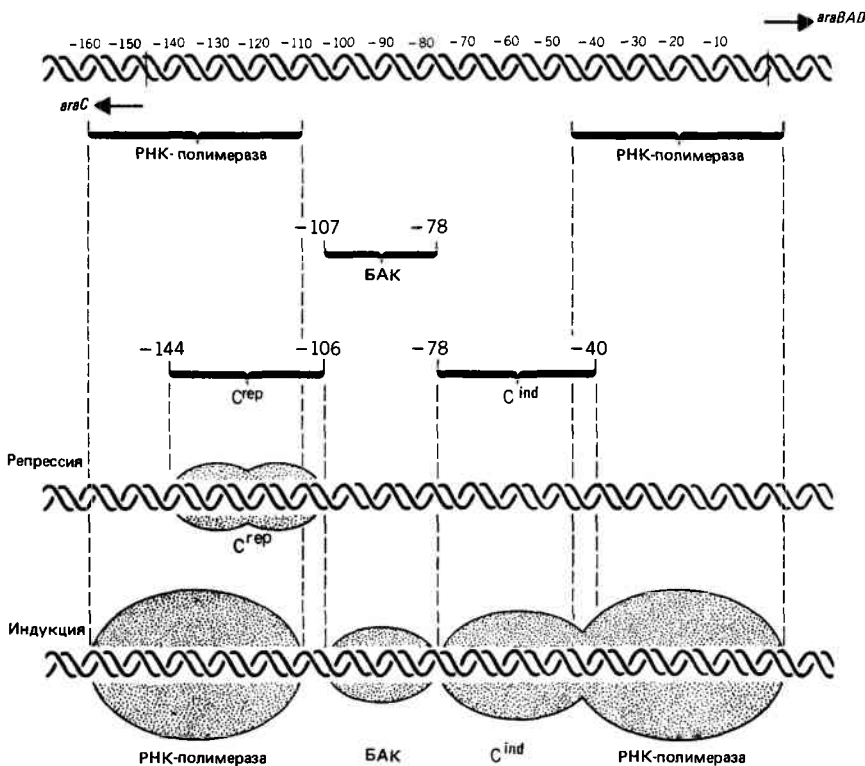


Рис. 15.14. Индукция *ara*-оперона включает множество актов связывания белков с ДНК. *Наверху* приведена карта регуляторной области *ara*-оперона с сайтами связывания белка С, белка БАК и РНК-полимеразы (предположительно). В условиях репрессии только белок С^{rep} связывается, как показано в центре. В условиях индукции белки БАК, С^{ind} и РНК-полимераза находятся в смежных сайтах вблизи генов *araBAD*, тогда как другая РНК-полимераза связывается с промотором *araC*-гена.

Двойной промотор галактозного оперона

Три фермента, кодируемые галактозным опероном, ответственны за превращение галактозы в галактозо-1-фосфат. Гены *galKTE* транскрибируются с одной из двух стартовых точек в полицистронную мРНК, которая последовательно транслируется с образованием трех ферментов. Процесс транскрипции подвержен классической отрицательной регуляции: репрессор, кодируемый несцепленным геном *galR*, инактивируется при добавлении галактозы. Для транскрипции необходимо присутствие активированного БАК-белка.

На рис. 15.15 показана нуклеотидная последовательность контролирующей области *gal*-оперона. Использование других ее стартовых точек при транскрипции определяется условиями инициации. В присутствии активного белка БАК инициация осуществляется в точке, обозначаемой обычно +1. (Все положения оснований в случае обоих промоторов обозначаются относительно этой точки.) Эта стартовая точка использована для идентификации промотора *galP1*, который содержит последовательность Прибнова между основаниями в положениях -12 и -16 и не содержит канонической последовательности -35. В отсутствие белка БАК оперон еще может транскрибироваться, но инициация происходит в положении -5. Эта точка соответствует промотору *galP2*, в котором имеется последовательность Прибнова между основаниями в положениях -17 и -11.

Таким образом, существует два перекрывающихся промотора, у которых разница между стартовыми точками составляет 5 пар оснований. Не известно, являются ли они независимыми в том смысле, что РНК-полимераза осуществляет полностью различный ряд контактов в каждом случае, или существует один связывающий сайт с одинаковыми контактами, производимыми в области,

расположенной выше против хода транскрипции, и стартовую точку определяет регуляция. Существование *galP1* и *galP2* устанавливает тот общий принцип, что выражение ряда структурных генов может быть иницировано более чем одним способом, т.е. обеспечивается возможность осуществления контроля независимыми системами.

Сайт связывания белка БАК расположен достаточно близко от стартовой точки в пределах расстояния, обычно перекрываемого РНК-полимеразой. Когда белок БАК связан с этой областью, он, по-видимому, непосредственно взаимодействует с РНК-полимеразой, влияя на ее конформацию таким образом, что она использует промотор *galP1*.

Оператор сцеплен с сайтом связывания белка БАК, но расположен дальше против хода транскрипции. Связывание с репрессором подавляет использование как *galP1*, так и *galP2*. При данном расстоянии оператора до стартовой точки и данном промежуточном положении сайта связывания БАК остается непонятным, как работает репрессор. Не исключено, что он только управляет проявлением стерических помех, предотвращающих связывание РНК-полимеразы с промотором, и это пока единственное объяснение его подавляющего эффекта в случае промотора *galP2*. Репрессор может препятствовать использованию промотора *galP1* непрямым способом, мешая связыванию БАК. Это та же самая модель, которую мы уже рассматривали при обсуждении случая репрессии генов *araBAD*.

При сравнении способов регуляции нескольких оперонов можно видеть, что используется ряд общих механизмов. Иницирование транскрипции контролируется белками, которые связываются со специфическими сайтами в ДНК. Присутствие белков в этих сайтах либо стимулирует, либо ингибирует связывание других белков в перекрывающихся или сцепленных сайтах. В каждом случае основным результатом является контроль способности

Таблица 15.1

Гены рибосомных белков, факторов белкового синтеза и субъединиц РНК-полимеразы разбросаны по нескольким оперонам¹

Оперон	Гены и белки (в направлении от промотора)	Регулятор
<i>str</i>	<i>rpsL- rpsG- fusA- tufA</i> S12 S7 EF-G EF-Tu	S7
<i>spc</i>	<i>rplN- rplX- rplE- rpsN- rpsH- rplF- rplR- rpsE- rplO- rpmD</i> L14 L24 L5 S14 S8 L6 L18 S5 L15 L30	S8
S10	<i>rpsJ- rplC- rplB- rplD- rplW- rplS- rplV- rpsC- rpsQ- rplP- rpmC</i> S10 L3 L2 L4 L23 S19 L22 S3 S17 L16 L29	L4
α	<i>rpsM- rpsK- rpsD- rpoA- rplQ</i> S13 S11 S4 α L17	S4
L11	<i>rplK- rplA</i> L11 L1	L1
<i>rif</i>	<i>rplJ- rplL- rpoB- rpoC</i> L10 L7/ β β' L12	L10

¹ Идентифицированы две главные группы генов, содержащих эти опероны. Одна содержит опероны: *str*-промежуток-S10-*spc*- α (где «промежуток» = 14 000 пар оснований); другая группа включает сцепленные опероны *L11-rif*. Каждый оперон записан таким образом, что его промотор находится в левом конце. Регуляторный белок указан справа. Белки, подверженные регуляции, подчеркнуты (коричневая линия). Гены, участие которых в системе регуляции точно не установлено, подчеркнуты прерывистой линией.

Почти каждый из этих белков детерминирован только одним геном в хромосоме *E. coli*. Гены, кодирующие рибосомные белки, факторы синтеза белков и субъединицы РНК-полимеразы, локализованы вперемежку и организованы в небольшое число оперонов. Данные об организации оперонов, которые охарактеризованы к настоящему времени, суммированы в табл. 15.1. Примерно половина генов рибосомных белков (часто сокращенно обозначаемых как **р-белки**) картируется в четырех оперонах, тесно сцепленных друг с другом. Они обозначаются как *str*, *spc*, S10 и α (каждый назван просто по первой из идентифицированных функций). Опероны *rif* и *L11* также сцеплены, но находятся в другой области.

Каждый оперон кодирует ряд функций. *str*-Оперон содержит гены рибосомных белков малой субчастицы, а также факторов EF-Tu и EF-G. В двух оперонах, *spc* и S10, расположены вперемежку гены белков как малой, так и большой рибосомных субчастиц. В α -опероне находятся гены белков обеих рибосомных субъединиц, а также ген α -субъединицы РНК-полимеразы. В состав *rif*-оперона входят гены белков большой субчастицы рибосом, а также β - и β' -субъединиц РНК-полимеразы.

В большинстве случаев не существует очевидной корреляции между локализацией и функциями белков. Белки 30S-субчастицы безусловно не кодируются в каком-либо порядке, отражающем их сборку в субъединицу. В то же время локализация группы белков 50S-субчастицы, кодируемых *rif*/L11-оперонами, проявляет корреляцию с порядком сборки, при котором L7/L12, L10 и L11 могут быть связанными. В этом случае белки, синтезируясь, могут объединяться, образуя некую субчастицу.

Все рибосомные белки, кроме одного, необходимы в эквимоллярных количествах, которые должны координироваться содержанием мРНК. (Вариации допустимы

и для тех рибосомных белков, которые проявляют дополнительные активности, например для белка S10, вовлекаемого в антитерминирующую функцию белка рN.) Разброс генов, продукты которых должны быть представлены в эквимоллярных количествах, и их перемежаемость с генами, продукты которых необходимы в различных количествах, порождают определенные интересные проблемы в связи с координированной регуляцией.

Исключение среди рибосомных белков представляет белок L7/L12, представленный, по-видимому, в виде четырех копий на рибосому. Другим исключением является фактор EF-Tu, присутствующий в количествах, примерно эквимоллярных количеству аминоксил-тРНК, т.е. примерно в 10 раз больших, чем количество рибосом. (Это один из случаев, когда имеется более одного гена, так что необходимость синтезировать сверхколичества распределяется между двумя генами, *tufA* и *tufB*.) Еще одно различие наблюдается между рибосомами и РНК-полимеразой, которая представлена в меньших количествах. Таким образом, какой-то механизм должен усиливать синтез L7/L12 и EF-Tu и ослаблять синтез субъединиц РНК-полимеразы относительно постоянного уровня рибосомных белков.

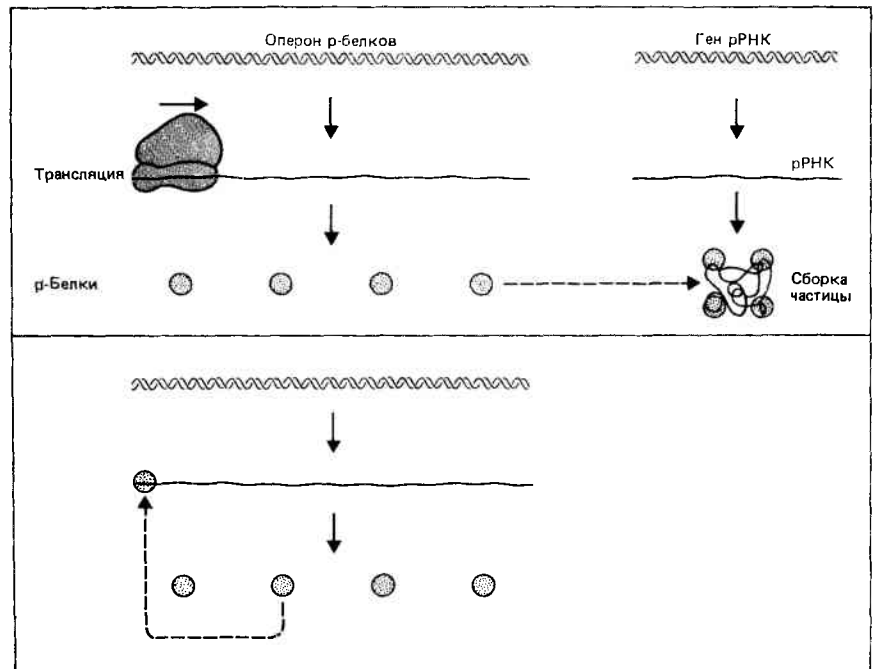
Общим свойством всех оперонов, указанных в табл. 15.1, является аутогенная регуляция некоторых из генов одним из продуктов. Обычно регуляторный белок подавляет выражение ряда смежных генов в пределах оперона и в том числе (всегда) выражение своего собственного гена.

Число генов, подчиняемых такому подавлению, варьирует от оперона к оперону. Так, L7 подавляет свой собственный синтез и синтез фактора EF-G, но не синтез факторов S12 или EF-Tu. Аналогично и S8 подавляет выражение всего своего оперона, за исключением первых

Рис. 15.16. Трансляция оперонов р-белков контролируется аутогенно и зависит от содержания рНК.

Вверху. При образовании рРНК происходит связывание с ней рибосомных белков; в результате пула свободных рРНК не образуется, и трансляция мРНК р-белков продолжается.

Внизу. После того как вся рРНК собрана в рибосомы и пула свободных рРНК нет, р-белки начинают накапливаться. Один из них связывается с мРНК р-белков и предотвращает дальнейшую трансляцию.



двух генов. С другой стороны, L4 и L1, вероятно, подавляют выражение всех генов в своих соответствующих оперонах. В оперонах, содержащих гены субъединиц РНК-полимеразы, белок S4 подавляет выражение генов всех других белков малой субъединицы рибосомы; подобным же образом L10 подавляет свое собственное образование, но не образование L7/L12 или субъединиц РНК-полимеразы.

В каждом случае имеет место накопление белка, который подавляет дальнейший синтез, детерминированный как собственным геном, так и любыми генами, подверженными репрессии. Эффект часто проявляется на уровне трансляции полицистронной мРНК и может быть в отдельных случаях воспроизведен *in vitro*. Таким образом, избыток свободного рибосомного белка служит сигналом для репрессии процесса трансляции.

Каждый из регуляторов является рибосомным белком, который связывается непосредственно с рРНК. Его влияние на трансляцию обусловлено способностью связываться также со своей собственной мРНК. В некоторых случаях были охарактеризованы сайты связывания. Например, в опероне L11 белок L1 связывается с сайтом, расположенным поблизости от иницирующего кодона первого гена. Это, по-видимому, подавляет связывание рибосом. Такое подавление влияет на трансляцию обоих генов оперона, вероятно, потому, что между генами L11 и L1 находится всего три нуклеотида. Следовательно, предположение, что гены могут быть транслированы только последовательно, вполне правдоподобно.

Использование р-белков, которые связываются с рРНК для установления немедленно аутогенной регуляции, свидетельствует, очевидно, о том, что это может служить механизмом для установления связи между синтезом р-белка и рРНК. Обобщенная модель приведена на рис. 15.6. Допустим, что сайты связывания аутогенно регуляторных р-белков на рРНК являются более сильными по сравнению с таковыми на мРНК. В этом случае ново-синтезируемые р-белки будут связываться с рРНК до тех пор, пока будет иметься свободная рРНК, что приводит

к сборке рибосом. Свободного р-белка, способного связываться с мРНК, не будет; в результате трансляция будет продолжаться. Однако, как только синтез рРНК замедлится или остановится, начнут накапливаться свободные р-белки. Они смогут связываться с мРНК и репрессировать дальнейшую трансляцию. Такой цикл гарантирует одинаковый ответ каждого оперона, детерминирующего синтез р-белка, на содержание рРНК.

С помощью такой регуляции оперонов р-белков достигаются две цели. Во-первых, концентрация р-белков устанавливается в соответствии с условиями роста клеток. Контролируя содержание рРНК, клетка регулирует продуцирование всех рибосомных компонентов. Во-вторых, другие белки, кодируемые этими оперонами, синтезируются со свойственной им скоростью независимо от трансляции генов р-белков. Синтез β -субъединицы РНК-полимеразы может быть подвержен своей собственной аутогенной регуляции. Как EF-Tu, так и L7/L12 транслируются с повышенной эффективностью. Следовательно, в пределах систем контроля этих оперонов существуют условия для достижения различных скоростей синтеза координированно регулируемых белков.

Аутогенный контроль трансляции — не единственный механизм, используемый для вариаций в пределах оперона. Некоторые опероны содержат дополнительные внутренние промоторы; кроме того, может использоваться аттенуация, может осуществляться регуляция на уровне процессинга мРНК.

Аутогенный контроль и сборка макромолекул

Возможно, аутогенный контроль — это общий тип контроля среди белков, участвующих в сборке макромолекулярных ансамблей. Преимущества такого контроля очевидны. Сама собираемая частица может быть непригодной как регулятор, так как она слишком большая, слишком многочисленная или имеет слишком ограниченную

локализацию. Однако потребность в синтезе ее компонентов может определяться пулом свободных субъединиц-предшественников. Если процесс сборки по какой-то причине блокирован, то субъединицы начинают накапливаться и дальнейший синтез компонентов, ставший теперь ненужным, прекращается.

Фактически первым примером аутогенной регуляции явились данные, полученные при изучении гена, детерминирующего синтез белка 32 у фага T4. Этот белок играет важную роль в процессах генетической рекомбинации, репарации и репликации ДНК, в которых его функция выражается благодаря его способности связываться с одноцепочечной ДНК. Доказательством того, что синтез белка гена р32 регулируется аутогенно, явился эффект нонсенс-мутации, ведущих к перепроизводству неактивного белка. Это означало, что в тех случаях, когда функция белка нарушена, он синтезируется в больших количествах. Этот эффект проявляется на уровне трансляции: мРНК гена 32 является стабильной и сохраняется независимо от поведения белкового продукта.

Если в инфицированной фагом клетке присутствует одноцепочечная ДНК, она будет связывать белок р32. Однако в отсутствие одноцепочечной ДНК или по крайней мере в условиях, при которых образуется избыток белка р32, белок превращает трансляцию своей собственной мРНК. Репрессия этого типа проявляет такой же кооперативный эффект, как и при связывании белка с ДНК, когда связывание одной белковой молекулы способствует более легкому связыванию другой. Следовательно, мы можем предполагать, что репрессия осуществляется путем связывания белка р32 с мРНК, что мешает инициации транскрипции. Вероятно, это происходит в протяженной одноцепочечной области вблизи сайта связывания с рибосомами.

Другая система, в которой наблюдается аутогенная регуляция, обнаружена в эукариотических клетках. Тубулин представляет собой мономер, из которого синтезируются микротрубочки – главная система микрофиламентов во всех эукариотических клетках. Образование мРНК тубулина контролируется пулом свободного тубулина. При достижении определенной концентрации в пуле дальнейшее образование тубулиновой мРНК прекращается. Снова используется тот же принцип: тубулин, удаленный из пула в процессе сборки макромолекул, не играет роли в регуляции, хотя размеры пула свободных предшественников определяют, будет ли продолжаться его синтез.

Неблагоприятные условия определяют строгий ответ

В тех случаях, когда бактерии находятся в таких жестких условиях роста, что они теряют возможность получить достаточное количество аминокислот для поддержания белкового синтеза, они включают целый ряд активностей. Такое явление получило название **строгий ответ**. Мы можем рассматривать его как механизм выживаемости в неблагоприятных условиях: бактерии экономят свои ресурсы, включая только минимум активностей до тех пор, пока условия питания не улучшатся. Тогда следует обратный эффект с включением полного набора метаболических активностей.

Строгий ответ вызывает значительное (в 10–20 раз) ослабление синтеза стабильных (рибосомных и транспортных) видов РНК. Одного этого достаточно, чтобы

уменьшить общий уровень синтеза РНК до 5–10% от прежнего уровня. Синтез лишь некоторых мРНК уменьшается; общее уменьшение синтеза мРНК достигает примерно трехкратного размера. Возрастает скорость деградации белка. Происходят многочисленные метаболические перестройки, о чем свидетельствует уменьшение синтезов нуклеотидов, углеводов, липидов и т.п.

Для инициации строгого ответа достаточно голодания по какой-либо одной аминокислоте или мутации, инактивирующей любую аминоацил-тРНК-синтетазу. Сигналом, запускающим всю серию событий в цепи, является присутствие ненагруженной тРНК в участке А рибосомы. Обычно при нормальных условиях только аминоацил-тРНК связывается с участком А с помощью фактора EF-Tu (гл. 6). Однако если нет подходящей аминоацил-тРНК, которая могла бы присоединиться в ответ на определенный кодон, способность присоединиться приобретает ненагруженная тРНК. Естественно, что это блокирует любое дальнейшее продвижение рибосомы, поэтому это явление описывается как **холостое взаимодействие**.

В тех случаях, когда клетка голодает по аминокислотам, она накапливает два необычных нуклеотида, первоначально получивших название **магические пятна I и II**. Теперь в результате изучения их структуры уже известно, что это ppGpp (гуанозинтетрафосфат с дифосфатами, присоединенными как в 5'-, так и 3'-положениях) и pppGpp (гуанозинпентафосфат с 5'-трифосфатной группой и 3'-дифосфатом). Во всех штаммах *E. coli* строгий ответ сопровождается накоплением ppGpp; что касается pppGpp, то его образование происходит не всегда. Эти нуклеотиды являются типичными низкомолекулярными эффекторами, активность которых проявляется благодаря их способности присоединяться к белку (белкам), вызывая изменение его конформации.

Компоненты, которые включаются в образование ppGpp и pppGpp, были идентифицированы с помощью мутаций, элиминирующих строгий ответ (stringent). Были выделены мутанты с **ослабленным контролем** (relaxed), у которых голодание по аминокислотам не вызывает какого-либо уменьшения в эффективности синтеза стабильной РНК или каких-либо других реакций, сопровождающих строгий ответ.

Чаще всего сайт мутаций, вызывающих ослабление контроля, локализован в гене *relA*, который кодирует белок, названный **фактором строгого контроля**. Этот фактор связывается с рибосомами, хотя его количество довольно незначительное – менее 1 молекулы на каждые 200 рибосом. В результате, по-видимому, только меньшинство рибосом способно дать строгий ответ.

Рибосомы, выделенные из бактерий, проявляющих строгий ответ, способны синтезировать ppGpp и pppGpp *in vitro* при условии, что участок А занят ненагруженной тРНК, *специфически соответствующей кодону*. Рибосомы, выделенные из мутантов с ослабленным контролем, не могут выполнять эту реакцию; однако они приобретают эту способность, если добавлен фактор строгого контроля.

На **рис. 15.17** показаны пути биосинтеза необычных гуаниновых нуклеотидов. Фактор строгого контроля – это фермент, катализирующий реакцию синтеза, в которой АТФ используется в качестве источника пирофосфатной группы, присоединяемой в 3'-положение либо 5'-GTP, либо GDP. Первый используется в качестве субстрата более часто. Однако pppGpp может быть превращен в ppGpp под действием ряда ферментов. Среди них способными

осуществлять такую реакцию дефосфорилирования являются факторы трансляции EF-Tu и EF-G. Образование $ppGpp$ из $pppGpp$ представляет собой наиболее распространенный путь.

Что представляет собой реакция холостого взаимодействия? Оказалось, что мутации в другом локусе, способном вызывать ослабленный тип контроля, лежат в пределах гена белка L11, входящего в состав 50S-субчастицы. Он локализован в непосредственной близости от участков A и P. Строгий ответ может включать TΨC-область тРНК, поскольку мутации в тРНК, особенно те, которые затрагивают эту область, способны влиять на реакцию. Не исключено, что эта область тРНК (локализованная в изгибе L-подобной третичной структуры) включается в передачу сигнала о том, что с кодоном спарилась правильная тРНК, позволяя, таким образом, рибосоме приступить к перемещению пептида. Тот же путь мог бы являться частью цепи событий, ведущих к холостому взаимодействию. Присутствие правильно спаренной ненагруженной тРНК в участке A может вызвать конформационное изменение в рибосоме; однако, поскольку тРНК не несет аминокислоты, имеет место холостое взаимодействие вместо транслокации полипептида от пептидил-тРНК.

Какова роль соединения $ppGpp$? Оно может служить эффектором, используемым для контроля отдельных реакций, в том числе для подавления транскрипции. Сообщалось о многих противоречивых эффектах, среди которых выделяются два. Существует специфическое подавление *иницирования* транскрипции в промоторах оперонов, кодирующих рРНК. Кроме того, добавление $ppGpp$ замедляет фазу элонгации транскрипции многих или большинства матриц *in vitro*. Объясняется это увеличением остановок фермента. Данный эффект обуславливает общее уменьшение эффективности транскрипции при добавлении $ppGpp$ *in vitro*. Мы еще не знаем специфичности такого ингибирования, но будет не удивительно, если обнаружатся вариации в степени его проявления от оперона к оперону, так что одни опероны будут подавляться сильнее других.

Интересно, что необычные нуклеотиды используются по крайней мере в двух системах контроля с координирующими функциями общего характера. Оба, по-видимому, являются специфичными для бактерий. Хотя в клетках эукариот циклический АМР является классическим

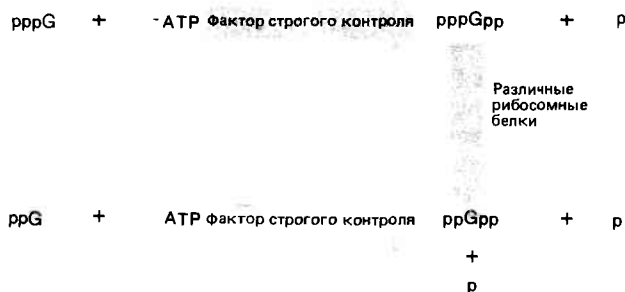


Рис. 15.17. Фактор строгого контроля катализирует синтез $pppGpp$ и $ppGpp$; рибосомные белки способны дефосфорилировать нуклеотид $pppGpp$, превращая его в $ppGpp$.

вторым посредником, не имеется никаких данных о том, что он играет какую-либо роль в специфическом контроле иницирования транскрипции. Интенсивные исследования, связанные с поиском необычного нуклеотида $ppGpp$ у эукариот, окончились неудачей.

Рекомендуемая литература

Функционирование отдельных оперонов рассматривается в книге Миллера и Резникова [Miller, Reznikoff (Eds), *The Operon*, Cold Spring Harbor Laboratory, New York, 1978]. Она включает обзоры: Platt — о *trp*-опероне (стр. 263–302), Lee — об *ara*-опероне (стр. 389–410), De Crombrughe, Pastan — о циклическом АМР и *gal*-опероне (стр. 303–324). Системы аттенуации *trp*-оперона рассматриваются Яновским (Yanofsky, *Nature*, **289**, 751–758, 1981) и Платтом (Platt, *Cell*, **24**, 10–23, 1981). Данные о функциях этого оперона были суммированы с точки зрения его последовательности Яновским и др. (Yanofsky et al., *Nuc. Acids Res.*, **9**, 6647–6668, 1981). Данные о способе контроля в *ara*-опероне могут быть найдены в статьях Огдена и др. (Ogden et al., *Proc. Nat. Acad. Sci.*, **77**, 3346–3350, 1980) и Ли и др. (Lee et al., *Proc. Nat. Acad. Sci.*, **78**, 752–756, 1981). Модель аутогенного контроля синтеза р-белка представлена Номуры и др. (Nomura et al., *Proc. Nat. Acad. Sci.*, **77**, 7084–7088, 1980). Данные о строгом контроле наиболее полно представлены в обзоре Галланта (Gallant, *Ann. Rev. Genet.*, **13**, 393–415, 1979).

Глава 16

ЛИТИЧЕСКИЙ КАСКАД И ЛИЗОГЕННАЯ РЕПРЕССИЯ

Некоторые фаги используют для выживания только одну линию поведения. Инфицируя чувствительную бактерию, они нарушают ее функции, с тем чтобы обеспечить образование большого числа фаговых частиц-потомков. В результате такой **литической инфекции** бактерия-хозяин погибает. В процессе типичного литического цикла фаговая ДНК (или РНК) проникает в клетку бактерии-хозяина, ее гены транскрибируются в установленном

порядке, генетический материал фага реплицируется, и в результате образуются белковые компоненты фаговых частиц. В конечном счете бактериальная клетка разрушается (**лизировается**), освобождая зрелые частицы фагового потомства.

Однако есть и такие фаги, для которых характерны две формы существования. Они способны воспроизводить себя путем такого же литического цикла, с помощью

которого обеспечивается быстрое образование большого числа копий фаговых частиц. Однако развитие этих фагов может пойти и по другому пути. В этом случае фаговый геном находится в бактериальной клетке в латентной форме, называемой **профагом**. Такое состояние получило название **лизогения**. В лизогенных бактериях профаг **интегрирован** с бактериальным геномом и наследуется, как любая другая его часть, тем же способом, как и бактериальные гены. Благодаря наличию профага лизогенная бактерия обладает **иммунитетом** против инфицирования другими фаговыми частицами того же типа. Поэтому в бактериальном геноме обычно содержится только одна копия профага любого определенного типа.

Переходы между лизогенным и литическим способами существования могут происходить в любом направлении. Если фаг, образованный в литическом цикле, проникает в клетку новой бактерии-хозяина, он может либо повторить литический цикл, либо перейти в лизогенное состояние. Выбор зависит от условий инфекции и генотипов фага и бактерии. Профаг может быть выведен из лизогенного состояния с помощью процесса, названного **индукцией**. В этом случае он **исключается** из бактериального генома и образует свободную фаговую ДНК, которая затем проходит через литический путь развития.

Другой тип существования в бактериальной клетке представлен **плазмидами**. Это автономные элементы, геномы которых существуют в клетке как **внехромосомные** единицы. Плазмиды—это самореплицирующиеся кольцевые молекулы ДНК, которые сохраняются в клетке стабильно и в характерном для каждого типа плазмид числе копий; иными словами, число их остается постоянным из поколения в поколение.

Некоторые из этих элементов также имеют альтернативные способы существования. Они могут находиться либо в автономном внехромосомном состоянии, либо могут быть интегрированы с бактериальной хромосомой и находиться в ней как ее часть, подобно любой другой последовательности. Такие элементы получили название **эписомы**. (Следует отметить, что с тех пор, как было дано первоначальное определение, термины «плазида» и «эписома» иногда свободно используются как взаимозаменяемые. Поэтому они не всегда точно соответствуют описанному здесь значению.)

Подобно лизогенизирующим фагам, плазмиды и эписомы эгоистически сохраняются в клетках своих бактерий-хозяев и часто препятствуют выживанию в них другого элемента того же типа. Такое явление названо **иммунитетом**. (Природа плазмидного иммунитета отличается от лизогенного иммунитета. См. гл. 31.)

Литический цикл состоит из отдельных стадий

Целью литического развития является образование возможно большего числа фаговых частиц потомства при инфицировании бактериальной клетки-хозяина. Инфекция начинается с прикрепления к бактерии одной фаговой частицы. **Икосаэдральные** фаги имеют **головку**, содержащую ДНК, и **хвостовой отросток**, с помощью которого фаг прикрепляется к бактериальной клетке. ДНК впрыскивается в бактериальную клетку, частица остается вне клетки. **Нитевидные** фаги имеют стержнеподобную структуру; их взаимодействие с клеткой бактерии изучено менее полно.

Фаговые геномы малы. В самом деле, как в случае всех вирусов, принципиальным ограничением является необходимость упаковки нуклеиновой кислоты внутри его белковой оболочки. Это обуславливает многое в стратегии вирусной репродукции. Обычно они используют аппарат клетки-хозяина, который вместо того, чтобы реплицировать и выражать бактериальные гены, реплицирует и выражает фаговые гены.

Обычно фаговые гены детерминируют функции, обеспечивающие предпочтительную репликацию фаговой ДНК. Эти функции могут быть связаны с иницированием репликации или даже с обеспечением клетки новой ДНК-полимеразой. Некоторые изменения всегда вносятся в способность клетки-хозяина осуществлять транскрипцию. Они могут выражаться в замещении РНК-полимеразы или модификации ее способности иницировать или терминировать процесс транскрипции. Результат всегда один и тот же: транскрибируется предпочтительно фаговая мРНК. Что касается белкового синтеза, обычно фаг удовлетворяется использованием аппарата клетки-хозяина, изменяя его активность в основном путем замещения бактериальной мРНК на фаговую мРНК.

Литическое развитие—это процесс, в котором фаговые функции выражаются в определенном порядке. Благодаря этому в соответствующий момент имеется нужное количество каждого компонента. Цикл литического развития можно разделить на две основные части, показанные на **рис. 16.1**. **Ранняя инфекция** включает период от момента проникновения ДНК в клетку до начала ее репликации. **Поздняя инфекция** включает период от начала репликации до конечной стадии лизиса бактериальной клетки с освобождением потомства фаговых частиц. При обычном порядке процесса ранняя фаза связана с образованием ферментов, участвующих в репродукции ДНК. Она включает в себя ферменты, связанные с синтезом ДНК, рекомбинацией и иногда модификацией. Благодаря этим ферментам в клетке накапливается **пул** фаговых геномов. В этом пуле геномы беспрерывно реплицируются и рекомбинируют таким образом, что события, происходящие даже в процессе одного литического цикла, имеют отношение к целой популяции фаговых геномов.

При наступлении поздней фазы синтезируются белковые компоненты фаговых частиц. Для образования головки и различных структур отростка требуется большое количество белков. Поэтому большая часть фагового генома кодирует поздние функции. Наряду с этим существуют также белки сборки, присутствие которых необходимо для конструирования частицы. Однако сами по себе эти белки не входят в состав фаговой частицы. К тому времени, когда белковые компоненты собираются в головки и отростки, репликация достигает своей максимальной скорости. Затем геномы внедряются в пустые белковые головки, к ним присоединяются отростки, и клетка лизируется.

Литическое развитие подвержено каскадной регуляции

Многие геномы фагов организованы таким образом, что их генетическая карта точно отражает последовательность литического развития. Принцип оперонной организации достигается при этом своего крайнего выражения, при котором гены, кодирующие белки с родственными функциями, сгруппированы для осуществления контроля

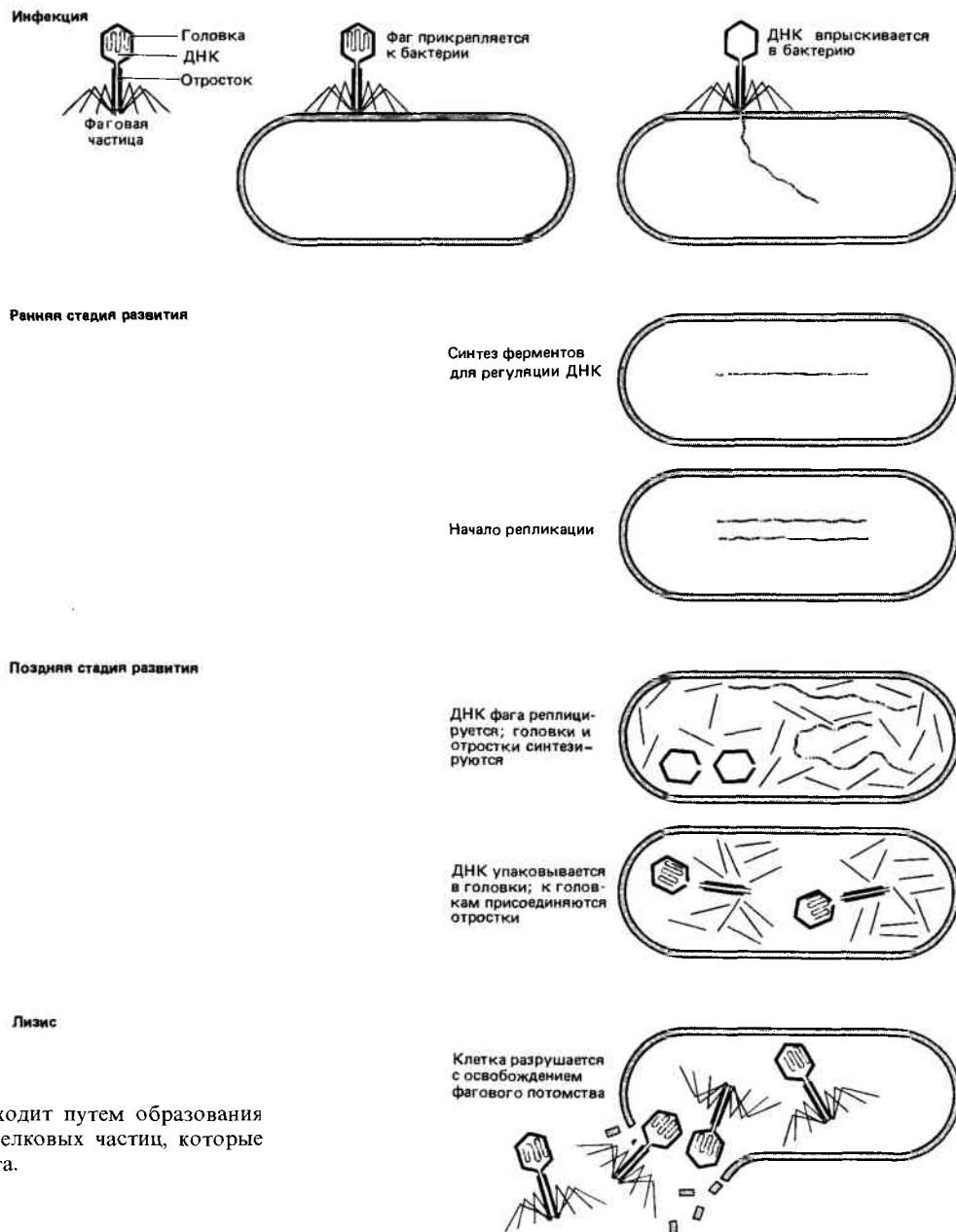


Рис. 16.1. Литическое развитие происходит путем образования большого числа фаговых геномов и белковых частиц, которые объединяются, образуя потомство фага.

с максимальной экономией. Это позволяет контролировать путь литического развития с помощью небольшого числа регуляторных переключателей.

Чтобы выражение фаговых генов происходило в определенном порядке, литический цикл находится под позитивным контролем. Каждая группа фаговых генов может выражаться лишь после того, как будет дан соответствующий сигнал.

Обычно с помощью аппарата клетки-хозяина может быть выражено лишь несколько фаговых генов. Их промоторы неотличимы от промоторов бактериальных генов. Название этого класса генов зависит от фага. В большинстве случаев их называют **ранними генами**. У фага лямбда они получили название **предранных генов**. Независимо от названия они контролируют только вступительную, представляющую самую начальную часть раннего периода. Иногда они связаны исключительно с переходом в следующий период. Во всяком случае, один из этих генов всегда кодирует белок, необходимый для транскрипции нового класса генов.

Второй класс генов называют либо **задержанно ранней**, либо **средней** группой. Как вытекает из самого названия, выражение генов этой группы начинается в раннем периоде, обычно как только появляется регуляторный белок. В зависимости от природы контролирующей системы первоначальный ряд генов может либо продолжать свое выражение, либо выключаться на этой стадии (см. рис. 13.5). Часто выражение генов клетки-хозяина бывает ослаблено. Обе группы ранних генов совместно отвечают за все фаговые функции, кроме тех, которые необходимы для сборки белковой оболочки частицы и лизиса клетки.

После начала репликации фаговой ДНК, осуществляемой в результате выражения ранних генов, наступает очередь для выражения **поздних генов**. Их транскрипция на этой стадии обеспечивается наличием другого регуляторного гена, входящего в состав предыдущего (задержанно раннего или среднего) набора генов. Таким регулятором может быть другой антитерминатор (как в случае фага лямбда) или другой сигма-фактор (как у фага SP01).

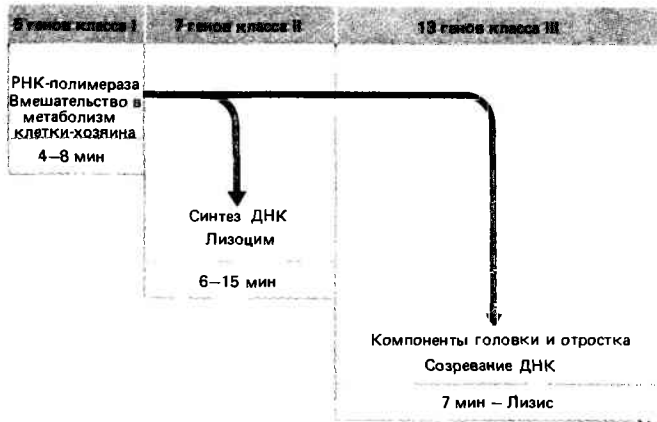


Рис. 16.2. Геном фага Т7 содержит три класса генов, выражающихся последовательно. Геном состоит примерно из 38 000 пар оснований.

Использование такого последовательного контроля с регуляторным геном в каждом ряду, необходимым для выражения следующего ряда генов, создает каскад, в котором группы генов включаются (или иногда выключаются) в определенное время. Каскадная регуляция у различных фагов может различаться в деталях, однако во всех случаях она приводит к одинаковому результату, как это будет видно из следующих разделов.

Образование кластеров генов с родственными функциями у фагов Т7 и Т4

В геноме фага Т7 можно выделить гены трех классов, каждый из которых включает в себя группу сцепленных локусов. Как показано на рис. 16.2, класс I представлен предранними генами, выражающимися с помощью РНК-полимеразы клетки-хозяина сразу же после внедрения в клетку фаговой ДНК. Среди продуктов этих генов имеются ферменты, препятствующие выражению генов клетки-хозяина (а следовательно, и своему собственному выражению), а также фаговая РНК-полимераза. Фаговый фермент ответствен за выражение генов класса II (связанных в основном с функциями синтеза ДНК) и генов класса III (связанных в основном со сборкой зрелых фаговых частиц). Чем обусловлено различие в выражении генов классов II и III, пока еще не разгадано.

Геном фага Т4 – один из наиболее крупных; для его организации характерно обширное группирование генов с родственными функциями. Генетическая карта фага Т4 приведена на рис. 16.3. Гены, которые на рисунке пронумерованы, имеют **жизненно важное значение**: мутация в любом из них предотвращает успешное завершение литического цикла. Гены, обозначенные трехбуквенными сокращениями, являются **несущественными**, по крайней мере при обычных условиях инфекции. Чем объясняется наличие в геноме такого большого количества несущественных генов, объяснить трудно. Но можно предположить, что они дают фагу преимущество при отборе. (В геноме большинства фагов подавляющая часть генов или даже все гены являются существенными.)

В экспрессии генов можно выделить три фазы; сводные данные о функциях генов, экспрессирующихся

в каждой фазе, приведены на рис. 16.4. Ранние гены транскрибируются РНК-полимеразой клетки-хозяина. Несколько позже транскрибируются квазипоздние гены, однако мы еще не понимаем, как регулируется экспрессия этих генов (их название отражает способ выражения, который не соответствует категориям, описанным для других фагов). Вместе взятые, ранние и квазипоздние гены обеспечивают практически все фаговые функции, связанные с синтезом ДНК, модификацией клеточной структуры или транскрипцией и трансляцией фаговых генов.

Два существенных гена из категории «транскрипции» выполняют регуляторную функцию: их продукты необходимы для экспрессии поздних генов. У фага Т4 существует прямая связь между репликацией и экспрессией поздних генов. Только активно реплицирующаяся ДНК может быть использована в качестве матрицы для транскрипции поздних генов. Это объясняется следующими причинами: бактериальная РНК-полимераза модифицируется таким путем, что ей становятся необходимы некоторые свойства, присущие только реплицирующейся ДНК (вероятно, наличие разрывов). Это обеспечивает образование белковых компонентов фага (которых очень много) в количествах, соответствующих пулу наличной ДНК.

О том, как фаг лямбда осуществляет свой литический каскад

Одна из наиболее сложных каскадных систем принадлежит фагу лямбда. Сам по себе каскад при литическом развитии является простым; в нем используются два регулятора, контролирующие последовательные стадии развития. Однако цикл литического развития смыкается с циклом установления лизогении, как это суммировано на рис. 16.5. При введении ДНК фага лямбда в новую клетку-хозяина литический и лизогенный циклы запускаются одним и тем же способом. Оба цикла требуют экспрессии предранних и задержанно ранних генов. Однако на следующем этапе эти циклы дивергируют. Литическое развитие происходит, если экспрессируются поздние гены; лизогенное состояние устанавливается, если синтезируется необходимый репрессор.

Фаг лямбда имеет только два предранних гена, транскрибируемых независимо РНК-полимеразой клетки-хозяина. Один из них, *cro*, выполняет двойную функцию: 1) предотвращает синтез лизогенного репрессора (условие, необходимое для осуществления литического цикла) и 2) выключает экспрессию предранних генов (ненужных на более поздних этапах литического цикла). Другой предранний ген – *N*, кодирующий, как мы видели в гл. 13, фактор антитерминации, благодаря которому транскрипция переходит в область, содержащую задержанно ранние гены.

Задержанно ранние гены включают в себя два гена репликации (необходимых для литической инфекции) и семь генов рекомбинации (некоторые из них отвечают за рекомбинацию при литической инфекции); два гена необходимы для интеграции ДНК фага лямбда с бактериальной хромосомой при лизогенизации). Функции двух генов-регуляторов, *cII/cIII*, необходимы для инициации синтеза лизогенного репрессора. Регуляторный ген *Q* кодирует фактор антитерминации, благодаря которому бактериальная РНК-полимераза получает возможность приступить к транскрипции поздних генов. Таким образом, задержанно ранние гены служат для двух целей: одни из них необходимы для установления фагом лизогенного со-



Рис. 16.3. Карта генома T4 является кольцевой. В геноме протяженностью 165 000 пар оснований наблюдается обширное группирование генов с родственными функциями.

стояния, другие обеспечивают осуществление литического цикла в определенном порядке.

Чтобы разобраться в особенностях двух путей, сначала мы рассмотрим литический цикл развития фага. На рис. 16.6 приведена карта ДНК фага лямбда. Группа генов, связанных с регуляцией, окружена генами, необходимыми для рекомбинации и репликации. В числе генов регуляторной группы находятся предранние гены *N* и *cro*. Они транскрибируются с разных цепей ДНК: *N* – по направлению влево, *cro* – вправо. В присутствии фактора антитерминации транскрипция продолжается влево от гена *N* в область генов рекомбинации и вправо от гена *cro* в область генов репликации (см. также рис. 13.7 и 13.8). Это проиллюстрировано на рис. 16.7, где приведено истинное состояние ДНК фага лямбда во время инфекции.

На карте, представленной на рис. 16.6, показана организация ДНК фага лямбда в фаговой частице. Однако вскоре после инфицирования концы ДНК соединяются, образуя кольцевую структуру. При этом поздние гены объединяются в одну группу, содержащую гены лизиса

S—R из правого конца линейной ДНК и гены *A—J* головки и отростка из левого конца (рис. 16.7).

Поздние гены выражаются как одна единица транскрипции, начинающаяся с промотора P_R' , который лежит между генами *Q* и *S*. В отсутствие продукта гена *Q* (являющегося последним геном в расположенной в правом конце группы задержанно ранних генов) эта транскрипция осуществляется конститутивно, но терминируется в сайте t_{R3} , расположенном рядом с промотором P_R' ; в результате транскрипции образуется 6S-РНК, содержащая 194 основания. (Вероятно, некоторые из РНК-полимераз, транскрибирующих задержанно ранние гены, способны переместиться в область поздних генов, но их число незначительно, и они не способны образовать нужное количество компонентов головки и отростка для поддержания литического развития.) Однако, если продукта гена *Q* становится достаточно, он супрессирует терминацию в сайте t_{R3} и 6S-РНК удлиняется; в результате выражение поздних генов резко усиливается.

Транскрипция поздних генов, по-видимому, не заканчивается в какой-то специфической точке, а продолжается

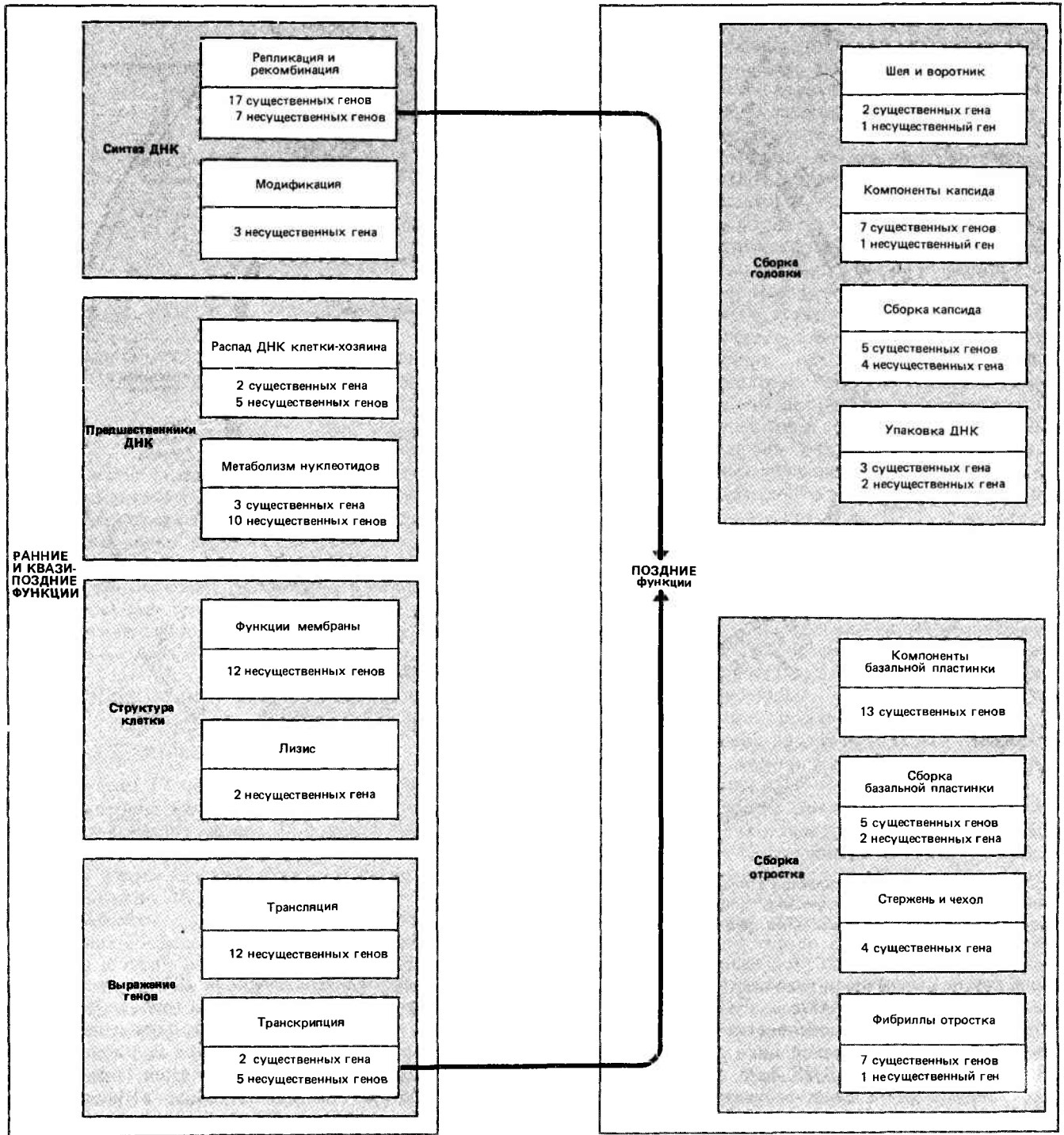


Рис. 16.4. Литический каскад фага T4 распадается на части: ранние и квазипоздние функции связаны с синтезом ДНК и вы-

ражением генов; поздние функции связаны со сборкой фаговых частей.

через всю область поздних генов, переходя в область, расположенную за ними. Подобное же событие происходит при транскрипции задержанноранних генов, идущей влево, которая продолжается после генов с рекомбинационными функциями. События транскрипции, идущие в каждом направлении, вероятно, заканчиваются раньше, чем может произойти столкновение полимераз.

Лизогения поддерживается благодаря аутогенному циклу

При рассмотрении литического каскада фага лямбда мы могли видеть, что вся программа запускается в действие путем инициирования транскрипции предранних генов *N* и *cro* в двух промоторах – P_L и P_R . Поскольку фаг

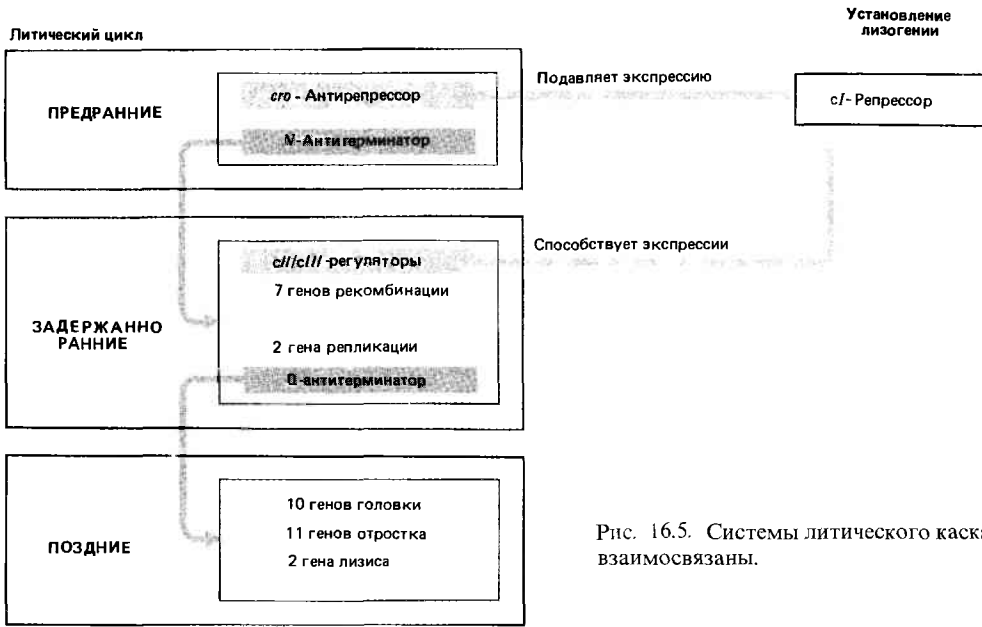


Рис. 16.5. Системы литического каскада и лизогении фага лямбда взаимосвязаны.

лямбда использует антитерминацию для перехода в следующую стадию (когда выражаются задержанно ранние гены), те же два промотора продолжают использоваться в течение всего раннего периода.

На рис. 16.6 приведена увеличенная карта регуляторной области, на которой показано, что промоторы P_L и P_R лежат по обе стороны от гена cI . С каждым промотором ассоциирован оператор (O_L , O_R), с которым связывается белок-репрессор для предотвращения инициирования транскрипции РНК-полимеразой. Последовательность каждого оператора перекрывается с последовательностью контролируемого промотора. Поэтому часто они описываются как контролирующие области P_L/O_L и P_R/O_R .

Поскольку литический каскад по своей природе включает несколько последовательных стадий, области контроля содержат «точки давления», в которых регулируется вступление в полный цикл. Затрудняя доступ РНК-полимеразы к промоторам, репрессор предотвращает инициацию литического цикла фаговым геномом. Тот факт, что операторы являются мишенями для действия репрессора,

был обнаружен с помощью **вирулентных** мутаций, изменяющих последовательность ДНК, которую узнает репрессор. Такие мутации препятствуют связыванию репрессора с O_L или O_R , что неминуемо приводит к переходу фага на литический путь развития. Вирулентные мутации у фагов эквивалентны оператор-конститутивным мутациям бактериальных оперонов.

Репрессорный белок кодируется геном cI . Мутанты по этому гену не способны поддерживать состояние лизогении и обречены всегда вступать в литический цикл. Это послужило основанием для того, чтобы ген получил название «с-гена» (от англ. *clear plaque*), отражающее фенотип образующихся при инфекции прозрачных пятен. При инфицировании бактериальной культуры фагом клетки лизируются, образуя области, которые при посеве культуры на чашку выглядят как **бляшки (стерильные пятна)**. При использовании фага дикого типа небольшая зона таких пятен является мутной, поскольку она содержит клетки, которые не лизировались, так как стали лизогенными. Мутации cI^- предотвращают лизогенизацию; в результате бляшки содержат только лизированные клетки и поэтому оказываются прозрачными.

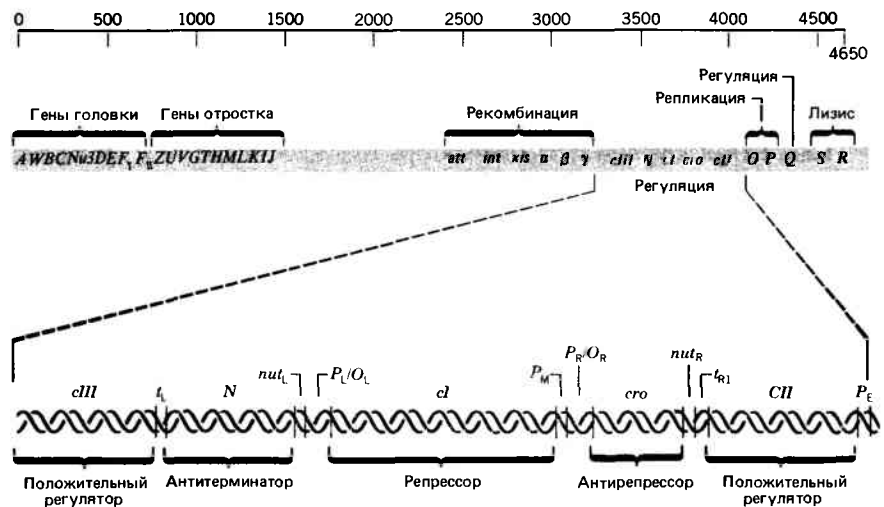


Рис. 16.6. Карта генома фага лямбда. Видно, что гены с родственными функциями сгруппированы в кластеры. Геном содержит около 46 500 пар оснований.

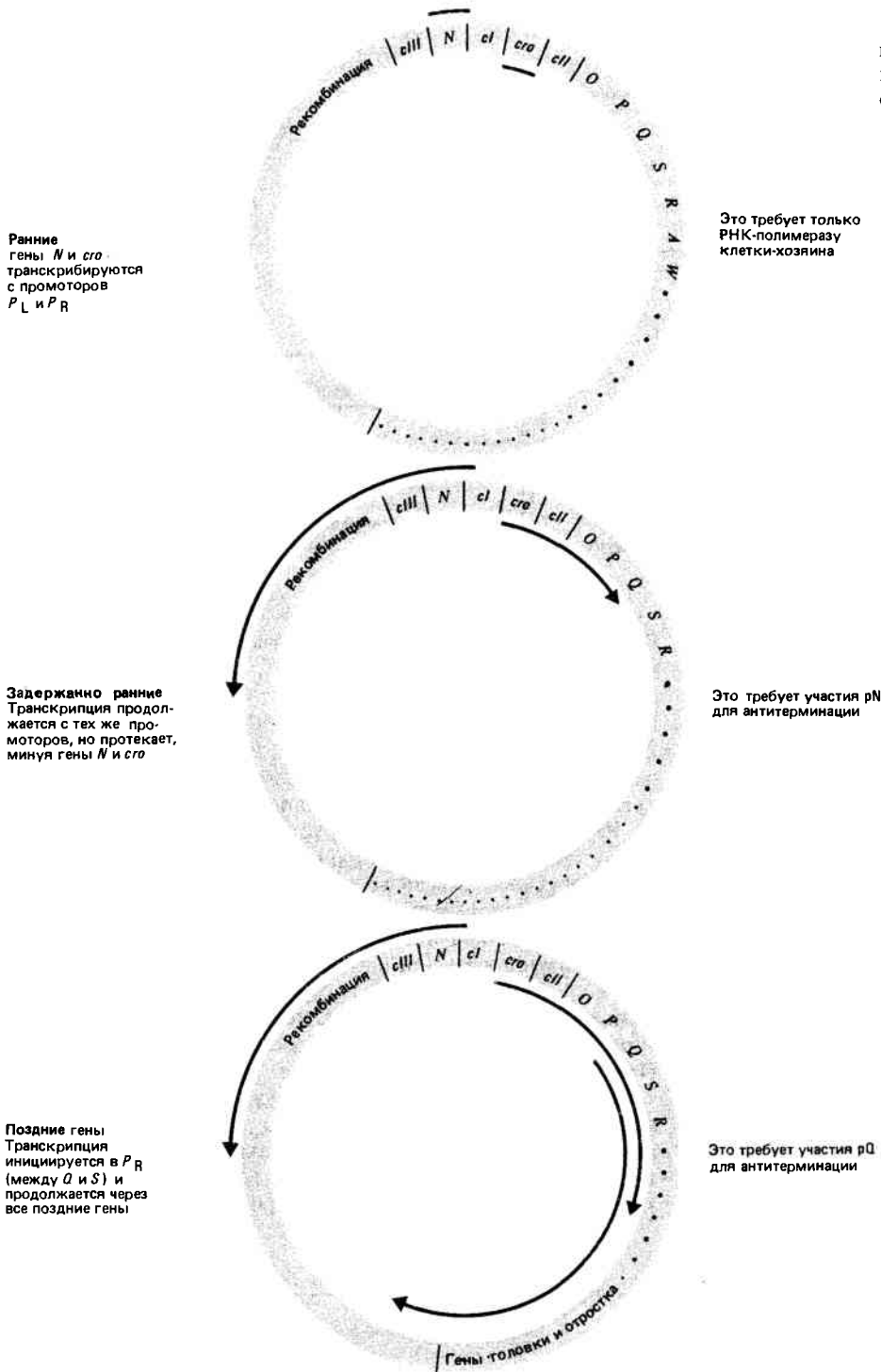


Рис. 16.7. ДНК фага лямбда в инфицированной клетке превращается в кольцевую; в результате группа поздних генов образует одну непрерывную единицу транскрипции.

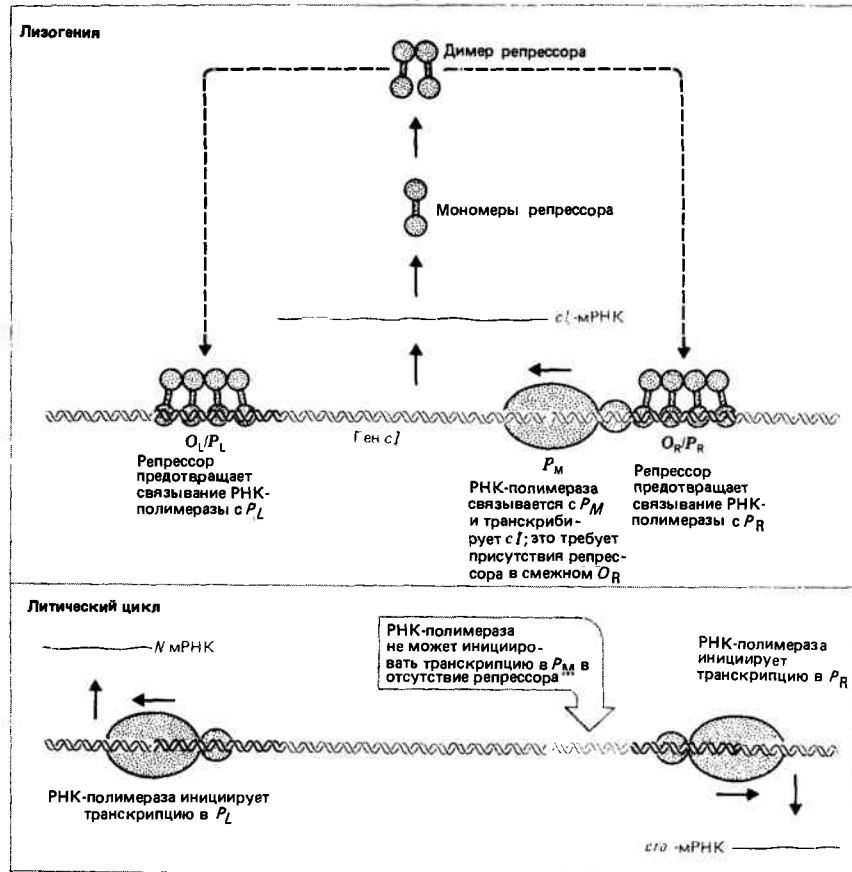
Ген *ci* транскрибируется с промотора P_M , иногда также называемого P_{RM} , который расположен в его правом конце (см. рис. 16.6). (Буква «М» от англ. maintenance – сохранение репрессора). Транскрипция терминируется в левом конце гена. мРНК начинается с кодона AUG, который используется для иницирования трансляции; из-за отсутствия обычного сайта связывания с рибосомой мРНК транслируется до некоторой степени неэффективно (гл. 9), образуя лишь низкое количество репрессорного белка.

Репрессор связывается независимо с двумя операторами. Он проявляет одну функцию, находясь в O_L , но две функции – находясь в O_R (рис. 16.8).

В O_L репрессор проявляет эффект, аналогичный тому, который мы уже обсуждали в случае других систем: он предотвращает иницирование транскрипции в P_L ферментом РНК-полимеразой. В результате выражение гена *N* останавливается. Поскольку промотор P_L используется для всей левосторонней транскрипции ранних генов, это действие препятствует выражению всей левосторонней единицы транскрипции. Таким образом литический цикл исключается, оказывается «загнанным в тупик» раньше, чем он может начаться.

Связывание репрессора с оператором O_R препятствует использованию промотора P_R . В результате экспрессия гена *cro* и других правосторонних ранних генов оказы-

Рис. 16.8. Лизогения сохраняется благодаря аутогенному циклу (вверху); при нарушении этого цикла начинается литический цикл (внизу).



вается невозможной. (Позднее мы увидим, почему предотвращение экспрессии гена *cro* имеет такое важное значение для поддержания лизогенного состояния.)

Однако присутствие репрессора в операторе O_R имеет также другой эффект. Промотор для синтеза репрессора, P_M , сцеплен с правосторонним оператором O_R . Оказывается, что РНК-полимераза способна иницировать транскрипцию в промоторе P_M только после того, как репрессор присоединился к оператору O_R . В этом случае репрессор действует как положительный регуляторный белок, необходимый для транскрипции гена *cI*. Поскольку он на самом деле является продуктом этого гена, то тут действует аутогенный цикл, в котором присутствие репрессора необходимо для поддержания его собственно продолжающегося синтеза.

Природа такого контролирующего цикла объясняет биологические особенности лизогенного состояния. Лизогения стабильна, так как контролирующая система обеспечивает непрерывное выражение гена *cI*, пока концентрация репрессора остается достаточной. Это приводит к тому, что операторы O_L и O_R остаются постоянно занятыми. В результате репрессии всего литического каскада это ведет к сохранению профага в его инертной форме.

Присутствие репрессора объясняет явление **иммунитета**. Если в лизогенную клетку проникает вторая молекула ДНК фага лямбда, белок-репрессор, синтезируемый геномом ранее интегрированного фага, будет немедленно связываться с операторами O_L и O_R нового генома. Это препятствует вступлению второго фага в литический цикл.

Индукция профага к переходу в литический цикл происходит в том случае, если лизогенная система нарушена.

Это случается при инактивации репрессора, присутствующего в клетке. Аутогенная природа цикла предопределяет высокую чувствительность ответа. Во-первых, отсутствие репрессора позволяет РНК-полимеразе связаться с промоторами P_L и P_R и начать литический цикл, как это показано на рис. 16.8. Во-вторых, так как присутствие репрессора необходимо для его собственного синтеза, выражение гена *cI* останавливается, как только существующий репрессор разрушается. Это препятствует дальнейшему синтезу репрессора, способного заместить поврежденные молекулы. В результате может начаться литический цикл без каких-либо вмешательств со стороны системы, поддерживающей лизогению.

Область, включающая левый и правый операторы, ген *cI*, кодирующий репрессор, и ген *cro*, кодирующий другой регуляторный белок, определяет иммунитет фага. Это значит, что любой фаг, обладающий такой областью, проявляет тот же тип иммунитета, так как эта область определяет репрессорный белок и сайты, на которые действует репрессор. Поэтому она получила название **области иммунитета**. Если мы говорим, что лизогенизирующий фаг вызывает иммунитет к любому другому фагу того же типа, мы имеем в виду, что иммунитет проявляется к любому другому фагу, имеющему ту же область иммунитета (безотносительно от различий в других областях).

Репрессор – димер с различными доменами

Репрессор фага лямбда функционирует как мультимерный белок с различными доменами. Субъединица репрессора представляет собой полипептид с мол. массой

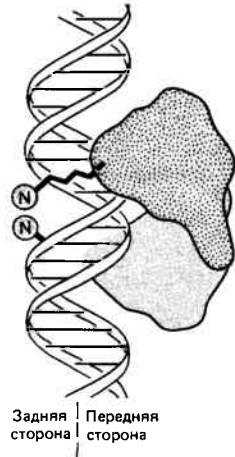


Рис. 16.9. Большая часть контактов репрессора устанавливается на одной стороне ДНК, а его N-концы контактируют с другой стороной.

27 000 дальтон, которая димеризуется в растворе. N-концевой домен, простирающийся от 1-го до 92-го аминокислотного остатка, соответствует сайту связывания с оператором. С-концевой домен протяженностью от 132-го до 236-го аминокислотного остатка ответствен за образование димеров. Два этих домена соединены коннектором из 40 аминокислотных остатков. При расщеплении выделенного белка протеазой оба домена освобождаются в виде фрагментов, относительно устойчивых к дальнейшей деградации.

Каждый домен способен выполнять свою функцию независимо от другого. С-концевой протеолитический фрагмент способен образовывать олигомеры. N-концевой фрагмент способен связаться с операторами, хотя и с более низкой эффективностью, чем интактный репрессор. Таким образом, информация для специфически контактирующей ДНК содержится в пределах N-концевого домена, но (так же как в случае с *lac*-репрессором) эффективность процесса повышается при присоединении С-концевого домена. Вероятно, это происходит из-за способности интактного репрессора димеризоваться, что обеспечивает одновременное связывание с ДНК его двух N-концевых доменов. Выделенная N-концевая область осуществляет те же контакты с ДНК, как и интактный репрессор. Однако при удалении последних трех N-концевых аминокислот способность к некоторым из контактов утрачивается. На основе этого наблюдения была предложена модель, показанная на рис. 16.9, согласно которой основная часть N-концевой области осуществляет контакты на одной стороне ДНК, тогда как три последние N-концевые аминокислоты образуют контакты на противоположной стороне ДНК.

Димерная структура репрессора имеет критическое значение для сохранения лизогенного состояния. Индукция лизогенного профага к вступлению в литический цикл вызывается разрезанием субъединицы репрессора в области коннектора, между аминокислотными остатками 111 и 112. (В известном смысле это соответствует аллостерическому изменению в конформации, возникающему при инактивации репрессора низкомолекулярным индуктором бактериального оперона, — способности, которой не обладает лизогенный репрессор.)

При расщеплении димерные С-концевые домены отделяются от N-концевых доменов, которые связаны с оператором. Как показано на рис. 16.10, N-концевые области после этого не обладают уже достаточным сродством, чтобы оставаться связанными с оператором. Они отделяются от ДНК, открывая тем самым путь для литической инфекции.

Таким образом, балансирование между лизогенией и литическим циклом зависит от концентрации репрессора. Интактный репрессор находится в клетке в концентрации, достаточной для обеспечения занятости операторов. Но, если репрессор расщепляется, его концентрация становится малой из-за низкого сродства отделенных N-концевых доменов к оператору. Слишком высокая концентрация репрессора делала бы невозможной индукцию литического цикла таким способом; слишком низкий уровень, естественно, делает невозможным сохранение лизогенного состояния.

Репрессор связывается кооперативно в каждом операторе

Каждый промотор содержит три сайта связывания с репрессором, как это можно видеть на рис. 16.11. Каждый сайт связывания представлен последовательностью из 17 пар оснований, проявляющих частичную симметрию относительно оси, проходящей через центральную пару оснований. Никакие два из шести отдельных сайтов связывания с репрессором не идентичны по последовательностям. Однако все они соответствуют канонической последовательности. Связывающие сайты внутри каждого оператора отделены спейсерами из 3–7 пар оснований, которые богаты А–Т-парами оснований. Сайты в каждом операторе пронумерованы, так что O_R

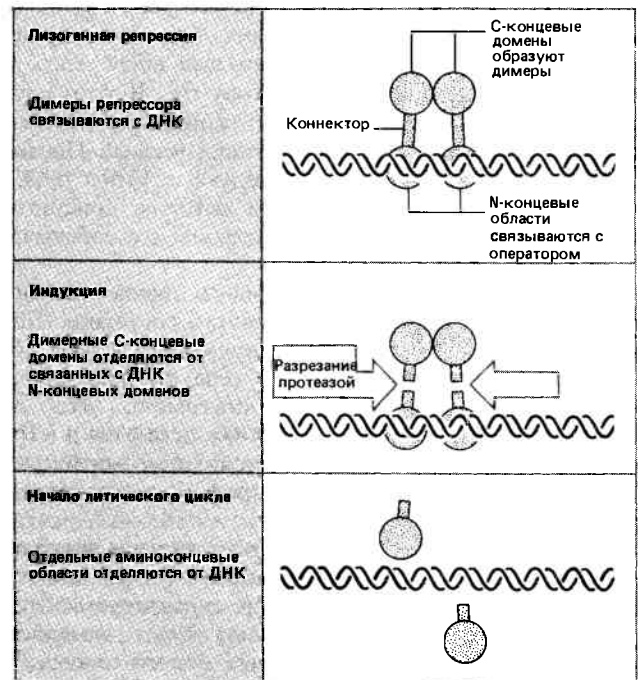


Рис. 16.10. Репрессор образует димеры, связывающиеся с оператором. Сродство N-концевых доменов репрессора с ДНК контролируется димеризацией С-концевых доменов.

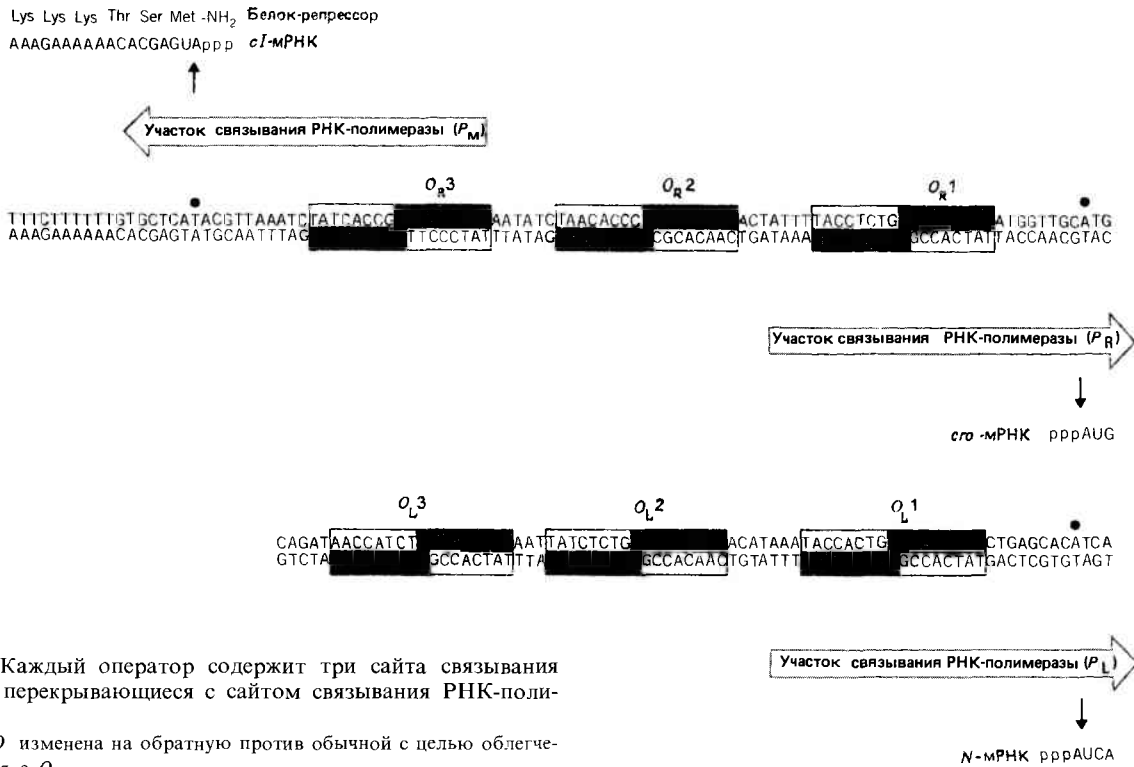


Рис. 16.11. Каждый оператор содержит три сайта связывания репрессора, перекрывающиеся с сайтом связывания РНК-полимеразы. Ориентация O изменена на обратную против обычной с целью облегчения сравнения с O_R .

содержит серию связывающих сайтов $O_{R1}-O_{R2}-O_{R3}$, а O_L содержит серию $O_{L1}-O_{L2}-O_{L3}$. В каждом случае сайт 1 лежит ближе к стартовой точке транскрипции в промоторе, чем сайты 2 и 3.

Каким образом репрессор определяет, в каком сайте начать связывание, сталкиваясь с такой трипликацией связывающих сайтов в каждом операторе? В каждом операторе сайт 1 имеет более высокое сродство (примерно в 10 раз) к репрессору по сравнению с другими сайтами. В результате репрессор всегда связывается сначала с O_{L1} или O_{R1} . Так, при низких концентрациях репрессора *in vitro* от действия нуклеаз в каждом операторе защищается фрагмент длиной около 25 пар оснований. Фрагмент соответствует 17 парам оснований сайта 1 плюс нескольким соседним нуклеотидам. При увеличении концентрации репрессора защищается фрагмент из 50 пар оснований, соответствующий операторам $O_{L1}-O_{L2}$ или $O_{R1}-O_{R2}$. Для связывания с сайтом 3 требуются большие концентрации репрессора, но, если это достигается, оказывается защищенным фрагмент из 80 пар оснований, соответствующий полному оператору.

Определение точек контакта с репрессором в каждом сайте связывания позволило предположить, что он связывается симметрично, так что каждый N-концевой домен димера контактирует с одинаковым набором оснований. Следовательно, отдельная N-концевая область контактирует лишь с половиной связывающего сайта. Участок сайта связывания, с которым осуществляется контакт репрессора, показан на рис. 16.11 (затененный). В двухцепочечной спиральной ДНК точки контакта лежат в основном вдоль большой бороздки на ДНК.

Некоторые связывающиеся с ДНК белки могут присоединяться к ДНК таким же способом. Общие свойства трех таких белков приведены в табл. 16.1. Несмотря на то, что каждый белок функционирует как димер, по своей общей организации они различны, так как у репрессора

фага λ ДНК-связывающая область находится на N-конце, у белка БАК — на С-конце, а белок Cto содержит только один домен. Однако в каждом случае активная область содержит некие короткие участки α -спирали, образующие центр контактов с ДНК.

Модель двуспиральных контактов предсказывает, что каждый мономер содержит область (названную α -спираль-3) из 9 аминокислот, которая лежит под углом к предшествующей области из 7 аминокислот, образующей α -спираль-2. У димера две противоположно направленные области α -спираль-3 находятся на расстоянии 34 Å; они могли бы разместиться в соседних больших бороздках ДНК. Области α -спираль-2 лежат под углом и поэтому они могут располагаться поперек бороздок.

Последовательности этих двух областей частично гомологичны у нескольких белков, связывающихся с ДНК, в том числе у БАК, *lac*-репрессора и нескольких фаговых репрессоров. Контакты между α -спиралью-3 и ДНК осуществляются с помощью водородных связей, устанавли-

Таблица 16.1
Области регуляторных белков, ответственные за связывание с ДНК, могут представлять различные участки этих белков

Белок	Длина мономера (в остатках)	Активная форма	ДНК-связывающая область	Другая область
Репрессор фага λ	236	Димер	N-концевая; 92 остатка	C-концевая (агрегаты)
Cto	66	—	Весь белок	Нет
Ак БАК	209	—	C-концевая; 74 остатка	N-концевая (связывается с cAMP)

ваемых между боковыми аминокислотными цепями и выступающими парами оснований. Различия в аминокислотах обуславливают, по-видимому, специфичность отдельных белков. Контакты α -спирали-2 и ДНК осуществляются с помощью водородных связей, устанавливаемых с фосфатным остовом. Наряду с этими контактами большая часть общей энергии взаимодействий с ДНК обеспечивается ионными взаимодействиями с фосфатной основой, которые не зависят от последовательности ДНК.

Репрессор связывается с каждым оператором кооперативно. Сначала репрессор-димер связывается с сайтом 1 из-за большего сродства с ним. Однако присутствие димера в сайте 1 приводит к сильному повышению сродства второго димера к сайту 2. Если оба сайта (1 и 2) заняты, такое взаимодействие не распространяется дальше на сайт 3. Из этого следует, что при концентрации репрессора, обычно обнаруживаемой в лизогенных бактериях, в каждом операторе заняты сайты 1 и 2, а сайт 3 не занят.

Если сайт 1 оказывается неактивным (в результате мутации), тогда репрессор кооперативно связывается с сайтами 2 и 3. Следовательно, связывание в сайте 2 облегчает связывание другого димера в сайте 3. Такое взаимодействие происходит непосредственно между димерами репрессора, а не через конформационные изменения в ДНК. Вероятно, область коннектора первого репрессора таким образом ориентирует С-концевые области димера, что они контактируют с С-концевыми областями второго димера (который поэтому лежит в таком положении, что не может способствовать присоединению третьего димера).

Исходя из последовательностей, приведенных на рис. 16.11, можно видеть, что O_L1 и O_R1 лежат более или менее в центре сайтов связывания с РНК-полимеразой — P_L и P_R соответственно. Занятость сайтов операторов O_L1-O_L2 и O_R1-O_R2 , таким образом, полностью блокирует доступ РНК-полимеразы к соответствующему промотору.

Взаимосвязь иного типа существует между оператором O_R и промотором P_M , используемым при транскрипции гена cI . Сайт связывания РНК-полимеразы расположен рядом с оператором O_R2 . Это делает понятным, каким образом репрессор аутогенно регулирует свой собственный синтез. Если два димера связаны с операторами O_R1-O_R2 , димер, находящийся в операторе O_R2 , взаимодействует с РНК-полимеразой, вероятно, посредством взаимодействия белок—белок. В отличие от взаимодействия между репрессорами этот эффект обеспечивается аминоконцевым доменом, который может стимулировать использование промотора P_M , даже находясь в состоянии независимого фрагмента. Эти и другие взаимодействия в O_R/P_R иллюстрированы на рис. 16.12.

Следует отметить, что, поскольку именно оператор O_R2 должен быть занят репрессором-димером, чтобы способствовать связыванию РНК-полимеразы с промотором P_M , может оказаться, что кооперативная зависимость от димера, первоначально связывающегося с O_R1 , играет важную роль в поддержании аутогенного цикла. У мутантов, утративших активность O_R1 , один димер репрессора, связанный оператором с O_R2 , может подавить иницирование транскрипции в промоторе P_R , но стимулировать ее в промоторе P_M . Сайт O_R2 только на одну пару оснований расположен ближе к стартовой точке

промотора P_R , чем к стартовой точке промотора P_M ; тем не менее различия в наблюдаемых эффектах столь значительны.

Но что произойдет, если димер репрессора свяжется с оператором O_R3 ? Этот сайт перекрывается с сайтом связывания РНК-полимеразы в промоторе P_M . Таким образом, если концентрация репрессора становится достаточно большой, чтобы занять оператор O_R3 , транскрипция гена cI предотвращается. Это в свою очередь приведет к уменьшению концентрации репрессора; в результате оператор O_R3 становится свободным, и аутогенный цикл может быть опять пущен в ход, так как оператор O_R2 остается занятым. Этим обеспечивается механизм, предотвращающий накопление репрессора в слишком большой концентрации. (Однако концентрация репрессора в лизогенных бактериях не является достаточно высокой для достижения такого эффекта. Поэтому его значение *in vivo* остается неясным.) Формально репрессор является аутогенным регулятором своего собственного выражения, который функционирует как положительный регулятор при низких концентрациях и как отрицательный — при высоких.

В сайтах 1 и 2 операторов O_L и O_R были найдены вирулентные мутации. В обоих случаях наблюдается изменение состава оснований. Эти мутации варьируют по степени вирулентности — в зависимости от степени уменьшения сродства сайта связывания с репрессором и от взаимоотношения поврежденного сайта с промотором. В соответствии с выводом о том, что операторы O_R3 и O_L3 обычно не связываются с репрессором, в этих сайтах не обнаружено вирулентных мутаций.

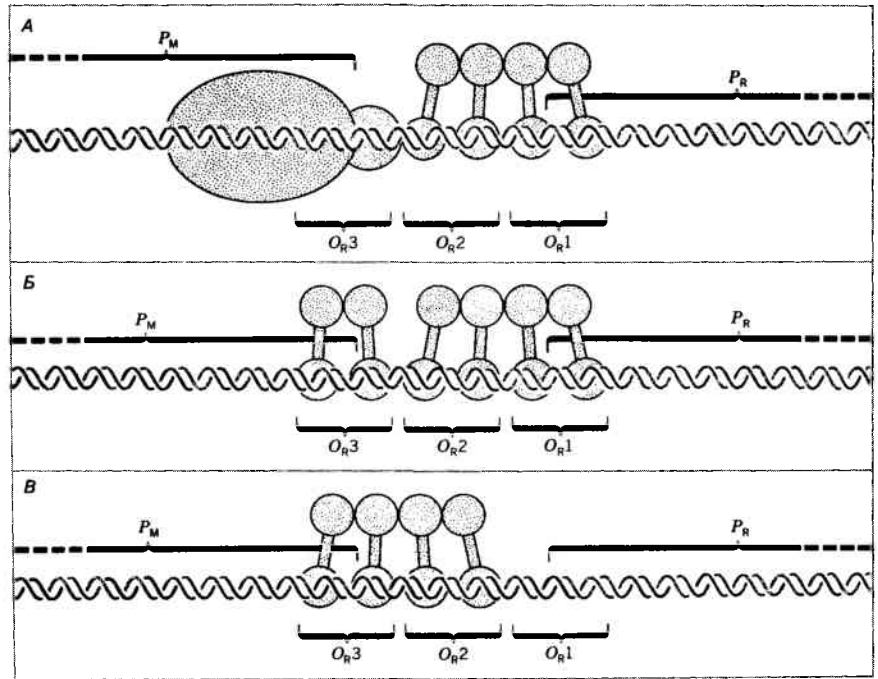
Как запускается синтез репрессора?

Контролирующая система, поддерживающая лизогенное состояние, представляет собой парадокс. Присутствие белка-репрессора необходимо для его собственного синтеза. Это объясняет, как сохраняется лизогенное состояние. Однако как осуществляется первоначальный запуск синтеза репрессора? При проникновении ДНК фага лямбда в новую клетку-хозяина бактериальная РНК-полимераза не способна транскрибировать ген cI , так как в клетке отсутствует репрессор, способствующий ее связыванию с промотором P_M . Но то же отсутствие репрессора означает, что промоторы P_R и P_L оказываются доступными. В результате первым событием при внедрении ДНК фага лямбда в бактериальную клетку является транскрипция генов N и cro . Затем под действием белка рN транскрипция захватывает последующие области. В результате ген $cIII$ (и другие гены) начинают транскрибироваться в процессе левосторонней, а ген cII (и другие гены) — в процессе правосторонней транскрипции (см. рис. 16.6).

Гены cII и $cIII$ подобны гену cI в том отношении, что мутации по этим генам вызывают образование прозрачных бляшек. Но существует и различие между этими генами. cI^- -мутанты не могут ни установить, ни поддерживать лизогенное состояние. В случае cII^- - и $cIII^-$ -мутантов установление лизогении сопряжено с определенными трудностями, но, если уж лизогения установлена, они способны поддерживать это состояние с помощью той же аутогенной системы, которую мы уже обсуждали. Из этого следует, что гены cII и $cIII$ вовлекаются в процесс как регуляторы, продукты которых нуж-

Рис. 16.12. Взаимодействия между димером репрессора и РНК-полимеразой контролируют использование трех промоторов в правой контролирующей области фага лямбда.

А. В случае лизогении димер репрессора контактирует с промоторами O_{R2} и O_{R1} , блокируя выражение генов через промотор P_R , но стимулируя использование промотора P_M . Б. Димер, связанный с оператором O_{R1} , взаимодействует с димером, связанным с оператором O_{R2} , однако связывание в операторе O_{R3} является независимым. При занятости всех трех сайтов ни P_R , ни P_M не используются. В. Если оператор O_{R1} поврежден мутацией, димеры, связанные с O_{R2} и O_{R3} , могут взаимодействовать. В этом случае не могут использоваться ни P_R , ни P_M .



ны для альтернативной системы синтеза репрессора. Данной системе необходимо *только* инициировать выражение гена *cI* для того, чтобы обойти неспособность аутогенной системы начать синтез de novo.

Между генами *cro* и *cII* имеется другой промотор, называемый P_E или иногда P_{RE} . (Буква «Е» указывает на его использование для установления репрессора — repressor establishment.) Этот промотор может узнаваться РНК-полимеразой только в присутствии продуктов генов *cII* и *cIII*, которые, следовательно, являются положительными регуляторными белками. Их действие показано на рис. 16.13.

Промотор P_E имеет необычную последовательность, утратившую обе кононические последовательности. Возможно, именно этой нехваткой объясняется его зависимость от генов *cII* и *cIII*. О значении областей -10 и -35 для узнавания РНК-полимеразой говорит существование мутаций *cU*. Их эффект подобен эффекту мутаций *cII⁻* и *cIII⁻*, предотвращающих установление лизогении; однако в отличие от последних они являются *cis*-действующими, а не *trans*-действующими. Они распределяются на две группы, локализующиеся в промоторе P_E вокруг положений -35 и -10 . Следовательно, они являются промоторными мутациями. Согласно существующим предположениям, продукты генов *cII/cIII* могут действовать в сайте -35 , разрешая РНК-полимеразе связываться с сайтом -10 .

Транскрипт промотора P_E содержит **антисмысловую** последовательность гена *cro*, который обычно транскрибируется в противоположном направлении (т.е. в направлении от P_R). Антисмысловая последовательность *cro* не транслируется с P_E -транскрипта, но кодирующая область *cI* транслируется очень эффективно (по сравнению со слабой трансляцией P_M -транскрипта, упоминавшейся ранее). В действительности репрессор синтезируется примерно в семь или восемь раз более эффективно при экспрессии с промотора P_E , чем с промотора P_M .

Возможность одновременной транскрипции гена в противоположных направлениях довольно долго казалась загадочной. Мы и в самом деле не знаем, что происходит, когда две РНК-полимеразы, двигаясь по направлению друг к другу, встречаются. Существует предположение, что при этом образуется клубок, блокирующий транскрипцию; однако не исключено, что полимеразы способны каким-то образом, который нам непонятен, продвигаться дальше. Возможно, что благодаря тому, что промотор P_E сильнее, его использование подавляет использование промотора P_R ; это могло бы объяснить эффект генов *cII* и *cIII*, выражающийся в ослаблении экспрессии правосторонних ранних генов.

Теперь мы можем перейти к обсуждению вопроса о том, как устанавливается лизогенное состояние при новой инфекции. На рис. 16.14 представлены обобщенные данные. Показаны первоначальные стадии, включающие транскрибирование генов *N* и *cro* и распространение этой транскрипции под действием белка pN на гены *cII* и *cIII*. Далее присутствие продуктов генов *cII/cIII* позволяет использовать промотор P_E для транскрипции, проходящей через ген *cI*. С этого транскрипта в больших количествах синтезируется белок-репрессор. Он немедленно связывается с операторами O_I и O_R .

Присоединение репрессора выключает выражение всех фаговых генов путем непосредственного подавления любой дальнейшей транскрипции, идущей от промоторов P_L и P_R . Происходит остановка синтеза *cII*- и *cIII*-белков, которые, будучи нестабильными, быстро распадаются; в результате промотор P_E не может больше использоваться. Таким образом, при включении рассмотренной цепи событий синтез репрессора прекращается.

Однако репрессор теперь находится в промоторе O_R . Это обуславливает включение системы поддержания экспрессии генов с промотора P_M . В результате синтез репрессора продолжается, хотя на этот раз на более низком уровне, типичном для функции P_M . Таким образом, установление лизогении начинается с синтеза репрессора на

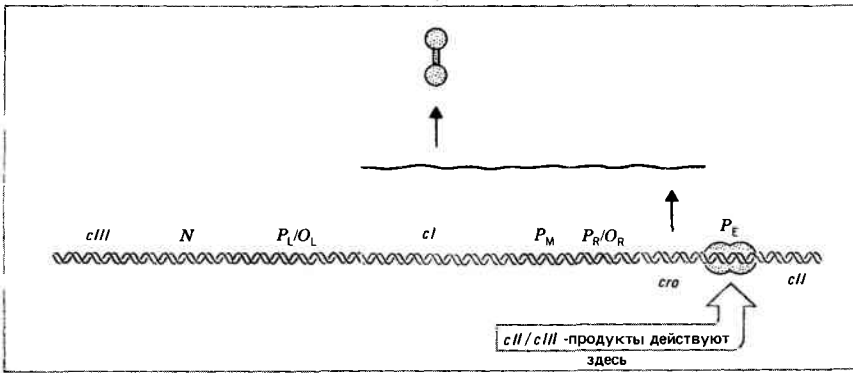


Рис. 16.13. Синтез репрессора запускается в P_E при участии продуктов генов cII и $cIII$, которые способствуют иницированию транскрипции, происходящей на антисмысловой цепи от гена cro к гену cl .

высоком уровне. Затем репрессор выключает все другие функции; при этом включается система поддержания, функционирующая на низком уровне, достаточном для сохранения лизогенного состояния.

Мы не будем сейчас детально рассматривать другие функции, необходимые для установления лизогенного состояния, и лишь заметим, что инфицирующая ДНК фага лямбда должна быть интегрирована с бактериальным геномом (гл. 35). Для интеграции необходимо образование продукта гена int , который экспрессируется со своего собственного промотора P_I . Для иницирования транскрипции необходимы продукты генов $cII/cIII$. Последовательность промотора P_I проявляет гомологию с последовательностью P_E . Функции, необходимые для установления лизогенного контролирующего цикла, подчиняются тому же контролю, что и функция, физически необходимая для манипулирования ДНК. Таким образом, завершение процесса лизогенизации находится под контролем, обеспечивающим осуществление всех необходимых событий в согласованной последовательности.

Для литической инфекции необходим антирепрессор

В начале этой главы мы говорили, что фаг лямбда способен вести себя двояко, осуществляя лизогенизацию или вступая на путь литического развития. Мы уже видели, что лизогенное состояние устанавливается с помощью программы, обеспечивающей запуск аутогенно контролируемой системы сохранения, подавляющей весь литический каскад «давлением» в двух точках. Цикл, связанный с установлением лизогении, в действительности включает в себя некоторые из тех же событий, которые уже рассматривались ранее в разделе, посвященном литическому каскаду (экспрессия задержанно ранних генов зависит от экспрессии гена N). Теперь мы рассмотрим другую проблему. Каким образом фаг вступает в литический цикл развития?

До сих пор мы не принимали во внимание роль гена cro , кодирующего другой регуляторный белок, иногда называемый **антирепрессором**. Данный регулятор отвечает за предотвращение синтеза белка-репрессора; это в свою очередь исключает возможность установления лизогении. Мутанты cro^- -типа обычно устанавливают лизогению, а не вступают на литический путь развития, поскольку они утратили способность переключать события от экспрессии репрессора в другое русло.

Продукт гена cro представляет собой небольшой белок (размером 9000 дальтон), образующий димер, ко-

торый действует в пределах области иммунности. Он проявляет два эффекта: предотвращает синтез репрессора, действуя на систему поддержания его синтеза, т.е. предотвращает транскрипцию с использованием P_M . Кроме того, он подавляет выражение ранних генов, иницируемое в промоторах P_L и P_R . Это означает, что при вступлении фага на литический путь развития белок Cro обуславливает как предотвращение синтеза репрессора, так и подавление экспрессии ранних генов.

Для выполнения своей функции белок Cro связывается с теми же сайтами, с которыми связывается репрессор. Он также связывается симметрично в этих сайтах, осуществляя контакты, подобные контактам репрессора. Каким образом два белка могут действовать в одних и тех же сайтах, вызывая такие противоположные эффекты? Ответом на этот вопрос служит различное сродство, проявляемое каждым из белков по отношению к определенным сайтам связывания в пределах оператора.

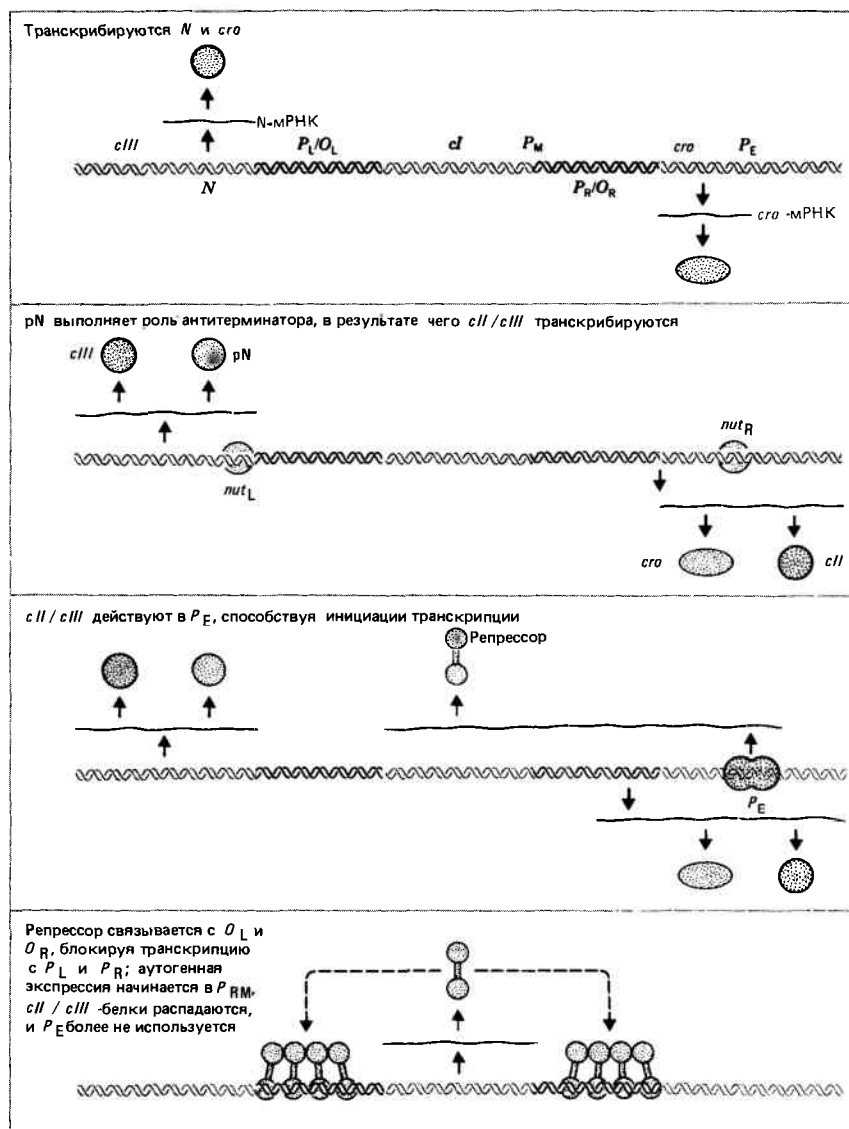
Рассмотрим это на примере наиболее изученного в этом отношении оператора O_R , где белок Cro проявляет оба своих эффекта. Ряд происходящих при этом событий показан на рис. 16.15.

Белок Cro имеет более высокое сродство к оператору O_R3 , чем к O_R2 или O_R1 . Таким образом, он первоначально связывается с O_R3 . Это подавляет связывание РНК-полимеразы с промотором P_M . Следовательно, первый эффект белка Cro состоит в предотвращении введения в действие системы поддержания лизогении.

Затем белок Cro связывается с оператором O_R2 или O_R1 . Его сродство к этим сайтам одинаково, и кооперативного эффекта не наблюдается. Присутствие этого белка в любом сайте оказывается достаточным для того, чтобы помешать РНК-полимеразе использовать промотор P_R . Это в свою очередь останавливает выражение ранних функций (включая сам ген cro). Поскольку продукты гена $cII/cIII$ нестабильны, использование промотора P_E полностью прекращается. Так две функции белка Cro полностью блокируют образование репрессора.

Что касается литического цикла, то белок Cro подавляет (хотя и не полностью элиминирует) выражение ранних генов. Его неполный эффект объясняется тем обстоятельством, что сродство этого белка к операторам O_R1 и O_R2 примерно в восемь раз ниже сродства репрессора. Этот эффект Cro не проявляется до тех пор, пока функционирование ранних генов не становится более или менее излишним из-за присутствия рQ. К этому времени начинается выражение поздних генов фага, и процесс сосредоточивается на образовании потомства фаговых частиц.

Рис. 16.14. Для установления лизогении необходима каскадная регуляция, но затем эта система выключается и заменяется аутогенной системой, сохраняющей репрессор.



Чувствительный баланс: лизогения против лизиса

Программы лизогенного и литического путей настолько тесно связаны между собой, что невозможно предсказать судьбу отдельного фагового генома, который вводится в клетку новой бактерии-хозяина. Будет ли антагонизм между репрессором и белком σ_{50} разрешен путем установления системы поддержания, показанной на рис. 16.14, или путем выключения синтеза репрессора и вступления в позднюю стадию развития, показанную на рис. 16.14? Один и тот же путь следует в обоих случаях вплоть до момента «принятия решения». В обоих случаях включается выражение предранних генов и распространение выражения на задержанно-ранние гены. Различие между ними возникнет при решении вопроса, что присоединится к двум операторам — репрессор или белок σ_{50} .

Ранняя фаза, в течение которой принимается решение, имеет ограниченную продолжительность в любом случае. Независимо от того, какой путь будет выбран фагом, экспрессия всех ранних генов будет предотвращаться, когда промоторы P_L и P_R репрессированы, а в результате ис-

чезновения продуктов генов cII и $cIII$ будет прекращено образование репрессора с использованием промотора P_E . Критическое значение приобретает вопрос: последует ли за прекращением транскрипции, инициируемой в промоторе P_E , активация промотора P_M с установлением лизогении, или P_M окажется не способным стать активным и регулятор pQ приведет фаг к литическому развитию?

Исходное событие при установлении лизогении — это связывание репрессора с O_L1 и O_R1 . Связывание в первых сайтах будет быстро способствовать кооперативному связыванию другого димера репрессора с операторами O_L2 и O_R2 . Это выключит синтез белка σ_{50} и включит синтез репрессора с использованием промотора P_M .

Исходным событием при вступлении в литический цикл является связывание белка σ_{50} в операторе O_R3 . Это блокирует систему сохранения лизогении на стадии инициирования транскрипции в промоторе P_M . Затем продукт σ_{50} должен связаться с оператором O_R1 или O_R2 и с O_L1 или O_L2 , чтобы подавить выражение ранних генов. В результате выражение генов cII и $cIII$ будет приостановлено, а это, в свою очередь, приведет к пре-

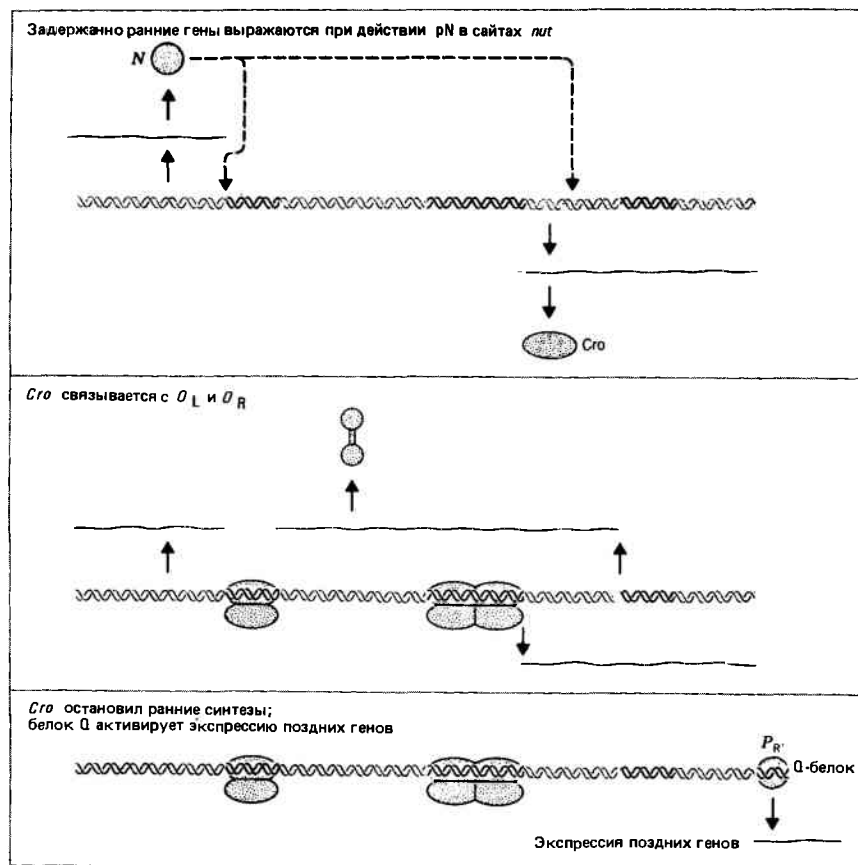


Рис. 16.15. Литический каскад требует участия белка Cro , непосредственно предотвращающего поддержание репрессора при использовании промотора P_M ; он также включает выражение задержанно ранних генов, предотвращающих непрямым путем установление репрессора через промотор P_E .

кращению синтеза репрессора с использованием промотора P_E . Выключение процесса установления репрессии происходит, когда нестабильные белки cII и $cIII$ распадаются.

Таким образом, существует цепь событий, конкурирующих между собой за одну и ту же область. РНК-полимераза, репрессор и белок Cro связываются в левой и правой контролируемых областях, используя способы взаимопомощи или взаимоисключения. Какой из регуляторов восторжествует, будет зависеть от чувствительного количественного и временного баланса. Флуктуации в согласованности и частоте транскрипции и трансляции в отдельных клетках могут определить, какой из путей последует в каждом случае.

Рекомендуемая литература

Основание для такого представления о фаге лямбда детально проанализировано в книге Льюина (*Lewin, Gene Expression, 3, Plasmids and Phages*, John Wiley, New York, 1978, pp. 274–535), в которой приведены также данные о фаге T4 (pp. 536–681), фагах T3 и T7 (pp. 682–723) и других фагах. Фаг лямбда рассматривается также в двух главах книги Миллера и Резникова [*The Operon* (Miller, Reznikoff, Eds.) Cold Spring Harbor Laboratory, New York, 1978]: *Rosenberg et al* (pp. 345–371) – о регуляторных системах и *Ptashne* (pp. 325–343) – о действии репрессора. Данные о взаимодействии между репрессором и белком Cro были изложены Пташне и др. (*Ptashne et al., Cell*, 19, 1–11, 1980).

Часть V

СТРОЕНИЕ ГЕНОМА ЭУКАРИОТ

В соответствии с основными положениями современной генетики совокупность признаков организма есть результат взаимодействия всех генов организма и среды, в которой происходит его развитие. Изменение любого гена, называемое мутацией, приводит к нарушению равновесия этой системы и проявляется как некое свойство организма, обычно аномальное и в некотором отношении худшее по сравнению с организмом дикого типа... Изучение нехваток, т. е. материала, характеризующегося отсутствием одного или нескольких генов, показало, что большинство нехваток оказываются летальными для организма, если они находятся в гомозиготном состоянии. Это указывает на то, что наличие по крайней мере большинства генов необходимо для выживания организма. Более того, изучение нехваток у *D. melanogaster* показало, что многие из них оказываются летальными на клеточном уровне, т. е. даже небольшая группа клеток, расположенных в нормальных тканях, но содержащая гомозиготную нехватку, существовать не может. Из этого следует, что гены активно функционируют во всех клетках и что, вероятно, основная их часть выполняет там весьма важные функции, необходимые для жизнедеятельности клеток.

Милицлав Демерец, 1935

ГЕНОМЫ ЭУКАРИОТ: МНОЖЕСТВО ПОСЛЕДОВАТЕЛЬНОСТЕЙ

«Эукариотический геном» – несколько расплывчатое понятие. Основной признак эукариот – наличие ядра, в котором заключена основная часть генома и которое отделено от цитоплазмы ядерной мембраной. В целом общее количество ядерной ДНК очень сильно различается у разных организмов; число хромосом, ее содержащих, сильно варьирует; характеристики относительно небольшой части генома, заключенной в оргanelлах, также колеблются в широких пределах; кроме того, могут обнаруживаться значительные различия в типах последовательностей, образующих ядерный геном.

В то же время эукариотическому геному присущи черты, отсутствующие в геноме прокариот. Индивидуальные гены могут иметь прерывистое строение, геном может содержать множественные и (иногда) идентичные копии определенных последовательностей, а также большие участки ДНК, не кодирующие белков. Вследствие пространственной разобщенности ядра и цитоплазмы механизмы экспрессии генов у эукариот и прокариот должны неизбежно различаться.

Отражают ли эти черты основополагающие принципы организации эукариотической ДНК? Не следует забывать, что царство эукариот чрезвычайно велико, а в настоящее время мы имеем подробную информацию об организации генома только у некоторых видов. Рассматривая свойства эукариотической ДНК, мы обнаруживаем такие особенности ее организации, которые могут быть в различной степени присущи разным индивидуальным геномам. Поэтому нас будет интересовать выяснение конкретных вариантов организации эукариотической ДНК, а не гипотетической «типичной» ее структуры.

Парадокс величины C характеризует различия в размерах геномов

Общее количество ДНК, приходящееся на (гаплоидный) геном, – характеристика каждого существующего вида организмов, известная под названием **величина C** . Эта величина может быть измерена химически и выражена в пикограммах (пкг) ДНК или определена методом исследования кинетики реассоциации ДНК (см. ниже), когда ее обычно выражают в парах нуклеотидных оснований или дальтонах. Соотношение между этими величинами таково: $1 \text{ пкг} = 0,965 \cdot 10^9 \text{ п. н.} = 6,1 \cdot 10^{11} \text{ дальтон}$. Величина C колеблется в огромных пределах от такого маленького значения, как всего лишь 10^4 п. н. у микоплазмы, до такого большого, как $\sim 10^{11}$ п. н. у некоторых растений и амфибий.

Так называемый **парадокс величины C** – результат нашей неспособности объяснить содержание ДНК в геноме с точки зрения осуществляемых им функций. Этот парадокс имеет два аспекта. Во-первых, наблюдаются огромные различия в величинах C у видов, сложность организации которых так сильно не различается; существенные различия могут наблюдаться даже между некоторыми близкородственными видами.

На **рис. 17.1** приведен диапазон значений величины C для организмов, находящихся на различных ступенях эволюционного развития. На каждой ступени с увеличением сложности организации наблюдается некоторое увеличение минимального размера генома. Это правило не является абсолютным, но очевидно, что одноклеточные эукариоты (чей жизненный цикл несколько напоминает прокариотический) имеют геном небольшого размера, превосходящий, однако, по размеру геному бактерий. Сама по себе принадлежность к эукариотам не означает значительного увеличения размеров генома; например, размер генома у дрожжей *S. cerevisiae* составляет около $2,3 \cdot 10^7$ п. н., что только в 5 раз превышает размер генома бактерии *E. coli*.

Дальнейшее небольшое увеличение размеров генома (немногим более, чем в два раза) оказывается достаточным для того, чтобы слизистые грибы (миксомицеты) *D. discoideum* могли существовать в виде как одноклеточных, так и многоклеточных организмов. Для возникновения первых полностью многоклеточных организмов необходимо дальнейшее увеличение сложности генома. Например, содержание ДНК у круглых червей *C. elegans* составляет $8 \cdot 10^7$ п. н. При дальнейшем эволюционировании четкой связи между сложностью организма и содержанием ДНК не наблюдается, хотя для появления насекомых понадобилось, чтобы размер генома превысил 10^8 п. н., для появления иглокожих – $4 \cdot 10^8$ п. н., для появления птиц или амфибий – $8 \cdot 10^8$ п. н. и $2 \cdot 10^9$ п. н. – для появления млекопитающих.

В некоторых случаях диапазон различий в размерах генома совсем невелик. Например, у птиц, рептилий и млекопитающих обнаруживаются небольшие различия внутри класса, причем размеры генома в каждом случае могут колебаться примерно в два раза. В других же случаях наблюдается широкий диапазон колебаний размеров генома – часто более чем в десять раз.

Эти факты свидетельствуют о некотором удивительном несоответствии между размерами генома и сложностью организма.

Необычный диапазон колебаний величины C обнаружен у амфибий, у которых минимальный размер генома не превышает 10^9 п. н., в то время как максимальный размер составляет почти 10^{11} п. н. Трудно поверить, что такие различия могут соответствовать 100-кратным различиям в количестве генов, определяющих специфические особенности разных амфибий. Подобный скептицизм усиливается фактом существования нескольких довольно близкородственных видов, у которых обнаруживаются удивительные различия в размерах генома в целом. Так, например, два вида амфибий, имеющих весьма сходное морфологическое строение, могут различаться, скажем, в 10 раз по относительному содержанию ДНК. Также кажется маловероятным существование 10-кратных различий в числе генов. Если все-таки число генов у разных организмов примерно одинаково, то основная часть ДНК у видов с большим размером генома не кодирует белков. Какова же тогда ее функция?

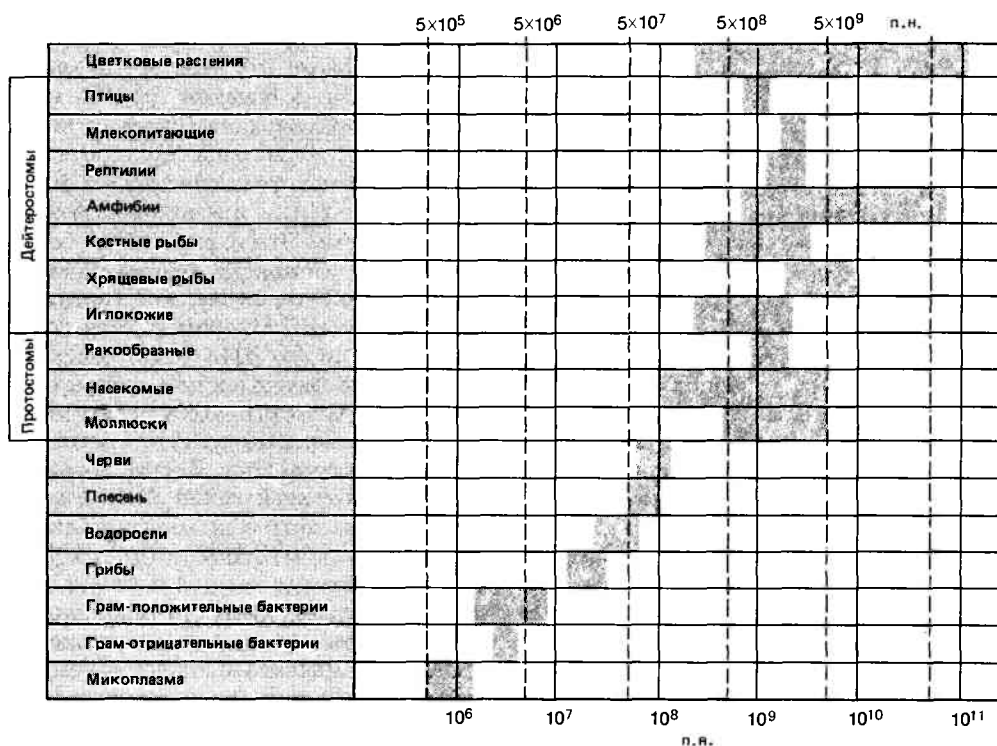


Рис. 17.1. Содержание ДНК, приходящееся на гаплоидный геном, непосредственно не связано с морфологической сложностью вида. Размер генома увеличивается при переходе от прокариот к протостомам, но варьирует в широких пределах

у дейтеростомных организмов. Диапазоны колебаний размеров ДНК в пределах систематической единицы обозначены заштрихованными областями.

Эти факты подводят нас ко второму аспекту парадокса величины C – существованию явного избытка ДНК по сравнению с ее количеством, необходимым для кодирования белков. В самом деле, об этом часто говорят как о проблеме избыточности эукариотической ДНК. Сейчас известно, что избыточность частично объясняется тем, что размер генов существенно больше размера последовательностей, необходимых для кодирования белков (в основном из-за наличия промежуточных последовательностей, разбивающих кодирующую область гена на несколько частей). Однако этого недостаточно для решения проблемы. Рассмотрим два примера.

Даже если предположить, что средний размер гена млекопитающих достигает 10 000 п.н. (что на самом деле превышает размеры почти всех известных прерывистых генов), число генов в геноме млекопитающих должно составить 300 000. Возможно ли это?

По-видимому, нет; поскольку, хотя число генов точно не известно, имеются два указания на то, что число генов измеряется в десятках тысяч, а не в сотнях тысяч.

Одно прямое доказательство получено при подсчете числа генов путем определения числа различных мРНК (гл. 18). Хотя это нельзя проделать для каждого типа клеток и определить тем самым общее число генов организма, полагают, что число генов, экспрессирующихся в клетках определенного типа, составляет примерно 10 000. Большинство из них экспрессируется во всех или почти во всех клетках организма. Таким образом, общее число экспрессирующихся генов, вероятно, находится в пределах этой величины с коэффициентом (примерно) 2-4.

Более косвенные данные получены при попытках установить число наиболее важных генов путем идентифика-

ции всех локусов, которые могут подвергаться мутациям. Большая часть этих экспериментов была проведена на *D. melanogaster*, для которой общее число важнейших генов можно считать равным примерно 5000. Если принять средний размер гена насекомого за 2000 п.н., то общая длина всех генов составит 10^7 п.н., что в 10 раз меньше, чем длина реальной ДНК.

Разумеется, с помощью мутаций можно выявить только те гены, повреждения которых приводят к видимым изменениям или летальному эффекту. К этому классу, очевидно, относятся лишь некоторые гены. Это означает, что по меньшей мере значительная часть, возможно, даже большинство генов кодируют белки, не имеющие существенного значения для выживания организма (по крайней мере в том смысле, что их мутационное повреждение не вызывает какого-либо заметного эффекта).

Таким образом, некоторые ключевые вопросы остались нерешенными. Какая часть геномной ДНК действительно связана с образованием белков в том смысле, что она соответствует гену, либо его кодирующей области, либо промежуточным или транскрибируемым фланкирующим¹ последовательностям? Какое количество генов жизненно необходимо, а какое несущественно для выживания? Какова функция (если она существует) ДНК, которая не входит в состав генов? Какое влияние на функционирование генома оказывает значительное изменение его общего размера, как в случае родственных представителей класса амфибий?

¹ Расположенным по обе стороны от гена. – Прим. перев.

Кинетика реассоциации зависит от генетической сложности последовательностей ДНК

Реассоциация двух комплементарных последовательностей ДНК происходит путем спаривания оснований — в отличие от процесса денатурации, при котором они разделяются (см. рис. 2.15). Мы уже рассмотрели многие эксперименты, в основе которых лежит использование этого метода для выделения индивидуальных последовательностей ДНК или РНК и их способности гибридизоваться со специфическим зондом. Сейчас мы увидим, что кинетика реассоциации отражает разнообразие присутствующих в клетке последовательностей, причем эта реакция может быть использована для определения количества генов и их РНК-продуктов. Такие реакции при их проведении в растворе называются **гибридизацией в растворе**.

Ренатурация ДНК зависит от случайных столкновений комплементарных цепей; таким образом, этот процесс подчиняется кинетике реакций второго порядка. Это означает, что скорость реакции определяется уравнением

$$\frac{dC}{dt} = -kC^2, \quad (1)$$

где C — концентрация одноцепочечной ДНК в момент времени t , а k — константа скорости реассоциации.

Проинтегрировав это уравнение в пределах от исходной концентрации ДНК, C_0 , в момент времени $t=0$ до концентрации молекул C , остающихся одноцепочечными к моменту времени t , протекание реакции можно описать следующим образом:

$$\frac{C}{C_0} = \frac{1}{1 + k \cdot C_0 t}. \quad (2)$$

Таким образом, когда реакция прошла наполовину, в момент времени $t_{1/2}$

$$\frac{C}{C_0} = \frac{1}{2} = \frac{1}{1 + k \cdot C_0 t_{1/2}}, \quad (3)$$

откуда

$$C_0 t_{1/2} = \frac{1}{k}. \quad (4)$$

Реассоциация любой специфической ДНК может быть описана с помощью константы скорости реакции k (моль нуклеотидов⁻¹·л·с⁻¹) или обратной ей величины $C_0 t_{1/2}$ (в молях нуклеотидов·с/л).

Из этого уравнения следует, что основным параметром реакции реассоциации является произведение концентрации ДНК (C_0) на время инкубации (t). Эта величина часто обозначается просто как $C_0 t$. Аналогичным образом величина, соответствующая осуществлению реакции реассоциации наполовину, обозначается как $C_0 t_{1/2}$. Большие значения $C_0 t_{1/2}$ указывают на более медленное протекание реакции, поскольку эта величина — произведение концентрации на время, необходимое для осуществления реакции наполовину.

Реассоциацию ДНК обычно изображают в виде **кривой** $C_0 t$, где значениям $\log C_0 t$ соответствует процентное содержание в смеси фракции реассоциировавшей ДНК ($1 - C/C_0$). На **рис. 17.2** приведены кривые $C_0 t$ для нескольких геномов. Форма кривых одинакова, ренатурация осуществляется в диапазоне изменения значений $C_0 t$

в 100 раз, и кривая располагается между точками, соответствующими 10%-ной и 90%-ной ренатурации. Но значения $C_0 t$ в каждом случае сильно различаются, и процесс реассоциации может быть охарактеризован величиной $C_0 t_{1/2}$ (значение $C_0 t$, соответствующее осуществлению реакции реассоциации на 50%).

Значение $C_0 t_{1/2}$ прямо пропорционально количеству ДНК в геноме. Это отражает то обстоятельство, что при увеличении сложности генома уменьшается количество копий каждой специфической последовательности в общей массе ДНК. Например, если величина C_0 ДНК составляет 12 пг, то ДНК будет содержать 3000 копий каждой последовательности в случае бактериального генома, размер которого — 0,004 пг, но лишь 4 копии каждой последовательности, присутствующей в эукариотическом геноме размером 3 пг. Таким образом, при одинаковой абсолютной концентрации ДНК, измеряемой в молях нуклеотидов на литр (C_0), концентрация каждой эукариотической последовательности будет в 750 раз ($3000/4$) ниже, чем концентрация каждой бактериальной последовательности. Поскольку скорость реассоциации зависит от концентрации комплементарных последовательностей, для достижения одинаковой *относительной* концентрации эукариотических и бактериальных последовательностей необходимо иметь в 750 раз больше эукариотической ДНК (либо инкубировать то же ее количество в 750 раз дольше). Соответственно $C_0 t_{1/2}$ реакции реассоциации в случае эукариотической ДНК в 750 раз больше, чем в случае бактериальной ДНК.

Таким образом, величина $C_0 t_{1/2}$ реакции указывает на *суммарную длину различных* присутствующих последовательностей, которая называется **сложностью генома** (*complexity*). Обычно эту величину выражают в парах нуклеотидных оснований, но она может быть выражена в дальтонах или любых других единицах массы.

При ренатурации ДНК любого генома (или части генома) величина $C_0 t_{1/2}$ будет пропорциональна его сложности. Поэтому сложность любой ДНК можно определить путем сравнения ее величины $C_0 t_{1/2}$ с соответствующим значением стандартной ДНК известной сложности. Обычно в качестве стандарта используют ДНК *E. coli*. Ее сложность принимают равной длине генома (считая, что каждая последовательность генома *E. coli* размером $4,2 \cdot 10^6$ п. н. уникальна). Тогда можно написать следующее соотношение:

$$\frac{C_0 t_{1/2} \text{ (ДНК любого генома)}}{C_0 t_{1/2} \text{ (ДНК } E. coli)} = \frac{\text{Сложность любого генома}}{4,2 \cdot 10^6 \text{ п. н.}} \quad (5)$$

Эукариотические геномы состоят из последовательностей нескольких типов

Когда ДНК эукариотического генома характеризуют с помощью кинетики реассоциации, значения $C_0 t$ реакции обычно находятся в пределах, различающихся на 8 порядков. Это гораздо более широкий диапазон, чем 100-кратный диапазон, которого можно было бы ожидать, исходя из уравнения 2, и который показан на **рис. 17.2**. Причина состоит в том, что уравнение описы-

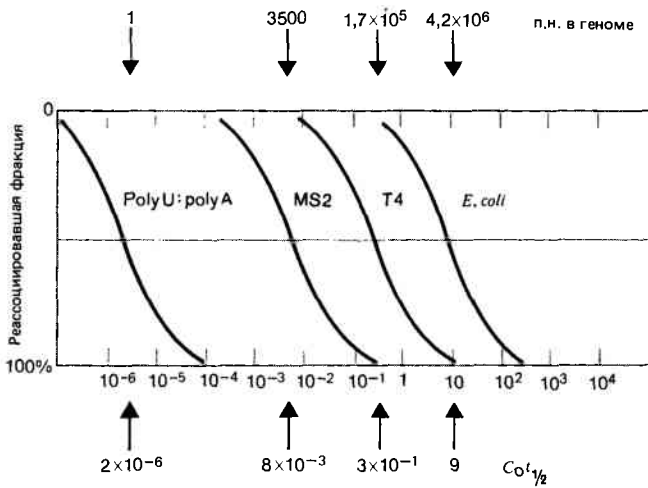


Рис. 17.2. Скорость реассоциации обратно пропорциональна длине реассоциирующей ДНК.

вает реассоциацию одного **кинетически чистого** компонента. В действительности же в состав генома может входить несколько таких компонентов, каждый из которых характеризуется своей кинетикой.

На рис. 17.3 показана реассоциация гипотетического эукариотического генома, начинающаяся при значении C_0t , равном 10^{-4} , и заканчивающаяся при значении $C_0t = 10^4$. В этой реакции можно выделить три фазы, обозначенные на рисунке закрашенными прямоугольниками. Видно, что первые две фазы разделяет плато, а вторая и третья слегка перекрываются. Каждая из этих фаз соответствует определенному кинетическому компоненту генома.

Фракцию ДНК, реассоциирующую первой, называют **быстро ренатурирующей**. В приведенном примере она составляет 25% от всей ДНК и ренатурирует при значениях C_0t от 10^{-4} до примерно $2 \cdot 10^{-2}$ и значении $C_0t_{1/2}$, равном 0,0013.

Следующая фракция называется **промежуточной**. Она составляет 30% от всей ДНК и ренатурирует при значениях C_0t от примерно 0,2 до 100 и значении $C_0t_{1/2}$, равном 1,9.

Фракция, реассоциирующая последней, называется **медленно ренатурирующей**. Она составляет 45% от всей ДНК и ренатурирует в диапазоне значений C_0t от < 100 до примерно 10000 при значении $C_0t_{1/2}$, равном 630.

При определении сложности этих фракций каждую из них следует рассматривать как независимый кинетический компонент, реассоциацию которого сравнивают с реассоциацией стандартной ДНК. Медленно ренатурирующая фракция составляет 45% от всей ДНК, поэтому ее концентрация в реакции реассоциации соответствует 0,45 измеренной величины C_0 (относящейся к общему количеству участвующей в реакции ДНК). Соответственно величина $C_0t_{1/2}$, относящаяся только к медленно ренатурирующей фракции, составляет $0,45 \cdot 630 = 283$. Таким образом, если бы ДНК этой фракции была выделена в виде чистого компонента, свободного от ДНК других фракций, то она ренатурировала бы при значении $C_0t_{1/2}$, равном 283. Предположим, что в таких же условиях ДНК *E. coli* реассоциирует при значении $C_0t_{1/2}$, равном 4,0. Подставив эти две величины в уравнение 5, мы увидим, что сложность этой фракции составляет $3,0 \cdot 10^8$ п.н. Ана-

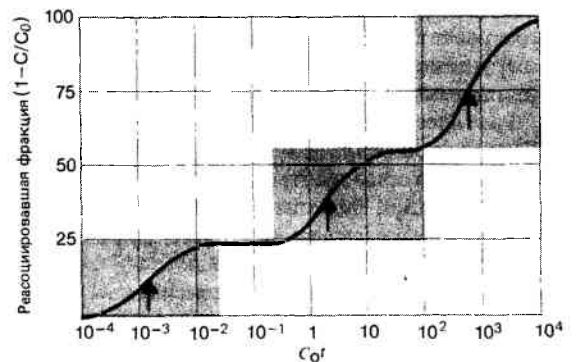
логичным образом можно определить, что сложность промежуточной фракции составляет $6 \cdot 10^5$ п.н., а сложность быстро ренатурирующей фракции — только 340 п.н. Этот расчет служит количественным обоснованием нашего утверждения о том, что, чем быстрее компонент реассоциирует, тем ниже его сложность.

Наоборот, если мы возьмем три препарата ДНК, каждый из которых содержит уникальную последовательность определенной длины (340 п.н., $6 \cdot 10^5$ п.н. и $3 \cdot 10^8$ п.н. соответственно), и смешаем их в соотношении 25:30:45, то каждый будет ренатурировать в соответствии с уравнением 2, как будто он является единственным компонентом. Кинетика ренатурации всей смеси целиком будет выглядеть так же, как кинетика ренатурации всего генома, приведенная на рис. 17.3.

Размер генома можно оценивать по сложности неповторяющейся ДНК

Сложность медленно ренатурирующей фракции ДНК соответствует ее размеру. Предположим, что определенное химическое содержание ДНК в гаплоидном геноме, реассоциация которого показана на рис. 17.3, составляет $7,0 \cdot 10^8$ п.н. Тогда 45% этой величины составят $3,15 \cdot 10^8$ п.н., что только ненамного превышает величину $3,0 \cdot 10^8$ п.н., определенную в соответствии с кинетикой реассоциации. Действительно, с учетом ошибок измерения каждого из этих двух методов можно утверждать, что значения сложности медленно ренатурирующей фракции, определяемые как химическим, так и кинетическим методом, одинаковы. Соответствующие величины называются **химической сложностью** и **кинетической сложностью**.

Совпадение этих величин означает, что медленно ренатурирующая фракция состоит из неповторяющихся (уникальных) последовательностей генома: после денатурации каждая одноцепочечная последовательность может



	Быстро ренатурирующая фракция	Промежуточная фракция	Медленно ренатурирующая фракция
Содержание в геноме, %	25	30	45
$C_0t_{1/2}$	0,0013	1,9	630
Сложность, п.н.	340	$6,0 \times 10^5$	$3,0 \times 10^8$
Частота повторяемости (количество копий на геноме)	500 000	350	1

Рис. 17.3. При исследовании кинетики реассоциации эукариотической ДНК в ней обнаруживаются три типа компонентов (обозначенных закрашенными областями).

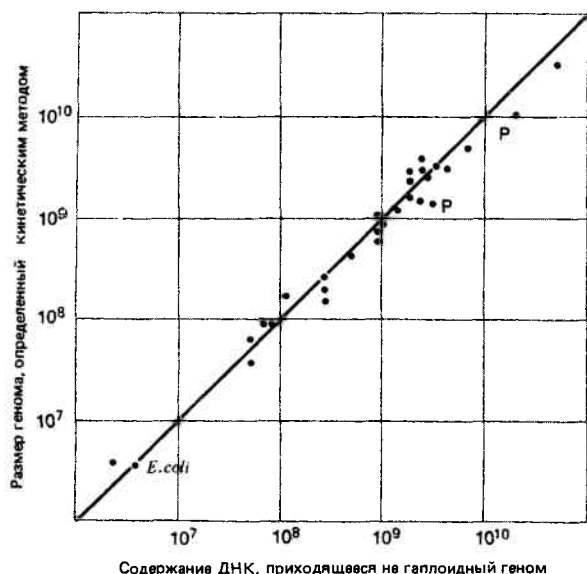


Рис. 17.4. Кинетическая сложность эукариотических геномов хорошо коррелирует с их химической сложностью, за исключением полиплоидных геномов (обозначенных буквой P).

ренатурировать *только* с соответствующей ей комплементарной последовательностью. Эта часть генома — единственный компонент прокариотической ДНК (о чем свидетельствуют приведенные на рис. 17.2 примеры и использование ДНК *E. coli* в качестве стандарта); этот компонент обычно является наибольшим в эукариотической ДНК. Он называется **неповторяющейся** (уникальной) ДНК.

Кинетическую сложность уникальной ДНК можно использовать для оценки сложности генома в целом. В этом случае делают расчет, обратный приведенному ранее. Для примера, рассмотренного на рис. 17.3, сложность неповторяющейся ДНК составляет $3,0 \cdot 10^8$ п. н. Если эта фракция уникальна и составляет 45% генома, то весь геном будет иметь размер $3,0 \cdot 10^8 / 0,45 = 6,6 \cdot 10^8$ п. н. Этот расчет дает независимую оценку размеров генома, которую можно сравнить с химической сложностью, составляющей в данном случае $7,0 \cdot 10^8$ п. н.

Фактически в составе каждого эукариотического генома имеется неповторяющаяся ДНК. Единственное исключение — некоторые растения, возникшие в результате полиплоидизации, вследствие которой каждая последовательность оказывается представленной более чем одной копией.

На рис. 17.4 приведена зависимость между размером генома, определенным по кинетике реассоциации неповторяющейся ДНК, и содержанием ДНК на гаплоидный геном, определенным с помощью химического анализа. Данные хорошо согласуются. Это подтверждает высказанное выше предположение о том, что неповторяющаяся ДНК действительно состоит из индивидуальных последовательностей, присутствующих в геноме в виде только одной копии.

Это важный факт в свете существующего парадокса величины C . Наличие неповторяющейся ДНК означает, что геномы больших размеров не образуются путем простого увеличения числа копий последовательностей, присутствующих в геномах меньшего размера. В этом случае геномы большего размера вели бы себя как поли-

плоидные; как уже было упомянуто, это привело бы к отсутствию неповторяющейся ДНК. Таким образом, данные по кинетике реассоциации не согласуются с моделями, объясняющими парадоксальные значения величины C в результате простого увеличения числа копий каждого гена в гаплоидном геноме. Это означает, что мы должны объяснять различия в размерах геномов тем, что геномам больших размеров действительно присуще большее разнообразие последовательностей.

Геномы эукариот содержат повторяющиеся последовательности

Какова природа фракций, ренатурирующих быстрее, чем неповторяющаяся (медленно ренатурирующая) ДНК? В приведенном на рис. 17.3 примере промежуточные последовательности составляют 30% генома, что (исходя из химической сложности) должно соответствовать общему количеству материала, равному $0,3 \cdot 7 \cdot 10^8 = 2,1 \cdot 10^8$ п. н. Но значение кинетической сложности составляет только $6 \cdot 10^5$ п. н.

Таким образом, длина уникальной ДНК, соответствующая $C_{0t_{1/2}}$ реассоциации, существенно меньше общей длины ДНК, соответствующей по химическому составу этой фракции. Иными словами, промежуточный компонент ведет себя так, как будто он включает последовательность длиной $6 \cdot 10^5$ п. н., представленную 350 копиями на геном (поскольку $350 \cdot 6 \cdot 10^5 = 2,1 \cdot 10^8$). После денатурации одиночные цепи каждой из этих копий могут реассоциировать с комплементарными им цепями любой из 350 копий. Это значительно повышает концентрацию последовательностей, участвующих в реакции реассоциации, и объясняет, почему эта фракция ренатурирует при более низких значениях $C_{0t_{1/2}}$.

Последовательности, представленные в геноме более чем одной копией, называются **повторяющейся ДНК**. Количество копий на геном называют **частотой повторяемости** (f). Формально частота повторяемости выражается отношением

$$f = \frac{\text{Химическая сложность}}{\text{Кинетическая сложность}} \quad (6)$$

Сложность и частота повторяемости вместе характеризуют свойства фракций последовательностей генома. Например, если геном состоит из одной последовательности длиной A п. н., 10 копий последовательности длиной B п. н. и 100 копий последовательности длиной V п. н., то его сложность равняется $A + B + V$. Частоты повторяемости последовательностей A , B и V составляют соответственно 1, 10 и 100. Последовательность A — неповторяющаяся, последовательности B и V — повторяющиеся.

Фракция повторяющейся ДНК включает все последовательности, частоты повторяемости которых существенно выше, чем 1. На практике неповторяющимися считают последовательности, частота повторяемости которых составляет от 1 до 2. (Последовательность не может быть представлена в гаплоидном геноме нецелое число раз!). Для примера, приведенного на рис. 17.3, уравнение 6 дает величину f , равную $3,15 \cdot 10^8 / 3,0 \cdot 10^8 = 1,05$. Обычно диапазон значений этой величины для неповторяющейся ДНК реальных геномов составляет от 0,8 до 1,6.

Повторяющаяся ДНК любого генома может состоять из нескольких компонентов, но в зависимости от диапа-

зона значений C_0t , при которых происходит реассоциация, различают два основных типа повторов. Они примерно соответствуют промежуточной и быстро ренатурирующей фракциям, изображенным на рис. 17.3. Говорят, что промежуточная фракция, обычно ренатурирующая при значениях C_0t в диапазоне от 10^{-2} до значений, соответствующих неповторяющейся ДНК, состоит из **умеренно повторяющейся ДНК**, а быстро ренатурирующая (при значениях C_0t менее 10^{-2}) фракция состоит из **высокоповторяющейся ДНК**.

Высокоповторяющаяся ДНК обязана своим названием очень большому числу копий основной реассоциирующей последовательности. Обычно эта последовательность довольно короткая. В примере, приведенном на рис. 17.3, быстро ренатурирующая фракция состоит из последовательности длиной только 340 п.н., присутствующей в количестве 500 000 копий на геном. Из-за небольшой длины такие последовательности иногда называют **простыми последовательностями ДНК** (*simple sequence DNA*).

Вместо того чтобы проводить описанную выше серию вычислений, частоту повторяемости любого компонента кривой реассоциации можно определить непосредственно путем сравнения с компонентом неповторяющейся ДНК. Зависимость между частотами повторяемости любых двух компонентов одной и той же кривой реассоциации обратно пропорциональна отношению их значений $C_0t_{1/2}$. Поэтому, если допустить, что неповторяющаяся ДНК действительно уникальна (ее величина f равна 1,0), можно написать другое уравнение для определения частоты любого повторяющегося компонента:

$$f = \frac{C_0t_{1/2} \text{ компонента неповторяющейся ДНК}}{C_0t_{1/2} \text{ компонента повторяющейся ДНК}} \quad (7)$$

К примеру, среднее значение величины f для промежуточного компонента кривой реассоциации, приведенной на рис. 17.3, составляет $630/1,9 = 332$. (Это значение несколько отличается от значения, вычисленного по уравнению 6 и соответствующего значению f для неповторяющейся ДНК, равному 1,05)

Умеренно повторяющаяся ДНК состоит из множества различных последовательностей

Фракция повторяющейся ДНК ведет себя в соответствии с усредненными характеристиками, полезными при описании последовательностей, входящих в ее состав. Соответствующие параметры не обязательно характеризуют свойства какой-либо определенной последовательности. В состав фракции умеренно повторяющейся ДНК, представленной на рис. 17.3, входит ДНК общей длиной $6 \cdot 10^5$ п.н., повторенная в геноме около 350 раз. Но эта цифра не соответствует длине целостной, реально существующей, непрерывной ДНК. Напротив, в состав этой фракции входит множество индивидуальных последовательностей, каждая из которых намного короче, а общая длина достигает $6 \cdot 10^5$ п.н. Средняя повторяемость этих последовательностей составляет 350, но одни из них представлены большим числом копий, а другие — меньшим.

При изучении кинетики реассоциации ДНК эукариотического генома фракции индивидуальных последовательностей редко разделяются так же хорошо, как это показана

но на рис. 17.3. В действительности они часто слишком перекрываются, чтобы их можно было различить на глаз. На рис. 17.5 приведен пример такого рода. ДНК генома морского ежа *S. purpuratus* ренатурирует в пределах значений C_0t , различающихся на восемь порядков; при этом видны точки перегиба, разделяющие разные компоненты, но процентное содержание каждого и его $C_0t_{1/2}$ сразу определить нельзя.

Метод, применяемый для идентификации индивидуальных компонентов, основан на использовании программы для ЭВМ, с помощью которой пытаются подобрать индивидуальные кривые (для гипотетических, кинетически однородных фракций) для каждой области кривой реассоциации. Эти индивидуальные кривые перекрываются и в сумме образуют кривую, отражающую наблюдаемое поведение последовательностей генома. Как отмечалось ранее, если приготовить смесь препаратов ДНК с определенными значениями сложности, ее кривая реассоциации совпадет с экспериментальной кривой для геномной ДНК.

На рис. 17.5 приведен подтверждающий такое заключение пример, где три индивидуальные кривые представляют компоненты, составляющие 19, 27 и 50% всей ДНК (в соответствии со скоростью их реассоциации). В этом случае два первых компонента следует отнести к умеренно повторяющейся ДНК со значениями $C_0t_{1/2}$, равными 0,53 и 8,3; эти значения соответствуют сложности, равной $1,0 \cdot 10^6$ п.н., повторенных 160 раз, и сложности, равной $2,3 \cdot 10^7$ п.н., повторенных 10 раз. Третий компонент кривой — неповторяющаяся ДНК.

Обычный подход при компьютерном анализе состоит в поиске минимального числа индивидуальных кинетических компонентов, суммирование кривых реассоциации которых обеспечит достаточно хорошее приближение к экспериментальной кривой. Однако это не обязательно отражает биологическую структуру генома. Часто оказывается возможным получить столь же хорошее приближение и при большем числе компонентов. Например, данные, приведенные на рис. 17.5, можно также объяснить наличием трех умеренно повторяющихся компонентов, составляющих 10, 27 и 25% ДНК генома с частотами повторяемости 8000, 240 и 19. При таком анализе также получают меньшие значения размеров компонента неповторяющейся ДНК.

Реально существующая система может, таким образом, включать множество повторяющихся компонентов, реассоциирующих в диапазоне значений C_0t , значительно (от 10 до 20 000 раз) превышающем этот диапазон для неповторяющегося компонента.

Кинетические компоненты, подбираемые в ходе анализа кривой, не соответствуют отдельным фракциям повторяющихся последовательностей генома; это только полезное приближение для его описания.

Для большинства эукариотических геномов достаточно подобрать два или три компонента, чтобы описать весь набор повторяющихся последовательностей. Часть генома, приходящаяся на долю повторяющейся ДНК, варьирует в широких пределах. Для низших эукариот большая часть ДНК представлена неповторяющимися последовательностями и только 10–20% ДНК приходится на одну или более фракций умеренно повторяющейся ДНК. В клетках животных до половины всей ДНК составляют фракции умеренно повторяющейся и высокоповторяющейся ДНК. У растений и амфибий доля повторов может быть даже выше, иногда составляя большую часть

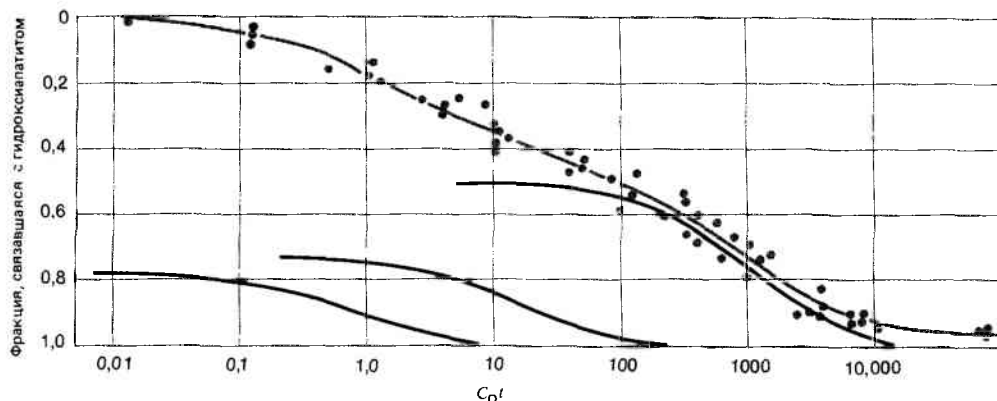


Рис. 17.5. При исследовании кинетики реассоциации ДНК *S. purpuratus* можно выделить несколько компонентов. Экспериментальная кривая проведена через точки, соответствующие из-

меренным значениям; кривые без экспериментальных точек соответствуют трем рассчитанным компонентам (минимальное число), которые в сумме могут образовать экспериментальную кривую.

генома, где на долю умеренно и высокоповторяющихся последовательностей приходится до 80% ДНК. Длина неповторяющейся ДНК имеет тенденцию к увеличению при возрастании размеров генома в целом, хотя прямо с этим не связана.

Это несколько уменьшает парадоксальность значений величины C в том смысле, что диапазон значений сложности неповторяющейся ДНК меньше диапазона значений размеров генома. Если предположить, что гены сосредоточены в пределах неповторяющейся ДНК, остается справедливым то, что следует по-прежнему найти объяснение существенным различиям в сложностях геномов, которые, по-видимому, не соответствуют морфологическим особенностям организма, но степень различий которых становится меньше. Например, сложность неповторяющейся ДНК плодовой мушки составляет $1 \cdot 10^8$ п. н., а ДНК комнатной мухи $2,8 \cdot 10^8$ п. н.; таким образом, наблюдается различие только в 3 раза, размеры же геномов различаются в 6 раз. Сложность неповторяющейся ДНК млекопитающих составляет около $9 \cdot 10^8$ п. н., лягушки *X. laevis* $1,6 \cdot 10^9$ п. н., а у растений и амфибий достигает примерно $9 \cdot 10^9$ п. н.

Наконец, следует подчеркнуть, что не существует одного определенного свойства, которое характеризовало бы все эукариотические геномы. Уже упоминалось, что в клетках растений, ставших полиплоидными недавно (по эволюционной шкале), уникальная ДНК может отсутствовать, поэтому наиболее медленно реассоцирующая фракция имеет частоту повторяемости 2–3 и более. (Действительно, это, по-видимому, временная стадия в эволюции, которая обычно приводит к дивергенции и появлению вновь уникальной ДНК.) В геномах ракообразных может совсем отсутствовать умеренно повторяющаяся ДНК, а присутствовать только высокоповторяющаяся и уникальная ДНК. У низших эукариот может отсутствовать высокоповторяющаяся ДНК.

Члены семейств повторяющихся последовательностей сходны, но не идентичны

Различные компоненты эукариотической ДНК могут быть выделены путем фракционирования двухцепочечных молекул в соответствии со значениями C_0t , необходимыми

для их ренатурации. Суть метода состоит в проведении реакции ренатурации только до достижения определенной величины C_0t . В этот момент ренатурировавшую ДНК выделяют из смеси на основе ее двухцепочечного строения; остальная часть ДНК остается одноцепочечной. Обычно двухцепочечные и одноцепочечные молекулы разделяют на колонках с гидроксипатитом, который преимущественно связывает двухцепочечную ДНК.

Поведение ренатурировавшей неповторяющейся ДНК очень похоже на поведение исходного препарата ДНК до его денатурации. При проведении денатурации ренатурировавший препарат быстро плавится при $T_{пл}$, значение которой чуть ниже значения $T_{пл}$ исходной нативной ДНК. Это означает, что реассоциация цепей прошла точно. Каждая уникальная последовательность сгибридовалась с полностью комплементарным ей участком.

Совершенно иначе ведет себя ренатурировавшая повторяющаяся ДНК, которая медленно плавится в пределах широкого интервала температур, как показано на рис. 17.6. Это означает, что она состоит из молекул, спаренных не совсем точно. Напротив, в этом случае в ренатурировавших двухцепочечных молекулах должно быть значительное количество ошибок спаривания. Действительно, чем больше имеется ошибок спаривания в молекуле, тем меньше водородных связей должно быть разрушено при плавлении и поэтому тем ниже $T_{пл}$. Ширина кривой плавления свидетельствует о том, что ренатурировавшая повторяющаяся ДНК содержит спектр последовательностей, начиная с последовательностей, образовавшихся при реассоциации лишь частично комплементарных участков, и кончая последовательностями, образовавшимися при реассоциации почти или даже полностью комплементарных участков ДНК. Как это может произойти?

Фракция повторяющейся ДНК состоит из семейств сходных, но не идентичных последовательностей. В состав каждого семейства входит набор нуклеотидных последовательностей, имеющих достаточное сходство, чтобы ренатурировать друг с другом. Различия же обусловлены заменами оснований, вставками и делециями, приводящими к появлению внутри сходных последовательностей точек, в которых не может происходить спаривание оснований комплементарных цепей. Число таких

изменений определяет степень взаимодействия между двумя любыми последовательностями. При ренатурации двух членов семейства, обладающих большим сходством, образуется дуплекс с высокой $T_{пл}$; при ренатурации последовательностей с меньшим сходством образуется дуплекс с более низкой $T_{пл}$.

Способность сходных, но не идентичных комплементарных последовательностей узнавать друг друга может определяться жесткостью условий проведения реассоциации. Более жесткие условия обеспечиваются, например, повышением температуры, препятствующей спариванию не полностью комплементарных участков ДНК. Так, проводя реакцию гибридизации при более высоких температурах, можно добиться того, что будут реассоциировать только члены семейства, обладающие довольно значительным сходством. При более низких температурах может происходить отжиг и менее сходных членов семейства. Обратите также внимание на тот факт, что при увеличении жесткости условий гибридизации возрастает также доля последовательностей генома, ведущих себя как неповторяющаяся ДНК.

При исследовании семейств повторов обнаруживается большое разнообразие в их строении. В некоторых случаях можно четко выделить семейства, члены которых имеют большое сходство. Такие семейства выявляются даже при увеличении жесткости условий гибридизации. В других случаях обнаруживается множество последовательностей, обладающих различной степенью сходства, так что при увеличении жесткости условий гибридизации размер семейства постоянно уменьшается. Определение размера таких семейств повторяющихся последовательностей является произвольным, поскольку зависит от конкретных условий гибридизации. Эти данные говорят о том, что геномы эукариот содержат сходные последовательности, произошедшие, по всей вероятности, путем дивергенции из некоторой общей исходной последовательности, количество копий которой в геноме увеличилось во много раз; степень сходства копий, однако, очень сильно варьирует.

Реассоциация двух сходных, но не идентичных последовательностей происходит медленнее, чем реассоциация идентичных последовательностей. Это означает, что для реассоциации сходных последовательностей требуется большее значение C_0t . Поэтому значения $C_0t_{1/2}$, наблюдаемые для фракций повторов, могут быть выше, чем это соответствует их действительной сложности. Таким образом, большинство фракций повторяющихся последовательностей имеют более низкие значения сложности и большую частоту повторяемости, чем это следует из кинетики их реассоциации. Мы увидим, что это особенно характерно для высокоповторяющейся ДНК (гл. 24).

Участки умеренно повторяющейся ДНК чередуются с участками неповторяющейся ДНК

В большинстве геномов ни умеренно повторяющаяся, ни уникальная ДНК не существуют в виде длинной непрерывной последовательности, не прерываемой другими

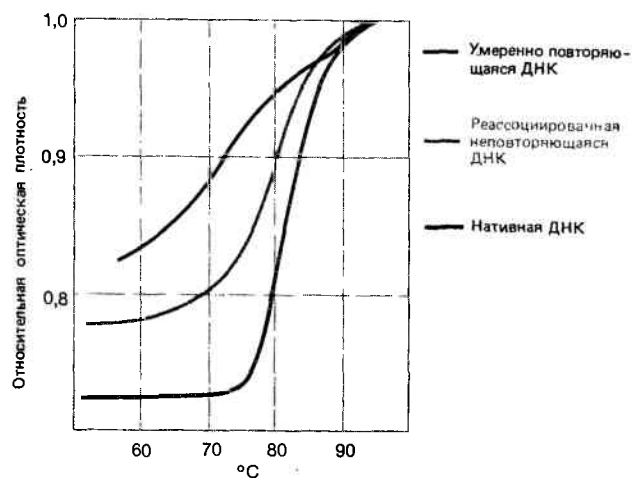


Рис. 17.6. Денатурация реассоциировавшей неповторяющейся ДНК происходит в узком интервале температур, близком к таковому для нативной ДНК, но реассоциировавшая умеренно повторяющаяся ДНК плавится в пределах широкого температурного интервала.

компонентами генома. Напротив, оба этих компонента присутствуют в виде отдельных последовательностей, чередующихся друг с другом. Во многих случаях их чередование характеризуется определенной регулярностью.

Наиболее распространено чередование коротких последовательностей, при котором умеренно повторяющаяся ДНК присутствует в геноме в виде отдельных последовательностей со средней длиной около 300 п.н. Эти последовательности перемежаются с неповторяющимися последовательностями, средняя длина которых колеблется в различных геномах от 800 до 1500 п.н. Последовательности, расположенные таким образом, могут составлять около половины всего генома и обычно включают большую часть как умеренно повторяющейся, так и уникальной ДНК.

В некоторых геномах обнаруживается чередование длинных последовательностей, когда также наблюдается чередование участков умеренно повторяющейся и неповторяющейся ДНК, но последовательности каждого типа имеют существенно большую длину. Ни при одном из способов чередования не обнаруживается каких-либо важных свойств, которые объяснялись бы именно этим способом.

Рекомендуемая литература

Основным источником данных, приведенных в этой главе, послужила книга Б. Льюина (*Lewin, Gene expression, 2, Eucaryotic Chromosomes, Wiley, New York, 1980, pp. 479–530*). Технику проведения реассоциации разработали Бриттен и Кон (*Britten, Kohne, Science, 1968, т. 161, pp. 529–540*), а последующие результаты приведены в обзорах Бриттена и Дэвидсона (*Britten, Davidson, Quart. Rev. Biol., 1971, 46, pp. 111–133* и *Quart. Rev. Biol., 1973, 48, pp. 565–613*).

Глава 18

СТРУКТУРНЫЕ ГЕНЫ:
КАК ОНИ ПРЕДСТАВЛЕНЫ В мРНК

Наследование простых признаков по Менделю предполагает существование в гаплоидном геноме единственной копии каждого детерминирующего фактора. Фактор может располагаться в определенном локусе, и самое простое предположение состоит в том, что каждый такой локус представляет собой последовательность ДНК, кодирующую единственный белок; примером этого служит определение гена, данное в гл. 2. Вот классический взгляд на структурный ген: уникальный компонент генома – единственная последовательность, кодирующая свой белковый продукт, которую поэтому можно идентифицировать с помощью мутаций, нарушающих функцию белка.

Но имеется также и другой класс генов, существующих в виде множества последовательностей, кодирующих один и тот же белок (или иногда белки, обладающие большим сходством). Такие гены бывает трудно или даже невозможно выявить с помощью мутаций, поскольку инактивация одной из копий гена не нарушит работу остальных. Генетические данные, таким образом, относятся скорее к генам первого типа, поэтому нам следует обратиться к прямому анализу ДНК генома для определения числа и процентного содержания в геноме уникальных и повторяющихся генов.

Для выявления структурных генов и определения их свойств идеальным посредником является мРНК. С одной стороны, белок, которому она соответствует, может быть получен путем ее трансляции. С другой стороны, гибридизуя мРНК с геномной ДНК, можно выделить ген, с которого она считана. Индивидуальная мРНК служит как бы рукояткой для поворота исследований в обратном направлении: от белка к гену. Набор последовательностей мРНК, обнаруживаемых в полисомах, соответствует полному набору генов, экспрессирующихся в клетке или ткани. Таким образом, состав популяции мРНК отражает как природу, так и число структурных генов. Используя метод гибридизации нуклеиновых кислот, можно прийти к решению ряда основных вопросов. Сколько существует копий каждого гена? Сколько генов экспрессируется в клетках определенного типа? Какова степень перекрытия наборов генов, экспрессия которых определяет различия между клетками разных типов?

Являются ли структурные гены
уникальными или повторяющимися?

В зависимости от входящих в состав эукариотического генома структурных генов можно представить себе три возможных способа его организации.

Структурный ген может быть уникальным и состоять из последовательности ДНК, представляющей в геноме специфический белок и не похожей ни на какую другую последовательность генома. Такая последовательность, несомненно, будет входить в состав фракции неповторяющейся ДНК.

В другом случае структурный ген может быть уникальным в том смысле, что он действительно представляет собой последовательность ДНК, кодирующую специфический белковый продукт, но в геноме могут присутствовать и другие последовательности, сходные с этим геном и кодирующие сходные белки. Такая ситуация наблюдается, когда кодирующая последовательность дублируется с последующей дивергенцией ее копий и появлением разных белков. В зависимости от степени сходства между последовательностями ДНК, числа копий и точно не предсказуемых последствий ошибок спаривания при реассоциации такие последовательности могут обнаруживаться во фракциях как неповторяющейся, так и умеренно повторяющейся ДНК. Еще один возможный способ организации генома эукариот состоит в том, что структурные гены могут повторяться: может существовать более одной копии последовательности, кодирующей специфический белок. Вследствие вырожденности генетического кода последовательности ДНК, кодирующие один и тот же белок, в действительности могут быть неидентичными. Однако любая последовательность, представленная небольшим числом копий (более двух или трех), будет вести себя как составная часть фракции умеренно повторяющейся ДНК. (Понятно, что такие последовательности могут составлять только часть этой фракции, поскольку известно, что основная часть умеренно повторяющейся ДНК состоит из отдельных последовательностей, слишком коротких, чтобы кодировать белки, и чередующихся с уникальной ДНК.)

Исследуя кинетику гибридизации, можно определить число копий каждого гена, соответствующее числу видов мРНК в популяции или числу индивидуальных мРНК. Говоря точнее, такой анализ дает ответ на вопрос: какие последовательности генома, уникальные или повторяющиеся, представлены в мРНК?

Поскольку этот метод имеет ограниченную разрешающую способность, структурные гены, обнаруживаемые в составе неповторяющейся ДНК, не обязательно являются уникальными. Они могут быть представлены небольшим числом копий, но это число несомненно должно быть меньше, чем, скажем, три или четыре. Для определения точного числа копий необходимо выделить ДНК, соответствующую индивидуальным мРНК. Это требует перехода от метода исследования кинетики гибридизации к методам, описанным в следующей главе.

Для структурных генов, относящихся к повторяющейся ДНК, частота повторяемости последовательности ДНК позволяет только примерно оценить число генов. Для определения числа генов и их взаимосвязей опять-таки необходимо выделить индивидуальные члены семейства. Свойства повторяющейся ДНК в целом соответствуют свойствам семейств, состоящих из сходных, но не идентичных последовательностей, что, однако, не исключает наличия в составе этой фракции некоторых семейств последовательностей, все члены которых идентичны или обладают очень большим сходством.

Большая часть структурных генов относится к неповторяющейся ДНК

Используя РНК в качестве зонда в эксперименте по реассоциации, можно определить, какие последовательности генома представлены в мРНК. Очень небольшое количество радиоактивной РНК (или ДНК) вносят в реакцию вместе с существенно большим количеством клеточной ДНК. Поскольку ДНК клетки присутствует в огромном избытке, ее часть, образовавшая гибриды с РНК-зондом, не изменит концентрации последовательностей одноцепочечной ДНК. Определяющим фактором реакции является реассоциация комплементарных цепей ДНК клетки (т.е. как будто РНК-зонд отсутствует). Такие реакции называют **реакциями, определяемыми концентрацией ДНК**. Реассоциацию всей ДНК определяют обычными способами (по изменению оптической плотности раствора или по связыванию с гидроксипапитом).

Степень гибридизации зонда определяют, измеряя радиоактивность двухцепочечных молекул.

РНК- (или ДНК-) зонд участвует в реакции так, как если бы он просто входил в состав фракции последовательностей, с которой он транскрибировался. Эту фракцию выявляют, сравнивая кривую гибридизации радиоактивного зонда с кривой реассоциации всей ДНК. Поэтому значения C_0t , при которых гибридизуется меченая РНК, могут быть использованы для определения частот повторяемости соответствующих последовательностей генома. Когда зондом служит индивидуальная мРНК, она гибридизуется при единственном значении $C_0t_{1/2}$, зависящем от частоты повторяемости гена или генов, представляющих ее в геноме. Для уникального гена эта величина будет такой же, как для неповторяющейся ДНК, в то время как для повторяющихся генов будет наблюдаться соответствующее уменьшение значения $C_0t_{1/2}$. При использовании популяции молекул мРНК каждая мРНК гибридизуется при определенном значении $C_0t_{1/2}$ так, что кривая в целом представляет собой сумму кривых гибридизации индивидуальных компонентов. Такую кривую можно разложить на составляющие ее компоненты так же, как и кривую реассоциации самой ДНК генома.

Типичная кривая гибридизации при использовании популяции молекул мРНК выглядит так, как показано на **рис. 18.1**. Небольшая часть РНК, обычно 10% или меньше, гибридизуется при значении $C_0t_{1/2}$, характерном для умеренно повторяющихся последовательностей. Основная часть РНК гибридизуется при значениях $C_0t_{1/2}$ идентичных или очень близких к значениям $C_0t_{1/2}$ для уникальной ДНК. Обычно эта часть составляет до 50% всей РНК.

Какая связь существует между последовательностями мРНК, гибридизующимися с неповторяющейся и повторяющейся ДНК? Такие последовательности могут быть разделены на разные классы (в зависимости от того, какая РНК гибридизуется первой). Это указывает на независимое участие в гибридизации молекул разных классов: один класс соответствует генам, относящимся к неповторяющейся ДНК, а другой — генам, относящимся к повторяющейся ДНК.

Основная сложность при интерпретации этих данных связана с выяснением природы вообще не гибридизующихся последовательностей, которые часто составляют до половины всей РНК. Возможно, что большая часть

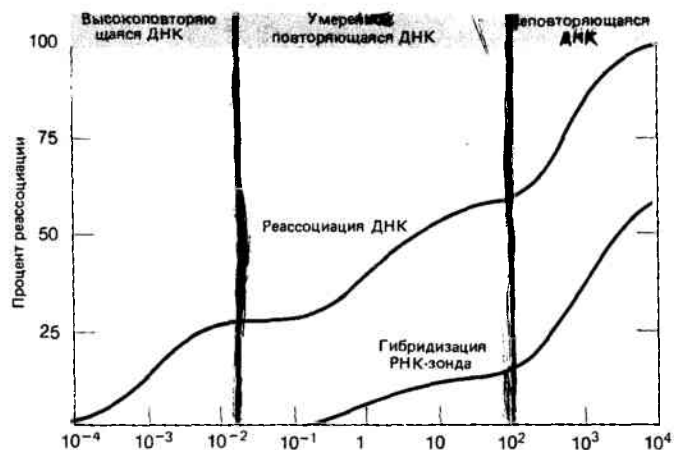


Рис. 18.1. Сопоставление кривой гибридизации мРНК-зонда с кривой реассоциации указывает на то, что основная часть последовательностей мРНК считается с неповторяющейся ДНК, а остальные последовательности — с умеренно повторяющейся ДНК. На высокоповторяющейся ДНК мРНК не синтезируются.

такого материала соответствует уникальным последовательностям. Одна из причин, по которой эти молекулы не вступают в гибридизацию, может состоять в том, что они относятся к мРНК, присутствующей в очень больших количествах, достаточных, чтобы предположение о том, что в зонде их мало по сравнению с ДНК, оказалось неверным. Иными словами, если в реакции, определяемой концентрацией ДНК, нет достаточного избытка ДНК, некоторые РНК могут не вступить в гибридизацию.

Из этих данных следует, что большую часть мРНК, возможно до 80%, составляют последовательности, гибридизующиеся с последовательностями неповторяющейся ДНК. Поскольку очень низкую степень повторяемости генов трудно обнаружить (особенно когда их копии сходны, но не идентичны, так что реассоциация замедляется из-за ошибок спаривания), то можно утверждать только, что они присутствуют в геноме в количестве менее трех или четырех копий, но не то, что они абсолютны уникальны. Лишь небольшую часть генов можно безусловно отнести к повторяющейся ДНК; являются ли такие множественные копии идентичными или просто сходными, при исследовании кинетики гибридизации выявить не удастся. Обратите внимание на то, что, поскольку эти эксперименты рассматриваются с точки зрения *общего количества мРНК*, а разные гены экспрессируются с образованием разных количеств мРНК, этим методом нельзя подтвердить, что 80% генов относится к фракции неповторяющейся ДНК, хотя, по-видимому, большая их часть к этой фракции относится.

Была исследована кинетика гибридизации индивидуальных мРНК, соответствующих некоторым генам; наиболее интересные результаты получены для α - и β -глобиновых генов нескольких видов млекопитающих. Полученные данные обычно указывают на существование одной или двух копий каждого гена. На самом деле эти данные дают заниженную оценку числа сходных последовательностей в геноме, поскольку в экспериментах по прямому выделению генов часто обнаруживаются дополнительные сходные с ними последовательности (гл. 21).

Сколько уникальных генов экспрессируется?

Число последовательностей неповторяющейся ДНК, представленных в виде РНК, может быть определено непосредственно как доля ДНК, способной гибридизоваться с РНК. При гибридизации небольшого количества одноцепочечной ДНК с большим количеством РНК все последовательности ДНК, комплементарные РНК, вступят в реакцию с образованием гибрида РНК–ДНК. Такой метод называется **методом насыщающей гибридизации**, поскольку его основной признак – наличие избытка РНК, достаточно большого для того, чтобы каждая присутствующая комплементарная последовательность ДНК вступила в гибридизацию. Основным параметр такой реакции, **определяемой концентрацией РНК**, – произведение концентрации РНК на время, называемое R_0t . Этот параметр в точности аналогичен величине C_0t , используемой для описания реакций, определяемых концентрацией ДНК.

Обычно реакцию насыщающей гибридизации изображают в виде графика, на котором процентам гибридизации соответствуют значения R_0t . Такого рода кривая представлена на **рис. 18.2**. В данном случае реакция завершается при значении R_0t , равном примерно 300, но ее необходимо проводить и дальше, чтобы быть уверенными в достижении плато. При дальнейшем увеличении R_0t до 1200 ДНК более не вступает в гибридизацию; это показывает, что гибридизовалась действительно вся ДНК, которая могла вступить в реакцию.

Зная процент гибридизации ДНК, можно определить число генов, представленных в популяции мРНК и участвующих в реакции. При достижении насыщения в гибридизацию вступает 1,35% присутствующей неповторяющейся ДНК. Поскольку только одна цепь ДНК транскрибируется с образованием РНК, лишь половина ДНК в принципе может гибридизоваться с РНК (другая половина идентична РНК по последовательности оснований). Поэтому мРНК представляет 2,7% всех последовательностей неповторяющейся ДНК.

Собственно неповторяющаяся ДНК составляет 75% генома размером $8,1 \cdot 10^8$ п.н. Следовательно, сложность ДНК, соответствующей популяции молекул РНК, состав-

ляет $0,027 \cdot 0,75 \cdot 8,1 \cdot 10^8 = 1,7 \cdot 10^7$ п.н. Популяция молекул РНК представлена препаратом мРНК полисом, средняя длина которой составляет 2000 оснований. Таким образом, общее число различных мРНК составляет $1,7 \cdot 10^7 / 2000 = 8500$. Это соответствует общему числу генов, экспрессирующихся в ткани, из которой была выделена мРНК, – из эмбриона морского ежа, находящегося на стадии гастрюлы.

Аналогичные эксперименты были проведены для большого числа организмов. Для низших эукариот, например дрожжей, общее число экспрессирующихся генов относительно невелико – около 4000. Для соматических клеток высших эукариот эта величина составляет от 10 000 до 15 000 генов. Так обстоит дело и у позвоночных, и у растений. Следовательно, общее количество ДНК, представленное в виде мРНК, обычно составляет очень небольшую часть генома, порядка 1–2%. Важное исключение составляет мозг млекопитающих, в котором, по-видимому, экспрессируется значительно большее число генов, хотя точно это число не определено.

Такие эксперименты можно проводить только с неповторяющейся ДНК, когда каждая последовательность может гибридизоваться только со считанной с нее мРНК. Из-за присутствия во фракции умеренно повторяющейся ДНК множества копий идентичных или сходных последовательностей, РНК, транскрибированные с этой ДНК, кроме специфической последовательности, с которой они были считаны, могут гибридизоваться и с другими последовательностями генома. Поэтому при достижении насыщения реакции гибридизации в гибриде может содержаться большое количество дополнительных последовательностей ДНК, что приведет к завышенной оценке числа экспрессирующихся генов. Действительно, для того чтобы быть уверенными, что этого не произошло, ДНК, гибридизовавшуюся с РНК, обычно выделяют (разрушая РНК) и затем используют в экспериментах по реассоциации, чтобы убедиться в том, что значение $C_0t_{1/2}$ действительно соответствует ожидаемому для неповторяющейся ДНК. Такой эксперимент называют **прямым воспроизведением**. Таким образом, число экспрессирующихся генов, определенное этим методом, относят к основной массе генов, последовательности которых представлены в неповторяющейся ДНК.

Оценка числа генов по кинетике реакции, определяемой концентрацией РНК

Поскольку в реакциях, определяемых концентрацией РНК, РНК находится в избытке по сравнению с ДНК, связывание небольшого количества РНК при образовании гибрида в ходе реакции существенно не изменяет ее концентрации. Это означает, что кинетика реакции насыщающей гибридизации соответствует кинетике реакций первого порядка, описываемой уравнением

$$\frac{R}{R_0} = e^{-k \cdot R_0t}, \quad (8)$$

так что, когда реакция завершена наполовину и $R/R_0 = 0,5$,

$$k = \frac{\ln 2}{R_0t_{1/2}}, \quad (9)$$

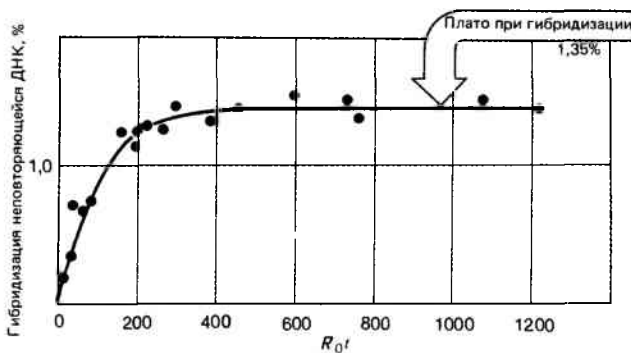


Рис. 18.2. При гибридизации избытка мРНК с неповторяющейся ДНК до достижения насыщения реакции гибридизации обнаруживается, что только небольшая часть ДНК представлена в мРНК.

Эксперимент проводили с использованием очищенной неповторяющейся ДНК, несущей радиоактивную метку; в процессе реакции определяли процент метки, включившейся в гибрид. На рисунке приведены данные по гибридизации неповторяющейся ДНК морского ежа *S. purpuratus* с мРНК, выделенной из полисом эмбриона на стадии гастрюлы.

Однако величину $R_{0t_{1/2}}$, полученную в реакции насыщающей гибридизации, определяемой концентрацией РНК, нельзя использовать для определения сложности популяции РНК. Причина состоит в том, что различные последовательности РНК присутствуют в разных концентрациях, определяемых уровнем экспрессии соответствующих генов. Это означает, что значение R_0 , определенное для всей массы РНК, неприменимо к каждой последовательности в отдельности. В действительности весьма значительная часть всей мРНК обычно представлена всего лишь несколькими типами последовательностей. Эти последовательности очень быстро гибридизуются со всеми комплементарными им последовательностями в ДНК, после чего ход реакции определяется концентрацией оставшихся последовательностей. Таким образом, реальное количество РНК, определяющее скорость реакции, гораздо меньше измеренного значения R_0 . Это означает, что измеренное значение $R_{0t_{1/2}}$ сильно завышено: оно включает большое количество РНК, не участвующей в реакции, поскольку комплементарные ей последовательности очень быстро включаются в состав гибрида.

Для использования данных по кинетике гибридизации, определяемой концентрацией РНК, при оценке сложности был разработан другой метод. Последовательность операций этого метода приведена на рис. 18.3. Он включает использование меченой кДНК, полученной на матрице этой мРНК в реакции обратной транскрипции. Мы рассмотрим этот метод в гл. 19; все, что необходимо знать сейчас, — это то, что кДНК представляет собой одноцепочечную ДНК, комплементарную мРНК.

Реакция гибридизации происходит между избытком мРНК и меченой кДНК, предварительно полученной на матрице этой мРНК. Каждая последовательность мРНК представлена в популяции молекул кДНК в количестве, соответствующем ее процентному содержанию в РНК. Поскольку все последовательности кДНК получены на матрице мРНК, вся меченая кДНК должна участвовать в образовании гибридов. Для каждой отдельной последовательности величина $R_{0t_{1/2}}$ реакции пропорциональна сложности популяции молекул РНК, аналогично тому как величина $C_{0t_{1/2}}$ зависит от сложности генома в реакции реассоциации. Поэтому сложность некоторой популяции молекул РНК может быть определена путем сравнения значения $R_{0t_{1/2}}$ со значением $R_{0t_{1/2}}$ стандартной реакции. В качестве стандарта, например, часто используют реакцию избытка глобиновой мРНК с ее кДНК. В этом случае можно составить следующее уравнение:

$$\frac{\text{Сложность любой РНК}}{\text{Сложность глобиновой мРНК}} = \frac{R_{0t_{1/2}} \text{ любой РНК}}{R_{0t_{1/2}} \text{ глобиновой мРНК}} \quad (10)$$

Это уравнение полностью аналогично уравнению 5 в гл. 17, определяющему сложность препаратов ДНК путем сравнения значений $C_{0t_{1/2}}$ исследуемых и стандартных препаратов ДНК. При использовании R_{0t} -анализа рассчитанная сложность также соответствует общей длине различных последовательностей без учета организации индивидуальных последовательностей в геноме.

Так же как и в случае кривой реассоциации ДНК, каждый компонент гибридизуется в пределах значений

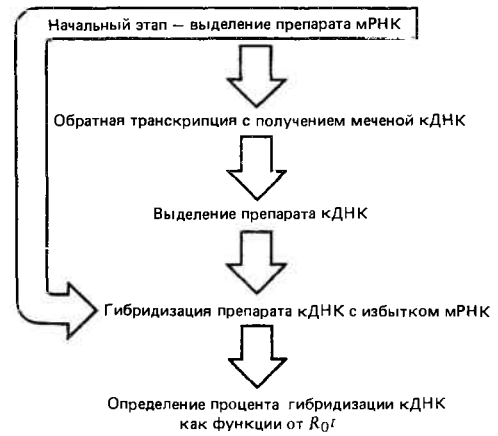


Рис. 18.3 Гибридизацию избытка мРНК с кДНК используют для определения сложности мРНК.

R_{0t} , различающихся на два порядка, и для кривой реакции, осуществляющейся в более широком диапазоне значений этой величины, необходимо применение компьютера, чтобы разложить ее на индивидуальные компоненты. И опять кривая гибридизации отражает существование целого спектра частот встречаемости последовательностей. При определении сложности каждого компонента он рассматривается как единственный в смеси.

Для этого измеренное значение $R_{0t_{1/2}}$ приводят к значению $R_{0t_{1/2}}$, ожидаемому для чистого компонента: полученное в реакции значение $R_{0t_{1/2}}$ умножают на процент содержания данного компонента в реакционной смеси. Фактически этим корректируется значение R_0 так, чтобы оно соответствовало количеству РНК, представленному только рассматриваемым компонентом.

Пример реакции гибридизации избытка мРНК с кДНК, в которой образуются три четко различимые фракции, приведен на рис. 18.4. Контролем в этой реакции служила гибридизация чистой овальбуминовой мРНК (длиной 2000 п. н.) с овальбуминовой кДНК, значение $R_{0t_{1/2}}$ реакции составляло 0,0008. Это означает, что сложности, равной 1000 п. н., соответствует значение $R_{0t_{1/2}}$, равное 0,0004.

Первый компонент составляет 50% от всей РНК, участвующей в реакции, и значение $R_{0t_{1/2}}$ для него равняется 0,0015. Тогда значение $R_{0t_{1/2}}$, соответствующее чистому первому компоненту, составит $0,5 \cdot 0,0015 = 0,00075$, что почти точно совпадает со значением $R_{0t_{1/2}}$ для мРНК овальбумина. Это означает, что первый компонент фактически представляет собой овальбуминовую мРНК! Последнее хорошо согласуется с известными свойствами ткани яйцеводов, где овальбумин действительно является основным продуктом экспрессии генов, и овальбуминовая мРНК составляет около половины всех мРНК клетки.

Для второго компонента значение $R_{0t_{1/2}}$ составляет 0,04, и на его долю приходится 15% участвующей в реакции РНК. Следовательно, значение $R_{0t_{1/2}}$ для чистого второго компонента составит $0,15 \cdot 0,04 = 0,006$. Таким образом, его сложность равна $0,006/0,0004 \cdot 1000 = 15000$ оснований. Это соответствует наличию в составе второго компонента 7–8 видов мРНК со средней длиной 2000 оснований.

Последний компонент имеет значение $R_{0t_{1/2}}$, равное 30, и на его долю приходится 35% РНК, так что значение

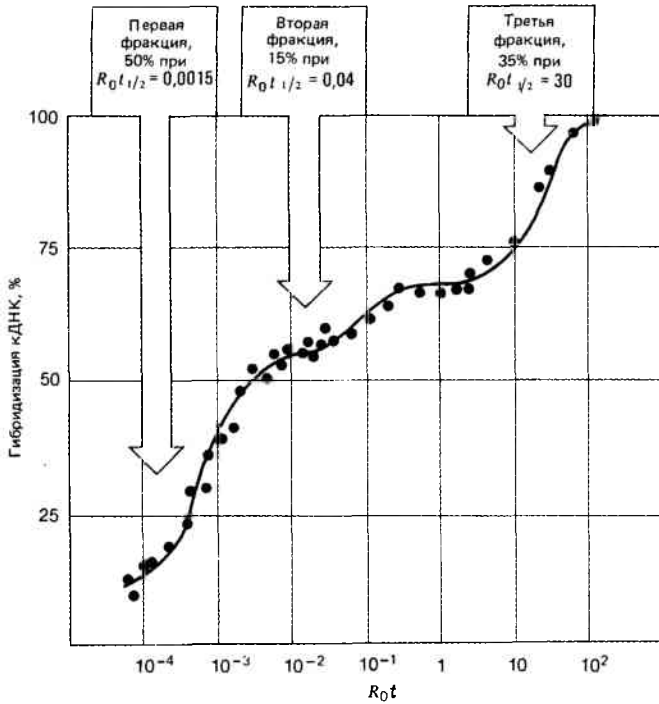


Рис. 18.4. При гибридизации избытка мРНК с кДНК обнаруживается несколько фракций РНК, каждая из которых характеризуется величиной $R_0 t_{1/2}$ реакции.

Степень гибридизации определяется по конечному количеству вступившей в реакцию кДНК. Обычно эта величина близка к 100% и почти всегда превышает 90%. На рисунке приведены данные по гибридизации мРНК, полученной из яйцевода кур.

$R_0 t_{1/2}$, соответствующее чистому компоненту, составляет $0,35 \cdot 0,30 = 10,5$, а его сложность равняется $10,5 / 0,0004 \cdot 1000 = 26\,000\,000$ оснований. Это соответствует примерно 13 000 видов мРНК со средней длиной 2000 оснований.

Такой анализ показывает, что около половины всех мРНК клетки представлено одним видом мРНК, около 15% — не более чем 7–8 видами мРНК и около 35% представлено большим числом (13 000) видов мРНК. Отсюда ясно, что количества мРНК, относящихся к разным компонентам, должны сильно различаться.

Уровни экспрессии генов сильно различаются

Среднее число молекул каждого вида мРНК, приходящееся на клетку, называется ее **копийностью** (abundance) или **представленностью**. Эту величину очень просто определить, если известно общее количество РНК в клетке. Для каждого кинетического компонента (фракции) общее количество РНК = Представленность · Сложность, поэтому в общем виде уравнение выглядит следующим образом:

$$\text{Представленность} = \frac{\text{г мРНК на клетку} \cdot \text{Доля РНК в компоненте} \cdot 6 \cdot 10^{23}}{\text{Сложность компонента в дальтонах}}$$

(11)

Обычно уравнение имеет именно такой вид, поскольку общее количество мРНК, определенное химически, выра-

жают в пикограммах, в то время как сложность (определяемую с помощью гибридизационных методов) выражают в основаниях или дальтонах.

В примере, приведенном на рис. 18.4, количество мРНК на клетку составляет 0,275 пг. Это соответствует 100 000 копий первого компонента (овальбуминовой мРНК), 4000 копий каждого из 7–8 видов мРНК второго компонента и только примерно 5 копиям каждого из 13 000 видов мРНК, входящих в состав третьего компонента.

Яйцевод — в некотором роде исключительный случай, когда такое значительное количество мРНК представлено всего лишь одним видом мРНК. Но и в большинстве других клеток большую часть всех мРНК составляет фракция мРНК, представленная всего несколькими видами мРНК. Такая **многокопийная** фракция обычно состоит из менее чем 100 различных видов мРНК, присутствующих в количестве более 1000 или даже 10 000 копий на клетку. Она часто включает наиболее значительную часть всей мРНК, составляя до 50% ее общего количества. В этой фракции можно выделить два компонента (как в случае яйцевода), или можно считать, что она представляет собой один компонент, однако для активно транскрибирующихся мРНК каждый соответствующий им ген, по-видимому, имеет свой характерный уровень экспрессии, отличающийся от уровня экспрессии других генов.

Во всех случаях около половины общего количества мРНК состоит из большого числа последовательностей (порядка 10 000), каждая из которых представлена только небольшим числом копий, скажем менее 10. Это класс **редких** или **сложных** мРНК. (Именно этот класс определяет скорость протекания реакции насыщающей гибридизации.)

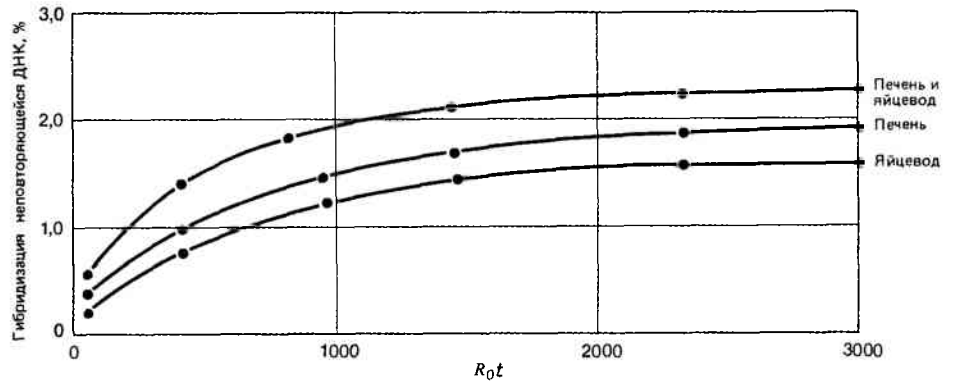
Оценки общего числа экспрессирующихся генов, полученные методом насыщающей гибридизации и кинетическим методом, обычно хорошо согласуются. Кинетический метод дает более низкие значения (поскольку некоторые, очень редкие последовательности могут не вступать в гибридизацию.) Метод насыщающей гибридизации дает более высокие значения (поскольку в гибридизации всегда могут участвовать некоторые последовательности, присутствующие в количестве более одной копии на геном). Поэтому для яйцевода птиц кинетическим методом обнаруживают около 13 000 видов мРНК, в то время как методом насыщающей гибридизации — 15 000 видов.

Перекрытие популяций мРНК

Число экспрессирующихся генов во многих соматических клетках высших эукариот колеблется в пределах от 10 000 до 20 000. Насколько перекрываются гены, экспрессирующиеся в разных тканях? Например, в печени цыпленка число экспрессирующихся генов составляет от 11 000 до 17 000, а в ткани яйцевода, как было только что упомянуто, — от 13 000 до 15 000. Сколько генов из этих двух наборов идентичны? Сколько генов специфичны для каждой ткани?

Сразу же можно сказать, что, по-видимому, имеются существенные различия между генами, экспрессирующимися с образованием класса мРНК с малым разнообразием видов («многокопийного» класса). Овальбумин, например, синтезируется только в яйцеводе и совсем не синтезируется в печени. Это означает, что 50% всей

Рис. 18.5. Перекрестная насыщающая гибридизация позволяет определить число различающихся последовательностей, присутствующих в двух препаратах мРНК.



мРНК яйцевода специфично для этой ткани. Если посмотреть шире, во многих случаях многокопийные мРНК транскрибируются с генов, кодирующих белки, которые являются основным продуктом клеток данного типа. Таким образом, могут обнаруживаться существенные различия в количестве мРНК и синтезируемого белка между клетками разных типов. Но многокопийные мРНК неизменно представляют только небольшую часть экспрессирующихся генов. Что касается общего числа генов организма и различий в транскрипции, которые должны существовать между клетками различных типов, то необходимо знать число генов, экспрессирующихся с образованием классов редких мРНК в клетках с разными фенотипами.

Для определения степени перекрытия популяций мРНК в целом не существует достаточно чувствительных методов, однако путем модификации методов, используемых для определения сложности популяций, удалось получить некоторые ориентировочные оценки.

Эксперименты по перекрестной насыщающей гибридизации позволяют определить количество последовательностей, различных для двух популяций мРНК. На рис. 18.5 в качестве примера приведены данные таких экспериментов с РНК печени и яйцевода цыпленка. При гибридизации препаратов мРНК, выделенных из этих тканей, с неповторяющейся ДНК насыщение реакции достигается, когда в гибридизацию вступают соответственно 2,05 и 1,80% ДНК. Если бы две популяции молекул мРНК были представлены только различающимися последовательностями, то вместе они насыщали бы 2,05 + 1,80 = 3,85% неповторяющейся ДНК. Однако реальное значение составляет 2,4%, что лишь ненамного выше значения, полученного для печени. Таким образом, около 75% последовательностей ДНК, экспрессирующихся в печени и яйцевоме, одинаковы (хотя, поскольку это эксперимент по насыщающей гибридизации, на основе приведенных данных не ясно, различается ли число копий этих последовательностей в двух тканях). Иными словами, около 12 000 генов экспрессируется и в печени, и в яйцевоме, еще около 5000 генов экспрессируется только в печени и около 3000 генов — только в яйцевоме.

Другой способ определения степени перекрытия мРНК в разных тканях начинается с гибридизации мРНК какой-либо ткани с неповторяющейся ДНК. Вступившую в реакцию ДНК затем выделяют, получая препарат мДНК, в то время как ДНК, не вступившая в гибридизацию, образует препарат «нулевой» ДНК. Далее полученные препараты ДНК по отдельности гибридизуют с избытком мРНК из какой-либо другой ткани. Доля сгибридизовавшейся мДНК соответствует доле ге-

нов, экспрессирующихся как во второй, так и в первой ткани. Количество вступившей в гибридизацию «нулевой» ДНК указывает на число генов, экспрессирующихся во второй, но не в первой ткани.

Эксперименты по перекрестной гибридизации можно также проводить, используя избыток мРНК из одного объекта для реакции с кДНК, полученной на матрице мРНК из другого объекта. Такие эксперименты свидетельствуют о том, что последовательности, представленные большим числом копий в одной ткани, могут не относиться к этому классу в другой ткани и даже обнаруживаться в классе редких мРНК. В печени и почках мыши около 90% редких мРНК являются общими, различия же между этими тканями по числу экспрессирующихся генов составляют всего 1000–2000 генов. Общий вывод, который можно сделать на основе нескольких сравнений такого рода, состоит в том, что только около 10% последовательностей мРНК уникальны для клетки. Большая часть последовательностей является общей для многих, возможно, даже для всех типов клеток.

На основе этих данных можно предположить, что общий набор экспрессирующихся генов, насчитывающий у млекопитающих, вероятно, около 10 000 генов, обеспечивает осуществление функций, необходимых для клеток всех типов. Иногда гены, обеспечивающие осуществление такого рода функций, называют генами «домашнего хозяйства» (housekeeping) или конститутивными генами. Они являются противоположностью генов, обеспечивающих осуществление специализированных функций (например, выполняемых овальбумином или глобином), необходимых только для клеток с определенным фенотипом. Такие гены иногда называют генами «роскоши». Конечно, если принять во внимание все разнообразие фенотипов клеток организма, наверняка обнаружится столько же генов «роскоши», сколько и генов «домашнего хозяйства». И все-таки общее число генов (по крайней мере относящихся к уникальной ДНК), по-видимому, превышает не более чем, скажем в 2–3 раза число генов «домашнего хозяйства» и находится в пределах от 20 000 до 40 000.

Некоторые специализированные типы клеток могут функционировать при наличии относительно небольшого числа экспрессирующихся генов. При детальном исследовании экспрессирующихся генов морского ежа *S. purpuratus* было обнаружено, что в его ооците содержится около 18 000 последовательностей мРНК. В ходе эмбриогенеза число экспрессирующихся генов уменьшается до 13 000 в бластуле, до 8500 — в гастрале и 7000 — в плутеусе (личинке). Однако в некоторых тканях взрослого организма — амбулакральной ножке, кишечнике, целомоците — чис-

ло экспрессирующихся генов падает до величины всего лишь 2500–3000. Таким образом, жизнедеятельность этих клеток может поддерживаться путем функционирования менее чем 10 000 генов «домашнего хозяйства», необходимых в случае клеток млекопитающих.

Особенно важная черта таких клеток – наличие очень значительного перекрытия последовательностей экспрессирующихся генов, так что развитие организма осуществляется за счет постепенного уменьшения числа экспрессирующихся генов. Эксперименты с препаратами мДНК/«нулевая» ДНК показали, что большая часть генов, экспрессирующихся на поздней стадии, экспрессировалась также и на ранних эмбриональных стадиях. Возможно даже, что все структурные гены организма

экспрессируются в ооците, откуда следует, что геном морского ежа состоит из менее чем 20 000 генов, кодирующие последовательности которых составляют только 3% всей ДНК гаплоидного генома.

Рекомендуемая литература

Подобный обзор работ по определению числа генов содержится в книге Льюина Б. (*Lewin B., Gene expression, 2, Eucaryotic genomes, Wiley, New York, 1980, pp. 694–727*). Более ранние экспериментальные данные суммированы в обзоре Льюина Б. (*Lewin B., Cell, 4, 77–93, 1975*).

Глава 19

ИССЛЕДОВАНИЕ ДНК

Могли ли мы представить себе, что за последние годы методы исследования ДНК получают такое колоссальное развитие. К замечательным методам, с помощью которых в настоящее время можно изучать ДНК, относятся методы создания молекул ДНК путем соединения последовательностей, имеющих совершенно разное происхождение. Получаемый продукт часто называют **рекомбинантной ДНК**, а методы (более популярный термин) – **генетической инженерией**. Эти методы в равной степени применимы к прокариотам и к эукариотам, хотя их возможности особенно очевидны при изучении эукариотических геномов. Используя такие методы, можно выделить и охарактеризовать гены, недоступные для изучения другими способами.

В основе этих методов лежит способность ферментов рестрикции (рестриктаз) расщеплять ДНК на отдельные, довольно короткие нуклеотидные последовательности. В гл. 3 мы уже обсуждали использование рестриктаз для составления физической карты ДНК и получения фрагментов ДНК, нуклеотидная последовательность которых может быть определена непосредственно. Таким образом, если исследователь располагает участком ДНК, соответствующим, например, определенному гену, то определение его нуклеотидной последовательности сейчас уже представляется вполне рутинной, чтобы не сказать рядовой, процедурой. Используя эти данные, с помощью некоторых химических и биохимических методов можно разрывая молекулу ДНК и снова восстанавливать ее целостность практически в любом месте.

В этой главе мы рассмотрим первые этапы исследования индивидуальных генов. Как выделить цитоплазматическую РНК или геномную ДНК, соответствующую определенному гену? Как, идентифицировав такую ДНК, получить ее в достаточном для исследования количестве? Можно ли использовать этот материал для синтеза продукта гена *in vitro* или *in vivo*? Можно ли вносить в полученную последовательность изменения, влияющие на ее экспрессию и способствующие выяснению природы регуляторных сигналов?

Любая последовательность ДНК может быть клонирована в бактериях

Клонирование фрагмента ДНК позволяет получить любое его количество из одной-единственной молекулы. (**Клоном** называется большое число клеток или молекул, идентичных исходной родительской клетке или молекуле). Клонирование ДНК возможно благодаря способности бактериальных плазмид и фагов продолжать нормально функционировать после встраивания в их геном дополнительных последовательностей ДНК.

Встраивание приводит к образованию **гибридных**, или **химерных**, плазмид или фагов, состоящих из ДНК их собственного генома и дополнительных «чужеродных» последовательностей. Такие гибридные последовательности реплицируются в бактериях точно так же, как и исходные плазмиды или фаги, и поэтому могут быть получены в больших количествах. Копии чужеродных фрагментов могут быть затем снова выделены в чистом виде из клеток потомства. Поскольку, за редкими исключениями, отдельные встроенные в геном чужеродные последовательности не влияют на свойства химерных штаммов, практически любая последовательность ДНК может быть клонирована таким способом. Так как фаг или плазмиды обеспечивают «перенос» чужеродной ДНК в качестве инертной части генома, их часто называют **клонировующими векторами**.

Основное требование, предъявляемое к клонирующему вектору, состоит в том, что он должен содержать участок, куда чужеродная ДНК может быть встроена без нарушения какой-либо важной функции. Самый простой способ решения этой проблемы – подобрать фермент рестрикции, вносящий единственный разрыв в такой участок векторной ДНК. Далее необходимо суметь отделить химерный геном от исходного вектора, так как иначе после клонирования будет трудно выделить химерный геном из всей массы материала.

Плазмиды – небольшие элементы бактерий, реплицирующиеся независимо от бактериальной хромосомы (гл. 31). Геном плазмид всегда представляет собой кольцевую двухцепочечную ДНК и имеет систему контроля репликации, которая поддерживает их количество в бактериальной клетке на определенном уровне. Существует два основных типа плазмид. В случае **однокопийных** плазмид на хромосому клетки-хозяина приходится 1 молекула плазмидной ДНК. **Мультикопийные** плазмиды присутствуют в клетке в большем количестве, обычно составляющем около 10–20 плазмидных геномов на клетку. Некоторые плазмиды находятся под **ослабленным контролем репликации**, в результате чего при прекращении роста бактерий они накапливаются в очень больших количествах (~1000 плазмидных геномов на клетку). Такие плазмиды часто используют для получения клонирующих векторов, поскольку они обеспечивают более высокий выход материала.

Внесение разрыва в один из участков кольцевой плазмидной ДНК превращает ее в линейную молекулу, как показано на **рис. 19.1**. Затем два конца этой молекулы могут быть присоединены к концам линейной молекулы чужеродной ДНК, что приводит к образованию кольцевой гибридной плазмиды. Формально не существует ограничений на длину встраиваемой в плазмиду чужеродной ДНК. Гибридная плаزمида может существовать в бактериальной клетке неограниченно долгое время. Для выделения гибридной плазмиды из клетки, например с помощью гель-электрофореза, можно использовать такие ее свойства, как размер или кольцевое строение.

Многие плазмиды несут гены, обуславливающие устойчивость к антибиотикам. Это оказывается полезным при создании клонирующих систем. Обычно используют плазмиду, несущую два гена, обеспечивающих устойчивость к разным антибиотикам. Один из них служит просто для идентификации бактерий, несущих плазмиду, путем отбора клеток, устойчивых к антибиотику. Другой служит для того, чтобы отличить гибридную плазмиду от родительского вектора. Если сайт встраивания чужеродной ДНК находится внутри такого гена, гибридная плазмиды теряет устойчивость к антибиотику. Таким образом, селекция родительского вектора может производиться по признаку устойчивости к двум антибиотикам, в то время как отбор гибридной плазмиды может основываться на сохранении устойчивости к одному антибиотику и чувствительности – к другому.

Часто природные плазмиды обладают многими, но не всеми необходимыми свойствами клонирующего вектора. В целях создания лучших векторов из исходного природного материала было разработано несколько подходов. Эти подходы могут заключаться во внесении изменений в систему контроля репликации или в добавлении в плазмиду генов, определяющих устойчивость к специфическим антибиотикам. Один из стандартных клонирующих векторов, наиболее широко используемых в настоящее время, *pBR322*, был получен путем нескольких последовательных изменений ранее существующих клонирующих векторов. Это мультикопийная плазмиды, несущая гены устойчивости к тетрациклину и ампициллину и имеющая в удобных участках сайты рестрикции для нескольких рестриктаз.

Фаги представляют собой другой тип векторных систем. Обычно фаг – это линейная молекула ДНК, поэтому внесение одного разрыва рестриктазой приводит к образованию двух фрагментов. Эти фрагменты сшивают с чужеродной ДНК с образованием химерного фага, как показано на **рис. 19.2**. Для выделения геномов химерных фагов необходимо, чтобы фаг прошел цикл литической инфекции с образованием фаговых частиц. Преимущество этого метода состоит в простоте получения химерных геномов. Однако он накладывает ограничения на длину чужеродной ДНК, которая может быть клонирована, поскольку размер головки фага препятствует упаковке в размножившихся частицах слишком длинных молекул ДНК.

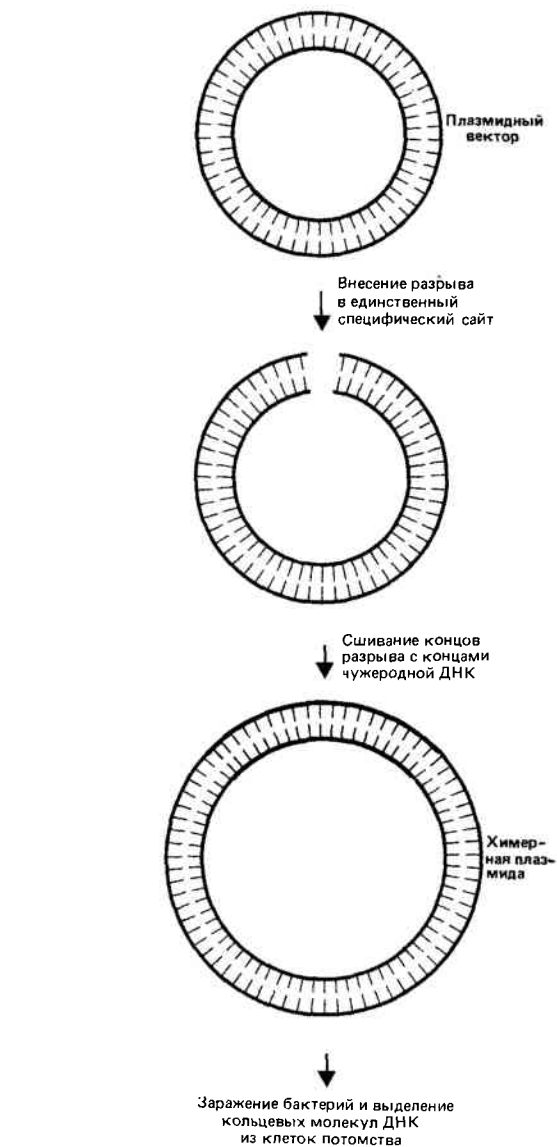


Рис. 19.1. Плазмидные векторы могут быть использованы для клонирования любого фрагмента ДНК, встраиваемого в соответствующий участок.

Часто для решения этой проблемы фрагмент вектора, не несущий необходимых для жизнедеятельности фага генов, *заменяют* на чужеродную ДНК вместо ее простого встраивания в качестве дополнительного материала. Такой подход дал блестящие результаты в случае фага λ , где путем преобразования ДНК был получен более короткий геном (содержащий только жизненно важные гены), включающий только один сайт рестрикции для рестриктазы *EcoRI*. Сам по себе этот клонирующий вектор *слишком короток* для его упаковки в головке фага, кото-

рого для решения этой проблемы фрагмент вектора, не несущий необходимых для жизнедеятельности фага генов, *заменяют* на чужеродную ДНК вместо ее простого встраивания в качестве дополнительного материала. Такой подход дал блестящие результаты в случае фага λ , где путем преобразования ДНК был получен более короткий геном (содержащий только жизненно важные гены), включающий только один сайт рестрикции для рестриктазы *EcoRI*. Сам по себе этот клонирующий вектор *слишком короток* для его упаковки в головке фага, кото-

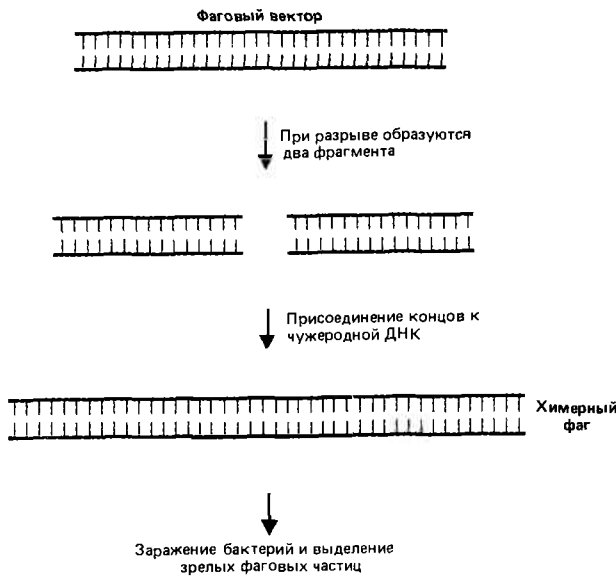


Рис. 19.2. Фаговые векторы могут быть использованы для клонирования чужеродной ДНК, встраиваемой в область генома фага, не содержащую жизненно необходимые гены.

рая накладывает ограничения не только на максимальную, но и на минимальную длину генома. Таким образом, чтобы получить фаг, способный размножиться с образованием зрелых фаговых частиц, в разрезанный родительский вектор должен быть непременно встроено фрагмент чужеродной ДНК. Это приводит к созданию автоматической селективной системы для получения химерных фаговых геномов.

Ценность таких векторов возросла с появлением систем упаковки ДНК в фаговую частицу *in vitro*. Попытка объединить некоторые преимущества плазмид и фагов привела к созданию космид. Это плазмиды со встроенными специфическими последовательностями ДНК (*cos*-сайтами), отвечающими за упаковку ДНК фага λ в фаговую частицу. Такие векторы по-прежнему могут существовать в бактериях в виде плазмид, но могут быть выделены и в чистом виде путем их упаковки в фаговые частицы *in vitro*. На длину этих векторов также накладывалось ограничение, обусловленное размером головки фага, но такой геном не должен включать гены, необходимые для литического цикла.

Мы рассмотрели клонирующие векторы применительно к использованию бактериальной клетки-хозяина. Иногда в качестве клетки-хозяина бывает удобно использовать эукариотическую клетку. Единственные известные природные эукариотические плазмиды обнаружены у дрожжей. Путем реконструкции ДНК были получены бирепликонные (*dual-purpore*), или «челночные», плазмиды, в состав которых входят последовательности, необходимые для их существования в клетках как *E. coli*, так и *S. cerevisiae*. Таким образом, один и тот же вектор может быть использован как с одним, так и с другим хозяином.

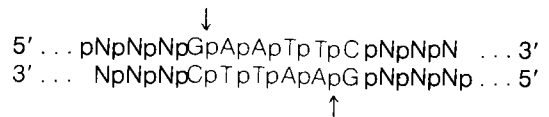
Получение химерной ДНК

Для присоединения фрагмента чужеродной ДНК к клонирующему вектору необходимо осуществление реакции взаимодействия между концами фрагмента и вектора. Это можно обеспечить путем образования кон-

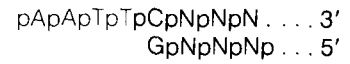
цевых комплементарных последовательностей у фрагмента и у вектора так, чтобы при смешивании они смогли образовать химерную ДНК.

Наиболее распространенный метод состоит в использовании рестриктаз, вносящих ступенчатые разрывы с образованием коротких комплементарных одноцепочечных липких концов. Классический пример такого фермента – рестриктаза *EcoRI*, разрывающая каждую из цепей двухцепочечной ДНК, но в разных точках. Эти точки располагаются по обе стороны от короткого палиндрома, входящего в состав сайта узнавания фермента.

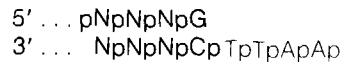
Для изображенной ниже последовательности ДНК, где N обозначает основания, не узнаваемые ферментом, *EcoRI* вносит разрыв в свой сайт узнавания (выделен красным цветом) в точках, отмеченных вертикальными стрелками:



Фрагменты ДНК по обе стороны от сайта-мишени отсоединяются, и вследствие взаимного сдвига сайтов рестрикции каждой цепи ДНК на каждом из фрагментов образуются выступающие, комплементарные друг другу одноцепочечные участки – липкие концы (выделены красным цветом):



и



Комплементарность липких концов позволяет им реассоциировать за счет спаривания оснований. Когда *EcoRI* «разрезает» две разные молекулы, на каждой из них образуются одинаковые липкие концы. Это делает возможной реассоциацию молекул, что показано на рис. 19.3. Такой метод позволяет получить химерную плазмиду, интактную в целом, но с отсутствием ковалентных связей между векторной и чужеродной ДНК. Отсутствующие связи образуются при действии фермента ДНК-лигазы *in vitro*. Этот метод рекомбинации двух молекул ДНК имеет свои достоинства и недостатки.

Достоинство метода состоит в том, что химерная плаزمиды имеет восстановленные сайты узнавания для *EcoRI* по обе стороны от встроенной ДНК. Поэтому фрагмент чужеродной ДНК относительно легко может быть выделен из клонированных копий химерного вектора с помощью расщепления рестриктазой *EcoRI*.

Недостаток метода состоит в том, что все полученные с помощью *EcoRI* липкие концы могут реассоциировать друг с другом. Из-за этого часть векторов восстанавливается в результате непосредственного взаимодействия своих концов без присоединения к встраиваемому фрагменту ДНК, тогда как другие векторы могут присоединяться к вставке, состоящей из нескольких последовательно соединенных фрагментов чужеродной ДНК. Поэтому необходимо осуществлять селекцию химерных плазмид, несущих только один встроенный фрагмент.

Сложность метода, основывающегося на использовании только таких рестриктаз, с помощью которых получают концы, участвующие в реакции сшивания, состоит

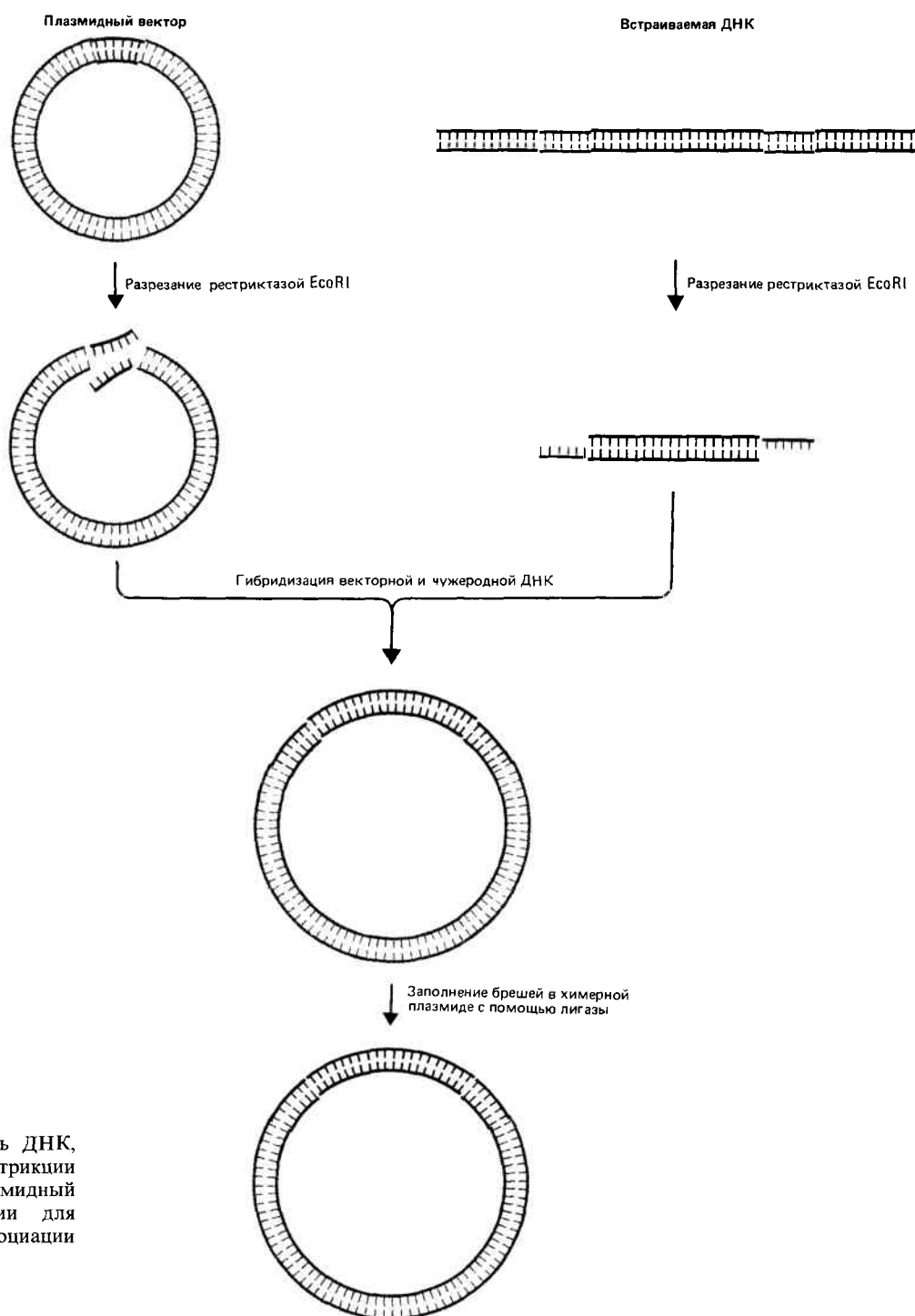


Рис. 19.3. Любая последовательность ДНК, находящаяся между сайтами рестрикции EcoRI, может быть встроена в плазмидный вектор, имеющий сайт рестрикции для EcoRI, путем разрезания и реассоциации двух молекул ДНК.

в том, что сайты узнавания рестриктаз могут находиться в неудобных точках (особенно в случае последовательности чужеродной ДНК). Существует другой метод, позволяющий использовать любой конец ДНК для рекомбинации с плазмидой.

Плазмиду разрезают, как описано ранее, ферментом, узнающим единственный сайт в нужном положении, но при этом не обязательно образуются зигзагообразные концы. Затем, используя фермент **терминальную трансферазу** и предшественник dATP, к двум 3'-концам плазмидной ДНК добавляют фрагмент poly(dA). Аналогичным образом присоединяют poly(dT) к 3'-концам молекулы чужеродной ДНК, которую нужно встроить; эти концы мо-

гут быть получены любым удобным способом. Как показано на **рис. 19.4**, poly(dA) плазмиды может реассоциировать *только* с poly(dT) встраиваемого фрагмента. Таким образом, в этом случае возможно осуществление только одной реакции: встраивания одного чужеродного фрагмента в вектор.

Недостаток этого метода состоит в сложности выделения встроеного фрагмента из клонированной химерной плазмиды, поскольку в исходном векторе после встраивания чужеродной ДНК не стало сайта узнавания рестриктазой. Встроенная последовательность с каждой стороны окружена участками спаренных последовательностей poly(dA) : poly(dT). При этом методе могут исполь-

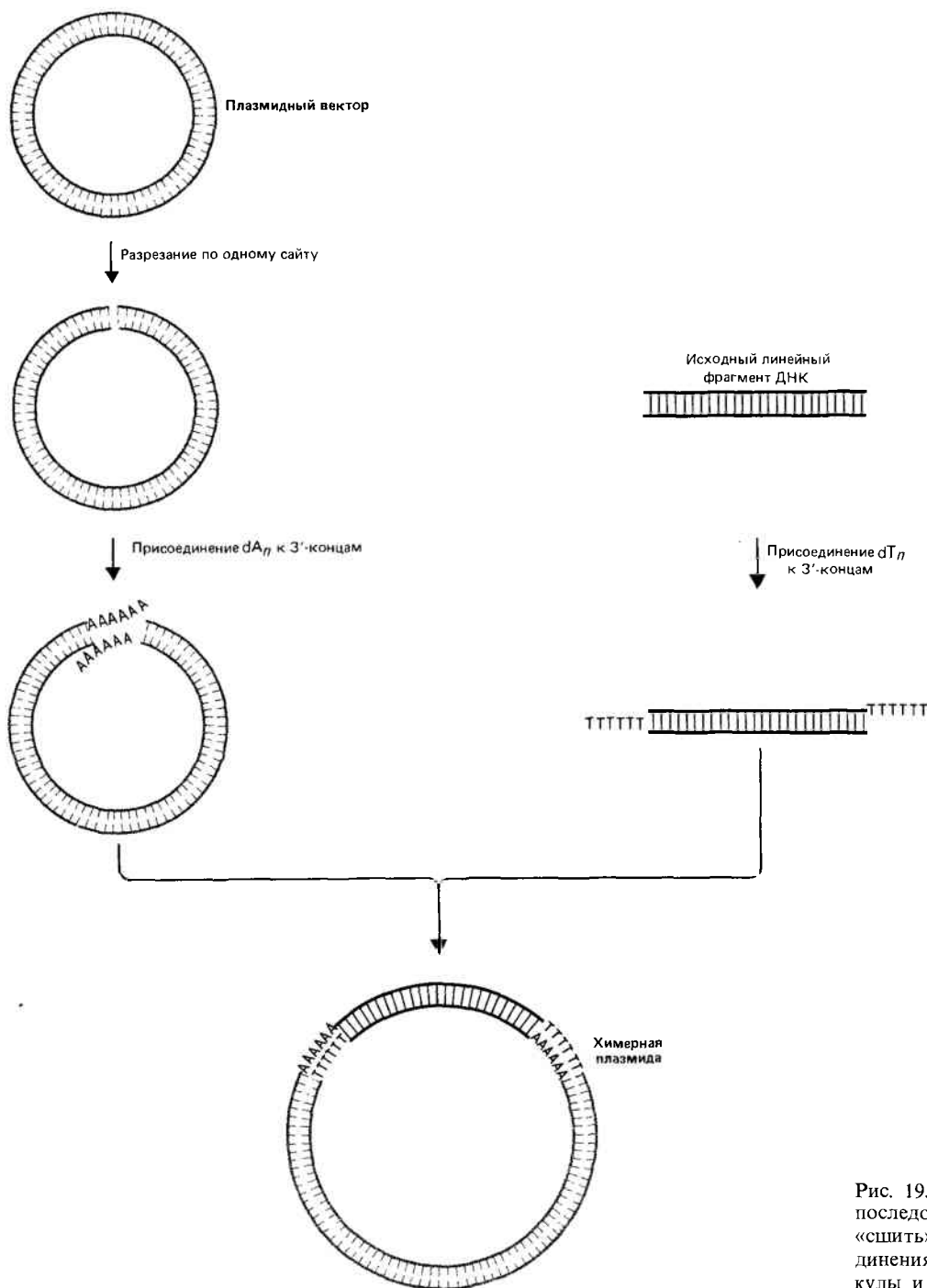


Рис. 19.4. Метод концевое присоединения последовательностей poly(dA : dT) позволяет «сшить» две молекулы ДНК путем присоединения poly(dA) к 3'-концам одной молекулы и poly(dT) – к 3'-концам другой.

зваться также и последовательности poly(dC) и poly(dG). Преимущество использования этих последовательностей состоит в том, что при расщеплении плазмиды рестриктазой PstI присоединение poly(dC): poly(dG) приводит к восстановлению сайта рестрикции для PstI по обе стороны от встроенного фрагмента.

Другой удобный метод – «сшивание» (лигирование) **тупых концов**. В его основе лежит способность ДНК-лигазы фага Т4 сшивать две молекулы ДНК, имеющие тупые концы, то есть молекулы, у которых отсутствуют выступающие одноцепочечные участки. (Эта реакция – дополнительное проявление активности фермента помимо сшивания разорванных связей внутри двухцепочечной молекулы.) Сшивание тупых концов приводит к непосред-

ственному соединению фрагментов, образованных при расщеплении ДНК рестриктазой, разрезающей обе цепи ДНК поперек в одной точке. Большое преимущество этого метода состоит в том, что может быть «сшита» любая пара концов независимо от их нуклеотидной последовательности. Метод особенно полезен, когда мы хотим сшить две определенные последовательности без встраивания между ними какого-либо дополнительного материала. Трудности, возникающие при использовании этого метода, связаны с отсутствием контроля над сшиванием определенных пар тупых концов, поэтому необходимо сначала провести реакцию сшивания, а затем отделить нужные продукты реакции от других ее продуктов.

Существует множество вариаций этих методов.

В одном методе используют короткие двухцепочечные молекулы ДНК («линкеры»), содержащие палиндром, узнаваемый *EcoRI* (или аналогичной рестриктазой). Такие линкеры могут быть синтезированы химически и затем ковалентно присоединены к концам плазмиды или встраиваемого фрагмента путем сшивания тупых концов. Это позволяет затем выделить встроенную ДНК с помощью расщепления рестриктазой *EcoRI*, но без наложения ограничений на исходный выбор сайтов рестрикции, используемых для создания концов. В настоящее время имеется достаточное число методик, позволяющих встраивать любой фрагмент чужеродной ДНК в любой специфический сайт вектора и при необходимости выделять этот фрагмент в чистом виде.

Фрагмент чужеродной ДНК может быть встроен в плазмиду в любой ориентации, т.е. любой его конец может быть соединен с любым концом плазмиды. Это не имеет значения, когда цель клонирования состоит просто в амплификации встроенной последовательности. Однако ее ориентация может оказаться существенной, когда эксперимент проводится с целью получения экспрессии чужеродной ДНК. В этом случае в результате случайных встраиваний могут быть получены популяции плазмид, ориентированных и в том, и в другом направлении, после чего путем рестрикционного картирования выявляется нужный класс. Можно также провести эксперимент таким образом, чтобы встраивание осуществлялось только в одной ориентации. Например, обе ДНК, векторная и встраиваемая, могут быть разрезаны двумя рестриктазами, образующими разные липкие концы, так что каждая молекула ДНК будет иметь такую структуру:

Конец 1 ————— Конец 2

В этом случае, если реассоциировать могут только две концевые последовательности, образующие конец 1, и только две последовательности, образующие конец 2, встраивание будет происходить только в одном направлении с образованием химерной плазмиды:

Встроенный
Конец 1—фрагмент— Конец 2
Конец 1—Плазида— Конец 2

Встроенный фрагмент замещает утерянный участок исходной плазмиды, находившийся между сайтами, рестрикция которых привела к образованию конца 1 и конца 2.

Получение ДНК-копий на матрице мРНК

Идеальный подход для получения последовательности ДНК, соответствующей индивидуальному белку, состоит в том, чтобы в качестве отправной точки использовать мРНК, которая, в конце концов, служит матрицей для синтеза белка *in vivo*. Синтез двухцепочечной ДНК на матрице мРНК оказывается возможным благодаря обратной транскрипции. Ее особенно легко осуществить для мРНК, имеющих poly(A)-последовательность на 3'-конце, как показано на рис. 19.5.

Прежде всего poly(A) гибридизуют с затравкой (праймером). Это короткая последовательность oligo(dT), пред-

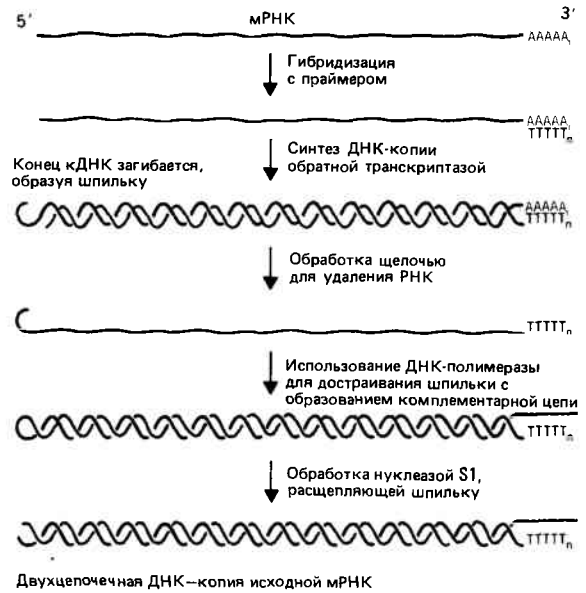


Рис. 19.5. На мРНК может быть синтезирована комплементарная ей двухцепочечная ДНК.

назначенная для создания свободного 3'-конца, который может быть достроен ферментом обратной транскриптазой. (Как и другие ферменты, синтезирующие ДНК, обратная транскриптаза не может начать синтез полинуклеотидной цепи без наличия праймера.) Фермент осуществляет обычный синтез в направлении 5'–3', добавляя дезоксинуклеотиды по одному; процесс происходит путем комплементарного спаривания оснований с мРНК-матрицей.

Продукт реакции – гибридная молекула, состоящая из цепи РНК-матрицы, спаренной с комплементарной цепью ДНК. Единственная практическая трудность при проведении этой реакции заключается в том, что *in vitro* обратная транскриптаза может останавливаться в разных точках до достижения 5'-конца мРНК. Образующийся при этом обратный транскрипт оказывается укороченным и соответствует только части молекулы мРНК, поскольку он лишен некоторых последовательностей, комплементарных 5'-концу. Однако при подборе оптимальных условий эксперимента нередко удается добиться полного считывания матрицы обратной транскриптазой.

Обратную транскрипцию часто завершает полезная реакция. Достигнув конца мРНК, фермент может образовать петлю на конце синтезируемой цепи ДНК, используя несколько последних оснований обратного транскрипта в качестве матрицы для синтеза комплементарного участка. Иными словами, конец комплементарной ДНК используется для прямого синтеза короткой последовательности, идентичной концевому участку мРНК и замещающей его. Это приводит к образованию на 3'-конце обратного транскрипта короткой «шпильки» длиной обычно 10–20 п.н.

На этой стадии мРНК-матрицу разрушают обработкой щелочью (на ДНК щелочь не влияет). В результате образуется одноцепочечная ДНК, комплементарная мРНК; она называется кДНК. Шпилька на ее 3'-конце служит естественным праймером для следующего этапа – использования фермента ДНК-полимеразы I *E. coli* для превращения одноцепочечной кДНК в двухцепочеч-

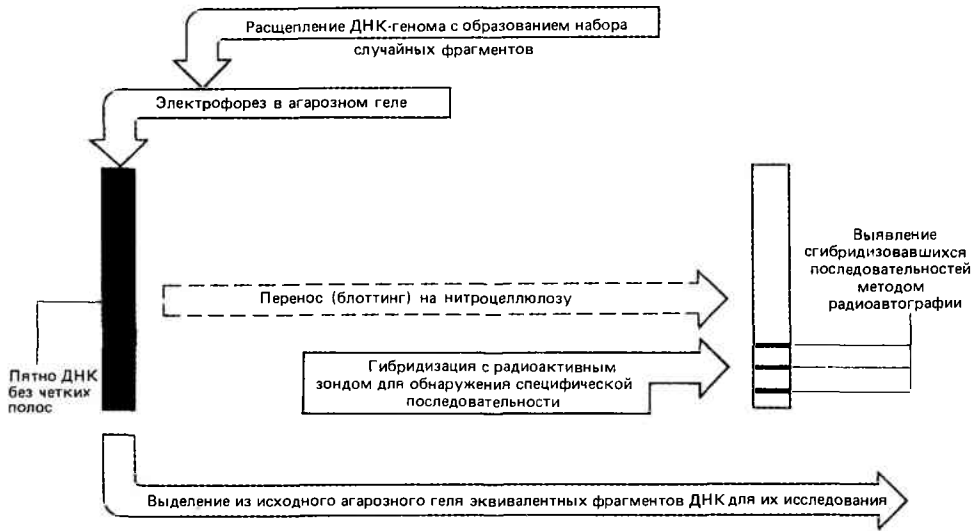


Рис. 19.6. Блоттинг по Саузеру позволяет осуществить прямое выделение фрагментов ДНК, соответствующих специфическому зонду, из смеси рестриктов ДНК генома.

ную путем синтеза комплементарной цепи. В этой реакции фермент использует кДНК в качестве матрицы для синтеза последовательности, идентичной исходной мРНК. Продукт реакции – двухцепочечная молекула со шпилькой на одном конце. Шпильку разрезают ферментом нуклеазой S1 (специфически расщепляющей одноцепочечную ДНК) с образованием обычной двухцепочечной ДНК.

Двухцепочечную ДНК можно клонировать для получения больших количеств синтетического гена, который представляет последовательность мРНК в виде двухцепочечной ДНК. Это называется клоном кДНК. (При использовании этой терминологии термин «кДНК» стал несколько неточно использоваться для обозначения двухцепочечного встроенного фрагмента, а не исходного одноцепочечного обратного транскрипта.)

Выделение из генома индивидуальных генов

Одно из основных применений техники клонирования – выделение индивидуальных генов непосредственно из генетического материала. Каждый индивидуальный ген представляет собой только очень небольшую часть эукариотического генома. Например, размер генома типичной клетки млекопитающих составляет около 10^9 п. н., так что один ген размером, скажем, 5000 п. н. составляет только 0,00005% всей ядерной ДНК. Для идентификации столь незначительного количества материала необходимо иметь высокоспецифичный зонд, взаимодействующий только с определенными, интересующими исследователя последовательностями, чтобы выделить их из огромного количества других последовательностей. Обычно применяемый метод состоит в использовании высокоочищенных радиоактивных РНК- или ДНК-зондов, гибридизация которых с геном регистрируется с помощью радиоавтографии.

Возникающее при этом затруднение связано с получением мРНК, соответствующей специфическому белку. Оно может оказаться крайне трудным, когда белковый продукт присутствует в клетке в незначительном количестве. Существует несколько методов выделения мРНК, основанных на использовании свойств ее продукта, в том числе чувствительный метод обнаружения продуктов син-

теза после инъекции РНК в ооциты ксенопуса. Эффективный метод выделения ДНК, комплементарной мРНК, – трансляция, подавляемая гибридом; в основе метода лежит способность кДНК подавлять трансляцию мРНК, гибридуясь с ней, в результате чего продукт реакции исчезает из системы трансляции *in vitro*.

Ранее сама мРНК использовалась в качестве зонда для выделения генов. Этот метод имел ряд практических недостатков. мРНК никогда не бывает абсолютно чистой, и ее трудно получить в достаточном количестве. Действительно, метод хорошо подходит для выделения генов, мРНК которых представлены большим числом копий, но не может быть использован в случае слабо экспрессирующихся генов. Другая трудность состоит в том, что часто не удается получить мРНК с достаточно высокой радиоактивной меткой.

Поэтому вместо мРНК в качестве зонда стали использовать клонированную кДНК-копию мРНК. С помощью уже описанных ранее методов могут быть получены большие количества очищенного материала. Этот материал может быть помечен *in vitro* с получением очень высокой специфической радиоактивности.

Первый шаг при идентификации гена, соответствующего специфическому зонду, состоит в дроблении ДНК генома на фрагменты удобного для работы размера. Желательно, чтобы ген находился в как можно меньшем количестве фрагментов (в идеальном случае – только в одном). Обычно максимальная длина фрагментов генома, с которыми можно работать, находится в пределах 15–20 т. п. н. Иногда оказывается невозможным получить ген в виде одного фрагмента, и тогда для определения его структуры необходимо собрать вместе данные, полученные при анализе его разных фрагментов.

Для фрагментирования генома могут быть использованы два метода. Один из них – расщепление рестриктазами. При этом каждый фрагмент заканчивается сайтом узнавания специфической рестриктазы. Другой метод состоит в механическом дроблении ДНК путем внесения в нее разрывов с частотой, определяемой условиями фрагментирования. В этом случае места разрывов располагаются совершенно случайно.

Когда рестриктазой расщепляют всю ДНК генома, частота разрывов определяется длиной последовательности узнавания фермента. Чем длиннее эта последовательность, тем реже происходят случайные разрывы. Напри-

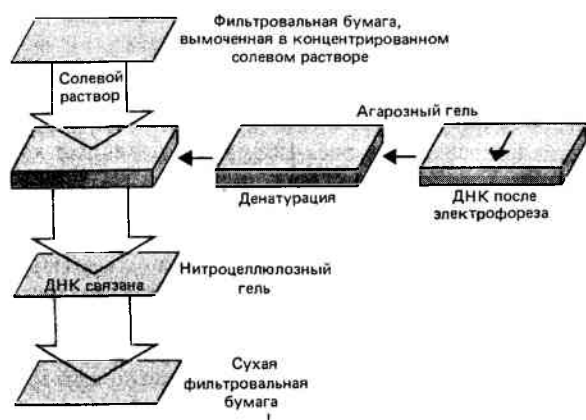


Рис. 19.7 При блоттинге по Саузерну фрагменты ДНК переносят с агарозного геля на нитроцеллюлозный фильтр.

мер, вероятность наличия определенной последовательности длиной 4 п.н. составляет $0,25^4 = 1/256$, поэтому фермент, узнающий такую короткую последовательность, будет вносить разрывы в ДНК довольно часто. Частота уменьшается до $1/1000$ для последовательности длиной 5 п.н. и до $1/4000$ для последовательности длиной 6 п.н. (Этот расчет основан на предположении о том, что каждое основание представлено в ДНК с равной частотой, чего обычно не бывает. Частота внесения разрывов может быть уменьшена при использовании фермента, чья последовательность-мишень содержит пары оснований, встречающихся в ДНК реже других, и наоборот.)

Распределение сайтов узнавания фермента носит случайный характер по отношению к геному в целом. Поэтому обработка эукариотической ДНК рестриктазами приводит к образованию множества фрагментов. При электрофорезе в геле эти фрагменты образуют пятно, в котором неразличимы отдельные полосы (за исключением ряда полос, соответствующих некоторым повторяющимся последовательностям, рассмотренным в гл. 24).

Однако при наличии специфического зонда в пятне рестриктов можно обнаружить соответствующие ему последовательности. Последовательность операций такого метода приведена на рис. 19.6. Ключевой момент — денатурация ДНК с образованием одноцепочечных фрагментов, которые переносят из агарозного геля на нитроцеллюлозный фильтр, где они иммобилизуются. Процесс переноса ДНК несколько похож на промокание (по-английски — блоттинг), и в разговорной речи этот термин используется для его обозначения. При работе с ДНК такой метод известен под названием **блоттинга по Саузерну** (Southern blotting) (по фамилии автора метода).

На рис. 19.7 схематически изображена последовательность операций по переносу фрагментов ДНК. Агарозный гель помещают на фильтровальную бумагу, замоченную в концентрированном солевом растворе. Затем на гель накладывают нитроцеллюлозный фильтр и сверху помещают сухую фильтровальную бумагу. Солевой раствор впитывается в сухую фильтровальную бумагу. Чтобы это произошло, он должен пройти сквозь агарозный гель и затем через нитроцеллюлозный фильтр. ДНК переносится вместе с раствором, но задерживается нитроцеллюлозой.

Иммобилизованную нитроцеллюлозой ДНК можно гибридизовать *in situ* с радиоактивным зондом. Со специ-

фическим зондом будут гибридизоваться только комплементарные ему фрагменты. Поскольку зонд радиоактивен, гибридизацию можно обнаружить с помощью радиоавтографии. Каждая комплементарная последовательность проявляется в виде радиоактивной полосы, местоположение которой определяется размером фрагмента ДНК.

Метод также можно применять и для РНК. Для переноса РНК с агарозы в среду, пригодную для проведения гибридизации, его необходимо несколько видоизменить. Такой метод известен под названием **Нозерн-блоттинга**¹.

Клонирование всей ДНК генома («шотган») с образованием библиотек генов

Прямое выделение фрагментов генома, соответствующих зонду, сопряжено с практическими трудностями, и для выделения отдельных генов используют эффективный метод с обратным порядком стадий, когда вначале проводят клонирование генома. Затем проводят отбор клон(ов), содержащего специфическую последовательность. Векторы, несущие ДНК, полученную из генома, называют **клонами геномной** или **хромосомной ДНК** (в отличие от клонов кДНК, являющихся копиями мРНК).

Клонирование всего генома (в противоположность клонированию специфических фрагментов) часто называют **«шотган»-экспериментами** (shotgun experiment — метод дробовика). Для их осуществления весь геном разделяют на фрагменты удобного размера. Затем фрагменты встраивают в клонирующий вектор с образованием популяции химерных векторов. Набор таких клонированных фрагментов называют **библиотекой генома**. Библиотека, однажды полученная с помощью фагового или плазмидного вектора (чаще — фагового вектора, поскольку в таком виде легче хранить необходимые большие количества химерных ДНК), может храниться неограниченно долгое время и при появлении нового зонда быстро может быть использована для поиска специфического фрагмента.

Клонирование фрагментов, полученных путем механического дробления, сопряжено с рядом технических трудностей; в вектор легче встроить рестрикт. Однако сложность такого встраивания состоит в том, что сайты рестрикции могут находиться в неудобных местах, например в середине гена, который надо клонировать. Один из способов преодоления этой трудности состоит в использовании более чем одной рестриктазы, т.е. в повторении эксперимента с различными ферментами, сайты узнавания которых занимают разные положения. Но это требует дополнительного времени, и при работе с длинной последовательностью бывает сложно найти фермент, не разрезающий ее.

Поэтому для создания удобной библиотеки путем клонирования рестриктов пользуются приемом, так регулирующим частоту внесения разрывов, чтобы получились более длинные фрагменты. Для этого используют фермент с короткой (4 п.н.) последовательностью узнавания в условиях, обеспечивающих *частичную* рестрикцию

¹ Southern (Саузерн), по имени которого назван оригинальный метод, по англ. означает южный. По аналогии видоизмененный метод назван Northern (нозерн), т.е. северный. — Прим. ред.

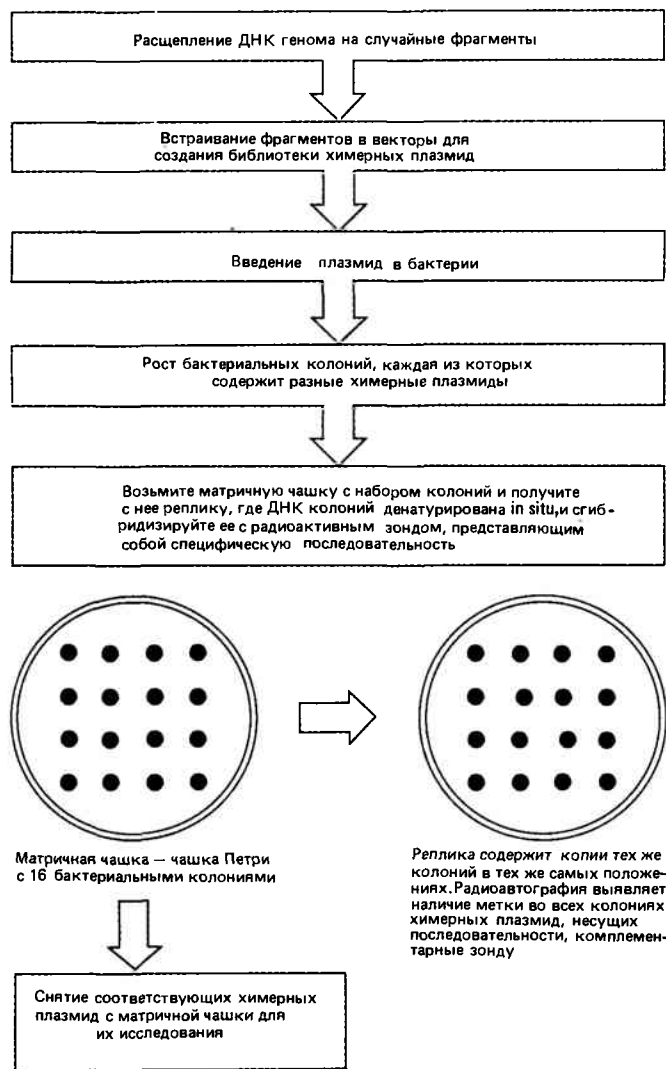


Рис. 19.8. Гибридизация колоний позволяет осуществлять селекцию химерных плазмид, несущих специфическую последовательность, за счет ее комплементарности с радиоактивным зондом.

ДНК. Любой специфический сайт-мишень разрезается совершенно случайно, поэтому присутствие сайта-мишени в некоторой последовательности не обязательно приведет к ее разрыву. Малая вероятность разрезания по каждому сайту в сочетании с частым расположением сайтов означает, что распределение фрагментов обеспечивается практически полностью случайным характером внесения разрывов в геном. При этом каждый фрагмент заканчивается одной и той же последовательностью, которая может иметь липкий конец и поэтому удобна для клонирования.

Число случайно полученных фрагментов, которые должны быть клонированы для того, чтобы обеспечить высокую вероятность наличия *каждой* последовательности генома хотя бы в одной химерной плазмиде, уменьшается с ростом размеров фрагмента и увеличивается с ростом размеров генома и желаемой вероятности. Для 99%-ной вероятности необходимо 1500 клонов с фрагментами ДНК *E. coli*; для дрожжей размер библиотеки возрастает до 4600 клонов, для *D. melanogaster* — до 48 000 и до 800 000 — для млекопитающих. Все эти библиотеки клонированных фрагментов, где достигается такая ве-

роятность представленности последовательностей генома, были получены.

Как осуществляют селекцию определенного клона геномной ДНК из библиотеки? Один из методов называется **гибридизацией колоний**. Последовательность операций этого метода приведена на рис. 19.8. Бактериальные колонии, несущие химерные векторы, лизируют на нитроцеллюлозных фильтрах. Затем их ДНК денатурируют и фиксируют на фильтре. Фильтр гибридизуют с радиоактивным зондом (обычно это клонированная кДНК), соответствующим интересующей исследователя последовательности. Все колонии, с которыми гибридизуется зонд, при автордиографии выявляются в виде темных пятен. После этого соответствующие химерные векторы могут быть выделены из исходной библиотеки.

Для гибридизации с зондом достаточно, чтобы клоны геномной ДНК содержали только часть комплементарной ему последовательности (обычно считают, что минимальный размер клонированной последовательности должен составлять около 50 п.н.). На практике, когда эукариотический ген имеет значительный размер, при создании библиотеки он может оказаться фрагментированным, так что разные участки гена окажутся в разных клонах, гибридизующихся с зондом. Полную нуклеотидную последовательность, соответствующую зонду, может не содержать ни один из клонов геномной ДНК. В этом случае необходимо реконструировать исходную геномную последовательность, используя наличие перекрывающих индивидуальных фрагментов, которые, по всей видимости, имеют разные концы.

Важное допущение, лежащее в основе этого подхода, состоит в том, что клонированная эукариотическая последовательность будет сохраняться в бактериальной клетке в неизменном виде. Основанием для такого предположения служат имеющиеся данные. Однако существует ряд исключений из этого правила, когда при заражении бактериальной клетки-хозяина происходят делеции клонированной последовательности.

Эукариотические гены могут экспрессироваться в бактериях с образованием белка

Вследствие универсальности генетического кода определенная кодирующая последовательность всегда будет содержать одну и ту же информацию. (Единственное исключение — митохондриальные гены, где имеются отличия в генетическом коде, как описано в гл. 4.) Поэтому при встраивании в вектор интактной последовательности, кодирующей эукариотический белок, возможна транскрипция этой последовательности с образованием мРНК, которая может транслироваться в бактерии-хозяине. Единственные отличия состоят в том, что в синтезированном белке могут отсутствовать модификации, имеющиеся в природном клеточном белке, и, конечно, всегда существует риск, что полипептидная цепь в бактериальной клетке окажется нестабильной. Однако при наличии в клетке подходящих условий любая эукариотическая последовательность может быть транслирована с образованием соответствующего белка.

Какие условия необходимы для эффективной экспрессии эукариотических генов в бактериях? Поскольку нуклеотидные последовательности эукариотических и прокариотических промоторов различаются, для того чтобы

ген экспрессировался, он должен быть помещен под контроль бактериального промотора. Это достигается путем включения эукариотической кодирующей последовательности в уже имеющуюся транскрипционную единицу, но обычно этого недостаточно для эффективной трансляции. Поэтому часто эукариотический фрагмент встраивают рядом с промотором, который обеспечивает его транскрипцию с образованием моноцистронной мРНК. Это можно осуществить, если найти подходящий сайт для встраивания фрагмента в клонирующий вектор; в качестве промотора широко используется промотор гена устойчивости к ампициллину, локализованного в векторе pBR322 (обычный продукт гена – фермент β-лактамаза). Другой метод состоит во встраивании в клонирующий вектор природного бактериального промотора с целью экспрессии чужеродного гена. Популярным промотором, используемым для этой цели, является *lac*-промотор (промотор лактозного оперона) *E. coli*, причем чаще используют его мутантную форму.

Для того чтобы синтезированная мРНК транслировалась, в непосредственной близости от иницирующего кодона AUG должен находиться участок связывания рибосомы. Были сконструированы клонирующие векторы, у которых сайт рестрикции для встраивания чужеродной ДНК находится рядом с участком связывания рибосомы. Обычно это достигается благодаря тому, что вставка содержит кодон AUG. На рис. 19.9 показано, как любая последовательность чужеродной ДНК (прокариотической или эукариотической), начинающаяся с кодона AUG, может быть встроена в этот участок и будет транскрибироваться и транслироваться в бактериях. В результате синтезируется белок, точно соответствующий кодирующему участку вставки.

Другой метод состоит во встраивании кодирующей последовательности в участок, находящийся непосредственно за местом начала синтеза бактериального белка. В результате образуется гибридный белок с N-концевой областью, соответствующей бактериальному белку, последовательность которого была прервана, и остальной частью белка, соответствующей встроеной чужеродной ДНК. N-концевая бактериальная последовательность может быть очень короткой и полезной. Например, она может представлять собой сигнальную последовательность, обеспечивающую секрецию гибридного белка в окружающую среду; она может также служить для защиты гибридного белка от деградации в бактериальной клетке.

Каковы возможные источники получения эукариотической кодирующей последовательности? Поскольку мно-

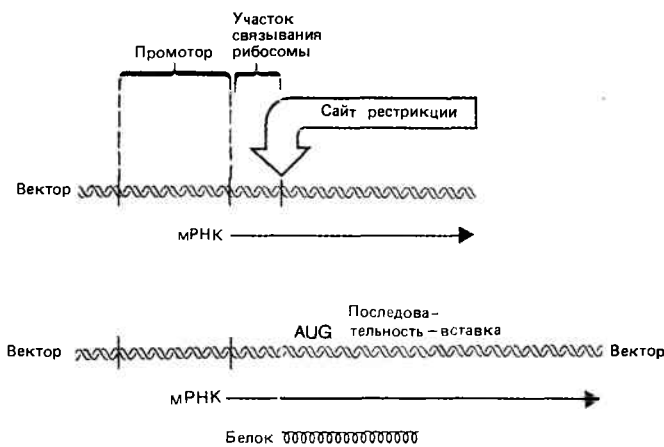


Рис. 19.9. Любая эукариотическая кодирующая последовательность может экспрессироваться с образованием белка при ее встраивании в подходящий участок клонирующего вектора.

гие эукариотические гены имеют прерывистое строение, кодирующая последовательность в клоне геномной ДНК также может прерываться. В таких случаях белок, соответствующий геномной ДНК, в бактериях не синтезируется. Поэтому кодирующую белок последовательность лучше всего выделять из клона кДНК, полученного на матрице мРНК, с непрерывной кодирующей последовательностью. Тогда она может не включать 5'- и 3'-нетранслируемые участки; достаточно кодирующей последовательности, начинающейся от инициаторного AUG-кодона и заканчивающейся терминирующим кодоном.

Иногда возникает необходимость охарактеризовать бактериальный промотор, но продукты находящихся под его контролем генов проанализировать трудно. Эта трудность преодолима при использовании обратного порядка операций – с присоединением к промотору генов лактозного оперона, продукты которых можно легко обнаружить. В этом случае об активности промотора судят на основе стандартных определений β-галактозидазы.

Рекомендуемая литература

Вопросы, затронутые в этой главе, подробно изложены в книге Льюина Б. (*Lewin B., Gene expression, 2, Eucaryotic chromosomes, Wiley, New York, 1980, pp. 761–785*). Более ранний обзор составлен Синшеймером (*Sinsheimer, Ann. Rev. Biochem., v. 46, 1977, pp. 415–438*).

Глава 20 СТРУКТУРНЫЕ ГЕНЫ: ВНУТРЕННЯЯ ОРГАНИЗАЦИЯ

В течение длительного времени существовали нечеткие представления о том, что эукариотические гены в чем-то необычны, что они могут принципиально отличаться от бактериальных генов. Причины существования такого взгляда крылись в явном несоответствии содержания ДНК сложности организма, получившем отражение

в виде так называемого парадокса величины С. Такая точка зрения была подкреплена результатами экспериментов по определению сложности генома, показавших, что только небольшая его часть может быть представлена в виде мРНК (гл. 18).

Одно из предположений, позволяющих найти выход

из этой ситуации, состоит в том, что значительная часть (даже большинство) последовательностей ДНК не относятся к структурным генам. Вопросы о количестве такой ДНК и функциях, ею осуществляемых (или не осуществляемых), предстоит выяснить.

В соответствии с другой гипотезой, структурный ген или по крайней мере транскрипционная единица имеет *гораздо больший размер, чем последовательность, представленная в виде мРНК*. Это предположение косвенно подтверждается данными о том, что РНК, обнаруживаемая в ядре (у высших эукариот), имеет значительно больший размер, чем мРНК. Большая часть избыточной длины РНК должна удаляться при процессинге РНК перед ее транспортом в цитоплазму.

Первоначально было выдвинуто предположение о том, что единица транскрипции может содержать протяженные последовательности, находящиеся по одну или другую сторону от области, представленной в мРНК (вероятнее всего, слева от 5'-конца мРНК), и, возможно, участвующие в регуляции. Также могут существовать и обширные нетранскрибирующиеся области, необходимые для регуляции транскрипции; таким образом, размер единицы экспрессии гена может значительно превышать размер мРНК.

Действительно, оказывается, что многие эукариотические гены гораздо длиннее своих мРНК. Но это связано с наличием в них *вставок*, разделяющих разные части кодирующего участка ДНК. По крайней мере этим частично объясняется разница в размерах мРНК и ядерной РНК: такие *вставочные последовательности (интроны)* входят в состав первичного транскрипта, но удаляются из него при созревании с образованием мРНК. Хотя прерывающие последовательности (интроны) входят в состав структурного гена, они не обладают функцией кодирования, и не ясно, выполняют ли они какую-либо структурную функцию, кроме их удаления при экспрессии генов.

Здесь мы должны вкратце вспомнить терминологию, использованную в гл. 3 при описании взаимоотношений между геном и его РНК-продуктом. Прерывистый ген состоит из различающихся между собой групп *экзонов* и *интронов*. Экзоны — это последовательности, представленные в РНК (мРНК, рРНК или тРНК, т.е. в одних случаях они осуществляют функцию кодирования белка, в других — нет). Интроны — это прерывающие ген вставочные последовательности, которые удаляются из первичного транскрипта и поэтому отсутствуют в зрелой РНК. Ген, конечно, должен начинаться и кончаться экзонами (соответствующими 5' и 3'-концам РНК), но внутри него может быть любой набор интронов. Экзоны и интроны часто обозначают цифрами или буквами в порядке их расположения вдоль гена.

Прерывание кодирующих областей характерно только для эукариот, но обнаружено не для всех эукариотических генов. Разумеется, к настоящему времени охарактеризовано еще недостаточное число генов для определения среднего соотношения между размером гена и мРНК. Поэтому до сих пор не понятно, в какой степени факт наличия интронов помогает объяснить парадокс величины С, но не похоже, чтобы он объяснил этот парадокс полностью.

Следующий вопрос после обсуждения структуры собственно гена касается определения его состава. Какое количество материала по обе стороны от гена участвует в его функционировании? На каком расстоянии от сосед-

него гена он расположен и какую часть генома составляют последовательности, находящиеся между транскрипционными единицами? Связаны ли между собой соседние гены и существует ли какой-либо механизм регуляции функционирования областей хромосом, имеющих большую длину, чем отдельные гены?

Обнаружение прерывистых генов

Существование прерывистых генов было обнаружено в результате экспериментов по выделению ДНК гена, соответствующего специфической мРНК. Когда эта работа начиналась, ее целью было не само по себе сравнение последовательностей двух нуклеиновых кислот, а локализация в геноме последовательности ДНК, соответствующей мРНК. Это в свою очередь позволяет идентифицировать и охарактеризовать фланкирующие последовательности, включая и возможные регуляторные зоны. Таким образом, основная идея этого подхода заключалась в том, чтобы проследовать от мРНК обратно к гену, определить строение целостной единицы транскрипции, включая промоторы и другие элементы, не обязательно представленные в мРНК. Выяснение характеристик окружения, в котором находится структурный ген, до сих пор остается целью таких экспериментов, в ряде случаев достигнутой, как мы увидим в гл. 21.

Вначале предполагали, что мРНК должна иметь такую же последовательность оснований, как ДНК, с которой она транскрибирована. Для доказательства того, что выделенная из генома последовательность ДНК действительно совпадала с последовательностью мРНК, использованной для ее выделения, их нуклеотидные последовательности сравнивали с помощью как электронной микроскопии, так и рестрикционного картирования. Но выявлены были как раз различия между их нуклеотидными последовательностями, заключающиеся в наличии дополнительных участков, присутствующих в ДНК генома, но отсутствующих в мРНК.

Обнаружение прерывистых генов с помощью электронной микроскопии

Электронная микроскопия может быть использована для выявления РНК—ДНК-гибридов двумя способами. В обоих случаях цель исследования состоит в том, чтобы отличить РНК—ДНК-гибриды от участков неспаренной ДНК. Тогда длина и относительное расположение гибридного участка указывают на местонахождение последовательности ДНК, представленной в мРНК.

РНК можно гибридизовать с одноцепочечной ДНК. В этом случае комплементарные области образуют двухцепочечный гибрид, в то время как другие области остаются одноцепочечными. Двухцепочечный участок имеет большую толщину, чем одноцепочечный, хотя иногда трудно точно разграничить расположение двух участков. На **рис. 20.1** показано, что метод позволяет обнаружить непрерывную последовательность ДНК, представленную также и в РНК, поскольку соответствующий участок имеет большую толщину, чем тонкая одиночная цепь ДНК, внутри которой он находится.

При использовании метода *картирования R-петель* РНК гибридизуют с двухцепочечной ДНК в условиях, при которых гибрид РНК—ДНК более стабилен, чем исходная двухцепочечная ДНК. Это позволяет заменить одну из цепей ДНК-дуплекса на РНК в той области, где

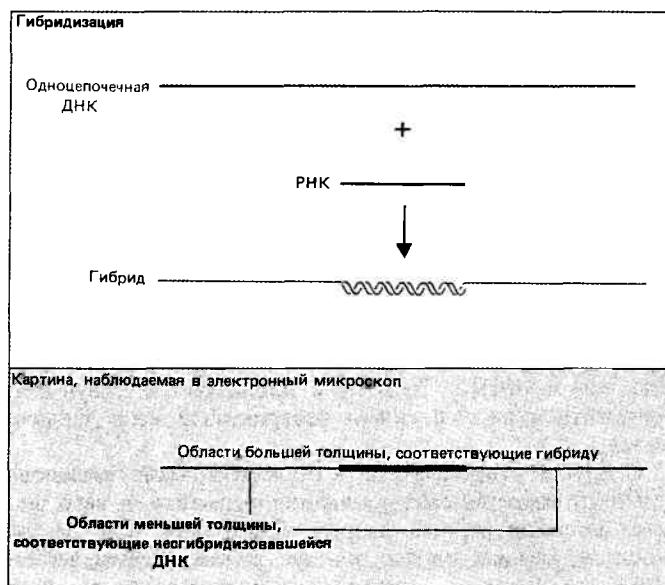


Рис. 20.1. Одноцепочечную ДНК гибридизуют с РНК с образованием структуры, в которой гибридную область можно обнаружить с помощью электронной микроскопии. Точная длина двухцепочечного участка зависит от условий, но 1 мкм примерно соответствует 3250 н. п.

В верхней части этого и следующих рисунков происходящие события иллюстрируются с точки зрения поведения отдельных цепей ДНК и РНК. В нижней части рисунков приведены действительно наблюдаемые в электронный микроскоп картины, где области большей и меньшей толщины соответствуют двух- и одноцепочечным участкам.

РНК может спариться с комплементарным ей участком. На рис. 20.2 показано, что замещенная на РНК цепь ДНК образует петлю, окруженную интактной двухцепочечной ДНК. Образование петли и дает название этому методу.

Основная особенность методов, приведенных на рис. 20.1 и 20.2, заключается в том, что, когда РНК соответствует непрерывной последовательности ДНК, обе

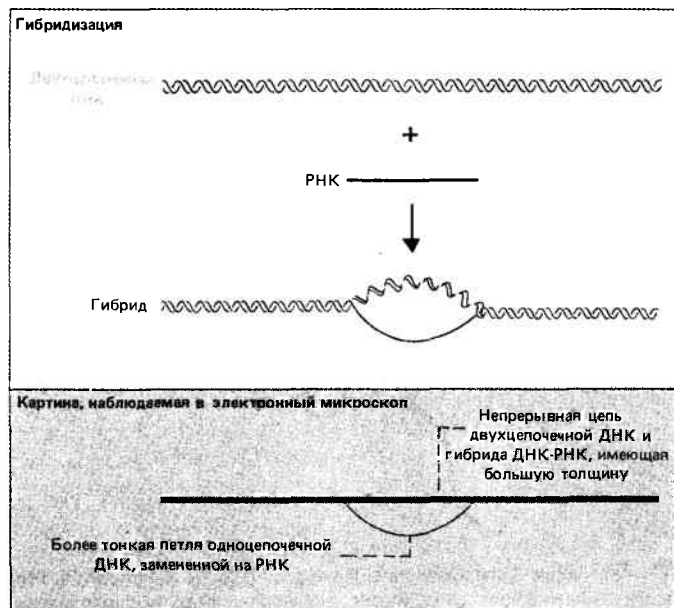


Рис. 20.2. В условиях образования R-петель РНК гибридизуется с ДНК, вытесняя одну из цепей двойной спирали.

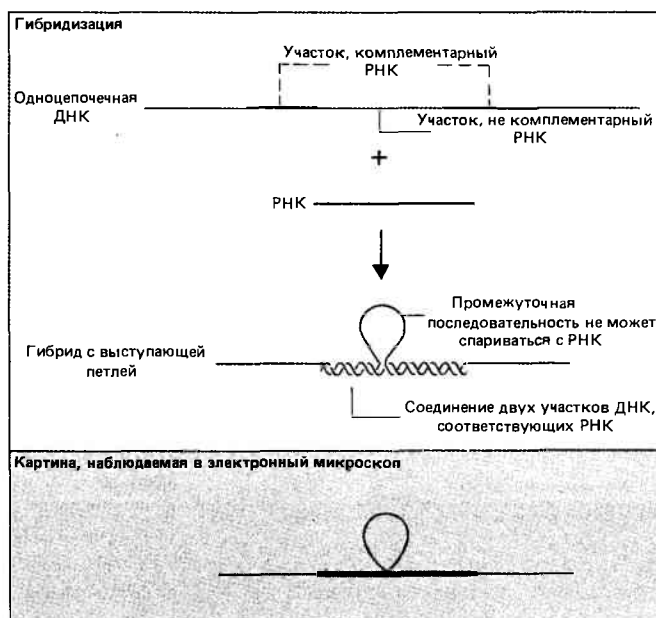


Рис. 20.3. При гибридизации РНК с двумя участками одноцепочечной ДНК, разделенными промежуточной последовательностью, место нахождения интрона выглядит как негибридизованная одноцепочечная петля ДНК, выступающая из гибрида.

молекулы колинеарны. Тогда внутри молекулы ДНК образуется единая непрерывная область гибрида.

Но когда в состав гена входят последовательности, отсутствующие в составе мРНК, эта часть ДНК не может гибридизоваться с РНК. Однако РНК может спариваться с последовательностями, находящимися по обе стороны от прерывающей последовательности, с образованием гибридных участков. Это приводит к возникновению характерных структур, выявляемых каждым из рассматриваемых электронно-микроскопических методов.

При гибридизации РНК с одноцепочечной ДНК промежуточная последовательность имеет вид одноцепочечной петли ДНК, выступающей из области РНК—ДНК-гибрида. Это показано на рис. 20.3. Длина и расположение петли внутри гибрида определяют размер и локализацию интрона внутри транскрипционной единицы.

При использовании метода картирования R-петель РНК замещает одну цепь ДНК, гибридизуясь с участками ДНК по обе стороны от промежуточной последовательности. Но сама промежуточная последовательность остается неизменной, сохраняя исходное двухцепочечное строение. В результате образуется структура, приведенная на рис. 20.4, где два участка, кодирующие РНК, в гибриде объединены, как это видно по двум вытесненным из гибрида одноцепочечным петлям ДНК. В месте соединения этих петель наружу вытесняется двухцепочечная петля ДНК, соответствующая промежуточной последовательности. В приведенном на рис. 20.5 примере виден единственный интрон β^{minor} -глобинового гена мыши. (В этом гене также имеется и второй интрон, размеры которого слишком малы для того, чтобы он был виден в электронный микроскоп; см. ниже.)

Первые результаты такого рода были получены для вирусных геномов: для аденовируса и вируса SV-40. В каждом случае при гибридизации поздних мРНК с вирусной ДНК основная часть мРНК образовывала гибриды с одним участком вирусной ДНК, но 5'-концевая область

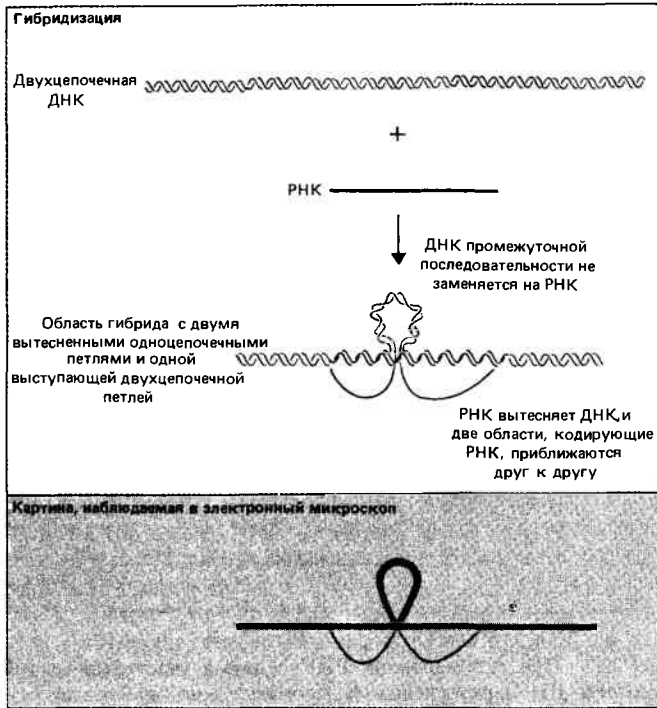


Рис. 20.4. При гибридизации РНК с прерывистым геном при картировании R-петель промежуточная последовательность гибридизоваться не может и остается в виде двухцепочечной ДНК. Она имеет вид двухцепочечной петли большей толщины, выступающей из области РНК—ДНК-гибрида. Последовательности ДНК, замененные на РНК, образуют одноцепочечные петли, соответствующие отдельным участкам, кодирующим РНК.

мРНК гибридизовалась с последовательностью или последовательностями, отстоявшими гораздо дальше от этого участка.

Для аденовируса показано (рис. 20.6), что 5'-конец поздней мРНК действительно образуется в результате объединения трех очень коротких последовательностей РНК, соответствующих далеко отстоящим участкам вирусного генома. В этом примере состоящий из трех частей 5'-участок соответствует нетранслируемой лидерной последовательности, в то время как кодирующая область представлена основной частью мРНК. На рис. 20.7 приведены данные электронной микроскопии по организации вирусного генома. Строение показанного там гибрида указывает на то, что при наличии нескольких промежуточных последовательностей структуры на микрофотографии становятся запутанными и сложными для интерпретации.

Электронно-микроскопическое картирование имеет хорошую разрешающую способность в диапазоне примерно 50–100 п.н., в зависимости от условий. Участки меньшего размера на электронно-микроскопической карте могут отсутствовать.

Рестрикционное картирование прерывистых генов

Рестрикционное картирование как мРНК, так и геномной ДНК наиболее легко осуществить, используя клонированные последовательности. Разумеется, при ра-

боте с мРНК необходимо в качестве ДНК-субстрата использовать клонированную кДНК. Работая с геном, можно использовать и другой способ—прямое расщепление геномной ДНК. Независимо от используемого материала наличие в гене последовательности, не входящей в состав соответствующей мРНК (кДНК), можно обнаружить одним из двух способов.

На рис. 20.8 показано, что если промежуточная последовательность не содержит сайта узнавания ни для одной используемой рестриктазы, длина соответствующего фрагмента изменяется. Наличие интрона означает, что между двумя сайтами рестрикции, расположенными по обе стороны от него, имеется большее расстояние в гене, чем в мРНК. Поэтому и длина соответствующего фрагмента в расщепленном рестриктазой гене увеличивается.

С другой стороны, если в промежуточной последовательности имеется сайт узнавания фермента, в него вносится дополнительный разрыв. Поскольку этот разрыв вносится только в ген, но не в последовательность мРНК, то при расщеплении гена получают два фрагмента вместо одного, образовавшегося из мРНК. Это показано на рис. 20.9.

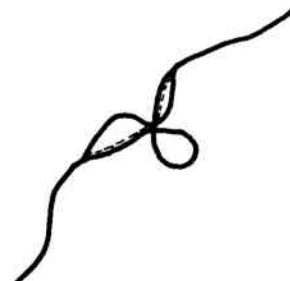
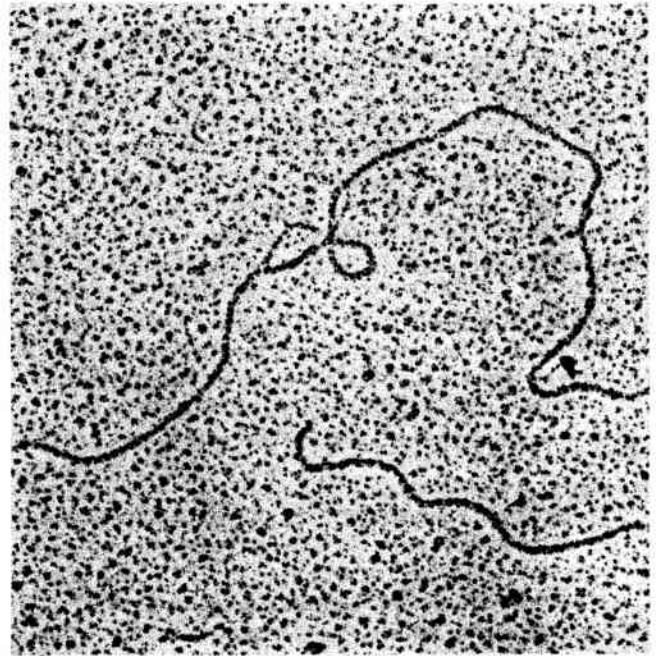
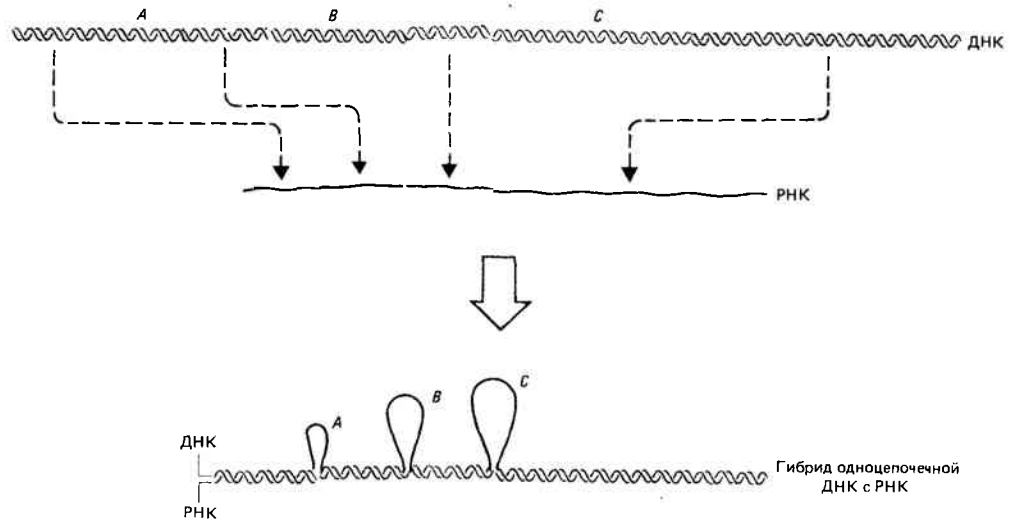


Рис. 20.5. При картировании R-петель β^{minor} -глобинового гена мышцы обнаруживается длинная промежуточная последовательность. Навверху приведена электронная микрофотография, а под ней—объясняющий ее рисунок.

Фотография любезно предоставлена Philip Leder

Рис. 20.6. Поздняя мРНК аденовируса образуется путем удаления трех промежуточных последовательностей.

В верхней части рисунка показана структура вирусного генома с промежуточными последовательностями А, В и С, которые удаляются из РНК. На расположенной ниже схеме приведен результат гибридизации мРНК с одноцепочечной ДНК, когда образуются участки гибридов, вытесняющие три одноцепочечные петли ДНК.



При картировании с использованием нескольких рестриктаз могут быть получены результаты и того, и другого типа. Разрывы, вносимые одними ферментами, не попадают в промежуточную последовательность; тогда длина соответствующих фрагментов увеличивается. При использовании других ферментов из той же промежуточной последовательности будут образовываться дополнительные фрагменты. Объединяя вместе эти данные, мы видим, что при создании полной рестрикционной карты гена, содержащего интроны, карта каждого конца гена соответствует карте каждого конца последовательности мРНК. Но в некоторой точке внутри гена карты различаются; там имеется дополнительная область, содержащая ряд сайтов узнавания, отсутствующих в мРНК. Разрешающая способность метода рестрикционного картирования позволяет обнаруживать участки гена размером до 20–30 п. н.

На рис. 20.10 приведено сравнение рестрикционных карт β-глобинового гена и глобиновой мРНК. При этом обнаруживаются два интрона. Интрон большего размера совпадает с интроном, обнаруженным с помощью электронной микроскопии и показанным на рис. 20.5. Меньший интрон присутствует только на рестрикционной карте.

Полезный метод исследования прерывистых генов состоит в изучении свойств гибридов, образованных между мРНК и ДНК, как это показано на рис. 20.11. Помимо гибридных РНК—ДНК-участков молекула содержит рас-

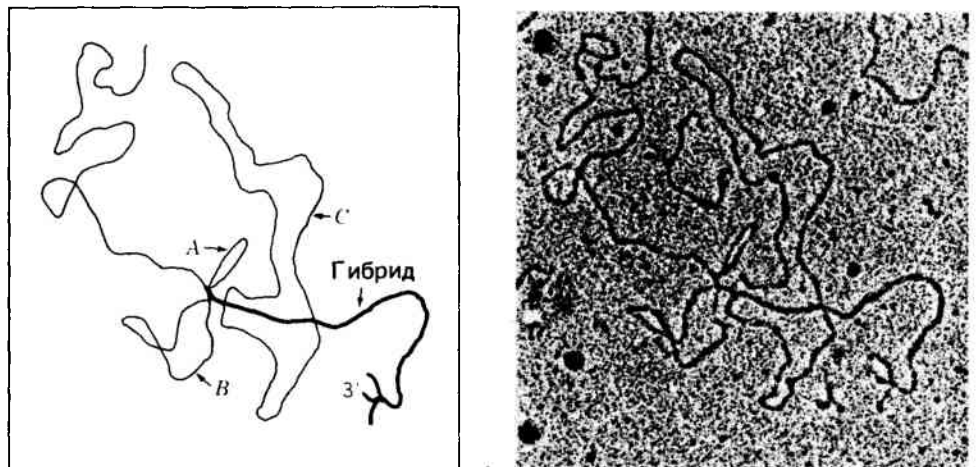
положенные по бокам от них (фланкирующие) участки одноцепочечной ДНК и вытесненные одноцепочечные интроны. Существует два способа обработки молекул.

Фланкирующую ДНК можно удалить обработкой экзонуклеазой VII. Затем с помощью щелочного гидролиза РНК легко выделить одноцепочечную ДНК, по длине соответствующую самому гену. Отдельные участки гена можно выделить обработкой нуклеазой S1, ферментом, специфически расщепляющим любую одноцепочечную ДНК, включая даже очень короткие одноцепочечные разрывы в двухцепочечной молекуле. В результате такой обработки удаляются фланкирующие области и интрон. Продукт реакции—РНК—ДНК-гибрид, по длине равный мРНК, с разрывом, нарушающим ковалентную непрерывность одной из цепей ДНК. РНК удаляют обработкой щелочью, в результате образуются индивидуальные фрагменты ДНК. Они соответствуют отдельным экзонам гена.

В таком виде, как это показано на рисунке, метод известен как «картирование методом Берка—Шарпа», по имени его авторов. В более общем виде он известен как картирование с помощью нуклеазы S1. Свойство этого фермента разрушать все связи в неспаренных участках ДНК оказывается очень полезным. Оно часто используется для локализации конца последовательности РНК на соответствующем участке ДНК. Например, ДНК гибридизуют с РНК и затем для удаления фланкирующих

Рис. 20.7. Одна из оригинальных электронных микрофотографий гибридизации РНК с одноцепочечной ДНК аденовируса, выявляющей мозаичную структуру транскрипционной единицы поздних генов.

Фотография любезно предоставлена Philip Sharp



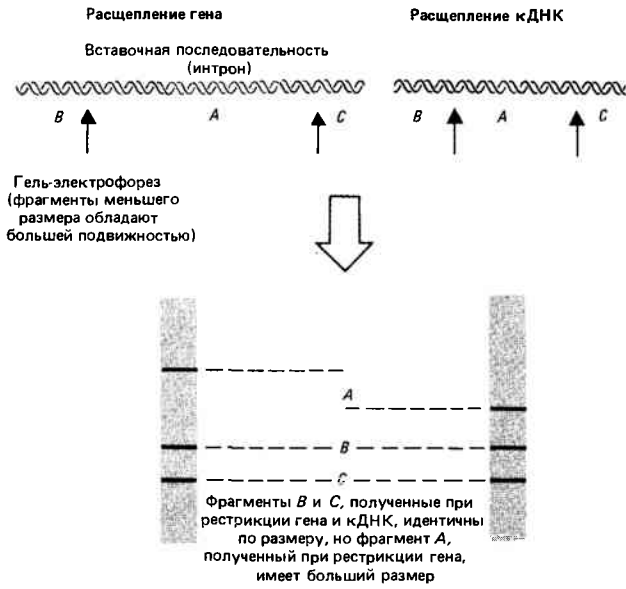


Рис. 20.8. При наличии интрона между двумя сайтами рестрикции размер соответствующего рестрикта увеличивается.

На рисунке приведено сравнение гена и кДНК, имеющих два сайта рестрикции. При рестрикции образуются три фрагмента, обозначенные как А, В и С в порядке уменьшения их размера. Промежуточная последовательность располагается между двумя сайтами, рестрикция которых приводит к образованию фрагмента А, поэтому этот фрагмент имеет больший размер в случае рестрикции гена.

участков используют нуклеазу S1. Тогда участок ДНК оканчивается точно в точке, соответствующей концу молекулы РНК (как было показано на рис. 11.2). При сравнении нуклеотидных последовательностей такого фрагмента ДНК и исходной ДНК можно точно установить местоположение начала мРНК.

Характеристика фрагментов геномной ДНК

При расщеплении ДНК генома рестриктазой, не разрезающей специфическую последовательность, соответствующую мРНК, разрывы вносятся по обе стороны от гена. Сами по себе такие разрывы приводят к образованию одного фрагмента. Какой же вывод можно сделать, когда в таком эксперименте в действительности обра-

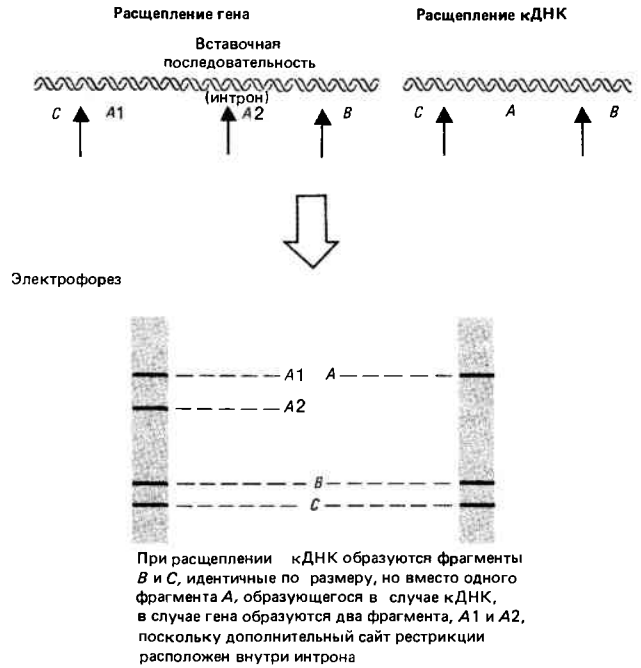


Рис. 20.9. Наличие вставочной последовательности приводит к образованию дополнительных рестриктов в том случае, когда она содержит сайт рестрикции для специфического фермента.

Наличие дополнительного сайта рестрикции в гене приводит к увеличению числа вносимых разрывов до трех вместо двух, обнаруживаемых в случае кДНК. Поэтому вместо фрагмента А, образующегося при рестрикции кДНК, при рестрикции геномной ДНК получают фрагменты А1 и А2. Суммарный размер А1 + А2 превышает размер А на длину вставочной последовательности.

зуются два (или более) фрагмента, гибридизующиеся с мРНК?

Одна из возможностей заключается в том, что в геноме имеется более одной копии гена, соответствующего мРНК. (Копии не обязательно должны быть идентичными, они просто могут обладать достаточным сходством, чтобы гибридизоваться друг с другом.) В этом случае нуклеотидные последовательности по обе стороны от каждой копии могут различаться, так что сайты рестрикции будут располагаться случайным образом по отношению к гену. Тогда все копии будут выделены в виде фрагментов рестрикции разных размеров.

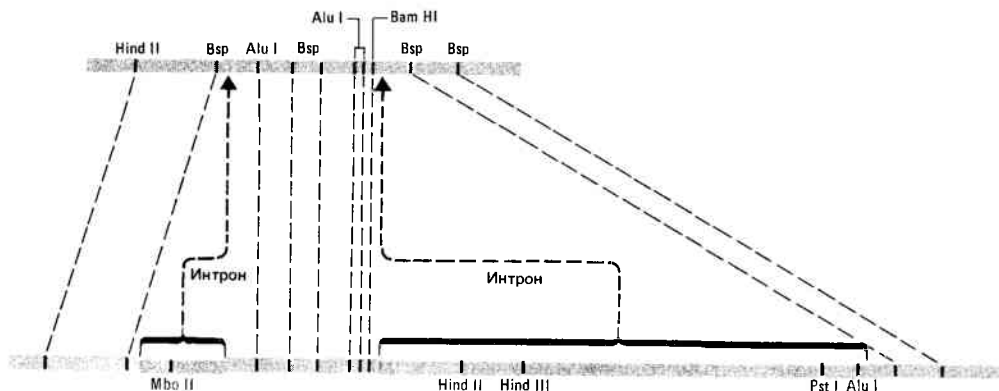


Рис. 20.10. При сравнении рестрикционных карт кДНК и геномной ДНК β-глобина мыши обнаруживается, что в гене

имеются два дополнительных участка, отсутствующие в кДНК. Другие области кДНК и гена полностью совпадают.

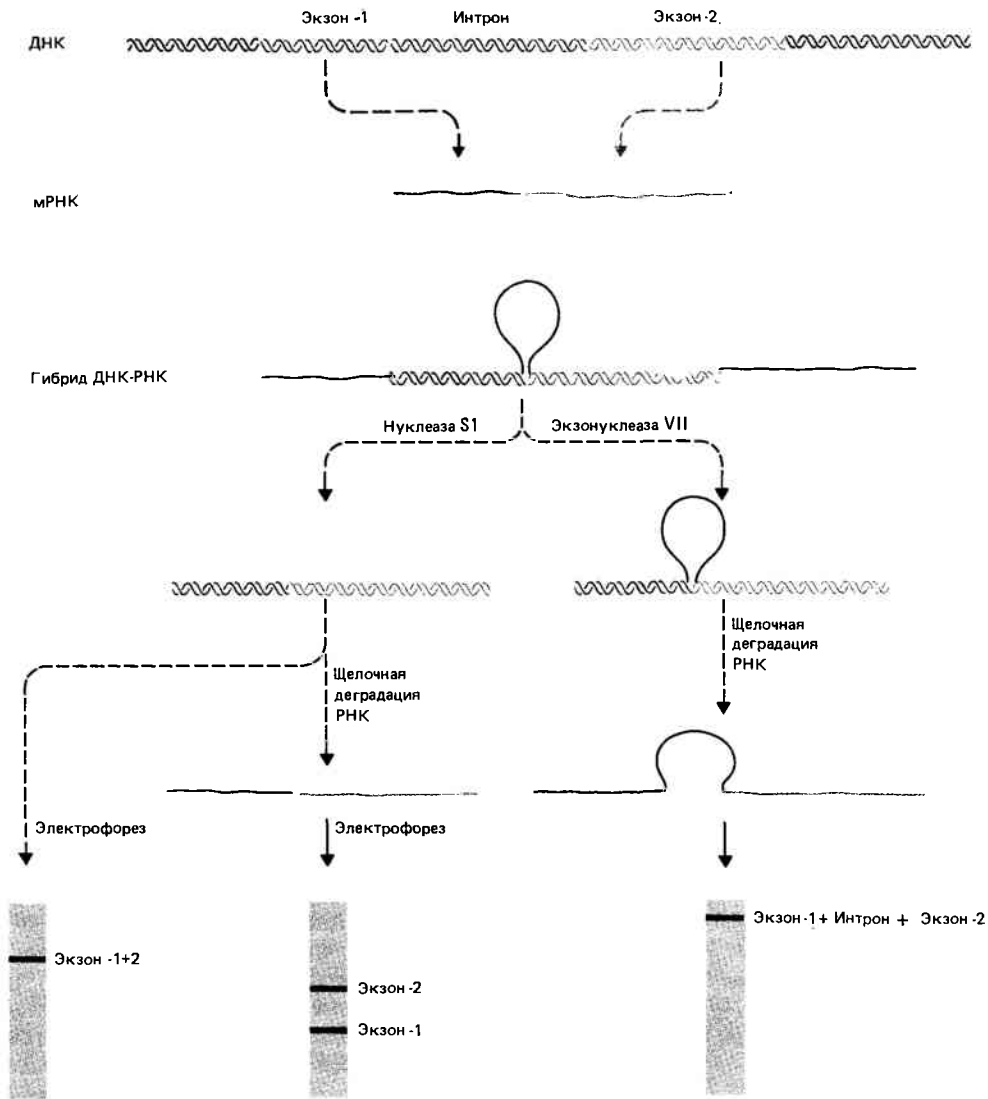


Рис. 20.11. Метод картирования с помощью нуклеазы S1 позволяет определить длину как интронов, так и экзонов.

Другая возможность состоит в том, что в состав гена входит промежуточная последовательность (или последовательности), имеющая сайт узнавания рестриктазы. Действительно, наличие каждого интрона, содержащего такой сайт, приведет к образованию дополнительного фрагмента. (Противоположный результат, когда получают один фрагмент, не доказывает отсутствия промежуточной последовательности, поскольку интрон не обязательно содержит сайт рестрикции. Это также не исключает возможности существования небольшого числа расположенных рядом копий.)

Как же определить, имеем ли мы дело с множеством копий одного гена или с прерывистым геном? Обычно для решения этой проблемы используют зонды, полученные из субклонов, содержащих последовательности, соответствующие только определенным областям мРНК. Такие последовательности получают путем разрезания клонированной кДНК на 5'- и 3'-области подходящей рестриктазой (т.е. такой, которая вносит разрыв в нужное место последовательности, соответствующей мРНК.) Затем такие специфичные по отношению к концам РНК зонды по отдельности повторно клонируют в других плаزمидях.

На рис. 20.12 показано, что если два фрагмента – разные копии одного гена, то каждый из них будет содержать его полную последовательность, а значит, оба фрагмента будут гибридизоваться с зондами, специфичными и для 5'-, и для 3'-концов.

На рис. 20.13 показано, что если два фрагмента генома представляют собой разные участки одного гена, образовавшиеся при внесении разрыва в интрон, то один из них будет гибридизоваться с кДНК-зондом для 5'-конца, а другой – с зондом для 3'-конца.

Как при получении фрагментов прерывистого гена можно быть уверенным в том, что обнаружены все интроны и все экзоны? Структуру гена устанавливают анализом *перекрывающихся* фрагментов, соответствующих различным его областям. Это общий принцип рестрикционного картирования и определения последовательности ДНК (описанных в гл. 3), и вероятность пропуска части гена в данном случае будет исключена.

Предположим, например, что ген содержит интрон, имеющий два сайта рестрикции для некоторого фермента, как показано на рис. 20.14. При разрезании гена по этим сайтам два фрагмента – левый (А) и правый (В) – будут состоять частично из последовательности, представ-

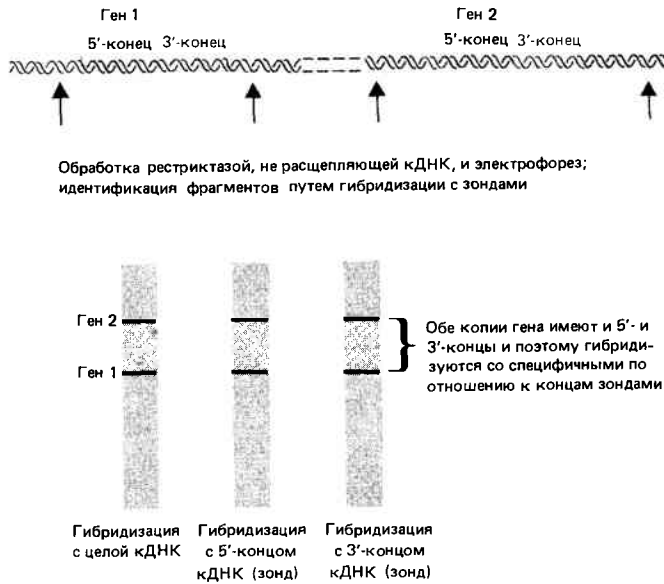


Рис. 20.12. При расщеплении ДНК генома рестриктазой, не вносящей разрывов в кДНК, в случае повторяющихся генов образуется множество интактных копий.

При расщеплении двух идентичных генов, содержащих 5'- и 3'-концевые участки, образуются различающиеся фрагменты, поскольку сайты рестрикции находятся в различных участках фланкирующей ДНК. Фрагменты идентифицируют с помощью блоттинга ДНК после электрофореза и гибридизации ее с меченой кДНК-зондом. Так как каждый фрагмент включает оба конца гена, он гибридизуется с кДНК-зондами, специфичными по отношению как к 5'-, так и к 3'-концевым последовательностям.

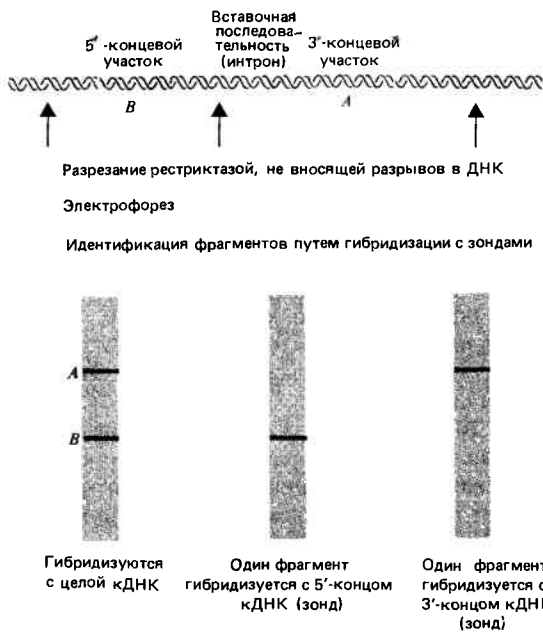


Рис. 20.13. При обработке ДНК генома рестриктазой, не вносящей разрывов в кДНК, из одного гена с интроном (интронами) может образоваться несколько фрагментов.

Если фермент вносит разрыв в интрон, образуются два фрагмента, которые могут гибридизоваться с меченой кДНК-зондом. Но при использовании зондов, специфичных по отношению к 5'- и 3'-концам, каждый фрагмент гибридизуется только с одним зондом.

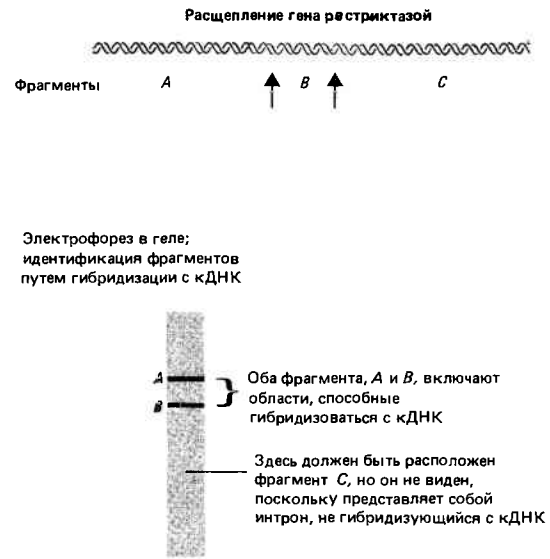


Рис. 20.14. При выявлении фрагментов генома путем гибридизации с кДНК или мРНК могут быть не обнаружены участки интронов, находящиеся между двумя сайтами рестрикции.

ленной в мРНК, и частично из интрона. Наличие первого компонента делает возможной гибридизацию общих фрагментов с зондом, комплементарным мРНК. Однако центральный компонент (С) состоит *только* из интронов. Поскольку в его состав не входит последовательность, соответствующая мРНК, он с зондом не гибридизуется.

Поэтому фрагмент С не будет обнаружен среди последовательностей, полученных при рестрикции ДНК данным ферментом. Такого рода случайности можно избежать, используя несколько ферментов и составляя карту таким образом, чтобы место соединения каждой пары соседних фрагментов также было бы целиком представлено в каком-либо другом фрагменте. Этот метод может включать использование в качестве зондов участков интронов.

В конечном итоге сравнение нуклеотидных последовательностей гена и клон кДНК точно выявляет местоположение и размер промежуточных последовательностей. Такой анализ на уровне последовательностей необходим, прежде чем мы сможем быть уверенными в том, что выявлены все участки гена. Действительно, в ряде случаев было неожиданно обнаружено, что геномная нуклеотидная последовательность содержит дополнительный короткий кодирующий участок. Когда кодирующий участок имеет длину менее 50 п.н., он может не гибридизоваться с кДНК-зондом и поэтому, попадая внутрь находящейся справа или слева от гена промежуточной последовательности, остаться незамеченным.

Гены имеют самое разнообразное строение и размеры

В настоящее время не существует единой схемы или объяснения существования многообразия структур эукариотических генов. Некоторые гены не прерываются, и в этом случае нуклеотидные последовательности генома коллинеарны последовательностям мРНК. Сейчас считают, что большинство генов имеет прерывистое строе-

ние, но вставочные последовательности очень сильно различаются как по числу, так и по размерам.

Некоторые основные черты характерны для всех прерывистых генов. Порядок расположения участков в прерывистом гене совпадает с их расположением в зрелом РНК-продукте. Поэтому гены являются скорее расщепленными, чем разбросанными. Интроны ядерных генов во всех рамках считывания обычно имеют терминирующие кодоны и поэтому не могут кодировать полипептиды. Прерывистые гены сохраняют одинаковую структуру во всех тканях, включая половые и соматические клетки, где эти гены экспрессируются или не экспрессируются. Таким образом, наличие вставочной последовательности – постоянное свойство гена.

Все классы генов могут иметь прерывистое строение: все гены, кодирующие белки, а также гены, кодирующие рРНК, и гены, кодирующие тРНК. Интроны обнаружены также в митохондриальных генах дрожжей и хлоропластных генах. Прерывистые гены, по-видимому, присутствуют в клетках эукариот всех классов, хотя их содержание варьирует. Например, их доля среди ядерных структурных генов позвоночных может превышать такую у грибов.

Промежуточные последовательности должны быть удалены до трансляции (или до того, как прерывистые предшественники рРНК или тРНК смогут осуществить свои функции). Они присутствуют в первичном транскрипте, но отсутствуют в зрелом РНК-продукте. Процесс удаления промежуточных последовательностей носит название **сплайсинга РНК**. Суть сплайсинга состоит в постепенном удалении всех промежуточных последовательностей из первичного продукта транскрипции со сшиванием концов участков РНК, находящихся по обе стороны от интрона, с образованием ковалентно-непрерывной РНК. Это показано на **рис. 20.15**. Механизм сплайсинга РНК будет обсужден в гл. 26.

Некоторые прерывистые гены имеют только один или всего несколько интронов. Пример хорошо изученного строения прерывистых генов – строение глобиновых генов (гл. 21). Два основных вида глобиновых генов, α и β , имеют общий тип структуры, при котором два интрона занимают постоянные положения относительно кодирующей последовательности. Для β -глобиновых генов млекопитающих характерно постоянство их строения, проиллюстрированное на **рис. 20.16**. В полипептидной цепи β -глобина мыши, состоящей из 146 аминокислот, интроны расположены между кодонами, занимающими положения 30/31 и 105/106.

Для всех известных активных генов глобина, включая гены ряда млекопитающих, птиц и лягушки, интроны занимают одни и те же положения. Первый интрон всегда довольно короткий, второй обычно имеет большую длину, но абсолютные длины интронов могут различаться. Большая часть различий в общей длине разных глобиновых генов обусловлена различиями в длине второго интрона. У мыши длина второго интрона α -глобинового

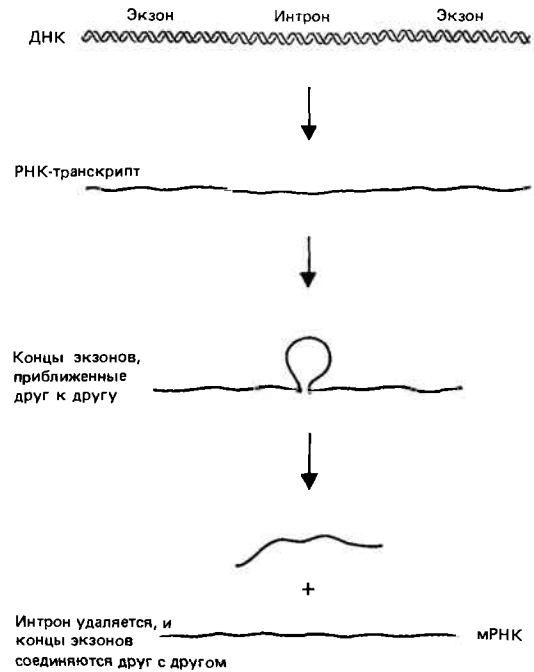


Рис. 20.15. Прерывистый ген состоит из чередующихся экзонов и интронов. Интроны удаляются при сплайсинге РНК-транскрипта, в результате чего образуется мРНК, состоящая только из последовательностей, комплементарных экзонам.

гена составляет только 150 п.н., поэтому общая длина гена равняется 850 п.н. по сравнению с 1382 п.н. β -глобинового гена. Таким образом, различия в длинах генов превышают различия в длинах мРНК (длина α -глобиновой мРНК – 585 нуклеотидов, β -глобиновой мРНК – 620 нуклеотидов). Центральный экзон каждого из этих генов имеет постоянную длину, поэтому различия в длинах мРНК – следствие различий в длинах первого и третьего экзона. Это вызвано главным образом различиями в длинах нетранслируемых участков (лидерных и концевых).

Другие гены могут иметь большее количество интронов. Два примера приведены на **рис. 20.17**. Гены овальбумина и кональбумина курицы содержат соответственно 8 и 17 экзона. В результате суммарные длины этих генов составляют 7500 и 10 000 п.н. по сравнению с длинами соответствующих им мРНК, равными 1859 и 2500 нуклеотидных оснований.

(В этом и в других случаях мы определяем размер структурного гена как расстояние в геноме между точками, соответствующими 5'- и 3'-концевым основаниям мРНК. Известно, что транскрипция, по всей вероятности, начинается с места расположения 5'-конца мРНК; однако окончательно не установлено, заканчивается ли она в месте, соответствующем 3'-концу мРНК, или этот конец образуется при расщеплении более длинного предше-

	Экзон 1	Интрон 1	Экзон 2	Интрон 2	Экзон 3
Длина, п. н.	142–145	116–130	222	573–904	216–255
Соответствует	5'-нетранслируемой области + аминокислотам 1–30		аминокислотам 31–104		аминокислотам 105-конца полипептидной цепи + 3'-нетранслируемой области

Рис. 20.16. Все функционирующие глобиновые гены имеют прерывистое строение и содержат три экзона. Размеры участков ге-

на, указанные на рисунке, относятся к β -глобиновым генам млекопитающих.

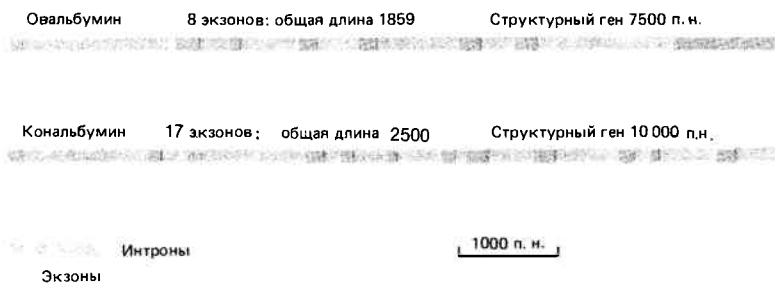


Рис. 20.17. Масштабные карты двух генов курицы показывают, что экзоны в основном представляют собой участки сравнительно небольшого размера (порядка 100 п.н.), связанные существенно более длинными интронами (хотя некоторые экзоны также имеют довольно большую длину).

ственника. В последнем случае ген будет иметь больший размер и включать дополнительные последовательности правее 3'-конца мРНК.)

Существуют гены, имеющие как большую длину, так и большую степень мозаичности по сравнению с приведенными примерами. В соответствии с имеющимися в настоящее время данными эти параметры достигают предельных значений в случае генов дигидрофоллат-редуктазы (ДГФР) мыши и α -коллагена цыпленка.

Ген ДГФР мыши включает 6 экзонов, соответствующих мРНК длиной 2000 нуклеотидных оснований. Но весь ген занимает участок ДНК размером более 31 000 п.н. ДНК. В данном случае интроны имеют чрезвычайно большую длину. Ген про- α 2-коллагена цыпленка разделен более чем на 50 экзонов, каждый из которых довольно короткий. Размеры охарактеризованных экзонов колеблются от 45 до 249 п.н. Одни интроны имеют небольшую длину, сравнимую с длиной экзонов, но другие существенно длиннее. В результате общая длина экзонов — около 5000 п.н. — рассеяна в пределах участка генома размером 40 000 п.н.

Таким образом, для некоторых эукариотических структурных генов экзоны составляют лишь небольшую часть их общей длины. Эти данные в сочетании с данными о небольших размерах экзонов дают представление о гене как о группе коротких экзонов, разбросанных по большому участку генома, который (возможно) выполняет всего лишь функцию поддержания целостности структурного гена. На карте таких генов их экзоны выглядят как прерывающиеся интроны участки, а не наоборот. Это ясно из рис. 20.17. Многие гены должны быть как бы старательно собраны путем соединения множества коротких экзонов, расположенных на значительных расстояниях друг от друга.

Из имеющихся в настоящее время данных нельзя сделать четкого заключения о том, что составляет большую часть генома (экзоны или интроны), но ясно, что мозаичное строение гена позволяет приблизиться к объяснению того факта, что в мРНК представлена небольшая часть ДНК. Насколько приблизиться — покажет время. Мы уже высказывали сомнение в том, что эти данные позволяют решить проблему выяснения всех функций, осуществляемых ДНК.

Интроны генов, кодирующих рРНК и тРНК

В ряде случаев в составе генов, кодирующих высокомолекулярные РНК, имеются интроны. Прерванные рибосомные гены были впервые обнаружены у *D. melanogaster*, у которой около двух третей генов, кодирующих 28S-рРНК, прерываются последовательностью длиной 5000 п.н. (Гены, кодирующие рРНК, всегда при-

сутствуют во множестве копий, обычно идентичных; см. гл. 24.) Но поскольку оставшаяся треть не имеет прерывистого строения и поскольку, по-видимому, прерывистые гены не используются для синтеза рРНК, сохраняется возможность того, что это поврежденные нефункционирующие гены. Действительно, их интрон может представлять собой вставку в генах, имевших ранее непрерывное строение (гл. 37).

Однако для некоторых низших эукариот все копии генов, кодирующих высокомолекулярную рРНК, являются прерванными. К ним относятся гены *Tetrahymena pigmentosa* (жгутиковые водоросли) и *Physarum polycephalum* (слизистые грибы). Таким образом, строение этих генов аналогично строению неповторяющихся ядерных генов, когда для получения активного продукта гена при сплайсинге должен быть удален интрон. Однако в клетках других низших эукариот, таких, как *S. cerevisiae* (дрожжи) и *D. discoideum* (слизистые грибы), гены рРНК не прерываются. В митохондриях *S. cerevisiae*, *N. crassa* (грибы) и в хлоропластах *C. reinhardtii* (водоросли) опять-таки ген высокомолекулярной рРНК прерывается, но он не прерывается в геноме митохондрий и хлоропластов других видов организмов. По-видимому, не существует определенных правил распределения в геноме прерывистых и непрерывистых генов рРНК. Однако почти во всех случаях, кроме одного, у них имеется только один интрон. Он всегда расположен на расстоянии, составляющем около двух третей длины гена от его начала. (Длина молекулы рРНК колеблется от 3000 до 4000 нуклеотидных оснований.) Может ли такое расположение совпадать для столь сильно различающихся организмов? Гены, кодирующие низкомолекулярные рРНК, не прерываются.

Имеется два случая полиморфизма в плане наличия интрона: некоторые штаммы организма имеют прерывистый ген для рРНК; в других штаммах ген не прерывается. Это наблюдается для ядерных генов *T. pigmentosa* и митохондриальных генов *S. cerevisiae*. Ни в одном случае не наблюдается различий в функционировании гена. Существование двух форм гена позволяет предположить, что интрон не является ни необходимым, ни вредным для работы гена.

Интроны были обнаружены в ядерных генах тРНК дрожжей. Интересная особенность в этом случае состоит в том, что интрон всегда занимает одно и то же относительное положение, располагаясь в начале петли антикодона. Это позволяет содержащим интрон предшественникам тРНК принимать конформацию, при которой антикодон спарен с частью интрона. Это может иметь отношение к механизму сплайсинга (см. рис. 26.1).

Длина интронов разных генов тРНК колеблется в широких пределах: от такой небольшой величины, как 14 п.н., до 46 п.н. И все-таки наличие интронов не является основной чертой генов дрожжевой ядерной

тРНК, поскольку некоторые гены тРНК не относятся к числу прерывистых. Ряд членов семейства генов тРНК могут иметь прерывистое строение, в то время как другие, родственные им гены являются непрерывными.

Интроны — неповторяющиеся и быстро эволюционирующие компоненты генома

Используя фрагменты рестрикции, соответствующие определенным участкам структурного гена, можно определить, имеют ли они сходство с другими нуклеотидными последовательностями генома. Клонированный рестрикт может быть использован в качестве зонда при гибридизации со всей ДНК генома для определения частоты его повторяемости или для обнаружения соответствующей последовательности в наборе рестриктов. Для этого можно использовать фрагменты, представляющие как экзоны, так и интроны.

Часто оказывается, что экзоны одного гена сходны с экзонами другого, и при использовании данного экзона в качестве зонда обнаруживаются фрагменты, входящие в состав другого гена или генов. На основе этих данных возникает предположение о том, что такие гены возникли в результате дупликации общего гена-предка, в копиях которого затем произошло накопление мутаций.

Обычно интроны не имеют сходства с другими последовательностями. А когда два гена имеют сходные экзоны, их интроны, по-видимому, обладают меньшим сходством, чем экзоны. Это означает, что, в то время как экзоны осуществляют жесткую функцию кодирования определенной аминокислотной последовательности, интроны ее не кодируют и поэтому в них может накапливаться больше мутаций.

Таким образом, в рамках первоначального представления о неповторяющейся ДНК как состоящей из уникальных последовательностей генома это свойство более всего связано с наличием интронов. Задаваясь вопросом, являются ли структурные гены неповторяющимися, мы, следовательно, видим, что ответ неоднозначен: ген как таковой в целом может быть уникальным, но его экзоны иногда имеют сходство с экзонами других генов. Поэтому, по крайней мере для некоторых генов, экзоны представляют собой слабо повторяющиеся последовательности, окруженные уникальными последовательностями интронов.

Для глобиновых генов было проведено детальное сопоставление соответствующих экзонов и интронов; более подробно это будет обсуждаться в гл. 21. Окончательный результат дивергенции экзонов выражается в различиях кодируемых ими белков. Эти различия в основном обусловлены заменами оснований. Однако многие замены не влияют на смысловое значение кодона, поскольку затрагивают третье основание кодона или находятся в не-транслируемых областях. Область гомологии простирается и за границы экзон—интрон и захватывает небольшой участок внутри интрона.

Последовательности вторых интронов β-глобиновых генов сильно дивергировали, что подтверждается почти двукратным различием их размеров. Процесс дивергенции вторых интронов включает как изменение размеров (вследствие делеций и вставок), так и замены оснований. Для некоторых пар β-глобиновых генов вторые интроны имеют такие большие различия, что соответствующие

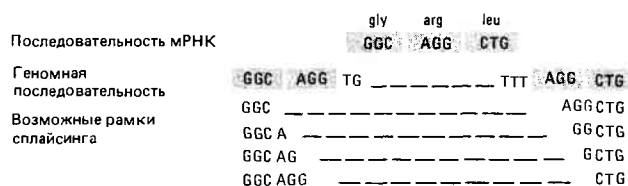


Рис. 20.18. Повторяемость тринуклеотида AGG на двух концах второго интрона β-глобинового гена мыши допускает существование нескольких возможных рамок сплайсинга.

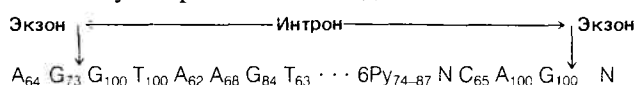
последовательности нельзя узнать. Иными словами, интроны эволюционируют гораздо быстрее, чем экзоны. Наличие таких изменений позволяет предположить, что последовательность интрона не несет специфической функции (или по крайней мере не обеспечивает осуществления одной и той же функции у разных видов организмов). Необходимо ли вообще его присутствие для функционирования гена, не известно, за исключением того факта, что функционирующие β-глобиновые гены, у которых отсутствовал бы второй интрон, обнаружены не были.

На границах экзон—интрон имеется каноническая последовательность

При сравнении нуклеотидных последовательностей мРНК и структурного гена можно выделить области границ между экзонами и интронами. Для них характерно наличие двух важных свойств (или их отсутствие). Во-первых, отсутствие сколько-нибудь значительной гомологии между двумя концами интрона. Это исключает возможность образования значительного по размеру участка со вторичной структурой, связывающего концы интрона вместе, что послужило бы предварительным этапом для сплайсинга. Во-вторых, оказалось, что на границе экзон—интрон имеется каноническая, присутствующая в разных генах, хотя и довольно короткая, последовательность; повсеместное присутствие этой последовательности вызывает предположение о ее участии в сплайсинге в ядре (гл. 26).

Во многих случаях невозможно однозначно определить расположение границы экзон-интрон, основываясь исключительно на сравнении мРНК и гена. Сложность состоит в том, что на каждом конце интрона повторяется короткая последовательность, обычно составляющая от 1 до 4 п.н. На рис. 20.18 приведен участок β-глобиновой последовательности мыши (или кролика), в которой любая из четырех пар сайтов может соответствовать концам интрона.

Несмотря на указанную трудность, концы каждого интрона можно обнаружить, основываясь на гомологии последовательностей, расположенных на границах экзон-интрон. Они могут быть изображены в виде участков, сходных с универсальной последовательностью:



Стрелки обозначают предполагаемые концы интрона. (В этом случае так же, как и в других, мы приводим ну-

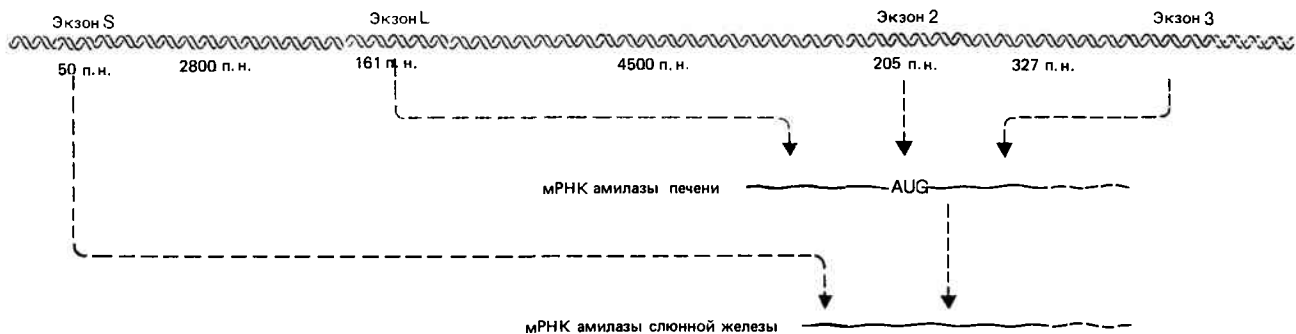


Рис. 20.19. Различия в сплайсинге обуславливают образование двух мРНК амилазы, транскрибируемых на одном и том же гене мыши.

клеотидную последовательность цепи ДНК, идентичную РНК-продукту.)

Индексы обозначают процент случаев, когда данное основание (или тип основания) обнаруживается на границе экзон-интрон. Действительно, высокий процент таких оснований обнаруживается только внутри интрона около предполагаемых границ экзон-интрон. Это и определяет истинные концы интрона:



Поскольку интрон, определенный таким способом, начинается с динуклеотида GT и заканчивается динуклеотидом AG, такую закономерность часто называют *правилом GT-AG*.

Обратите внимание на то, что две границы различаются по последовательности оснований, поэтому они однозначно определяют расположение концов интрона. Их обозначают, следуя слева направо вдоль интрона, как *левую* и *правую* точки сплайсинга. Иногда эти участки называют соответственно *донорным* и *акцепторным* сайтами, но мы будем избегать такого определения, поскольку оно предполагает наличие механизма, существование которого не доказано.

Универсальная последовательность обнаружена в ядерных генах многих эукариот. Насколько нам известно, она имеется у всех высших эукариот, поэтому можно предполагать существование общего механизма удаления интронов из РНК. Однако правило GT-AG не распространяется на интроны митохондрий и хлоропластов, а также на гены дрожжевой тРНК; по крайней мере в этих случаях механизмы сплайсинга могут иметь существенные различия (гл. 26).

Интрон одного гена может быть экзоном другого гена

До сих пор мы подходили к строению прерывистых генов как к определенному чередованию интронов и экзонов. Несмотря на то что наличие такой организации едва ли можно было ожидать на основании исходного определения понятия «цистрон», она не находится с ним в противоречии. На практике мозаичность структуры гена приводит к тому, что мутации, изменяющие белки, будут скапливаться в группах участков, соответствующих экзонам, а не распределяться непрерывно по ДНК. Никакие мутации в каждой такой группе не будут комплементарны друг другу или мутациям в других группах. Иными сло-

вами, мозаичный ген экспрессируется только посредством сплайсинга, объединяющего экзоны в одну молекулу РНК, что делает справедливой исходную концепцию *цис-транс-теста* (гл. 2).

Будут ли мутации, затрагивающие интроны, проявляться в нарушении функционирования гена? Известно, что, поскольку интроны быстро эволюционируют, их последовательности менее консервативны, из чего следует, что многие, а возможно, и большинство изменений в интроне останутся незамеченными. Более того, эти последовательности будут удалены при экспрессии гена. Имеется несколько примеров мутаций в интронах ядерных генов, которые влияют на узнавание РНК аппаратом сплайсинга. Обычно такие изменения затрагивают универсальные пограничные последовательности, препятствуя правильному удалению интронов (гл. 26). Такие мутации относятся к числу *цис-мутаций* и поэтому могут рассматриваться как дополнительные члены одной группы комплементации.

В отличие от обычной ситуации, когда ген мозаичен, но во всем остальном соответствует традиционным представлениям, имеется ряд примеров, когда однозначное соответствие между геном и белком отсутствует. При этом одному и тому же участку ДНК может соответствовать более чем одна последовательность мРНК. Поэтому информация, закодированная в специфическом участке ДНК, не является однозначной, а зависит от способа его экспрессии. В таком случае нельзя ожидать, что комплементация мутаций будет осуществляться обычным способом.

На рис. 20.19 суммированы тканеспецифические особенности образования мРНК амилазы мыши. Один и тот же белок, амилаза, синтезируется в печени и в слюнной железе мыши и кодируется одним и тем же геном. Кодировочные области мРНК, присутствующих в каждой ткани, идентичны, но начальные участки 5'-нетранслируемых лидерных последовательностей различаются. Лидерная последовательность мРНК печени длиннее и по последовательности оснований отличается от соответствующей последовательности мРНК слюнной железы.

Последовательность, кодирующая амилазу, начинается примерно с 50-й п.н. экзона 2 и формируется путем присоединения экзона 2 к экзону 3 и последующим экзонам. Этой последовательности может предшествовать один из двух экзонов. В печени первые 161 основание мРНК кодируются экзонами L, расположенным примерно за 4500 п.н. от экзона 2. В слюнной железе первые 50 оснований мРНК кодируются экзонами S, находящим-

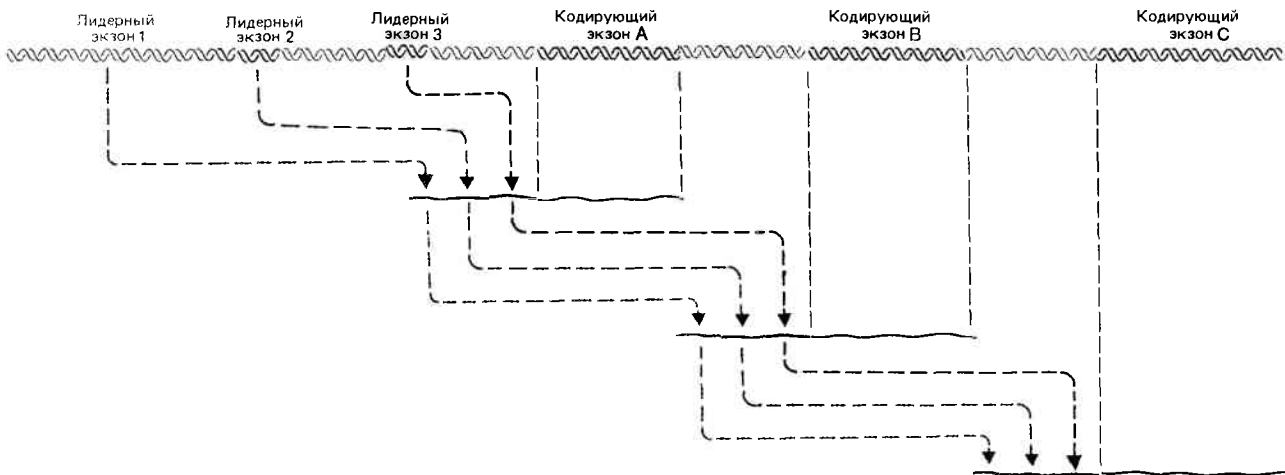


Рис. 20.20. При экспрессии поздних генов аденовируса одна и та же лидерная последовательность может быть присоединена к различным кодирующим участкам мРНК. На рисунке показаны только три кодирующих экзона, однако в действительности

имеется пять групп кодирующих последовательностей, определяемых наличием общих 3'-концов. Эти группы включают по меньшей мере девять различных индивидуальных кодирующих участков.

ся примерно за 7300 п.н. от экзона 2. Таким образом, экзоны S и L обуславливают различия в исходных нуклеотидных последовательностях амилазы. Последовательность экзона L фактически является частью длинного интрона, удаляемого при сплайсинге в клетках слюнной железы.

Преобразования другого рода происходят в ходе созревания мРНК при процессинге в случае иммуноглобулиновых генов. Они включают замену одного экзона другим на 3'-конце единицы транскрипции. Такая замена приводит к изменениям С-концевой аминокислотной последовательности белка; в результате область, ответственная за связывание белка с мембраной, заменяется на область, ответственную за секрецию. Это обсуждается в гл. 39.

При экспрессии поздних генов аденовируса одна и та же лидерная последовательность в результате сплайсинга может быть соединена с одной из нескольких различных кодирующих последовательностей. Происходящие при этом события схематично представлены на рис. 20.20. Здесь имеется одна единица транскрипции, экспрессия которой инициируется в одной-единственной точке. В начальной части транскрипционной единицы присутствуют три последовательности, которые при сплайсинге соединяются вместе, образуя нетранслируемую лидерную последовательность. (Такая ситуация показана на рис. 20.6 и 20.7.) Компоненты лидерной последовательности имеют небольшую длину и обозначены как лидерные экзоны 1, 2 и 3. Правее этих последовательностей находится несколько кодирующих участков, каждый из которых соответствует позднему белку вируса. Эти участки обозначены как кодирующие экзоны А, В, С и т.д. Для каждого первичного продукта транскрипции лидерная последовательность, состоящая из трех частей и образовавшаяся в результате первых двух этапов сплайсинга, затем может быть присоединена к одному из трех кодирующих участков.

3'-конец каждого кодирующего экзона образуется путем расщепления первичного транскрипта и полиаденилирования концевой участка. Последовательности, расположенные правее, затем удаляются. Далее 5'-конец кодирующего экзона присоединяется к состоящей из трех

компонентов лидерной последовательности. В случае первого интрона это не связано с особыми трудностями, поскольку интрон удаляется обычным способом. Однако для кодирующего экзона В «интрон» захватывает всю область между его 5'-концом и лидерным экзонам 3. Эта область включает и кодирующий экзон А. Аналогичным образом, когда к лидеру присоединяется кодирующий экзон С, в состав удаляемого «интрона» входят и экзон А, и экзон В. Таким образом, ни один из экзонов не является абсолютно необходимым для синтеза белка в молекуле специфического транскрипта. В действительности в каждой молекуле РНК может использоваться только один из кодирующих экзонов, другие же отбрасываются, если они находятся правее этого экзона, или удаляются при сплайсинге, если находятся левее его. По-видимому, то, какие экзоны будут при сплайсинге присоединены к трехкомпонентной лидерной последовательности, определяется местоположением 3'-конца РНК.

Другая ситуация наблюдается в случае небольших молекул вирусов SV-40 и полиоми, когда в результате различия протекания сплайсинга одновременно продуцируются белки с перекрывающимися аминокислотными последовательностями. Это происходит при транскрипции и ранних, и поздних генов.

При транскрипции ранних генов SV-40 синтезируются два белка: Т (называемый «большим Т-антигеном») и t (называемый «малым t-антигеном»). Эти белки имеют идентичные N-концевые последовательности, но различающиеся С-концевые последовательности. Кодируются белки одним и тем же участком ДНК, как это показано на рис. 20.21.

Участок ДНК, кодирующий Т-антиген, включает два экзона. Первый экзон кодирует 5'-нетранслируемую область мРНК и N-концевой участок белка; второй экзон кодирует оставшуюся часть белка и 3'-нетранслируемую область мРНК. Длина интрона, находящегося между ними, составляет 347 п.н., и на границах экзон-интрон он имеет обычные универсальные последовательности.

мРНК t-антигена также содержит один интрон. Он представляет собой часть интрона, удаляемого при образовании Т-антигена, но содержит только его последние 67 оснований. Поэтому расположение правых границ сплай-

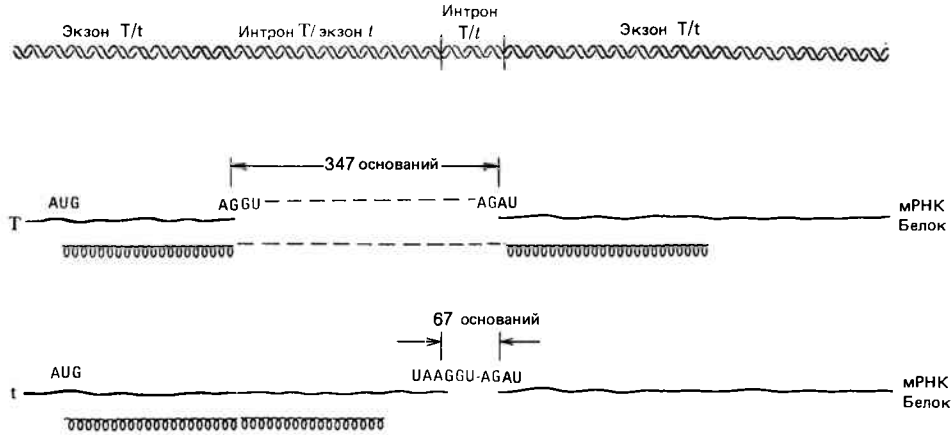


Рис. 20.21. При различных вариантах сплайсинга из одной и той же последовательности ДНК образуются Т- и t-антигены вируса SV-40.

синга идентично в двух случаях, а разница заключается в выборе левой границы. Местоположение левой границы сплайсинга при образовании t-мРНК также подчиняется универсальному правилу сплайсинга.

Трансляция при синтезе Т- и t-антигенов начинается с кодона AUG в общей для двух антигенов инициаторной последовательности. Аминокислотные последовательности антигенов различаются начиная с места присоединения первого экзона Т-последовательности ко второму при сплайсинге. В результате образуются два белка, имеющие одинаковые N-концевые, но разные С-концевые части молекулы. t-Антиген полностью кодируется первым экзоном t-мРНК. Терминирующий кодон находится непосредственно перед концом экзона. (Хотя второй экзон Т и присутствует в t-мРНК, он не транслируется, поскольку синтез белка заканчивается на t-UAA-кодоне; этот кодон не терминирует синтез Т-антигена, поскольку он удаляется как часть Т-интрона.) С учетом сказанного для этих генов не имеет большого смысла обозначать определенные участки ДНК как интроны и экзоны.

Интрон, который может кодировать регуляторный белок

Два прерывистых гена в митохондриях дрожжей обладают исключительными генетическими свойствами, которые находятся в таком противоречии с традиционным представлением о гене, что могут получить объяснение только при исследовании ДНК на молекулярном уровне. Структура этих генов приведено на рис. 20.22.

Локус *ox13* кодирует субъединицу 1 цитохромоксидазы. Кодирующая область гена имеет размер 1530 п.н., располагаясь на участке ДНК длиной 10000 п.н. В состав локуса входит восемь экзонов, два из которых имеют очень небольшую длину.

Ген, кодирующий цитохром *b*, известен под названием *box* или *cob*. Он существует в двух формах. Некоторые штаммы дрожжей имеют «длинный» ген, кодирующая область которого размером 1155 п.н. располагается на участке ДНК длиной около 6400 п.н. и имеет в своем составе шесть экзонов (обозначаемых как В1–В6), в том числе несколько очень коротких. Другие штаммы имеют «короткий» ген, по длине равный примерно половине

«длинного», у которого первые четыре экзона «длинного» гена объединены в один непрерывный экзон. Иными словами, три первых интрона гена либо все вместе присутствуют, либо все вместе отсутствуют. Обе формы гена одинаково хорошо экспрессируются. Для длинного гена был проведен генетический анализ, который мы обсудим. Интроны обозначают в соответствии с порядком их расположения в длинном гене как П1–П5.

Если выделить молекулы РНК, соответствующие гену *box*, то они по размеру распадаются на ряд дискретных классов. Наиболее короткие молекулы составляют в длину около 3000 оснований и соответствуют мРНК. Они длиннее кодирующей области и включают 5'- и 3'-нетранслируемые области, не показанные на карте, приведенной на рис. 20.22. Молекулы больших размеров содержат один или более интронов. Самые большие молекулы размером 8500 оснований, по-видимому, представляют собой интактный первичный продукт транскрипции. Другие молекулы — это промежуточные продукты сплайсинга, в которых одни интроны уже удалены, а другие — еще нет.

Мутации, приводящие к нарушениям синтеза цитохрома *b*, при картировании могут быть объединены в ряд кластеров. Каждый кластер занимает небольшой по длине участок, порядка 1% рекомбинационной карты. Расстояния между кластерами довольно велики и составляют 2–9% рекомбинационной карты. Каждый кластер обозначается как *box* с соответствующим номером. Кластеры распадаются на три класса.

Мутации в четырех кластерах непосредственно влияют на активность белка — цитохрома *b*. Все мутанты такого рода синтезируют нормальную мРНК. Мутации проявляются на уровне трансляции и выражаются в считывании матрицы с ошибками типа «миссенс» или «нонсенс». Ни одна из таких мутаций не комплементирует какую-либо другую в том же или другом кластере. По этому критерию все они находятся в одном и том же гене. Кластеры соответствуют некоторым экзонам, а именно: *box 4* = В1, *box 8* = В3, *box 1* = В4, *box 6* = В6 (рис. 20.22). В двух других экзонах мутации не обнаружены, возможно вследствие их малых размеров (В2 = 14 п.н., В5 = 51 п.н.). Рассмотренные группы мутаций проявляют в точности такие же свойства, какие можно ожидать от прерывистых генов. Фактически это един-

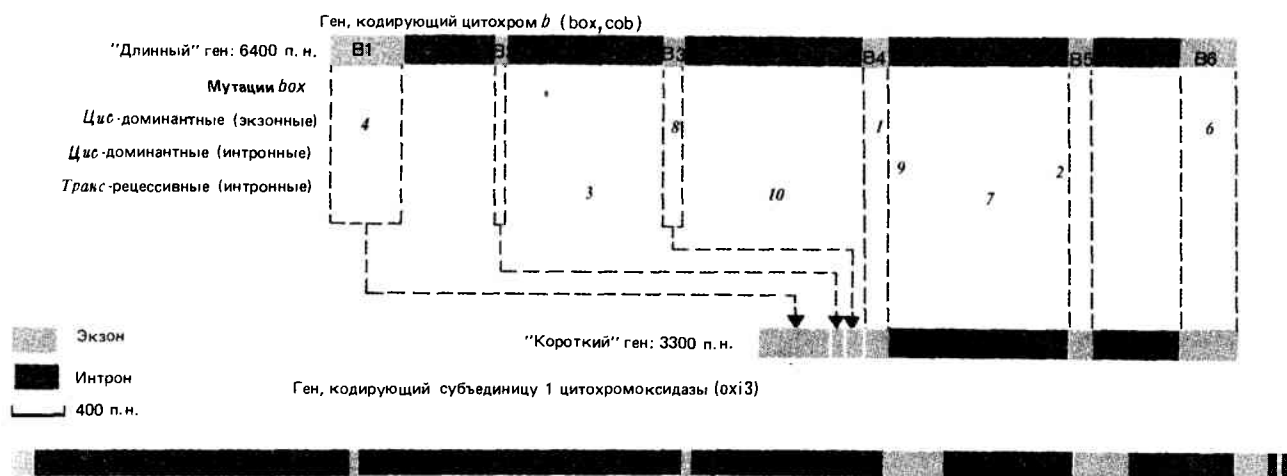


Рис. 20.22. Два гена митохондрий дрожжей имеют мозаичную структуру.

Ген *oxi3* имеет восемь экзонов, кодирующих субъединицу 1 цитохромоксидазы, состоящую, по-видимому, из 510 аминокислот. «Короткий» и «длинный» варианты гена кодируют один и тот же белок — цитохром *b*, состоящий из 385 аминокислот. Кластеры мутаций обозначены как *box*

с номером, который соответствует порядку, в котором они были обнаружены, а не их расположению на генетической карте. Положение кластеров на карте соответствует положению мутировавших экзонов или интронов. Оба гена изображены так, что они соответствуют кодирующим белок участкам мРНК и не содержат 5'- и 3'-нетранслируемых участков.

ственный известный на сегодня случай, когда можно изучать генетику мозаичного гена.

Мутации в кластерах *box 9* и *box 2* также не комплементируют мутации в других кластерах. Следовательно, по такому генетическому критерию они неотличимы от мутаций, затрагивающих экзоны. Однако их биохимические свойства различны, на что указывает нарушение синтеза соответствующей нормальной мРНК. При анализе нуклеотидной последовательности ДНК обнаруживается, что оба этих кластера находятся в области I4. Мутации кластера *box 9* затрагивают последовательность ДНК длиной 8 п. н., находящуюся на 350 п. н. правее границы с B4. Мутации кластера *box 2* смещены к другому концу интрона и находятся на расстоянии 25 п. н. левее границы с B5. Обе группы мутаций препятствуют объединению участков B4 и B5 в результате удаления области I4 при сплайсинге. Существование этих мутаций указывает на два важных обстоятельства общего характера. Во-первых, мутации, затрагивающие специфические сайты, могут препятствовать узнаванию определенных границ сплайсинга, причем такие сайты могут быть достаточно удалены от самих границ. Во-вторых, с помощью генетических методов анализа эти мутации нельзя отличить от мутаций, затрагивающих кодирующие белок участки. (Точно так же неразличимы классические *цис*-мутации, затрагивающие промоторы или операторы и их структурные гены; см. гл. 14.)

Каждая из трех других групп мутаций, *box 3*, *box 10* и *box 7*, представляет собой небольшой кластер мутаций, затрагивающих интрон. Их проявление при постановке комплементационных тестов свидетельствует о том, что это мутации нового типа. Мутации, входящие в каждый из перечисленных трех кластеров, образуют одну группу комплементации. Эти мутации не могут комплементировать другие мутации своего кластера, но могут комплементировать мутации любого другого кластера, как аналогичные, так и мутации в экзонах.

Формально это означает, что каждый из трех таких кластеров при мутациях по *транс*-типу затрагивает образование растворимого продукта, отличающегося от белковой части цитохрома *b*, но необходимого для его син-

теза. По существу, это означает, что в составе трех интронов имеются последовательности, функция которых заключается в кодировании некоторого независимого продукта, осуществляющего определенную регуляторную функцию при образовании цитохрома *b*.

На природу этой регуляторной функции указывает другое свойство таких мутаций. Все они блокируют образование мРНК цитохрома *b*, вызывая накопление РНК-предшественников. Размеры таких предшественников составляют: 7500 оснований для мутантов по *box 3*, 7100 оснований для мутантов по *box 10*, 3500 оснований для мутантов по *box 7*. Это указывает на то, что каждый кластер мутаций может блокировать определенную стадию созревания РНК путем инактивации растворимого продукта, вероятно необходимого для удаления определенного интрона. (При этом блокирование не может быть вызвано просто мутациями на границе экзон-интрон, поскольку такие мутации относились бы к *цис*-типу, подобно *box 2* и *box 9*, а не к *транс*-типу.)

Предположения о природе этой функции основываются на данных по секвенированию (определению последовательности) гена. Характерные особенности организации гена представлены на рис. 20.23. В первом экзоне (B1) присутствует 139 N-концевых кодонов (417 п. н.) цитохрома *b*. Интрон I1 (765 п. н.), заблокированный во всех рамках считывания, отделяет первый экзон от второго, очень короткого экзона B2 (5 кодонов). За этим экзоном расположен длинный второй интрон (I2). Важная особенность этого интрона состоит в том, что его первые 840 п. н. попадают в открытую рамку считывания, в точности совпадающую с рамкой считывания предыдущего экзона.

Мутации *box 3* находятся в этой области, и при определении нуклеотидной последовательности некоторых мутировавших участков выяснилось, что эти мутации приводят к возникновению бессмысленных кодонов в открытой рамке считывания. Таким образом, этот участок интрона может осуществлять функцию кодирования и соответствующий белок может осуществлять свою функцию при сплайсинге. Этот белок был назван РНК-матуразой. (Остальная часть интрона заблокирована во всех рамках считывания.)

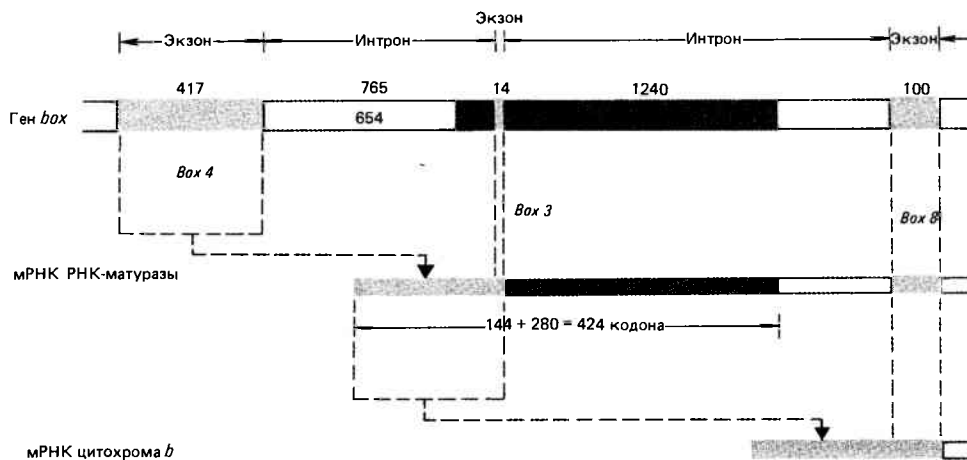


Рис. 20.23. При осуществлении последовательных этапов сплайсинга транскрипта гена *box* могут образовываться мРНК, кодирующие белки с перекрывающимися аминокислотными последовательностями. При удалении только первого интрона образуется мРНК РНК-матуразы; при удалении второго интрона образуется последовательность, кодирующая цитохром *b*.

Какова же структура РНК-матуразы? Имеется косвенное указание на то, что ее трансляция инициируется не внутри интрона, а при считывании информации со второго экзона. Это может происходить при трансляции показанного на рис. 20.23 первого промежуточного продукта сплайсинга. РНК образуется просто путем удаления первого интрона. Это приводит к объединению первого и второго экзонов и образованию рамки считывания, захватывающей также часть второго интрона. Полная длина рамки считывания — 424 кодона, и при ее трансляции будет образовываться белок, состоящий из 144 N-концевых аминокислот цитохрома *b* и 279 аминокислот, кодируемых интроном.

Если РНК-матураза, кодируемая такой последовательностью, предназначена специально для удаления второго интрона при сплайсинге, ее ферментативная активность обеспечивает наличие исключительно чувствительной отрицательной обратной связи, изображенной на рис. 20.24. Удаление второго интрона и присоединение первых двух экзонов к третьему приводит к разрушению последовательности, кодирующей матуразу. Таким образом, функционирование РНК-матуразы при сплайсинге приводит к прекращению ее синтеза. Поэтому устанавливается равновесие между содержанием двух форм РНК и РНК-матуразы (которая присутствует в клетке в чрезвычайно небольших количествах и которую поэтому очень трудно охарактеризовать).

Является ли эта схема общей моделью организации интронов в геноме митохондрий? Четкие аргументы в пользу этого отсутствуют, но для двух генов, *box* и *oxi3*, установлено, что основной способ организации интронов состоит в создании открытой рамки считывания, в точности совпадающей с рамкой считывания предыдущего экзона. Это не может быть совпадением. По-видимому, такие рамки считывания транслируются.

В отличие от ядерных структурных генов дрожжей границы экзон-интрон в митохондриях не подчиняются правилу GT-AG; на границах отсутствует также и кака-либо другая универсальная последовательность. Поэтому РНК-матураза, по-видимому, проявляет специфичность по отношению к определенному интрону или интронам. Возможная функция РНК-матуразы *box3* состоит в узнавании концов только второго интрона, поэтому окру-

жающие его экзоны могут объединяться при сплайсинге. В другом случае имеется указание на то, что кодируемый интроном продукт может участвовать в двух событиях сплайсинга, поскольку у мутантов *box7* (по четвертому интрону гена *box*) нарушен сплайсинг не только *box*-мРНК, но и *oxi3*-мРНК. Мутации *box7* приводят к появлению нонсенс-кодонов в открытой рамке считывания интрона.

Какую функцию может выполнять матураза? Мысль о том, что каждая рамка считывания интронов кодирует матуразу, обладающую специфичностью по отношению к данному интрону (или, по крайней мере по отношению к очень небольшому числу интронов), позволяет предположить, что роль матуразы, вероятнее всего, состоит в отдельных событиях узнавания, а не в ее каталитической активности *per se*. Например, матураза, возможно, обеспечивает создание специфической конформации комплекса, в котором какой-то другой компонент осуществляет сам процесс разрыва связей или создания новых связей.

Было обнаружено существование интересного сходства между митохондриальными интронами дрожжей и некоторыми генами рРНК. В их состав входит несколько довольно коротких сходных последовательностей. Они расположены на некотором расстоянии от границ экзон-интрон (на самих границах таких консервативных последовательностей нет). Некоторые последовательности причастны к сплайсингу, по крайней мере, в случае гена *box*, поскольку они обеспечивают создание сайтов, соответствующих *cis*-мутациям *box9* и *box2*, блокирующим сплайсинг.

По меньшей мере для нескольких генов, содержащих консервативные последовательности, можно представить себе вторичную структуру, необходимую для участия в образовании спаренных участков. Происходящее при этом спаривание оснований приводит к созданию вторичной структуры, при которой концы интрона располагаются один вблизи от другого. В соответствии с одной из существующих моделей такое сближение затем дает возможность основаниям другой последовательности («внутренней направляющей последовательности») спариться с основаниями концов экзонов, удерживая их рядом, пока происходит ферментативное разрезание и сшивание.

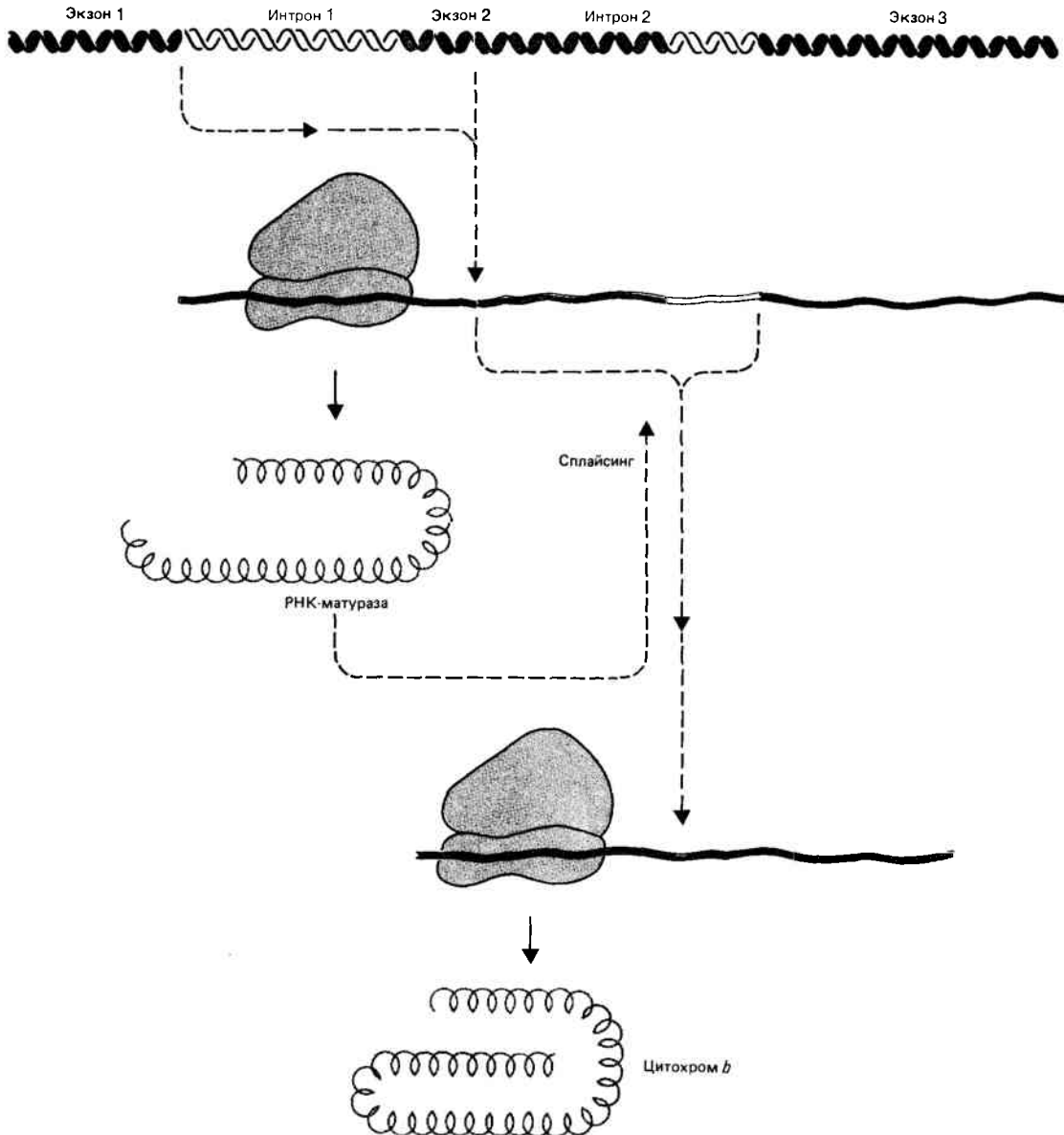


Рис. 20.24. Если при сплайсинге РНК-матураза избирательно удаляет интрон, кодирующий ее синтез, то содержание белка и двух мРНК будет регулироваться по принципу отрицательной обратной связи.

Затененные участки ДНК и РНК соответствуют открытым рамкам считывания; незатененные области заблокированы во всех рамках считывания.

По-видимому, все митохондриальные интроны дрожжей имеют такие канонические последовательности. Так, их обнаруживают в интронах, имеющих открытые кодирующие участки, и в других интронах, в которых все возможные рамки считывания заблокированы. К ядерным генам, содержащим канонические последовательности, относится ген *Tetrahymena*, кодирующий предшественник рРНК, подвергающийся автономному сплайсингу. Эти данные служат замечательной демонстрацией существования неожиданного эволюционного сходства ядерного и внеядерных геномов и позволяют сделать предположение о том, что, возможно, сплайсинг разных РНК протекает в соответствии с общим механизмом. Это заключение наводит на мысль об интересном парадоксе, касающемся механизмов сплайсинга.

С одной стороны, каноническая последовательность обнаруживается в составе интронов, кодирующих белок, необходимый для протекания сплайсинга; с другой сто-

роны, они имеются и в гене, РНК-транскрипт которого способен к автономному сплайсингу. Одна из моделей, доказывая универсальность характера сплайсинга, утверждает, что РНК-копиям генов обоих типов присуще внутреннее свойство образовывать определенную вторичную (третичную) структуру и катализировать разрыв и восстановление связей. Однако для стабилизации структуры необходимы белки. Несомненно справедливо, что нуклеиновые кислоты, имеющие такую же длину, как и рассматриваемые интроны, могут образовывать различные вторичные структуры; вероятно, требуется наличие некоторой внешней силы для того, чтобы всегда образовывалась структура, необходимая для осуществления сплайсинга.

В случае митохондриальных интронов дрожжей белки не только специфичны по отношению к индивидуальным интронам, но и кодируются интронами, на которые затем и воздействуют. (Эта модель не применима ко всем ми-

тохондриальным интронам дрожжей, поскольку некоторые из них не кодируют белков; кроме того, все необходимые вспомогательные белки должны кодироваться другими последовательностями.) Проверкой этой модели послужило бы выделение матуразы, соответствующей открытой рамке считывания интрона, и проверка ее активности *in vitro*, при которой обнаружилось бы, с каким участком интрона связывается матураза, обладает ли она каталитической активностью, и если обладает, то какой.

Существует ли какая-либо очевидная эволюционная взаимосвязь между разными функциями, кодируемыми интронами? Открытые рамки считывания интронов содержат ряд сходных последовательностей, но до сих пор не ясно, имеют ли они какое-либо значение. Во всех кодирующих областях интронов частота использования определенных кодонов несколько отличается от частоты использования кодонов в экзонах. Поскольку не ясно, какие преимущества митохондрии извлекают из сохранения такой сложной системы экспрессии своих прерывистых генов, можно предположить, что когда-то существовал отдельный белок, участвующий в сплайсинге второго интрона гена *box*. Позже соответствующий ему ген мог быть транслоцирован в интрон, что привело к возникновению существующей в настоящее время структуры. Тем не менее остается справедливым утверждение, что, по-видимому, единственная цель кодируемой интроном функции состоит в удалении кодирующей последовательности из мРНК.

Сложные локусы имеют очень большие размеры и участвуют в регуляции

Поведение нескольких локусов *D. melanogaster* не согласуется с предположением о том, что каждый из них содержит один-единственный ген, кодирующий уникальный продукт. Такие локусы называются **сложными локусами**.

Обычный ген, в том числе и прерывистый ген, на уровне построения генетической карты обнаруживается как кластер прочно связанных некомплементирующих мутаций. Признак сложного локуса — наличие расположенных на значительном расстоянии друг от друга мутаций, занимающих довольно большой участок на генетической карте и часто характеризующихся сложным способом комплементации. Например, мутации могут распадаться на несколько перекрывающихся групп комплементации. Отдельные мутации способны вызывать различные сложные морфологические изменения фенотипа.

По крайней мере некоторые сложные локусы, по-видимому, участвуют в регуляции формирования тела взрослого насекомого. Оно делится на ряд **сегментов**. Сегменты состоят из передней и задней частей, каждая из которых развивается из относительно небольшого числа эмбриональных клеток.

Мутанты, затрагивающие формообразовательные процессы, подразделяются на два основных класса. **Сегментационные мутанты** характеризуются изменением числа или полярности сегментов их тела. **Гомеозисные мутации** способны вызывать превращение части сегмента или всего сегмента в сегмент другого типа. Такие мутации затрагивают локусы, представляющие особый интерес в связи с тем, что они могут участвовать в определении типа строения частей тела насекомого.

Классический сложный локус *bithorax* характеризуется наличием нескольких групп гомеозисных мутаций, затрагивающих последовательные стадии развития грудных сегментов, что вызывает существенные морфологические изменения брюшка. При делеции всего локуса *bithorax* насекомое погибает на поздних стадиях эмбрионального развития. Однако среди мутаций в пределах этого сложного локуса могут быть и не летальные, но ведущие к изменению фенотипа определенных сегментов.

Генетическая карта локуса *bithorax* включает несколько типов мутаций, свойства которых суммированы в табл. 20.1. Локус делится на две части. Мишенями мутаций в домене *Ultrabithorax* служат грудные сегменты, тогда как мишенями мутаций домена *infraabdominal* — сегменты брюшка.

Наибольшее количество данных имеется о домене *Ultrabithorax*, в котором обнаружено несколько различных групп комплементации. Мутации в каждой группе образуют набор мутантных аллелей со сходным фенотипом разной степени выраженности, но характеризующимся одними и теми же типами трансформации тканей. Каждый набор соответствует гену, ответственному за определенный участок тела. *Внутри локуса bithorax, таким образом, имеется несколько индивидуальных локусов, продукты которых выполняют сходные функции в процессе развития мухи.*

Отдельным типам мутаций даны названия, образованные от *bithorax*. Первая обнаруженная мутация, называемая просто *bithorax* (так что одним и тем же словом обозначают как один из типов мутаций в сложном локусе, так и весь локус в целом), имеет критическое значение, поскольку ведет к превращению третьего грудного сегмента (несущего жужжальца, т.е. редуцированные крылья) в ткань, соответствующую второму сегменту (несущему крылья). При мутациях этого типа у мухи развивается четыре крыла вместо обычных двух.

Собственно мутации *bithorax* и мутации, сходные с ними, затрагивают грудные сегменты или приводят к превращению первого брюшного сегмента в грудной. Такие мутации связаны с доменом *Ultrabithorax*, занимающим левую часть генетической карты. Рядом с ними на карте располагается домен *infraabdominal*, с которым связаны

Таблица 20.1

Каждый тип мутаций в пределах локуса *bithorax* имеет характерное фенотипическое проявление

Мутация	Сегмент, на который она влияет	Фенотипическое проявление
<i>abx (anterobithorax)</i>	T3	Сходное (но не идентичное) с <i>bx</i>
<i>bx (bithorax)</i>	T3	Превращает передний T3 в передний T2
<i>cbx (Contrabithorax)</i>	T2	Превращает T2 в T3
<i>ubx (Ultrabithorax)</i>	T3	Сумма <i>Vx</i> , <i>Vxd</i> и <i>pVx</i>
<i>bxid (bithoraxoid)</i>	A1	Превращает передний A1 в T3
<i>pbx (postbithorax)</i>	T3	Превращает задний T3 в T2
<i>Hab (Hyperabdominal)</i>	Брюшной	Детально не описан
<i>Mcp (Miscadestral pigmentation)</i>	Брюшной	Детально не описан
<i>iab2 (infraabdominal)</i>	A2	Превращает A2 в сегмент, напоминающий A1

другие мутации, воздействующие непосредственно на брюшные сегменты.

Методом «прогулки по хромосоме» была построена генетическая карта участка размером более 100 т.п.н., занимающего левую половину сложного локуса.

Мутации типа *Vx* и *Vxd* захватывают большие, но явно дискретные участки хромосомы. Каждый участок может участвовать в кодировании белка, необходимого для развития определенного сегмента (или части сегмента).

Мутации *pbx* и *Cbx* комплементируют друг друга по своему фенотипическому эффекту и, как оказалось, связаны между собой на молекулярном уровне. При делеции *pbx*¹ теряется участок размером 17 т.п.н., который, встраиваясь в какое-то другое место (в противоположной ориентации) приводит к возникновению мутаций *Cbx*¹. Если этот участок несет информацию, определяющую специфическое развитие соответствующего участка тела (третий, т.е. задний, грудной сегмент), то его утрата у *pbx*-мутантов может привести к тому, что этот сегмент станет походить на предыдущий сегмент (второй грудной сегмент). В то же время встраивание рассматриваемого участка в сайт, расположенный на расстоянии 44 т.п.н. левее, скажется на развитии второго грудного сегмента, приводя к образованию на его месте структуры, характерной для третьего грудного сегмента.

Мутации типа *Ubx* проявляются так, как будто имеют место сразу три мутации, *bx*, *bxd* и *pbx*. Для большей части аллелей *Ubx* обнаруживаются цитологически выявляемые перестройки хромосомы, которые на молекулярном уровне соответствуют делециям, приводящим к удалению генетического материала разной длины в левой половине кластера. Некоторым другим мутациям *Ubx* соответствуют делеции небольших участков в пределах данного кластера. Нам до сих пор не ясна природа *Ubx*-мутаций, но, по-видимому, они указывают на то, что для проявления генотипа *Ubx*⁺ необходимо наличие всей левой половины локуса.

Расположение мутаций на генетической карте, по всей вероятности, коррелирует с их воздействием на фенотип. Перемещению мутаций вдоль генного локуса слева направо, как правило, соответствуют изменения тела мухи в направлении сверху вниз. Каждый сегмент мухи отличается от предшествующего синтезом дополнительного белка, кодируемого следующим на генетической карте геном. К сожалению, имеющаяся у нас информация относится в основном лишь к трем «грудным» локусам, тогда как большинство функций, связанных с брюшными участками, изучены плохо.

Сложность локуса *bithorax* в целом можно объяснить, если предположить наличие взаимосвязи между процессами экспрессии относящихся к нему индивидуальных генов или функциями их продуктов. Хотелось бы знать, происходит ли транскрипция этих генов независимо, или она каким-то образом скоординирована? Существует ли какая-нибудь регуляция функционирования локуса в целом? Какова география распределения продуктов локуса в эмбрионе?

Notch (*N*)—другой сложный локус, мутации в котором приводят к ряду явно не похожих между собой фенотипических проявлений. *N*-мутанты обладают доминантным фенотипом—укороченными крыльями—в случае гетерозигот *N/+*; рецессивный генотип летален, поскольку приводит к гибели гомозигот *N/N*. *N*-мутации не комплементируют друг друга.

В пределах локуса имеется несколько групп рецес-

сивных мутаций, по-разному влияющих на фенотип. Некоторые из них нарушают формирование крыльев, другие приводят к изменениям глаз и щетинок. Все такие мутации по признаку их неспособности комплементировать *N*-мутации относятся к типу *Notch*, однако в пределах самой этой группы имеется по меньшей мере четыре группы комплементации.

Локус *Notch* содержит единицу транскрипции, занимающую участок ДНК размером 37 т.п.н. и включающую 9 экзонов, длина которых варьирует от 130 п.н. до 7250 п.н. Длина соответствующей РНК составляет около 11,7 т.п.н. Большинство *N*-мутаций локализуется в экзонах; некоторые рецессивные мутации представляют собой вставки в интроны. Когда мы сможем охарактеризовать мутации, выяснив их действие на последовательность ДНК данного локуса, станет понятно, чем обусловлена сложная природа мутаций *Notch*.

Сложный локус *Antennapedia* включает по меньшей мере три гомеозисных гена: *proboscipedia* (*pb*) (вместо хоботков развиваются ноги), *Sex combs reduced* (*Sxr*) (редуцированные половые щетинок) и *Antennapedia* (*Antp*) (вместо антенн развиваются ноги). У некоторых таких мутантов имеются аллели, приводящие к тому, что вместо антенн развивается вторая пара ног или вместо второй и третьей пары—первая пара ног. И опять-таки этот локус на молекулярном уровне оказывается весьма протяженным; как показано с помощью мутационного картирования этого участка, мутации охватывают область размером > 100 т.п.н.

Данный локус обладает свойствами прерывистого гена, поскольку довольно далекие друг от друга участки представлены в непрерывных молекулах РНК существенно меньшего размера (< 5,0 т.п.н.). Такие транскрипты все еще изучены недостаточно полно.

Различия между сложными локусами и обычными генами *D. melanogaster* весьма существенны. Исследованные гены плодовой мушки в большинстве своем имеют довольно небольшие размеры; они либо не содержат интронов, либо входящие в их состав интроны невелики. Напротив, сложные локусы имеют огромные размеры, и при сплайсинге их первичного транскрипта образуются молекулы РНК, длина которых на порядок меньше длины локуса. Почему локусы, ответственные за регуляцию процесса развития взрослого насекомого из личинки (эмбриональная стадия), отличаются от генов, кодирующих обычные белки?

Одно из возможных объяснений состоит в том, что разные способы экспрессии генов сложного локуса приводят к образованию неодинаковых продуктов. Различия между этими продуктами могли бы быть обусловлены перестройками ДНК (что в настоящее время кажется маловероятным, поскольку каких-либо изменений в геноме обнаружено не было) или наличием разных способов протекания сплайсинга. Такие модели способны дать объяснение перекрывающимся, но близким способам комплементации, при которых мутации иногда затрагивают независимые гены, расположенные в пределах локуса, а иногда—гены, входящие в состав одного сложного локуса.

Экспрессию таких локусов исследовать сложно. Крайне трудно выделить молекулы РНК длиной ~ 100 т.п.н. и, в любом случае, если сплайсинг начинается до завершения транскрипции, интактного первичного продукта транскрипции существовать не может. Молекулы РНК, в разной степени подвергнувшиеся сплайсингу, могут

присутствовать в клетке в довольно низких концентрациях, что затрудняет выявление индивидуальных РНК.

Возможность существования связи между функциями гомеозисных генов была обнаружена в результате выявления гомологии между последовательностями гомеозисных локусов *Ultrabithorax*, *Antennapedia* и влияющим на развитие сегментов локусом *fushi tarazu (ftz)*, мутации в котором приводят к уменьшению числа сегментов наполовину. Каждый из таких генов экспрессируется в пределах ограниченного участка тела насекомого в процессе эмбрионального развития. Например, *ftz* необходим для развития каждого второго сегмента из соответствующего зачатка.

Во всех таких локусах имеется участок размером 180 п.н., расположенный около 3'-конца транскрипционной единицы. При попарном сравнении соответствующих последовательностей каждого локуса степень гомологии их нуклеотидных последовательностей составляет ~80%. Такой «гомео-бокс» представляет собой часть открытой рамки считывания, в которой изменения оснований во многих случаях происходят в третьем положении кодонов.

Еще более удивительно, что ДНК-зонд гомео-бокса гибридизуется с ДНК лягушки, мыши и человека. Был клонирован ген *X. laevis*, включающий гомео-бокс, который кодирует 55 из 60 аминокислот, кодируемых гомео-боксом локуса *Antp*. Данный ген экспрессируется на ранних стадиях развития, и на нем синтезируются три транскрипта. Гены мыши и человека также содержат высококонсервативные гомео-боксы.

Какие сходные функции могут кодировать гены, участвующие в регуляции развития плодовых мушек, лягушек и млекопитающих? Возможно, продукты генов служат универсальными «инструкциями», давая приказы клетке прекратить деление, делиться быстрее или остановиться на определенной стадии развития.

Считают, что гомеозисные гены регулируют переключение функций в процессе развития, действуя в пределах определенных участков тела особи, возможно, кодируя белки, взаимодействующие с определенной батареей генов, продукты которых придают клеткам каждого такого участка их неповторимую индивидуальность. Существуют ли сходные механизмы регуляции у позвоночных и беспозвоночных, несмотря на весьма существенные различия в их организации? Удалось ли нам обнаружить универсальный механизм регуляции эмбрионального развития?

Как появились прерывистые гены?

Какова была исходная организация генов, имеющих в настоящее время прерывистое строение? Представляли ли собой исходные, кодирующие белок единицы непрерывные последовательности ДНК, в которые впоследствии, в процессе их эволюционирования с образованием современной структуры, были встроены интроны? Или эти гены изначально возникли как прерывистые структуры, которые так и сохранили свое строение? В другой форме этот вопрос сводится к следующему: обусловлены ли различия между эукариотическими и прокариотическими генами приобретением интронов эукариотами или потерей интронов прокариотами?

Может ли мозаичное строение гена быть отголоском

происходивших ранее перестроек генов, обеспечивших создание новых белков? Предположим, что вначале в клетке присутствовал ряд отдельных последовательностей, кодирующих белки. Один из аспектов эволюции клетки, по-видимому, состоял в изменении организации и взаимного расположения разных полипептидных единиц с целью создания новых белков. Если кодирующая белок единица представляет собой непрерывный ряд кодонов, каждая такая перестройка потребует точной рекомбинации ДНК, чтобы две единицы, кодирующие белок, располагались рядом, конец в конец, в пределах одной и той же рамки считывания. Кроме того, неудачные комбинации приводили бы к гибели клетки вследствие утраты исходных, кодирующих белок единиц. Но если два участка ДНК, кодирующие белок, в результате неточной рекомбинации оказались бы внутри одной транскрипционной единицы, то в результате нескольких вариантов сплайсинга могло бы стать возможным объединение двух белков в единую полипептидную цепь. И если такие комбинации белков окажутся неудачными, исходные белок-кодирующие единицы сохраняются для дальнейших поисков. По существу, это позволяет клетке осуществлять регуляцию отбора делеций на уровне РНК, не подвергаясь угрозе разрушающей нестабильности, как в том случае, если бы такие преобразования происходили с ДНК.

Если существующие в настоящее время белки появились в результате создания комбинаций исходно различающихся белков, то это, по-видимому, происходило постепенно, в течение определенного периода времени, путем добавления экзонов по одному к генам, кодирующим уже существующие белки. В случае соответствия этой модели действительности возникает вопрос: можно ли в существующих в настоящее время генах выявить структуры, ранее кодировавшие различные функции, из которых сформировались современные гены? Иными словами, можно ли считать, что специфические функции существующих белков кодируются индивидуальными экзонами?

В некоторых случаях действительно имеется четкая связь между структурой гена и белка. Убедительный пример такого рода – иммуноглобулины, кодируемые генами, каждый экзон которых точно соответствует определенному функциональному домену белка (гл. 39).

Имеются и другие примеры, когда экзоны соответствуют определенным белковым функциям. Часто первый экзон, кодирующий N-концевую часть полипептида, соответствует области, содержащей сигнальную последовательность, участвующую в мембранной секреции. Пример такого рода – инсулин.

Иногда процесс эволюционирования гена может включать и дубликацию экзонов, в результате которой в составе белка оказываются повторяющиеся последовательности. Например, в гене коллагена цыпленка экзон размером 54 п.н., по-видимому, повторяется несколько раз. (В то же время в гене коллагена *D. melanogaster* имеется только два больших экзона.)

В некоторых случаях можно обнаружить экзоны, кодирующие участки, включающие место соединения двух доменов белка. Примеры такого рода обнаружены для коллагена. В случае гена алкоголь-дегидрогеназы (АДГ) *D. melanogaster* имеется любопытная двойственность, когда одни экзоны точно соответствуют доменам белка, в то время как другие кодируют одну из двух частей домена.

Известно несколько интересных случаев, когда гомологичные гены имеют разную структуру. Это может дать

информацию об их эволюции. Примером такого рода служит инсулин.

Млекопитающие и птицы имеют только один инсулиновый ген, кроме грызунов, у которых два гена. На рис. 20.25 приведено строение таких генов.

Единственный инсулиновый ген курицы содержит два интрона; один из двух генов крысы имеет аналогичную структуру. Наличие такого сходства структуры указывает на то, что инсулиновый ген-предок содержал два интрона. Однако в состав второго инсулинового гена крысы входит только один интрон. Этот ген, по-видимому, произошел от первого в результате процесса, при котором за дубликацией гена у грызунов последовало точное удаление интрона из одной копии.

Прекрасный пример эволюционной стабильности демонстрируют глобины, все гены которых содержат три экзона. Два интрона занимают постоянные положения относительно кодирующей последовательности. Центральный экзон, по-видимому, соответствует связывающему гем домену полипептидной цепи глобина. Активный белок представляет собой тетрамер, содержащий две глобиновые цепи α -типа и две цепи β -типа.

Возможен иной взгляд на эту структуру, связанный с существованием двух других типов белков, родственных глобину. **Миоглобин** – связывающий кислород белок-мономер животных, аминокислотная последовательность которого позволяет предположить наличие общего (хотя и древнего) происхождения с субъединицами глобина. **Леггемоглобины** – это связывающие кислород белки бобовых растений, так же как и миоглобин, являющиеся мономерами. Они также могут иметь общее происхождение с другими белками, связывающими гем.

Гены глобинов миоглобина и леггемоглобинов вместе образуют глобиновые суперсемейство – набор генов, происшедших от общего предка.

Гены леггемоглобина содержат три интрона, причем локализация первого и последнего по отношению к кодирующей последовательности гомологична локализации двух интронов глобиновых генов. Такое удивительное сходство указывает на то, что гены, кодирующие гем-содержащие белки, с весьма древних времен имеют прерывистое строение, показанное на рис. 20.26.

Центральный интрон гена леггемоглобина разделяет два экзона, которые вместе кодируют последовательность, соответствующую одному центральному экзону глобинового гена. Мог ли центральный экзон глобинового гена образоваться путем слияния двух центральных экзонов гена-предка, т.е. объединения последовательностей, кодирующих две части белковой цепи, вместе образующие структуру, связывающую гем? В свете этого

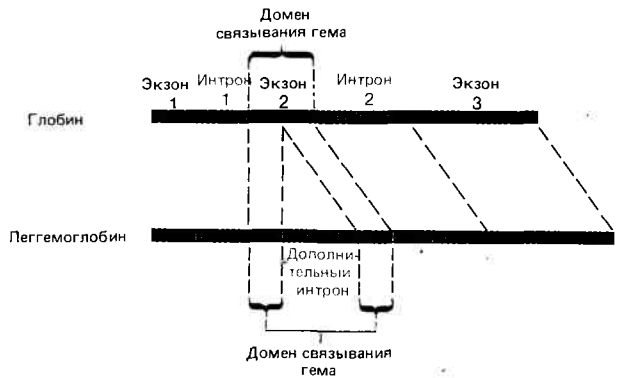


Рис. 20.26. Экзонная структура генов глобина соответствует функции белков, но в центральном домене леггемоглобина имеется дополнительный интрон.

определение структуры гена миоглобина, возможно, даст интересный информацию.

Соответствие по крайней мере некоторых экзонов белковым доменам подтверждает предположение о том, что оно имеет фундаментальное значение в эволюции генов. Ясно, что дубликации и слияние экзонов могли играть важную роль в эволюции. Мы не можем проследить за действительными событиями, происшедшими в ходе эволюции каждого гена. Имеется несколько примеров взаимоотношений между экзонами и белковыми доменами, когда отсутствует их простое соответствие, но это можно объяснить тем, что такие события, как слияние экзонов, изменили структуру гена-предка в процессе эволюции ядерных генов. Однако в ряде случаев мы сталкиваемся с большими несоответствиями между структурами генов и белков. Митохондриальные гены дрожжей и млекопитающих кодируют практически идентичные митохондриальные белки, несмотря на существенные различия в организации генов. Геном митохондрий позвоночных очень мал и имеет чрезвычайно компактную организацию нерасщепленных генов (гл. 22), тогда как митохондриальный геном дрожжей имеет большие размеры и включает ряд сложных прерывистых генов. Какая форма гена была исходной?

Противоречит ли наличие интронов в геноме митохондрий дрожжей предположению о том, что митохондрии появились в результате возникшего в древности симбиоза, когда прототип бактерии был встроен в цитоплазму эукариотической клетки? Нет, если предположить, что такое встраивание предшествовало потере интронов прокариотами, хотя при этом мы должны допустить независимую потерю способности к удалению интронов при эволюции бактерий и митохондрий позвоночных.

По-видимому, генам рРНК и тРНК свойствен полиморфизм, так что часто можно обнаружить разные формы генов, как содержащие, так и не содержащие интроны. В случае тРНК, когда все молекулы имеют одну сходную в общих чертах структуру, кажется маловероятным, что в результате эволюционного процесса произошло объединение двух участков гена. Ведь для спаривания оснований используются разные области молекулы, что имеет важное значение для создания определенного типа ее структуры, так что, возможно, в этом случае интроны были встроены в нерасщепленные гены.

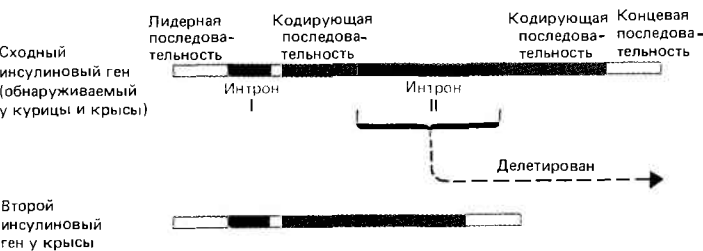


Рис. 20.25. Инсулиновый ген крысы, содержащий один интрон, произошел от гена-предка с двумя интронами в результате потери одного из этих интронов.

Рекомендуемая литература

Строение прерывистых генов подробно рассматривается в книге Б. Льюина (*Levin B., Gene expression, 2, Eucaryotic chromosomes, Wiley, New York, 1980, pp. 790–847*). Некоторые, более новые представления приводятся в журнале *Cell* (**22**, pp. 645–646, 1980). Первый содержательный обзор, посвященный структуре генов, со времени обнаружения прерывистых генов написан Брет-

нахом и Шамбоном (*Breathnach, Chambon, Ann. Rev. Biochem., v. 50, pp. 349–383, 1981*). Биологии bithorax посвящен обзор Лоренса (*Lawrence, Cell, 35, pp. 595–601, 1983*), а первым попыткам исследовать его молекулярные свойства—статья Бендера (*Bender, Science, 221, pp. 23–29, 1983*). Гомео-бокс был открыт Мак-Гиннисом и др. (*McGinnis et al., Cell, 37, pp. 403–408, 1984*), а также Карраско и др. (*Carrasco et al., Cell, 37, pp. 409–414, 1984*).

Часть VI

КЛАСТЕРЫ СХОДНЫХ ПОСЛЕДОВАТЕЛЬНОСТЕЙ ДНК

В генетической программе, таким образом, записан результат воспроизведения всех прошлых поколений. Это своего рода коллекция успехов, поскольку все следы неудач исчезли. Генетическая информация, программа жизнедеятельности современного организма, похожа на текст без автора, который корректор исправлял на протяжении более чем двух миллиардов лет, постоянно улучшая, очищая и дополняя его, постепенно устраняя все изъяны.

Франсуа Жакоб, 1973

СТРУКТУРНЫЕ ГЕНЫ: ОРГАНИЗАЦИЯ РОДСТВЕННЫХ ГЕНОВ

В настоящее время появляется все больше оснований считать, что лишь немногие гены присутствуют в эукариотическом геноме в виде единственной копии. Несмотря на то что большинство генов при гибридизации в растворе ведет себя как последовательности неповторяющейся ДНК, при более детальном их исследовании с помощью выделения соответствующих фрагментов генома обычно обнаруживаются другие последовательности, сходные с экзонами этих генов. Иногда такие последовательности представляют собой функционирующие гены, в большей или меньшей степени дивергировавшие в процессе эволюции; иногда это нефункционирующие **псевдогены**, реликты эволюции, происшедшие от генов, когда-то, по-видимому, функционировавших, но теперь ставших неактивными, накопившими значительное число мутационных изменений.

Рассматривая структуру прерывистых эукариотических генов, иногда занимающих весьма протяженные участки ДНК, можно представить себе эукариотический геном как море интронов (нуклеотидные последовательности которых в большинстве случаев, но не всегда, уникальны), в котором растянутые вереницей островки экзонов (иногда очень короткие) образуют отдельные архипелаги, представляющие собой гены. Сопоставляя соответствующие друг другу экзоны в родственных генах, можно обнаружить их сходство, что подчеркивает важную роль дубликаций в эволюции генов как механизма образования новых генов. Одна из копий может эволюционировать в результате мутаций, тогда как другая сохраняет свою первоначальную функцию, так что, по-видимому, история гена складывается из ряда событий, начинающихся с объединения экзонов, составляющих ген, с образованием кодирующего участка; вполне возможно, что позже вся группа экзонов и интронов, формирующих ген, дублировалась. За дубликацией могла последовать относительно небольшая дивергенция нуклеотидных последовательностей экзонов и более значительная дивергенция последовательностей интронов. Таким образом, нуклеотидные последовательности родственных генов позволяют нам восстановить историю их эволюции.

Набор генов, происшедших путем дубликаций и изменений от некоторого гена-предка, называется **семейством генов**. Члены одного семейства могут быть расположены рядом или разбросаны по разным хромосомам (или их расположение может представлять собой сочетание этих двух вариантов). **Кластеры генов** имеют различную организацию — начиная от ситуации, когда в результате дубликации возникают два расположенных рядом сходных гена, до случаев, когда сотни идентичных генов расположены в виде тандема. Ситуации, при которых родственные гены разбросаны по разным участкам генома, по-видимому, возникли в результате транслокации, происшедшей уже после дубликации.

Члены одного семейства структурных генов имеют сходные или идентичные функции, хотя они могут экспрессироваться в разное время или в клетках разных типов. В случаях когда имеется высокая тандемная повто-

ряемость гена или генов, продукты их экспрессии требуются в необычайно больших количествах. К примерам такого рода относятся гены рРНК или гистонов (гл. 23). Иногда для образования слегка различающихся белков, необходимых при определенных обстоятельствах, требуется меньшая степень повторяемости. Например, в эмбриональных и взрослых клетках красной крови синтезируются разные глобины, а в мышечных и нервных клетках функционируют различные актины. Иногда смысл повторяемости неясен; например, нет видимых различий в экспрессии или функционировании дублированных генов инсулина у грызунов (и мы знаем, что у других млекопитающих вполне достаточно одного инсулинового гена).

Множество типов глобинов

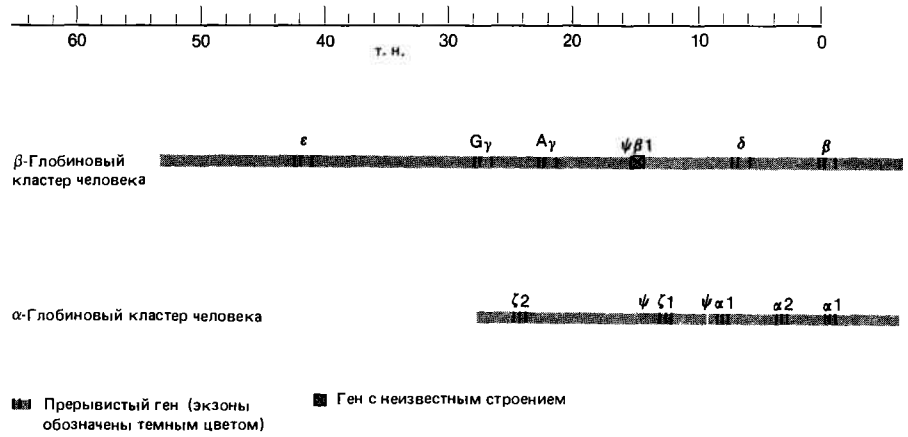
Транспорт кислорода по кровяному руслу — основная жизненная функция в царстве животных — кодируется древним семейством генов. Основной компонент эритроцита клетки — белок-тетрамер **глобин**, который при связывании с (железосодержащим) гемом образует **гемоглобин**. Для ряда организмов были исследованы гены, кодирующие полипептидные цепи глобина. Все функционально активные гены глобина имеют сходную в общих чертах структуру и включают три экзона, как было показано ранее на рис. 20.16. Мы пришли к выводу о том, что все гены глобина произошли от общего гена-предка (гл. 20). Поэтому, проследив за развитием индивидуальных генов глобина в пределах одного вида и у представителей разных видов, можно многое узнать об общих механизмах эволюции систем генов.

В клетках взрослого животного глобиновый тетрамер состоит из двух идентичных α -цепей и двух идентичных β -цепей. Гены α - и β -глобинов кодируются независимыми генетическими локусами, экспрессия которых должна быть скоординирована таким образом, чтобы обеспечить продуцирование одинаковых количеств каждого из двух полипептидов. Таким образом, эта система представляет собой пример того, когда для поддержания определенного фенотипа клетки необходимо осуществление одновременной регуляции работы генов, расположенных в разных участках генома.

Эмбриональные эритроциты содержат тетрамеры гемоглобина, отличающиеся от гемоглобина клеток взрослого организма. Такие эмбриональные тетрамеры состоят из двух идентичных α -подобных цепей и двух идентичных β -подобных цепей, каждая из которых сходна с полипептидной цепью взрослого организма и позднее заменяется на нее. Это пример генетического контроля развития, когда для обеспечения клетки разными продуктами, выполняющими одну и ту же функцию в разные его периоды, последовательно включаются и выключаются разные гены.

В деталях взаимоотношения между эмбриональными и взрослыми гемоглобинами различаются у разных организмов. У человека путь к образованию взрослого гемоглобина включает три стадии (которые до сих пор до

Рис. 21.1. Семейства α - и β -подобных глобиновых генов человека организованы в виде кластеров, включающих в себя функционально активные гены и псевдогены (ψ).



конца не изучены). На разных стадиях развития обнаруживаются следующие типы гемоглобинов:

Эмбриональные гемоглобины (до 8 нед развития)	$\zeta_2\epsilon_2$ и $\zeta_2\gamma_2$ и $\alpha_2\epsilon_2$
Гемоглобин плода	$\alpha_2\gamma_2$
Гемоглобины взрослого	$\alpha_2\delta_2$ и $\alpha_2\beta_2$

ζ - и α -цепи—две α -подобные цепи. ϵ -, γ -, δ - и β -цепи— β -подобные цепи. Применительно к этапам развития организма человека можно предположить наличие самого простого пути образования взрослого гемоглобина: первой синтезируется α -подобная ζ -цепь, которая вскоре заменяется на саму α -цепь. Что касается пути преобразования β -подобных цепей, то первыми синтезируются ϵ - и γ -цепи, которые позднее заменяются на δ - и β -цепи. У взрослого человека 97% всего гемоглобина составляет гемоглобин $\alpha_2\beta_2$; на долю гемоглобина $\alpha_2\delta_2$ приходится около 2%, и примерно 1% представлен гемоглобином плода $\alpha_2\gamma_2$.

Гены глобинов организованы в виде кластеров

Деление глобиновых цепей на α - и β -подобные обусловлено структурой глобина и организацией соответствующих генов. Все семейство β -подобных полипептидных цепей характеризуется способностью непосредственно взаимодействовать с α -цепями. ζ -Цепь относят к α -подобным на основе ее способности взаимодействовать с эмбриональными β -подобными цепями. Оказалось, что каждый тип глобина кодируется генами, организованными в виде единого кластера. У человека α -кластер располагается в 11-й хромосоме, а β -кластер—в 16-й.

Каждый кластер был обнаружен по способности соответствующих фрагментов генома гибридизоваться с глобиновой мРНК взрослого организма. Использование этого метода для анализа других участков ДНК позволило обнаружить новые гены; некоторые из них кодировали известные или неизвестные эмбриональные глобины, другие представляли собой нефункционирующие псевдогены.

Строение двух кластеров глобиновых генов в геноме человека приведено на рис. 21.1. Гены глобинов у других высших приматов по своей организации очень похожи на рассмотренные. Например, кластеры β -глобиновых генов гориллы и павиана практически не отличаются от соответствующих кластеров человека. Все гены транскрибируются слева направо и располагаются в порядке их экс-

прессии в процессе развития организма: 5'-конец—гены эмбриональных гемоглобинов—гены гемоглобина плода (в β -кластере)—гены гемоглобина взрослого—3'-конец.

Занимая участок ДНК размером более 50 т.п.н., β -кластер включает в себя пять функционально активных генов—ген ϵ , два гена γ , ген δ , ген β и один псевдоген. О существовании функционально активных генов было известно ранее. Продукты, кодируемые двумя генами γ -глобина, различаются только на одну аминокислоту: в глобине G в положении 136 находится глицин, а в глобине А—аланин.

Более компактно организованный α -кластер занимает участок ДНК размером более 28 т.п.н. и включает в себя один активный ген ζ , один псевдоген ζ , два гена α и один псевдоген α . О существовании двух генов α , кодирующих идентичные белки, ранее предполагали на основе генетического анализа. Два (или более) идентичных гена, прилегающие в одной хромосоме, называются **неаллельными** копиями.

Похожая в целом организация кластеров глобиновых генов обнаружена и у других позвоночных, но детали их строения и число генов могут различаться. Примеры организации некоторых β -глобиновых кластеров приведены на рис. 21.2.

У кролика были обнаружены четыре β -подобных гена; два гена эмбрионального гемоглобина, один псевдоген и один ген гемоглобина взрослого. И снова функционально активные гены располагаются в порядке их экспрессии, от 5'-конца к 3'-концу.

У козы имеется семь β -подобных генов глобина (не показанных на рисунке): два гена эмбриональных гемоглобинов, один ген гемоглобина плода, два гена гемоглобинов взрослых и два псевдогена.

У мыши, по-видимому, имеется по меньшей мере девять β -подобных генов глобина. К настоящему времени обнаружены семь генов. К ним относятся два β -гена гемоглобина взрослых, причем уровень экспрессии гена β^{major} выше, чем уровень экспрессии гена β^{minor} . Нуклеотидные последовательности генов $\beta h1$ и $\beta h0$ обладают большим сходством с последовательностями известных генов взрослых и эмбриональных гемоглобинов, так что, возможно, они эквивалентны β -подобным глобиновым генам плода человека (хотя их белковые продукты до сих пор не обнаружены). Структура гена $\epsilon\gamma 3$ хорошо соответствует известным эмбриональным белкам; похоже, что он может кодировать неизвестную β -подобную эмбриональную полипептидную цепь. Для двух эмбриональных β -подобных глобинов, $\epsilon\gamma 2$ и ϵz , гены пока не обнаружены.

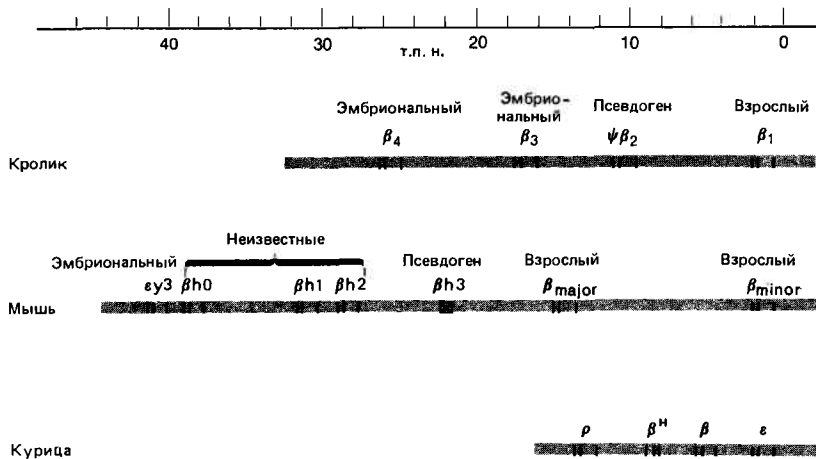


Рис. 21.2. Кластеры β -глобиновых генов и псевдогенов обнаруживаются и у других позвоночных.

Помимо пяти обнаруженных и двух необнаруженных функционально активных генов имеются два других гена, $\beta h3$ и $\beta h2$, нуклеотидные последовательности которых, по-видимому, не соответствуют известным глобинам мыши и которые, возможно, представляют собой псевдогены.

Кластеры генов глобина могут иметь меньшие размеры и не содержать псевдогенов. Кластер β -глобиновых генов цыпленка имеет в длину менее 14 т.п.н. и, по-видимому, включает только четыре функционально активных β -подобных гена. Два эмбриональных гена располагаются снаружи, а два взрослых — внутри, так что между порядком экспрессии гена и его расположением корреляции нет.

В результате анализа таких кластеров генов был сделан важный вывод общего характера. В семейство генов может входить большее число членов (как функционально активных, так и неактивных), чем можно предположить на основе анализа соответствующих белков. Дополнительные функционирующие гены могут либо представлять собой дубликации, кодирующие идентичные белки, либо быть сходными с генами для известных белков, хотя и отличаясь от них. Возможно, что эти гены экспрессируются в течение короткого времени или уровень их экспрессии низок. Для псевдогенов характерна различная степень родства с функционально активными генами, и их можно изучать только на уровне нуклеотидной последовательности ДНК. Принимая это во внимание, нельзя считать, что обнаружены все члены кластера, пока не будут тщательно исследованы фланкирующие области, прилегающие к крайним членам кластера.

Все ли гены глобина входят в состав кластеров? В каждом геноме все активные гены глобина относятся к α - и β -кластерам, иногда вместе с псевдогенами. В геноме мыши было обнаружено несколько дополнительных псевдогенов, не входящих в состав кластеров. α -Кластер состоит из одного эмбрионального и двух взрослых генов. Два таких псевдогена располагаются в разных хромосомах. Могут обнаруживаться также и другие псевдогены. Их возникновение, по-видимому, было обусловлено транслокациями, в результате которых произошло удаление гена от кластера.

Рассматривая вопрос о том, какое количество ДНК требуется для кодирования определенного белка, можно отметить, что у различных млекопитающих и птиц область ДНК, кодирующая β -подобные глобины, имеет размеры от 14 до 50 п.н. Не стоит и говорить о том, что

размер этой области много больше, чем можно было бы ожидать, исходя из размеров известных полипептидных цепей β -глобинов. По-видимому, никакие другие белки не кодируются последовательностями, входящими в состав β -кластера, но до сих пор не известно, какая часть некодирующей ДНК (включая и фланкирующие последовательности, и интроны) выполняет функции, необходимые для экспрессии генов β -глобина. И до тех пор пока не будут изучены многие другие гены или кластеры генов, кодирующие специфические белки, мы не сможем сказать, является ли сложное строение кластеров глобиновых генов исключительным или типичным.

Неравный кроссинговер приводит к перестройке кластеров генов

Существует много возможностей для осуществления перестроек внутри кластера родственных или идентичных генов. Результат таких перестроек можно увидеть, сравнивая кластеры генов β -глобинов млекопитающих, показанные на рис. 21.1 и 21.2. Несмотря на то что эти кластеры выполняют одну и ту же роль и имеют сходную в общих чертах организацию, различаясь по размеру, имеются вариации общего числа и типов β -глобиновых генов; различаются также и число, и структура псевдогенов. Все эти изменения произошли, по-видимому, после разделения млекопитающих на виды около 85 млн. лет назад (это был конец периода общего эволюционного развития всех млекопитающих). На основе сравнения различных кластеров можно сделать общее заключение о том, что дубликации, перестройки и изменения генов могут быть столь же важными факторами эволюции, как и медленное накопление точковых мутаций в индивидуальных генах с последующим отбором белков с улучшенными свойствами. Каковы же механизмы генных перестроек?

Размер кластера генов может возрасти или уменьшиться в результате **неравного кроссинговера**, когда происходит рекомбинация между неаллельными генами, как показано на рис. 21.3. Обычно рекомбинация происходит (как описано в гл. 1) между соответствующими друг другу последовательностями ДНК, расположенными точно одна напротив другой в двух гомологичных хромосомах. Однако, когда в каждой хромосоме имеются две копии гена, напротив друг друга могут случайно оказаться последовательности разных копий, что сделает возможным неправильное спаривание между ними. (Для этого необ-

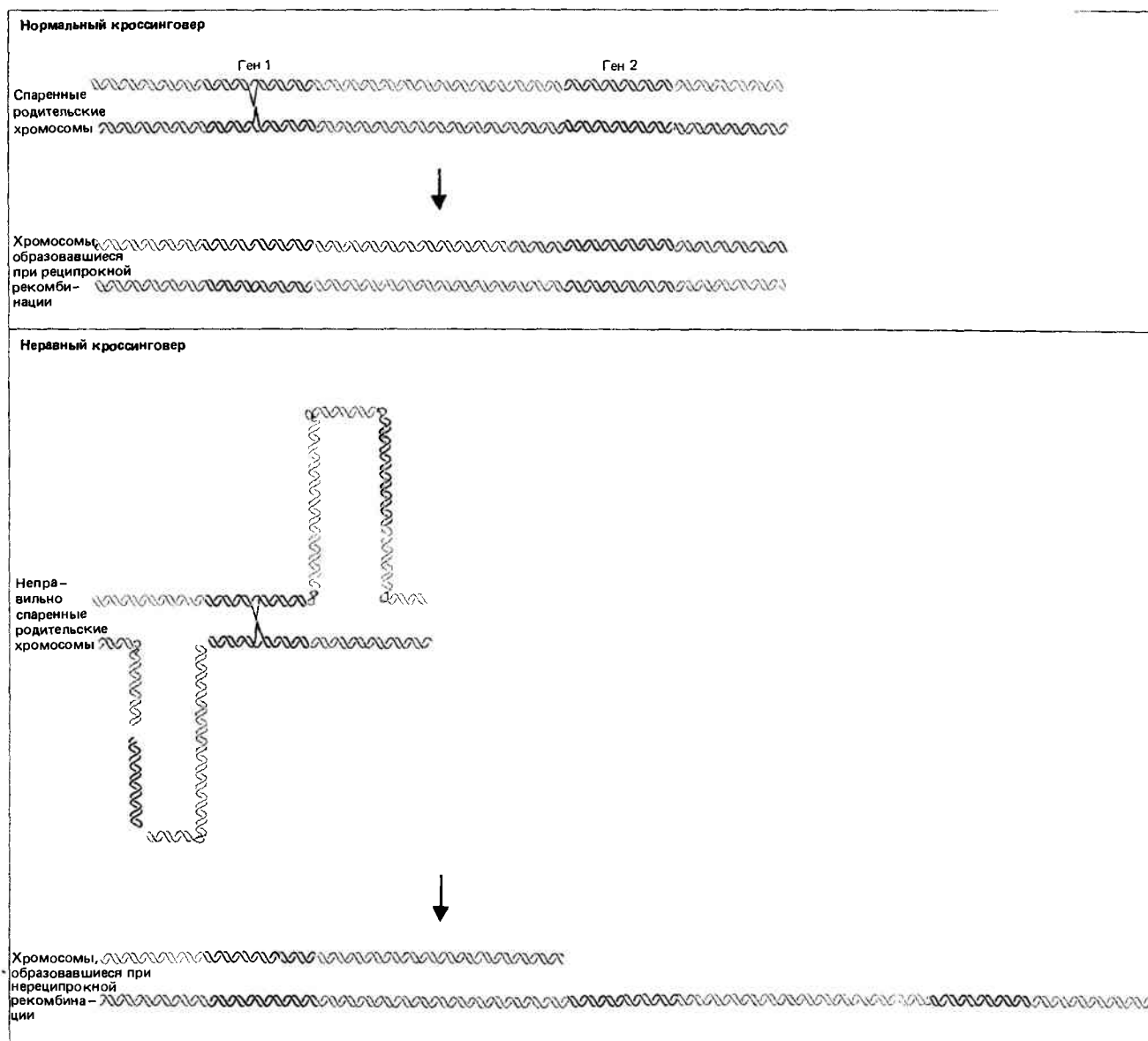


Рис. 21.3. Число генов может изменяться в результате неравного кроссинговера.

Если ген 1 одной хромосомы спаривается с геном 2 другой хромосомы, то другие копии гена не участвуют в спаривании; они показаны в виде

выступающих петель. При рекомбинации между неправильно спаренными генами образуется одна хромосома с одной (рекомбинантной) копией гена и одна хромосома с тремя копиями гена (по одной от каждого родителя и одна рекомбинантная).

ходимо, чтобы некоторые из соседних участков оставались неспаренными.)

Когда происходит рекомбинация между неправильно спаренными копиями генов, образуются **нереципрочные рекомбинантные хромосомы**, в одной из которых имеется дупликация данного гена, а в другой — его делеция. Дуплицируется весь материал, находящийся между точками неравной рекомбинации; делеция затрагивает участок между этими же точками.

Таким образом, в результате рекомбинации в первом случае число копий возрастает с двух до трех, а во втором уменьшается с двух до одной.

В приведенном примере мы рассматривали копии различающихся генов 1 и 2, как если бы они были полностью гомологичны. Однако неравный кроссинговер возможен и тогда, когда соседние гены неодинаковы, но имеют большое сходство (хотя в этом случае ве-

роятность возникновения ошибок спаривания, по-видимому, меньше, чем когда гены идентичны). Более серьезное препятствие для осуществления неравного кроссинговера представляет наличие у генов прерывистой структуры. В случае генов, таких, как гены глобинов, соответствующие друг другу экзоны соседних копий генов, по-видимому, имеют достаточно большое сходство, чтобы спариваться друг с другом; однако последовательности интронов могут оказаться сильно дивергировавшими. Ограничение способности экзонов к спариванию значительно уменьшает длину непрерывного участка ДНК, который может участвовать в кроссинговере и соответственно снижает вероятность неравного кроссинговера. Можно предположить, что наличие различий между интронами повышает стабильность кластера генов, препятствуя осуществлению неравного кроссинговера.

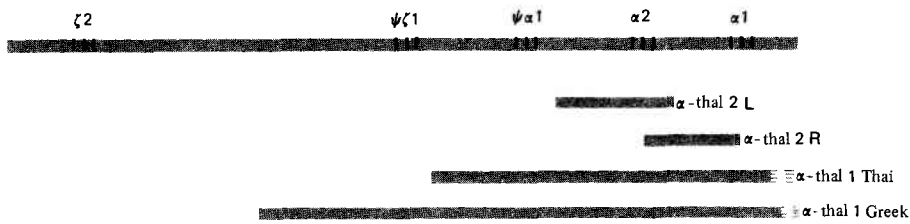


Рис. 21.4. Талассемии возникают в результате различных делеций в кластере генов α -глобина.

Многие формы α -талассемии – результат неравного кроссинговера

Осуществление неравного кроссинговера в кластере глобиновых генов человека подтверждается природой некоторых форм талассемии. Талассемия может возникнуть в результате любой мутации, нарушающей синтез либо α -, либо β -глобина. В зависимости от степени наблюдаемого нарушения талассемию подразделяют на α^0 - и β^0 -формы (когда отсутствует сколько-нибудь заметный синтез цепей одного из этих типов) или α^+ - и β^+ -формы (когда снижен уровень синтеза цепей). Талассемия была обнаружена у больных, страдающих анемиями (наиболее распространенными в определенных популяциях людей, например среди жителей Средиземноморья). Многие α^0 - и β^0 -формы талассемии возникают в результате делеций части соответствующего кластера генов глобина. По крайней мере в некоторых случаях концы участков, подвергающихся делециям, находятся в гомологичных областях, что в точности соответствует предположению об их возможном возникновении в результате неравного кроссинговера.

В геноме больных α -талассемией выделяют два основных типа делеций. В случае делеции α -thal-1 отсутствуют оба глобиновых гена, при делеции α -thal-2 – только один из α -глобиновых генов. Цифры обозначают хромосому, несущую делецию. В зависимости от диплоидной комбинации хромосом с мутациями у индивидуума при талассемии может обнаруживаться любое число α -цепей, от 0 до 4. Состав крови человека с двумя или тремя α -генами мало отличается от состава крови дикого типа (с четырьмя α -генами). При наличии только одного α -гена в результате избытка β -цепей образуется необычный тетрамер β_4 , обуславливающий такую болезнь, как **гемоглобинопатия Н**. Полное отсутствие α -генов приводит к возникновению **водянки плода**, который в результате погибает до или во время родов.

На рис. 21.4 показаны делеции, вызывающие возникновение α -талассемий. Все делеции α -thal-1 затрагивают длинные участки ДНК и различаются по локализации своих левых концов; положение их правых концов пока еще не установлено. Все делеции α -thal-2 затрагивают короткие участки ДНК. В случае L-формы при делеции удаляется фрагмент ДНК длиной 4,2 т. п. н., включающий в себя ген $\alpha 2$. По-видимому, эта делеция возникает в результате неравного кроссинговера, поскольку ее концы находятся в гомологичных участках хромосомы, расположенных соответственно правее генов $\psi\alpha 1$ и $\alpha 2$. Р-делеция возникает в результате удаления фрагмента ДНК размером 3,7 т. п. н., точно соответствующего расстоянию между $\alpha 1$ и $\alpha 2$. По-видимому, эта делеция возникла в результате неравного кроссинговера между самими $\alpha 1$ - и $\alpha 2$ -генами. Эта ситуация в точности соответствует изображенной на рис. 21.3.

В результате неравного кроссинговера, приводящего к образованию талассемической хромосомы, образуется также и хромосома с тремя α -генами. В ряде популяций были выявлены индивидуумы с такими хромосомами. В некоторых популяциях частота встречаемости локуса с тремя генами α -глобина приблизительно совпадает с частотой локуса с одним α -геном; в других локус с тремя α -генами распространен в значительно *меньшей* степени, чем локус с одним α -геном. Исходя из этого, можно предположить, что для поддержания определенного уровня генов в различных популяциях действуют неизвестные факторы отбора.

Во всех случаях колебания в числе α -генов обнаруживаются довольно часто. Это подтверждает тот факт, что неравный кроссинговер в α -кластере может быть весьма распространенным явлением. Создается впечатление, что в α -кластере неравный кроссинговер происходит чаще, чем в β -кластере. Одна из возможных причин этого состоит в том, что интроны α -генов гораздо короче и поэтому меньше препятствуют неправильному спариванию неомологичных генов.

Новые гены, образующиеся при β -формах талассемии

β -Формы талассемии интересны во многих отношениях, в особенности тем, что многие из них могут быть обусловлены дефектами регуляции. Это дает редкую возможность исследовать экспрессию семейства генов.

Известно много типов β^0 -талассемии. При некоторых формах β^0 -талассемии делеций не обнаруживается. Следовательно, полное отсутствие β -цепей, может быть, по-видимому, обусловлено мутациями, влияющими на транскрипцию или трансляцию. В одном случае мутация приводит к образованию нонсенс-кодона в положении 17, так что дефект проявляется на уровне трансляции. В другом случае нарушение затрагивает более раннюю стадию экспрессии гена; ядерная РНК синтезируется, но мРНК в цитоплазме отсутствует. Причина этого – мутация, нарушающая сплайсинг мРНК (гл. 26).

Делеции, приводящие к возникновению β^0 -форм талассемии, суммированы на рис. 21.5. В некоторых (редких) случаях затрагивается только ген β -глобина. Тогда делеция охватывает участок размером 600 п. н., начиная со второго интрона и кончая 3'-фланкирующими участками. В других случаях делеция затрагивает несколько генов.

Гемоглобин тина Lepore – классический пример делеции, возникающей в результате неравного кроссинговера между генами одного кластера. Нуклеотидные последовательности генов β - и δ -глобина различаются только на 7%. Неравная рекомбинация приводит к делеции части ДНК, находящейся между генами; в результате гены соединяются (см. рис. 21.3). Объединенный ген обуславливает синтез одной β -подобной цепи глобина, состоящей

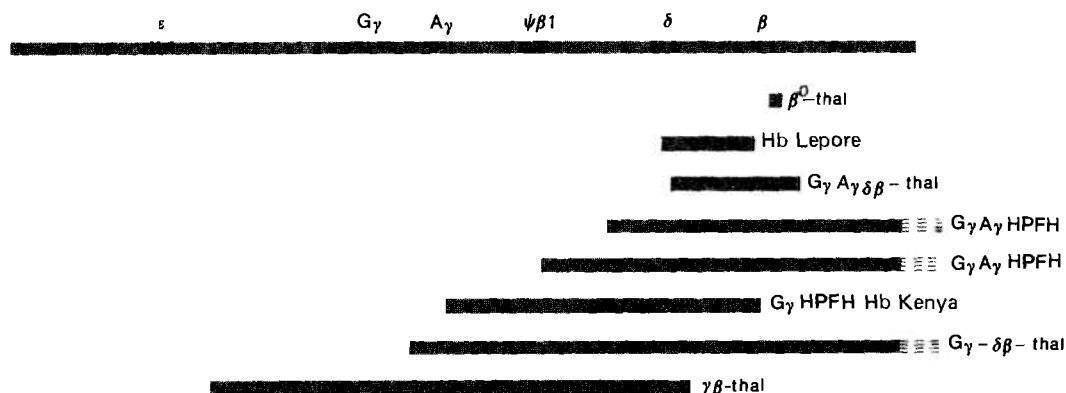


Рис. 21.5. Делеции в кластере генов β-глобина приводят к возникновению нескольких типов талассемии.

из N-концевой последовательности δ-цепи и C-концевой последовательности β-цепи.

В настоящее время известно несколько типов гемоглобинов *Lepore*, различающихся тем, в какой точке последовательность δ-цепи переходит в последовательность β-цепи. Поэтому при спаривании δ- и β-генов в случае неравного кроссинговера точка рекомбинации точно определяет место в аминокислотной цепи, где произойдет смена δ-последовательности на β-последовательность.

Явление, обратное описанному, было обнаружено в случае *гемоглобина анти-Lepore*; его ген обуславливает синтез белка с N-концевым участком β-цепи и C-концевым участком δ-цепи. Такой «слившийся» ген должен располагаться между нормальными δ- и β-генами.

В пользу предположения, что неравный кроссинговер может происходить между генами, имеющими меньшее сходство, свидетельствует обнаружение *гемоглобина Kenya* – гемоглобина другого типа, образующегося при слиянии генов. Этот гемоглобин содержит N-концевую последовательность $\Lambda\gamma$ -гена и C-концевую последовательность β-гена. Образование «слившегося» гемоглобина, по-видимому, произошло в результате неравного кроссинговера между $\Lambda\gamma$ - и β-генами, нуклеотидные последовательности которых различаются на 20%.

Отсутствие и δ-, и β-цепей проявляется в двух фенотипических формах. В случае НПФГ¹ клинические симптомы отсутствуют; болезнь возникает, поскольку синтез гемоглобина плода ($\alpha_2\gamma_2$) продолжается и после стадии развития организма, когда он в норме прекращается. В случае $\delta\beta$ -талассемии имеются симптомы анемии. Объясняется это следующими причинами: несмотря на то что экспрессия γ-гена может продолжаться и во взрослом состоянии, она происходит с меньшей эффективностью, чем в случае НПФГ.

Обнаружено много делеций, приводящих к нарушению синтеза δ- и β-цепей глобина. Эти делеции и соответствующие фенотипические проявления приведены на рис. 21.5. Одно из предположений, касающихся причин различий между талассемиями НПФГ и $\delta\beta$, состоит в том, что делеции, вызывающие гемоглобинопатию НПФГ, могут приводить к удалению некоторого регуляторного сайта, расположенного между $\Lambda\gamma$ - и δ-генами, в норме отвечающего за прекращение экспрессии γ-генов.

Аналогичное выключение может также существовать и в α-кластере. Объясняется это следующими причинами. Если в норме экспрессия ζ-гена у взрослого организма выключена, то при делеции двух α-генов выключения ζ-гена не происходит.

Кластеры генов подвержены постоянным перестройкам

На основе различий, существующих между кластерами генов глобина у разных млекопитающих, можно сделать вывод о том, что дупликации, за которыми следуют (иногда) изменения генов, играли важную роль в эволюции каждого кластера. Наличие делеций при различных формах талассемии человека свидетельствует о том, что неравный кроссинговер продолжает осуществляться в обоих кластерах глобиновых генов. В результате каждого такого события возникают как дупликация, так и делеция. Попытаемся проследить за судьбой двух рекомбинантных локусов в популяции. Делеции в принципе могут возникать в результате рекомбинации между гомологичными последовательностями, расположенными в *одной и той же хромосоме*. Это не влечет за собой соответствующей дупликации.

Определить действительную частоту таких событий трудно, поскольку в результате отбора уровни содержания кластеров в популяции быстро выравниваются. Возможно, существует грубая корреляция между вероятностью осуществления неравного кроссинговера и степенью сходства генов: чем больше сходство (включая и экзоны, и интроны), тем больше вероятность неправильного спаривания. (Однако иногда неравная рекомбинация происходит не между самими генами, а между расположенными рядом повторяющимися последовательностями.)

Обычно уменьшение числа генов, по-видимому, оказывается вредным и отбрасывается в процессе отбора. Однако в некоторых популяциях формы с делециями могут иметь некоторое преимущество; в результате частота встречаемости этих форм будет поддерживаться на некотором невысоком уровне.

Что происходит при увеличении числа копий генов? Единственные охарактеризованные примеры такого рода – локусы с тремя генами α-глобина и Hb анти-Lepore. У индивидуумов с пятью генами α-глобина (одним нормальным локусом и одним – с тремя α-генами) не про-

¹ НПФГ – наследственная персистенция фетального гемоглобина. – Прим. ред.

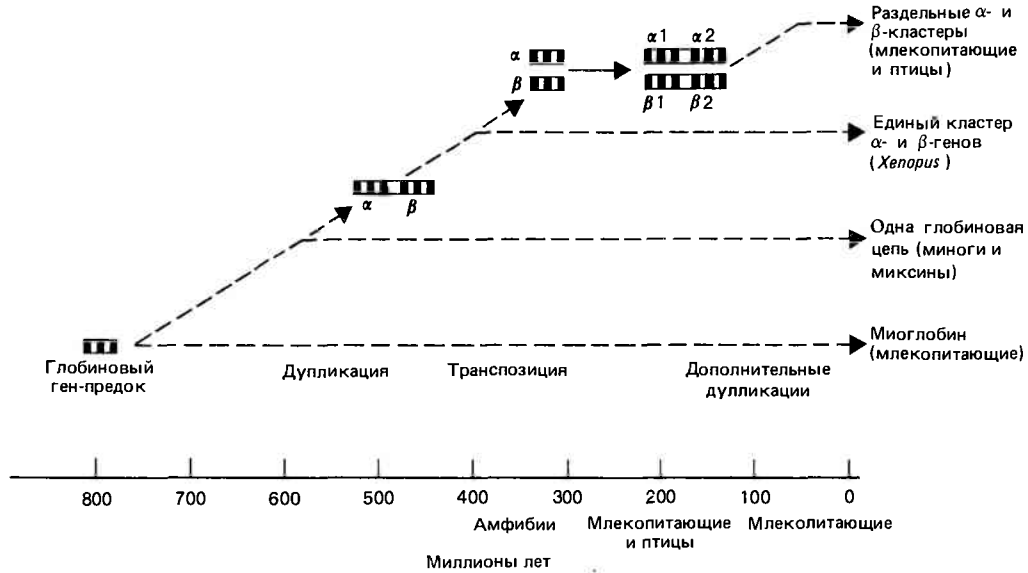


Рис. 21.6. Все глобиновые гены произошли от единого гена-предка в результате ряда дупликаций, транспозиций и мутаций.

является каких-либо изменений в синтез гемоглобина. Однако возможно, что дальнейшее увеличение числа генов (до гомозиготы по локусу с тремя генами) может оказаться вредным, поскольку будет приводить к нарушению баланса синтеза глобина в результате образования избыточных α -цепей. Индивидуумы с гемоглобином анти-Лероте помимо нормальных генов β - и δ -глобинов имеют слившийся ген $\beta\delta$. Возможно, что дополнительная полипептидная цепь оказывается вредной, поскольку она препятствует сборке нормального гемоглобина.

Таким образом, описанные изменения в числе генов, видимо, не дают таких преимуществ при отборе, которые способствовали бы их распространению в популяции. Однако в существующих в настоящее время кластерах глобиновых генов у человека обнаруживается несколько дупликаций, что свидетельствует о важности этих механизмов. К функционально активным генам относятся два α -гена, кодирующие один и тот же белок, довольно близкие по структуре β - и δ -гены и два практически идентичных γ -гена. Эти сравнительно недавно возникшие независимые дупликации сохранились в популяции, не говоря уже о более давних дупликациях, которые привели к образованию различных типов генов. Другие дупликации, возможно, привели к возникновению псевдогенов или были утрачены.

Известны два поразительных примера, когда попытки размножить кластеры генов путем клонирования их в клетках *E. coli* привели к повышению частоты возникновения делеций. Один из четырех генов β -глобинового кластера цыпленка было трудно выделить, поскольку при клонировании он, как правило, утрачивался, по-видимому, в результате рекомбинаций, в которых участвовали гомологичные последовательности по обе стороны от гена. В случае α -кластера человека два типа делеций, наблюдаемые при клонировании в клетках *E. coli*, точно соответствуют двум типам делеции α -thal-2, приведенным на рис. 21.4. Тот факт, что в бактериальной клетке происходят перестройки генома, приводящие к таким же результатам, которые наблюдаются и в природной человеческой популяции, указывает на то, что такие события —

неизбежное следствие механизмов, участвующих в рекомбинации ДНК. Таким образом, можно предполагать, что постоянно происходящие дупликации и делеции — характерное свойство всех генных кластеров.

Эволюционное древо глобиновых генов

Зная строение глобиновых генов у представителей различных видов, можно, наконец, проследить за эволюцией существующих в настоящее время глобиновых генов от одного гена-предка. Современное представление об эволюционном происхождении генов глобина проиллюстрировано на рис. 21.6.

Поскольку и у птиц, и у млекопитающих имеются раздельные кластеры генов α - и β -глобинов, α - и β -гены, должно быть, были физически разобщены в геноме и до дивергенции млекопитающих и птиц от их общего предка, которая, вероятно, произошла около 270 млн. лет назад.

Предшествующая этому стадия эволюции глобиновых генов представлена их организацией у лягушки *Xenopus laevis*, где α - и β -гены сцеплены, располагаясь на расстоянии только около 8 т. п. н. друг от друга. (Оба гена присутствуют в геноме в виде двух неидентичных копий, что, возможно, объясняется тетраплоидизацией генома организма-предка.) Поскольку амфибии отделились от ветви млекопитающих-птиц около 350 млн. лет назад, время разделения генов α - и β -глобинов, произошедшего предположительно путем транспозиции, можно приблизительно отнести к раннему периоду эволюции позвоночных.

В случае организмов, находящихся на предыдущей стадии эволюционного развития, мы пока располагаем данными только о белковых цепях, но не о генах. У некоторых «примитивных» рыб имеется только один тип глобиновых цепей, так что они, по-видимому, отделились от общего эволюционного пути еще до дупликации глобинового гена-предка, которая привела к образованию его α - и β -вариантов. Скорее всего это произошло около 500 млн. лет тому назад в процессе эволюции костных рыб.

Можно проследить и более ранние этапы эволюции глобиновых генов, поскольку последовательность аминокислот единственной цепи миоглобина млекопитающих свидетельствует о том, что он отделился от пути эволюции глобинов примерно 800 млн. лет назад. Мы все еще не располагаем данными об этом гене. Ген леггемоглобина растений, имеющий сходство с глобиновым геном, отделился еще раньше (гл. 20).

Дивергенция нуклеотидных последовательностей указывает на различие путей эволюции организмов

Большинство изменений в аминокислотной последовательности белков обусловлено мутациями небольших участков генома, медленно накапливающимися с течением времени. Точковые мутации и небольшие вставки и делеции возникают случайно, по-видимому, с более или менее равной вероятностью во всех участках генома, за исключением «горячих точек», где частота мутирования существенно выше. Многие мутации, изменяющие аминокислотную последовательность, оказываются *вредными* и довольно быстро отбрасываются в ходе естественного отбора (скорость этого процесса зависит от степени повреждающего эффекта). Меньшее число мутаций оказывается полезным, но эти мутации могут распространиться в популяции и в конце концов вытеснить исходную нуклеотидную последовательность. Когда мутантный вариант гена вытесняет исходный, говорят, что мутация *закрепилась* в популяции. Очень спорный вопрос: какая доля мутационных изменений в аминокислотной последовательности может оставаться *нейтральной*, т.е. не оказывать действия на функцию белка, и поэтому может накапливаться в результате *случайного дрейфа и закрепления*?

Скорость накопления мутационных изменений является характерной особенностью каждого белка, вероятно зависящей (по крайней мере частично) от его чувствительности к изменениям. В пределах вида белок эволюционирует в результате мутационных замен, которые удаляются или закрепляются в популяции. Присутствие в популяции двух (или более) аллельных вариантов называют *полиморфизмом*. Такая ситуация (полиморфизм) может быть стабильной в двух случаях: 1) если ни один из вариантов не имеет преимуществ перед другим, 2) если преимущества какого-либо варианта уравновешиваются его недостатками. Полиморфизм может быть и временным — как, по-видимому, в том случае, когда один вариант вытесняется другим. При тщательном изучении пула генов того или иного вида мы обнаруживаем только уцелевшие варианты.

При разделении одного вида на два новых каждый из них образует самостоятельный эволюционный пул. Сравнительно соответствующие белки двух видов, можно выявить различия между ними, накопившиеся с *того времени, когда их предки перестали скрещиваться между собой*. Некоторые белки *высококонсервативны* и мало меняются или не меняются от вида к виду. Это говорит о том, что практически любые изменения оказываются вредными и отбрасываются при естественном отборе.

Различия между двумя белками выражаются как их *дивергенция*, т.е. процент различающихся аминокислотных остатков. При сравнительном изучении опреде-

ленного белка у нескольких видов его дивергенция для каждой пары сравниваемых видов пропорциональна (более или менее) времени, прошедшему с момента их разделения. Тем самым мы получаем *эволюционные часы*, измеряющие, по всей видимости, равномерное накопление мутаций в процессе эволюции данного белка.

Скорость дивергенции может быть измерена как процент различий, накопившихся за миллион лет, или обратной величины — *единицы эволюционного времени (ЕЭВ)*, соответствующей времени (в миллионах лет), необходимому для осуществления дивергенции на 1%. Понятие «часы», принятое в отношении попарного сравнения видов (с учетом практических трудностей при определении действительного времени видообразования), можно распространить и на родственные гены *внутри* одного вида. Зная дивергенцию таких генов, можно вычислить время, прошедшее с момента их возникновения в результате дупликации.

Два типа дивергенций последовательностей ДНК

Дивергенция нуклеотидных последовательностей нуклеиновых кислот может отличаться от дивергенции соответствующих белков. Различия эти могут быть обусловлены тем, что каждая аминокислота кодируется триплетом нуклеотидных оснований, где третье основание часто не является значащим. Поэтому необходимо разделить нуклеотиды на *потенциальные сайты замещения и молчащие сайты*. Мутация в сайте замещения приводит к изменению аминокислоты, кодируемой данным триплетом. Эффект мутации (вредной, нейтральной или полезной) зависит от результата, к которому приводит замена аминокислоты. Мутация в молчащих сайтах приводит лишь к замене одного синонимичного кодона на другой, и, следовательно, изменения белка при этом не происходит. Как правило, сайты замещения составляют 75% кодирующих последовательностей, а молчащие сайты — 25%.

Несмотря на то что мутации молчащих сайтов нейтральны по отношению к белку, они могут влиять на экспрессию гена, вызывая изменения нуклеотидных последовательностей РНК. Это может привести к изменению вторичной структуры РНК, отражающейся на транскрипции, процессинге или трансляции. Другая возможность заключается в том, что изменение кодона на кодон-синоним потребует наличия другой тРНК, что может повлиять на эффективность трансляции. И наконец, кроме кодирующих последовательностей гена имеются и не транслируемые участки. Мутации в этих участках также потенциально нейтральны, не считая их влияния на вторичную структуру или регуляторные области (обычно довольно короткие).

Мутации в сайтах замещения должны были бы соответствовать дивергенции аминокислот, которая фактически выражается как процент замен. В действительности подсчитанная дивергенция не отражает всех происшедших в процессе эволюции изменений, поскольку в одном кодоне могло произойти несколько замен. Обычно на это делается поправка. Без поправки дивергенция нуклеиновой кислоты по сайтам замещения, равная 0,45%, соответствует дивергенции аминокислотных замен, равной 1% (считая, что среднее число сайтов замещения в кодоне составляет 2,25).

Рассмотрим для примера β - и δ -глобиновые цепи человека: из 146 аминокислотных остатков различаются 10,

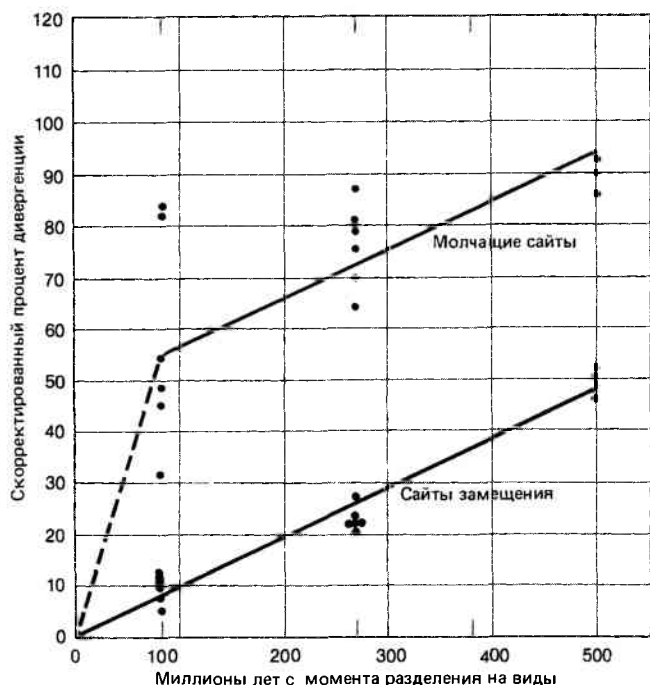


Рис. 21.7. Дивергенция нуклеотидных последовательностей ДНК зависит от расхождения видов в процессе эволюции. Каждая точка на графике соответствует степени дивергенции пары генов.

и дивергенция, следовательно, равна 6,9%. Из 441 нуклеотидов ДНК различается 31. Однако эти изменения очень неравномерно распределены между сайтами замещения и молчащими сайтами. Имеется 11 замен в 330 сайтах замещения, но 20 замен всего лишь в 111 молчащих сайтах. С учетом этого скорректированные оценки дивергенции составляют 3,7% для сайтов замещения и 32% для молчащих сайтов, различаясь почти на порядок.

Столь значительные различия в дивергенции сайтов замещения и молчащих сайтов непосредственно свидетельствуют о наличии гораздо больших ограничений на положения нуклеотидов, влияющих на строение белка, чем на положения нуклеотидов, не оказывающих такого действия. Поэтому маловероятно, что сколько-нибудь значительная часть аминокислотных замен нейтральна. Если принять, что скорость мутирования молчащих сайтов соответствует скорости закрепления мутаций (иными словами, если допустить, что молчащие сайты вообще не подвергаются отбору), то тогда за период, прошедший со времени дивергенции β - и δ -генов, должны были произойти замены в 32% из 330 сайтов замещения, т. е. в 105 сайтах. Из них сохранились только 11. Следовательно, 90% всех мутаций были отсеяны в ходе отбора.

Использование часов для изучения эволюции генов глобина

Сравнивая нуклеотидные последовательности гомологичных генов различных видов организмов, можно определить темпы дивергенции как сайтов замещения, так и молчащих сайтов.

При попарном сопоставлении видов обнаруживается, что для α - и β -глобиновых генов млекопитающих, дивергировавших со времени разделения млекопитающих на виды около 85 млн. лет назад, среднее значение диверген-

ции составляет 10%. Это соответствует скорости дивергенции сайтов замещения, равной 0,12% за 1 млн. лет, или ЕЭВ, равной 8,5.

Эта же скорость дивергенции сохраняется для генов, дивергировавших еще раньше. Например, среднее значение дивергенции сайтов замещения соответствующих глобиновых генов млекопитающих и цыплят составляет 23%. Считая, что эволюционное расхождение этих организмов произошло примерно 270 млн. лет назад, мы получим скорость дивергенции, равную 0,09% за 1 млн. лет.

Продвигаясь еще дальше назад по эволюционной лестнице, можно сравнивать α - и β -глобиновые гены внутри одного вида. Эти гены дивергировали не менее 500 млн. лет назад (см. рис. 21.6). Средняя дивергенция сайтов замещения для них составляет около 50%, что дает скорость 0,1% за 1 млн. лет.

Соответствующие данные приведены в виде графика на рис. 21.7, где показано, что часы дивергенции сайтов замещения генов глобина имеют среднюю скорость около 0,096% за 1 млн. лет или ЕЭВ, равную 10,4. С учетом неточностей в определении времени дивергенции видов результаты хорошо подтверждают предположение о линейности часов.

Данные по дивергенции молчащих сайтов характеризуются значительно меньшей точностью. Для каждого случая очевидно, что дивергенция молчащего сайта существенно выше дивергенции сайта замещения (в 2–10 раз). Но увеличение дивергенции молчащих сайтов при попарном сравнении настолько велико, что трудно понять, применимо ли в данном случае понятие часов.

Как видно из рис. 21.7, скорость дивергенции нелинейно возрастает во времени. Если принять, что нулевому времени расхождения видов соответствует нулевая дивергенция, то видно, что скорость дивергенции молчащих сайтов гораздо выше в течение примерно первых 100 млн. лет. Затем она уменьшается. Одно из объяснений этого состоит в том, что фракция, содержащая примерно половину молчащих сайтов, быстро (в течение 100 млн. лет) оказалась насыщенной мутациями; эта фракция ведет себя как содержащая нейтральные сайты. В другой фракции накопление мутаций происходит медленнее, примерно со скоростью их накопления сайтами замещения; эта фракция соответствует сайтам, молчащим по отношению к белку, но подвергающимся отбору по какой-либо другой причине. Во всех случаях наличие таких неопределенностей не позволяет использовать молчащие сайты в качестве эволюционных часов.

Можно, наоборот, использовать значение скорости дивергенции для определения того времени, когда произошло расхождение генов внутри вида. Для β - и δ -глобиновых генов человека различия в числе сайтов замещения составляют 3,7%. При значении ЕЭВ, равной 10,4, эти гены должны были дивергировать $10,4 \cdot 3,7 = 40$ млн. лет назад, примерно в то время, когда произошла дивергенция обезьян на обезьян Нового Света, Старого Света, крупных человекообразных и людей. У всех высших приматов имеются β - и δ -гены; из этого следует, что дивергенция данных генов началась непосредственно перед этим этапом эволюции.

Что касается еще более ранних событий, то дивергенция сайтов замещения β - и δ -генов равняется 10%; это соответствует времени расхождения генов около 100 млн. лет тому назад. Таким образом, расхождение глобиновых генов могло предшествовать разделению млекопитающих на виды или происходить с ним одновременно.

Дивергенция сайтов замещения β - и γ -глобиновых генов составляет 18%; соответственно расхождение генов должно было произойти примерно 200 млн. лет назад. Поэтому дивергенция между генами гемоглобина взрослого и эмбриона (плода), по-видимому, произошла спустя некоторое время после разделения млекопитающих и птиц (т.е. в процессе эволюции рептилий, предшествовавшем разделению млекопитающих на виды).

Таким образом, можно построить эволюционное древо для кластера глобиновых генов человека (рис. 21.8). Свойства, получившие развитие в процессе эволюции до разделения млекопитающих на виды, обнаруживаются у всех млекопитающих. Свойства, эволюционировавшие после этого, должны были эволюционировать независимо у различных видов млекопитающих. В каждом виде произошли сравнительно недавние изменения в структуре кластеров; так, мы наблюдаем различия в числе генов (один β -глобиновый ген у человека, два — у мыши) или их типе (до сих пор нет уверенности в том, имеются ли различные β -подобные глобины эмбриона и плода у кролика и мыши).

Механизмы, обеспечивающие сохранение в геноме функционально активных последовательностей

Дупликация гена, по-видимому, приводит к немедленному ослаблению давления отбора на его нуклеотидную последовательность. Теперь, когда имеются две идентичные копии гена, изменение нуклеотидной последовательности одной из них уже не вызовет потери организмом функционально активного белка, поскольку другая копия гена будет по-прежнему кодировать необходимую аминокислотную последовательность. Таким образом, при наличии двух генов давление отбора будет уменьшаться до тех пор, пока один из них не мутирует столь сильно по сравнению с исходным вариантом гена, что давление целиком будет направлено на другой ген. Вполне вероятно, что, сразу же после того как происходит дупликация гена, изменения быстрее накапливаются в одной из копий, что в конце концов приводит к появлению новой функции (или к ее исчезновению в случае появления псевдогена). При появлении новой функции эволюция данного гена далее может происходить с прежней, более медленной скоростью, характерной для гена с первоначальной функцией. По-видимому, именно такой механизм лежит в основе разделения функций генов эмбриональных и взрослых глобинов.

И все-таки в некоторых случаях дублицировавшиеся гены осуществляют одну и ту же функцию, кодируя идентичные или почти идентичные белки. У человека с двумя генами α -глобина кодируются идентичные белки, а два γ -глобина различаются только одной аминокислотой. Не парадоксальны ли эти факты в свете предположения о том, что за дубликацией гена должно следовать его изменение? Для таких случаев, по-видимому, возможно исключить единственное простое объяснение, состоящее в том, что оба гена необходимы организму и выдерживают давление естественного отбора, поскольку необходим синтез достаточного количества белка. Тогда каким образом отбор оказывает давление на оба гена?

Это может осуществляться с помощью механизмов двух типов. В их основе лежит предположение о том, что неаллельные гены наследуются не независимо, а снова

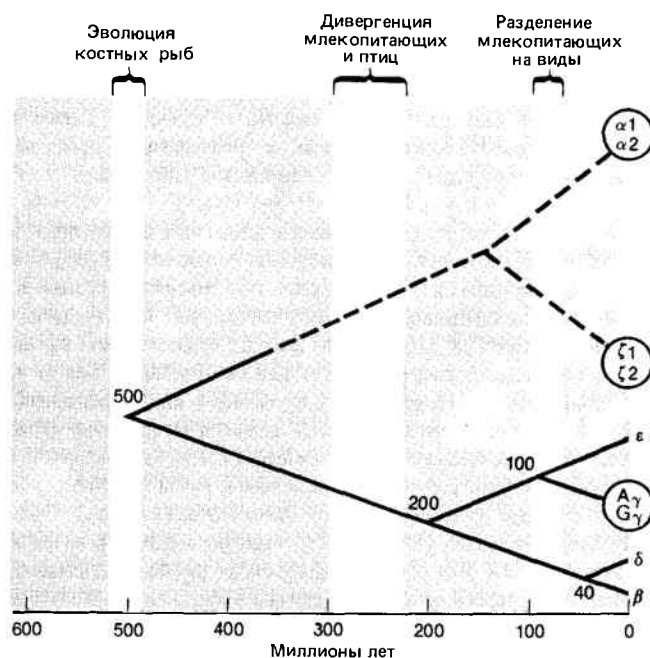


Рис. 21.8. Дивергенция сайтов замещения между парами β -глобиновых генов позволяет проследить процесс эволюции кластера глобиновых генов человека.

Это эволюционное древо объясняет, как произошло разделение генов глобина на классы. Причины возникновения дубликаций индивидуальных генов, обозначенных кружками, не известны. Время дивергенции генов не установлено.

и снова воспроизводятся *одной* из копий предыдущего поколения. Когда мутация затрагивает одну из копий, она либо случайно элиминируется (поскольку функцию этой копии принимает на себя другая копия), либо распространяется на обе копии (поскольку мутантная копия становится доминантной). Во втором случае мутация служит материалом для естественного отбора. В результате два гена эволюционируют вместе, как будто они представляют собой один локус. Такое явление называют **согласованной** или **сопряженной эволюцией** (иногда **козволонцией**).

В случае одного из механизмов предполагается, что нуклеотидные последовательности неаллельных генов непосредственно сравниваются друг с другом и приводятся к одному виду («гомогенизируются») под действием ферментов, распознающих любые различия в последовательности ДНК. Такая «гомогенизация» может происходить путем обмена между генами одиночными цепями ДНК с образованием генов, одна цепь которых происходит от одной копии, а вторая — от другой. Любые различия проявляются в неправильном спаривании оснований, что может привлекать ферменты, способные удалять основание и заменять его на другое так, чтобы оставались только пары А—Т и G—С. Такой процесс, называемый **генной конверсией**, связан с генетической рекомбинацией, как описано в гл. 35.

При использовании другого механизма обе копии гена должны физически воспроизводиться из одной копии. Это может быть результатом неравного кроссинговера. Например, в случае, изображенном на рис. 21.3, хромосома, содержащая локус из трех генов, может потерять один ген в результате делеции. Из двух оставшихся генов $1\frac{1}{2}$ последовательности включают последовательности

одной из исходных копий; от второй копии гена осталась лишь $1/2$ ее последовательности. Любая мутация первой области теперь будет затрагивать оба гена и попадет под давление естественного отбора. Такое явление называют **закреплением в результате кроссинговера**; оно обсуждается при описании эволюции сателлитной ДНК в гл. 24.

Мы можем определить рамки действия описанных механизмов, сравнивая нуклеотидные последовательности дублировавшихся генов. Если эти последовательности подвергались сопряженной эволюции, мы не увидим накопления замен в молчащих сайтах (поскольку процесс «гомогенизации» касается их в той же степени, как и сайтов замещения). Известно, что механизм, обеспечивающий «гомогенизацию» последовательностей, не обязательно распространяется за пределы гена, поскольку в некоторых случаях дублировавшиеся гены имеют совершенно разные фланкирующие последовательности. Действительно, можно увидеть, что границы концов «гомогенизированных» последовательностей резко выделяются. Кроме того, необходимо помнить, что наличие таких механизмов может обесценить определение хода эволюции таких генов на основании их *дивергенции, поскольку дивергенция свидетельствует только о времени, прошедшем с момента последнего события гомогенизации/регенерации, а не исходной дупликации генов.*

Псевдогены — тупики эволюции

Псевдогены называются так потому, что они содержат нуклеотидные последовательности, сходные с последовательностями функционально активных генов, но не могут экспрессироваться с образованием функционально активного белка.

Некоторые псевдогены имеют в целом такую же структуру, как и функционально активные гены, с обычным расположением последовательностей, соответствующих экзонам и интронам. Они становятся неактивными в результате мутаций, нарушающих одну или все стадии экспрессии гена. Эти изменения могут проявляться в виде нарушения иницирования транскрипции, препятствовать осуществлению сплайсинга на границах экзон—интрон или приводить к преждевременному терминированию трансляции. Обычно псевдоген несет несколько вредных мутаций, вероятно, потому, что ген, однажды перестав быть активным, стал объектом для дальнейшего накопления мутаций. Такие псевдогены были обнаружены во многих системах генов, включая гены глобинов, иммуноглобулинов, антигенов гистосовместимости и т. д.

Типичный пример псевдогена — псевдоген кролика $\psi\beta 2$ с обычной организацией экзонов и интронов, по строению близкий к функционально активному гену $\beta 1$. Но в кодоне 20 псевдогена $\psi\beta 2$ имеется делеция одной пары нуклеотидных оснований, вызывающая сдвиг рамки считывания, из-за которого трансляция терминируется вскоре после ее начала. В результате точковых мутаций оказались измененными несколько расположенных правее кодонов, кодирующих аминокислоты, имеющихся во всех β -глобинах. Ни один из двух интронов псевдогена $\psi\beta 2$ не сохранил пограничных последовательностей, удовлетворяющих правилу GT—AG. Поэтому, вероятно, интроны не могут быть удалены при сплайсинге, даже если бы ген и транскрибировался. Однако транскриптов, соответствующих $\psi\beta 2$, не обнаружено, возможно, вследствие изменений в его 5'-фланкирующей области.

Поскольку этот перечень дефектов включает мутации, потенциально препятствующие каждой стадии экспрессии генов, мы не в силах определить, какое именно событие привело к инактивации данного гена. Однако, исходя из степени дивергенции последовательностей псевдогена и функционально активного гена, можно предположительно установить время появления псевдогена и время начала накопления в нем мутаций.

Если псевдоген стал неактивным сразу же после его образования из гена $\beta 1$ в результате дупликации, можно было бы ожидать, что скорости дивергенции как сайтов замещения, так и молчащих сайтов будут одинаковы. (Причин для того, чтобы они различались, нет, поскольку ген не транслируется.) Но в действительности в сайтах замещения оказывается меньше замен, чем в молчащих сайтах. Это указывает на то, что вначале отбор препятствовал заменам в сайтах замещения. Сравнивая число замен в сайтах двух типов, можно подсчитать, что псевдоген $\psi\beta 2$, дивергировав от гена $\beta 1$ около 55 млн. лет назад, оставался функционально активным геном в течение 22 млн. лет, а в течение последних 33 млн. лет является псевдогеном.

Подобные расчеты можно проделать и в отношении других псевдогенов. Некоторые из них, по-видимому, также до того, как стали псевдогенами, активно функционировали. Другие, вероятно, были неактивными с момента своего возникновения. Общий вывод, который можно сделать, исходя из структуры таких псевдогенов, — это независимый характер эволюционирования каждого из них в процессе эволюции кластера глобиновых генов каждого вида организмов. Сказанное подтверждает предположение о том, что возникновение новых генов, за которым следует их закрепление в геноме в качестве функциональных копий, их изменение, приводящее к образованию новых функционально активных генов, или инактивация с образованием псевдогенов — процессы, происходящие в кластере постоянно.

Ген $\psi\alpha 3$ мыши имеет интересную особенность: у него полностью отсутствуют оба интрона. Его нуклеотидная последовательность соответствует (с учетом накопленных мутаций) α -глобиновой мРНК. Примерное время его инактивации совпадает с моментом возникновения дупликации. Исходя из этого, можно предположить, что его инактивация, очевидно, связана с потерей интронов. Ген, лишенный интронов, мог утратить способность к функционированию и сразу начать накапливать мутации. В данном случае основной вопрос заключается в следующем: каким образом утрата интронов произошла столь точно? Поскольку нет оснований считать, что ферментные системы для перестройки ДНК должны узнавать границы экзон—интрон, можно предположить, что в этом сыграло роль участие на определенной стадии самой мРНК. Оно могло бы обеспечивать, например, встраивание в геном обратного транскрипта (возможно, осуществляемого ретровирусом, для которого такое действие является естественным) или могло бы произойти спаривание мРНК с одной из цепей ДНК на одной из стадий процесса рекомбинации (гл. 38).

Псевдогены, которые имеют сходство с РНК-транскриптом, называются **процессированными псевдогенами**. В пользу предположения об их происхождении путем встраивания (в некоторое случайное положение гена) последовательности, синтезированной на РНК, говорит тот факт, что такие псевдогены могут быть расположены в любом участке генома, не обязательно даже на той же хромосоме, что и активный ген.

Насколько распространены псевдогены? Они являются членами большей части семейства генов. Обычно псевдогены составляют очень небольшую часть от общего числа генов. Однако имеется и исключение: рибосомный белок мыши кодируется одним активным геном, имеющим около 15 сходных с ним процессированных псевдогенов. О возможности существования таких случаев следует помнить, когда мы пытаемся рассчитать число генов, исходя из результатов, полученных в опытах по гибридизации.

Если псевдогены — это тупики эволюции, просто нежелательный побочный эффект перестроек функционально активных генов, то почему они до сих пор находятся в геноме? Выполняют ли они какую-нибудь функцию или совершенно бесполезны и, следовательно, не должны подвергаться воздействию естественного отбора? Необходимо помнить, что мы имеем дело с генами, сохранившимися в современных популяциях. Ранее могло быть отброшено любое количество других псевдогенов. Удаление псевдогена, возможно, происходит в результате случайной делеции последовательности псевдогена или накопления мутаций до такой степени, что его уже нельзя отнести к членам исходного семейства последовательностей (по-видимому, это удел любого псевдогена, избежавшего удаления из генома).

Даже реликты эволюции могут дублироваться. В случае генов β -глобина козы имеются два взрослых вида генов: β^A и β^C . Каждый из них имеет свой псевдоген, расположенный на несколько тысяч нуклеотидов левее (псевдогены называются соответственно $\psi\beta^Z$ и $\psi\beta^X$). Псевдогены обладают большим сходством друг с другом, чем со взрослыми генами β -глобина; в частности, у них имеется несколько одинаковых инактивирующих мутаций. Оба взрослых гена β -глобина также имеют большее сходство друг с другом, чем с псевдогенами. Это дает основание предположить, что исходно существовала структура $\psi\beta$ - β , в результате дубликации которой образовались два активных β -гена (далее дивергировавшие вместе с образованием β^A и β^C -генов) и два неактивных гена (дивергировавшие с образованием ныне существующих псевдогенов). Общее предположение, которое может объяснить это явление, состоит в том, что механизмы, ответственные за дубликации, делеции и перестройки генов, воздействуют на все последовательности, относящиеся к членам кластера, независимо от того, являются ли они функционально активными или нет. А какому из них оказать предпочтение в процессе эволюции — функция естественного отбора.

Семейства генов обычно кодируют белки, богато представленные в клетке

Из экспериментов по определению числа генов путем выделения фрагментов генома, соответствующих определенному мРНК- или кДНК-зонду, известно, что в ряде случаев зонду соответствует множество последовательностей генома. Конечно, до тех пор пока не будет установлена первичная структура каждой такой последовательности, мы не будем знать, все ли последовательности активно функционируют или некоторые из них относятся к псевдогенам. Обычно предполагают, что в большинстве своем гены, по-видимому, активны.

По такому критерию (соответствия последовательно-

стей специфичному зонду) обнаруживается множество генов, кодирующих синтез интерферона, актина и тубулина. Внутри семейства генов некоторые его члены могут повторяться чаще, чем другие. Например, существует, по-видимому, по меньшей мере восемь генов для α -типа (лейкоцитарного) интерферона человека, но только один для β -типа (фибробластного).

Множественные гены могут быть объединены в кластеры или рассеяны по геному. В случае интерферона ясно, что они объединены в кластер. В случае генов актина и тубулина *D. melanogaster* копии одного гена, по-видимому, расположены в разных участках хромосомы; следовательно, они явно не связаны.

Является ли постоянство структуры характерным признаком копий множественных генов? Общий план строения глобиновых генов консервативен. Гены интерферона, по-видимому, имеют сходную в общих чертах структуру, для которой характерно полное отсутствие интронов. Но гены актина имеют прерывистую структуру, сильно варьирующую у разных генов. У этих генов участки, кодирующие белок, обладают высокой степенью гомологии, но сходство между фланкирующими или даже нетранслируемыми областями в пределах одного вида организмов незначительно (или оно вообще отсутствует). Например, интроны генов актина *D. melanogaster* занимают разные положения. Ни один из этих генов не похож на единственный ген актина, обнаруженный в дрожжах. Они отличаются также от актиновых генов морского ежа; по меньшей мере некоторые из них объединены в кластер. Таким образом, если все эти актиновые гены произошли от общего гена-предка, расположение экзонов и интронов претерпело существенные изменения. Возможно, что ген-предок обладал большим числом интронов, и в разных копиях гена внутри вида или у разных видов были утрачены разные интроны. Несомненно, это предполагает более высокую скорость изменений, чем в случае глобиновых генов.

Такие данные выглядят несколько анекдотично, поскольку они говорят о том, что строение гена может варьировать в широких пределах, но ничего не говорят о том, что же следует принимать за норму.

Единственный вывод общего характера, который можно сделать на основе приведенного ограниченного объема данных, состоит в том, что функционирующие гены могут иметь прерывистое и непрерывное строение и что изменения в расположении интронов не обязательно влияют на активность генов. Гены, кодирующие одинаковые или близкие белки, не обязательно должны быть организованы в виде тандема, но могут быть рассеяны по геному в виде отдельных индивидуальных генов или, возможно, небольших кластеров.

Какое заключение можно сделать относительно общего числа генов? Пытаясь найти ген, соответствующий определенному белку, мы часто находим другие сходные с ним активные гены и (иногда) псевдогены. Мы не предполагаем достаточным количеством данных, чтобы по ним экстраполировать общее число генов, но можно задать вопрос: в какой мере обнаруживаемая степень повторяемости может помочь нам понять, чем объясняется наличие явно избыточного количества ДНК?

Повторение последовательностей приводит к увеличению количества ДНК, кодирующей определенную функцию. Возможно существование нескольких активных генов, кодирующих либо один и тот же белок, либо какие-то неожиданные его варианты. В тех случаях, когда

копии генов объединены в кластер, расстояния между генами могут быть весьма значительными, в связи с чем увеличивается общее количество ДНК, ответственное за осуществление данной функции. Псевдогены можно рассматривать как дополнительные последовательности, связанные с осуществлением функции, но не участвующие в ее реализации.

Было бы не удивительно, если бы количество ДНК, кодирующее определенную функцию, в 3–4 раза превышало ее ожидаемое количество в случае кодирования каждой функции одним-единственным геном. Анализируя число генов в гл. 18 и 19, мы пришли к выводу о наличии у млекопитающих примерно 30 000–40 000 кодируемых генами функций. Если допустить, что это, по-видимому, соответствует 150 000 генов, считая, что каждый ген в несколько раз длиннее своей кодирующей области и занимает участок ДНК, скажем, 5000 п.н., то мы получим величину около 10^9 п.н., что составляет около трети от размеров генома.

Первое предположение, касающееся этой части генома, состоит в том, что основная часть функции кодируется несколькими генами, а не уникальным геном, как казалось ранее. Второе предположение заключается в том, что значительное количество ДНК не несет кодирующей функции, находясь либо внутри генов (в виде интронов), либо между ними (поскольку кластеры генов могут быть достаточно протяженными и при этом находиться на значительном расстоянии друг от друга). Это чрезвычайно большое количество ДНК только косвенным образом связано с функционированием генов.

Мы ощутило продвинулись вперед в определении количества ДНК, допустив, что число генов просто больше предполагаемого, поскольку при изучении определенной функции мы чаще имеем дело с небольшими семействами, а не с индивидуальными генами. Принимая это во внимание, можно считать, что около $2 \cdot 10^8$ п.н. генома млекопитающих (около 10% ДНК) непосредственно кодируют белки.

С учетом дополнительной ДНК, находящейся внутри этих генов или в межгенных последовательностях, не кодирующих каких-либо других функций, мы получим в

3 раза большее количество ДНК, т.е. $6 \cdot 10^8$ п.н. Оценка количества ДНК, осуществляющей определенную функцию, будет заниженной, если размеры генов или расстояния между ними больше, чем мы предполагали. Может ли геном млекопитающих размером $3 \cdot 10^9$ п.н. целиком состоять только из генов, число которых предположительно составляет 150 000?

Мы разобрались с частью генома, допустив, что она как-то связана с кодирующими участками; хотя эта часть сама по себе не кодирует белков, она не выполняет какой-либо другой явной функции. Можно ли быть в этом уверенным? А что если что-то не так в системе с таким избыточным количеством ДНК; не пропустили ли мы какое-то важное свойство генома?

Даже если наши рассуждения о составе генома млекопитающих верны, они не объясняют парадоксальности значений величины C , т.е. того, что общее количество ДНК может варьировать в более широких пределах, чем можно было бы ожидать при существующей вариации числа генов. Означает ли это, что размеры самих генов или расстояния между генами не столь существенны для организма и подвержены целенаправленной коррекции?

Рекомендуемая литература

Обзор, посвященный структуре и функции гемоглобинов человека, написан Маниатисом и др. (*Maniatis et al.*, *Ann. Rev. Genet.*, **14**, 145–178, 1980). Основные черты эволюции генов и белков рассмотрены в обзоре Уилсона и др. (*Wilson et al.*, *Ann. Rev. Biochem.*, **46**, 573–639, 1977). Оригинальное сообщение о псевдогене, не имеющем интронов, было сделано Нишиока и др. (*Nishioka et al.*, *Proc. Nat. Acad. Sci. USA*, **77**, 2806–2809, 1980.) Псевдогенам посвящен обзор Литтла (*Little*, *Cell*, **28**, 683–684, 1982). Новые данные, касающиеся причин возникновения нарушений при талассемиях¹, изложены Уэзероллом и Клэггом (*Weatherall*, *Clegg*, *Cell*, **29**, pp. 7–9, 1982).

¹ Лимборская С. А. Системы глобиновых генов.—Итоги науки и техники, Молекулярная биология, ВИНТИ, **19**, 84–116, 1982.—Прим. ред.

Глава 22

ГЕНОМЫ КЛЕТОЧНЫХ ОРГАНЕЛЛ

В геноме высших эукариот не обнаружено точного аналога оперона прокариот, зато в ряде случаев имеется определенный участок ДНК, содержащий набор генов, каждый из которых вносит свой вклад в создание одного и того же элемента фенотипа клетки. ДНК-геномы митохондрий и хлоропластов кодируют образование всех видов РНК и некоторых видов белков, обеспечивающих осуществление функций органеллы. Несмотря на то что геном клеточной органеллы не есть кластер в смысле объединения тандемно повторяющихся генов, возникших в результате дупликации и изменения общего гена-предка, его можно считать кластером в том смысле, что все его гены кодируют белки с близкими функциями.

В некоторых отношениях организация генома орга-

неллы, возможно, имеет большее сходство со строением вирусного, а не клеточного генома, поскольку он кодирует лишь некоторые, а не все белки, необходимые для постоянного воспроизведения органеллы. Другие белки кодируются генами ядра, синтезируются с помощью цитоплазматического аппарата синтеза белка, а затем «импортируются» в органеллу. Взаимоотношения органеллы и клетки носят скорее характер симбиоза, а не паразитирования, поскольку органелла — постоянный компонент клетки, а не вирус, использующий ее. Фактически геном органеллы — часть ДНК, физически ограниченная определенным участком клетки и поэтому характеризующаяся своим собственным способом экспрессии генов и ее регуляции.

Гены органелл не подчиняются законам Менделя

Первым доводом в пользу наличия генов за пределами ядра послужили данные о неменделевском наследовании признаков у растений (обнаруженном в начале этого века сразу после перераскрытия менделевского наследования). Неменделевское наследование отличается в двух отношениях.

Во-первых, при скрещивании между мутантным видом и видом дикого типа наблюдается аномальное распределение признаков. Крайняя форма этого — *однородительское наследование*, когда наследуется генотип только одного из родителей, а генотип второго родителя постоянно теряется. В более умеренных случаях потомство с генотипом одного из родителей преобладает по сравнению с потомством, наследующим генотип другого родителя. Обычно предпочтительнее наследуется материнский генотип. Это явление иногда называют **материнским наследованием**. Здесь важно то, что в потомстве преобладает генотип родителя одного пола в отличие от законов генетики Менделя, когда при реципрокных скрещиваниях признаки обоих родителей наследуются в равной степени (гл. 1).

Во-вторых, в тех случаях, когда сохраняются оба генотипа, происходит сегрегация дикого и мутантного фенотипов во время соматического роста растения. Так у одного и того же растения некоторые ткани могут иметь фенотип одного из родителей, тогда как другие ткани имеют фенотип второго родителя. Это противоречит представлению о постоянстве фенотипа всех соматических клеток, определяемого ядерным генотипом, как следует из законов Менделя.

Традиционные генетические методы не могли быть применены к наследованию таких признаков; поэтому все, что можно было сказать, исходя из этих результатов, — это то, какой бы механизм их ни обуславливал, он не связан с хромосомами. Каждый такой ген необходимо было исследовать независимо, поскольку рекомбинация между признаками не наблюдалась; соответственно было невозможно построить карты сцепления. Спорным был вопрос о том, какие механизмы, генетические или эпигенетические, определяли такое наследование. (**Эпигенетические изменения** — изменения, влияющие на фенотип клетки, но не затрагивающие ее генотипа; например, стабильные изменения экспрессии гена.)

Когда стало ясно, что в этом процессе участвуют истинные генетические механизмы, для описания расположения соответствующих генов стали использовать термин **цитоплазматическая наследственность**. Однако мы не будем использовать этот термин, поскольку важно отличать процессы, происходящие в цитоплазме вообще, от процессов в специфических органеллах. Для генов, расположенных вне ядра, лучше использовать термин **внеядерный**, оставляя открытым в каждом конкретном случае вопрос о том, в какой органелле заключены гены (если они заключены в какой-либо органелле). Это позволяет использовать термин **цитоплазматический синтез белка** для описания конечной стадии экспрессии ядерных генов, тогда как гены органеллы транскрибируются и транслируются в *той же самой* органелле, в которой они расположены.

Наличие ДНК в митохондриях и хлоропластах было непосредственно продемонстрировано ее выделением из очищенных препаратов этих органелл. Особенности на-

следственного материала органелл объясняются локализацией и свойствами ДНК.

В какой степени генетическая информация, заключенная в органелле, представлена в зиготе, зависит от относительного вклада каждого из родителей или от различной степени стабильности генетической информации (или и от того и от другого). Этим она резко отличается от генетической информации, заключенной в ядре, в равной степени наследуемой от каждого из родителей.

Существование материнского наследования у высших животных согласуется с гипотезой о том, что источником митохондрий в зиготе, по-видимому, служит только яйцеклетка, а не сперматозоид. Таким образом, митохондриальные гены поступают целиком от матери; митохондриальные гены мужских клеток отбрасываются при каждой смене поколений.

Другой механизм однородительского наследования обнаружен у водоросли *Chlamydomonas reinhardtii*, у которой зигота образуется в результате слияния двух гамет одинакового размера. Оба родителя привносят в зиготу хлоропластную ДНК. Однако вскоре после образования зиготы ДНК одного из родителей разрушается, и в органелле остается только ДНК второго родительского организма.

В отличие от этих двух случаев в других ситуациях (например, у высших растений) невозможно определить причину преобладания одного родительского генотипа над другим. Даже в тех случаях, когда в зиготу поступают разные количества цитоплазмы от разных родителей, так, что, по-видимому, один из родителей «поставляет» большее число копий генома органеллы, чем другой, возможно, что в дальнейшем баланс изменяется под действием механизмов отбора.

Еще одно свойство неменделевского наследования состоит в том, что в тех случаях, когда сохраняются оба генотипа, во время роста растения одни ткани приобретают дикий, а другие — мутантный фенотип. Таким образом, у одного и того же гетерозиготного растения некоторые ткани имеют фенотип одного из родителей, в то время как другие обладают фенотипом второго родителя. Такое явление **соматической сегрегации** отличается от стабильности свойств, определяемых, согласно менделевской генетике, ядерным генотипом.

Связь соматической сегрегации с неменделевским наследованием свидетельствует, очевидно, о наличии множества копий генов органеллы в зиготе и о том, что во время деления соматической клетки они могут быть распределены по разным дочерним клеткам. Нам не известно, каким законам подчиняется деление ДНК органелл (в частности, имеет ли родительская клетка какое-либо влияние на этот процесс).

Вопрос о том, какая из родительских органелл является носителем определенного генетического признака, нетривиален. Поскольку известно, что единственная органелла в клетках высших животных, содержащая ДНК, — это митохондрия, то, вероятно, она и служит единственным хранилищем внеядерного генетического материала. Однако у растений и у некоторых одноклеточных эукариот имеются и хлоропласты, и митохондрии. И опять-таки это единственные органеллы, о которых известно, что они содержат ДНК; вероятно (хотя и не доказано), гены, кодирующие признаки, подчиняющиеся неменделевскому наследованию, располагаются в одной из этих органелл. Однако может быть не ясно, в какой именно из этих органелл локализован опреде-

ленный ген. В некоторых случаях его местонахождение можно определить по характеру влияния мутации на фенотип, однако в других случаях нарушение имеет более общий характер, и его нельзя однозначно связать с какой-либо органеллой.

Геномы органелл представляют собой кольцевые молекулы ДНК

Почти все охарактеризованные к настоящему времени геномы органелл представляют собой одну молекулу ДНК с уникальной последовательностью нуклеотидных оснований. Обычно ее можно выделить в виде кольцевой молекулы, хотя иногда при выделении возникают столь частые разрывы, что основная часть материала приобретает вид линейных фрагментов ДНК. Исключение составляют инфузории, у которых митохондриальная ДНК представлена линейной молекулой. Как правило, в каждой митохондрии содержится несколько копий ее генома. Поскольку в клетке имеется множество митохондрий, то на одну клетку может приходиться большое число геномов данной органеллы. Таким образом, несмотря на то что сам по себе геном органеллы уникален, он представляет собой повторяющуюся последовательность, относительно сходную с любой неповторяющейся ядерной последовательностью.

Геномы хлоропластов имеют сравнительно большие размеры — обычно около 140 т.п.н. у высших растений и до 200 т.п.н. у низших эукариот. Они сопоставимы по размеру с геномом большого бактериофага (например, фага T4), достигающим 110 т.п.н. В табл. 22.1 показано, что у двух видов низших эукариот в каждом хлоропласте содержится большое число кольцевых молекул ДНК. Общее количество хлоропластной ДНК составляет несколько процентов всей клеточной ДНК. К сожалению, мы не располагаем аналогичными цифрами для высших растений, но у них количество ядерной ДНК гораздо больше, и поэтому доля хлоропластной ДНК должна быть довольно мала.

Как видно из табл. 22.2, полные размеры митохондриальных геномов у разных видов могут различаться почти на порядок. Митохондриальные геномы млекопитающих малы, причем малы настолько, что удалось определить полные нуклеотидные последовательности ДНК митохондрий человека, мыши и коровы; все эти геномы имеют размер около 16,5 т.п.н. Данные о числе органелл на одну клетку имеются только для культивируемых линий клеток, и число это велико (несколько сотен). Общее количество митохондриальной ДНК по сравнению с ядерной ДНК мало — менее 1%. По-видимому, на одну митохондрию приходится более одного генома, но

Таблица 22.1
Сложность хлоропластной ДНК варьирует в узких пределах

Вид	Сложность хлоропластной ДНК, т.п.н.	Число геномов на одну органеллу	Число органелл на одну клетку	Соотношение хлоропластной ДНК и всей ДНК клетки
<i>C. reinhardtii</i>	195	70–100	1	14%
<i>E. gracilis</i>	135	40	15	3%
Высшие растения	140	Неизвестно	Неизвестно	Неизвестно

Таблица 22.2
Сложность митохондриальной ДНК варьирует в широких пределах

Вид	Сложность митохондриальной ДНК, т.п.н.	Число геномов на одну органеллу	Число органелл на одну клетку	Соотношение митохондриальной ДНК и всей ДНК клетки
Мышь (L-клетки)	16,2	2	500	0,2%
Человек (клетки HeLa)	16,6	10	800	1%
<i>D. melanogaster</i>	18,4	Неизвестно	Неизвестно	Неизвестно
<i>X. laevis</i>	18,4	»	»	»
<i>S. cerevisiae</i>	84	4	22	18%
<i>P. sativum</i>	110	Неизвестно	Неизвестно	Неизвестно

менее десяти. Размеры митохондриальных геномов несколько выше у плодовой мушки и лягушки, для которых мы не располагаем данными о числе геномов и органелл.

Геном митохондрий дрожжей существенно больше. Его размер у различных штаммов *S. cerevisiae* варьирует в широких пределах, составляя в среднем около 84 т.п.н. В дрожжевых клетках при стандартном определении обнаруживается около 22 митохондрий на одну клетку, что предполагает наличие примерно 4 копий генома в каждой органелле. В растущих клетках содержание митохондриальной ДНК может достигать 18% всей ДНК клетки. (Митохондриальные геномы других видов дрожжей могут иметь существенно большие размеры.)

У растений размеры ДНК митохондрий варьируют в широких пределах. Имеются некоторые данные в пользу того, что митохондриальный геном растений представляет собой не одну молекулу ДНК, а, возможно, состоит из нескольких молекул разного размера. Общий размер митохондриального генома у растений примерно сопоставим с размером хлоропластного генома, так что общий объем генетической информации вне ядра, по-видимому, достаточно велик.

В органеллах экспрессируются их собственные гены

Оба типа органелл, обеспечивающих преобразование энергии в клетке, имеют смешанное «хозяйство». Большая часть белков этих органелл импортируется из окружающей цитоплазмы, где их синтез представляет собой конечную стадию экспрессии ядерных генов. Но в каждой органелле происходит также и свой собственный синтез белка. В случае митохондрий этот процесс направлен на синтез небольшого числа белков, каждый из которых — компонент олигомерного комплекса, состоящего из нескольких импортируемых из цитоплазмы белковых субъединиц, как показано на рис. 22.1. В случае хлоропластов эндогенная экспрессия генов органеллы может приводить к синтезу большего числа белков.

Сам по себе аппарат белкового синтеза в органеллах имеет смешанное происхождение. Большая часть его белковых компонентов или все они транспортируются в органеллу из окружающей цитоплазмы. Однако ни в митохондриях, ни в хлоропластах, по-видимому, не происходит переноса нуклеиновых кислот через мембрану органеллы как в одном, так и в другом направлении.

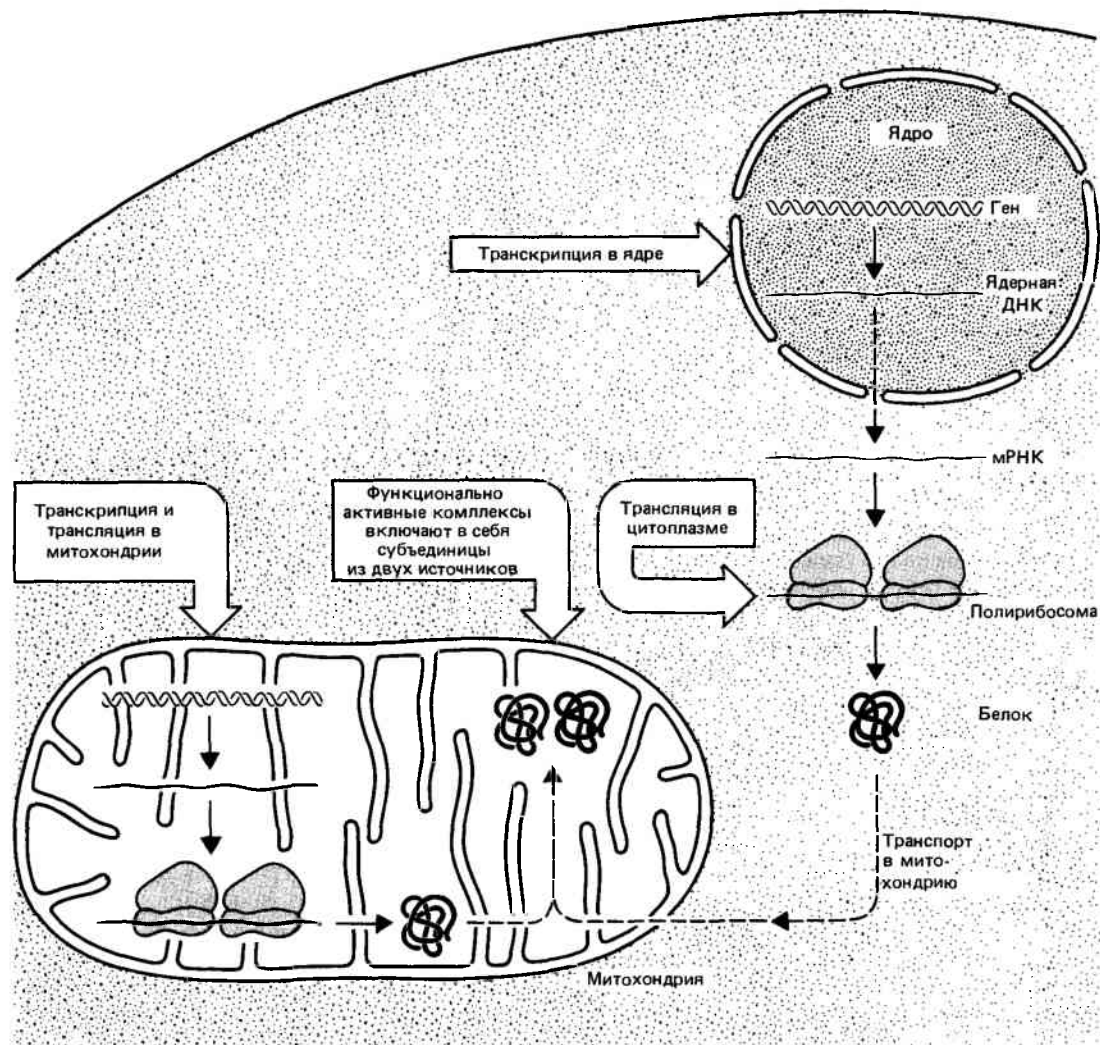


Рис. 22.1. Белковые комплексы митохондрий образуются из продуктов экспрессии ядерных и митохондриальных генов.

Поэтому все РНК, входящие в состав аппарата белково-го синтеза, продуцируются самой органеллой.

Данные о происхождении компонентов, необходимых для экспрессии генов в митохондриях дрожжей, приведены в табл. 22.3. В целом аналогичная картина наблюдается в митохондриях других видов организмов и в хлоропластах. В митохондриях дрожжевых клеток можно выделить восемь мРНК, и они транскрибируются с помощью РНК-полимеразы, которая (предположительно) кодируется ядерными генами. Только мРНК, транскрибируемые с митохондриальных генов, могут транслироваться внутри митохондрии, и, с другой стороны, это единственное место в клетке, где эти мРНК могут экспрессироваться.

Митохондриальная рибосома состоит из обычных большой и малой субъединиц, каждая из которых содержит одну рРНК, кодируемую митохондриальными генами, и множество белков, импортируемых из цитоплазмы. Однако структура этих рибосом у разных видов организмов различается (гл. 8). Набор тРНК меньше, чем это необходимо в цитоплазме, как описано в гл. 7. Аминоацил-тРНК-синтетазы, ответственные за присоединение аминокислот к тРНК, поступают из цитоплазмы. По-видимому, все белковые компоненты митохондриального

аппарата синтеза белков являются уникальными, т. е. они кодируются ядерными генами, отличными от генов, кодирующих компоненты с аналогичными функциями, необходимые для экспрессии генов в ядре и окружающей цитоплазме.

Восемью митохондриальным мРНК соответствуют восемь белков, обнаруживаемых в качестве продуктов синтетических процессов, происходящих в митохондриях.

Таблица 22.3
Митохондриальный аппарат синтеза белка включает в себя РНК, синтезированные в митохондриях, и белки, поступающие из цитоплазмы

Компонент	РНК, синтезированные в митохондриях	Белки, синтезированные в цитоплазме
Матрица	~ 8 мРНК	РНК-полимераза
Рибосома	2 рРНК	~ 75 рибосомных белков
Адаптор	22 тРНК	22 аминокислот-тРНК-синтетаза

Место синтеза белка можно установить *in vivo* по его чувствительности к воздействию веществ, преимущественно ингибирующих синтез белка в цитоплазме или органелле. Обычно синтез белка в митохондриях и хлоропластах чувствителен к антибиотикам, ингибирующим синтез белка в бактериях, например к эритромицину и хлорамфениколу. Эти антибиотики не влияют на синтез белка в цитоплазме, который, однако, обладает чувствительностью к циклогексимиду; на синтез белка в органелле циклогексимид не влияет. Происхождение белков органелл может быть установлено *in vitro* при исследовании продуктов, синтезируемых в препаратах изолированных органелл.

В табл. 22.4 приведены комплексы, содержащие белки, синтезированные в митохондриях дрожжей. АТРаза состоит из двух частей: мембранного фактора, образуемого двумя или более субъединицами, кодируемых митохондриальным геномом, и растворимой АТРаза F1, состоящей примерно из пяти субъединиц, синтезируемых в цитоплазме. Цитохром-с—оксидаза также состоит из субъединиц, происходящих из обоих источников. В состав комплекса цитохромов bc_1 входит один белок митохондриального происхождения, связанный с шестью субъединицами цитоплазматического происхождения. Малая субъединица рибосомы включает в себя один белок (Var 1), кодируемый митохондриальными генами. Были получены мутации, позволяющие идентифицировать почти все митохондриальные гены.

Также были обнаружены и мутации по ядерным генам, препятствующие образованию всех перечисленных комплексов. В мутантах по генам органелл и ядерным генам сборка комплексов осуществляется неправильно, так что некоторые субъединицы, по-видимому, отсутствуют. Это указывает на то, что сборка каждого из таких комплексов с образованием его обычной связанной с мембраной формы—сложный процесс, при котором отсутствие одного компонента может препятствовать сборке других. На практике это приводит к тому, что сложно связать мутации по ядерным генам с определенными субъединицами. Кроме того, имеется ряд мутаций по ядерным генам, препятствующих сборке комплекса, но, по-видимому, не затрагивающих отдельные субъединицы. Это может означать, что имеется несколько кодируемых ядерными генами продуктов, необходимых для сборки, но не входящих в состав комплекса в его окончательном виде.

Таблица 22.4

Белковые комплексы митохондрий состоят из белков, синтезированных в различных частях клетки

Белковый комплекс дрожжей	Субъединицы, синтезированные в митохондриях	Субъединицы, синтезированные в цитоплазме	Субъединицы неизвестного происхождения
АТРаза, чувствительная к олигомицину (340 000 Да)	АТРаза 9 (мембранный фактор)	АТРаза 2, 3, 4, 7 (АТРаза F1)	АТРаза 1, 5, 6, 8
Цитохром-с—оксидаза (137 000 Да)	ЦО 1, 2, 3	ЦО 4, 5, 6, 7	
Комплекс цитохрома bc_1 (160 000 Да)	Цитохром b , апо-протеин	6 субъединиц	

То, что определенная мутация по ядерным генам может нарушать синтез определенного белкового компонента, проще всего объяснить тем, что она затронула уникальный структурный ген. Из этого следует, что должны существовать количественные различия в представительстве белковых компонентов, кодируемых генами органелл и ядерными генами, поскольку копий митохондриального генома намного больше (см. табл. 22.2). По-видимому, ядерные гены экспрессируются более эффективно. Наличие таких различий было прямо подтверждено в случае одного из ферментов хлоропластов, рибулозобисфосфат-карбоксилазы кукурузы, большая субъединица которой кодируется геном органеллы, присутствующим в клетке в виде большого числа копий, а малая субъединица кодируется ядерным геном, представленным неповторяющейся ДНК.

Является ли специфическое разделение труда, демонстрируемое митохондриями и ядрами дрожжей, характерным и для других видов организмов? Общее строение агрегатов, состоящих из белков со смешанным происхождением, сходно по крайней мере у митохондрий других грибов и млекопитающих. В геноме митохондрий дрожжей и млекопитающих присутствуют гены, кодирующие одни и те же белки. Таким образом можно сделать общее предположение о наличии значительной консервативности кодирующих функций митохондриальных геномов у различных видов организмов.

Было обнаружено одно исключение из этого правила. Самая небольшая субъединица АТРаза кодируется митохондриальным геномом у дрожжей и у *Aspergillus nidulans*. Но у других грибов (у *Neurospora crassa*) соответствующий белок, по-видимому, синтезируется в цитоплазме. Соответствующий ген отсутствует и в митохондриальной ДНК млекопитающих. Если в этих случаях интересующая нас субъединица кодируется ядерным геном, то, значит, когда-то должен был произойти обмен этим геном между митохондриальным и ядерным геномами.

Уникальны ли геномы митохондрий и хлоропластов или некоторые последовательности присутствуют в геноме и тех и других органелл или в геноме одной из этих органелл и ядра? Основная часть генома органеллы уникальна, однако у хлоропластной и митохондриальной ДНК кукурузы найдены гомологичные участки размером 12 т.п.н. В некоторых последовательностях генома митохондрий морского ежа *S. purpuratus* имеется гомология с ядерными последовательностями. Обмен ДНК между органеллами или между органеллами и ядром—несомненно редкое событие, но, по-видимому, иногда он все-таки происходит. Мы до сих пор еще не знаем, каково функциональное значение таких гомологичных последовательностей.

Митохондриальный геном дрожжей имеет большие размеры

Пятикратные различия в размере геномов *S. cerevisiae* (84 т.п.н.) и млекопитающих (16 т.п.н.) уже сами по себе наводят на мысль о том, что в их генетической организации должны быть большие различия, несмотря на общность их функции. Мы уже упоминали об отсутствии данных о существенном увеличении у дрожжей количества эндогенно синтезируемых продуктов, ответственных за осуществление ферментативных процессов в митохондриях. Кодирует ли дополнительный генетический мате-

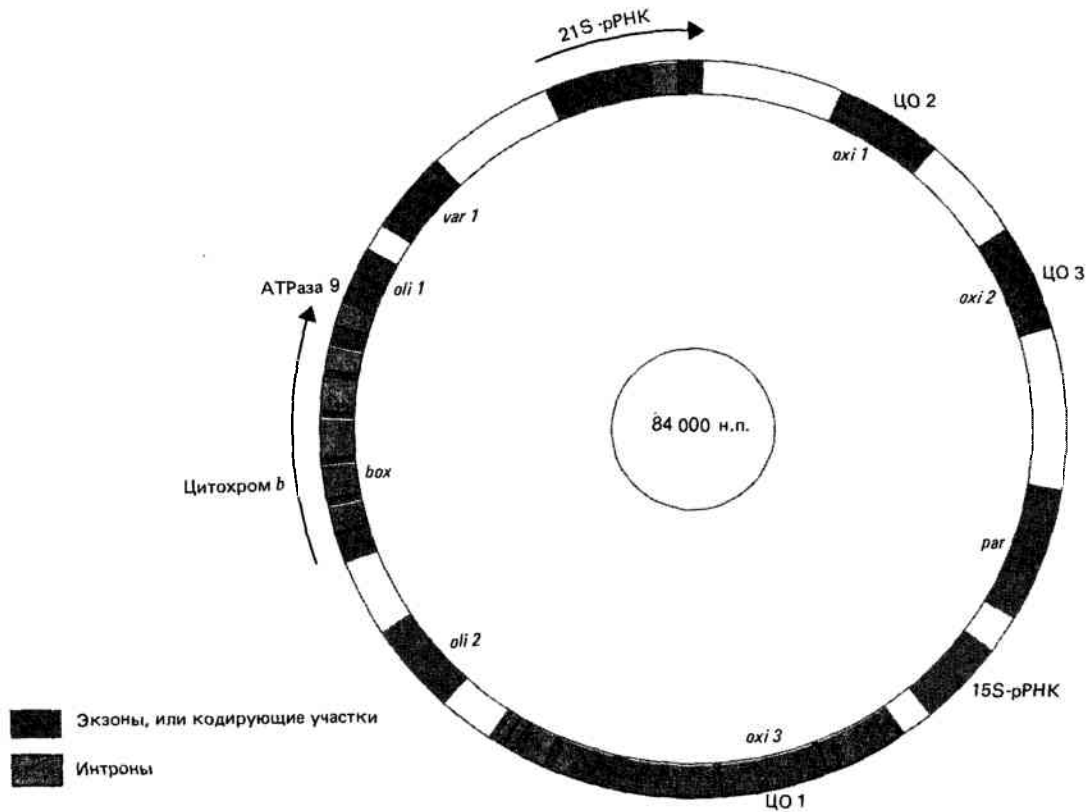


Рис. 22.2. В состав генома митохондрий дрожжей входят имеющие прерывистое и непрерывное строение гены, кодирующие белки, гены, кодирующие рРНК, и гены, кодирующие тРНК (положение тРНК-генов не указано). Стрелки показывают направление транскрипции.

oli-гены, кодирующие субъединицы олигомицин-чувствительной АТФазы; *oxi*-гены, кодирующие субъединицы цитохром-с-оксидазы; *box*-кодирует цитохром *b*; *par*-функции не известны; *var*-кодирует малую субъединицу рибосомного белка

риал в митохондриях дрожжевых клеток другие белки, возможно участвующие в регуляторных процессах, или же он не экспрессируется?

Существование генетической рекомбинации в митохондриях дрожжей (см. ниже) позволило создать генетическую карту генома митохондрий, которая была сопоставлена с физической картой при изучении возникающих естественным путем делеций митохондриального генома (называемых **петит-мутациями**). Участок, занимаемый каждым геном, определяли по положению на карте соответствующей мРНК.

На рис. 22.2 приведена карта митохондриального генома дрожжей, показывающая взаимное расположение генов для основной части РНК и белков (за исключением молекул тРНК, число видов которых составляет около 22 и гены для которых до сих пор полностью не картированы). Наиболее характерная особенность этой карты — пространственная разделенность локусов.

Отличительная особенность митохондриального генома дрожжей состоит в пространственной разобщенности генов, кодирующих рРНК, что встречается крайне редко. Ген, кодирующий 15S-рРНК, непрерывен и располагается на расстоянии около 25 000 п.н. от гена для 21S-рРНК. В некоторых штаммах дрожжей последний ген имеет один интрон (как показано на карте), в других штаммах он не прерывается.

Два наиболее интересных локуса — мозаичные гены *box* (кодирующий цитохром *b*) и *oxi3* (кодирующий субъединицу 1 цитохром-оксидазы), организация которых уже была рассмотрена в гл. 20. Суммарная длина двух этих

генов равна почти всей длине митохондриального генома млекопитающих! Многие длинные интроны этих генов имеют открытые рамки считывания, совпадающие с рамками считывания предыдущих экзонов; по меньшей мере в некоторых случаях показано, что интрон транслируется. Это обуславливает наличие в митохондриях дрожжей нескольких дополнительных белков, по-видимому синтезируемых в небольших количествах.

Остальные гены, по всей вероятности, непрерывны. Им соответствуют две другие кодируемые митохондриальными генами субъединицы цитохром-оксидазы, субъединица (субъединицы) АТФазы и (в случае *var1*) рибосомный белок митохондрий.

Весьма значительная часть митохондриального генома дрожжей (около 25%) состоит из коротких (А—Т)-богатых участков ДНК, вероятно не несущих функции кодирования. Однако значительная часть генетического материала все еще не изучена, и будет удивительно, если бы оставшихся незаполненными участков карты не будут обнаружены другие гены. Но, даже допуская такую возможность, можно все же считать, что общее число генов в митохондриях дрожжей вряд ли превышает 20.

Компактная организация генома митохондрий млекопитающих

Совершенно другая картина наблюдается в случае митохондриальной ДНК млекопитающих. Эта ДНК чрезвычайно компактно организована, интронов в ней нет, некоторые гены перекрываются, и почти каждая пара

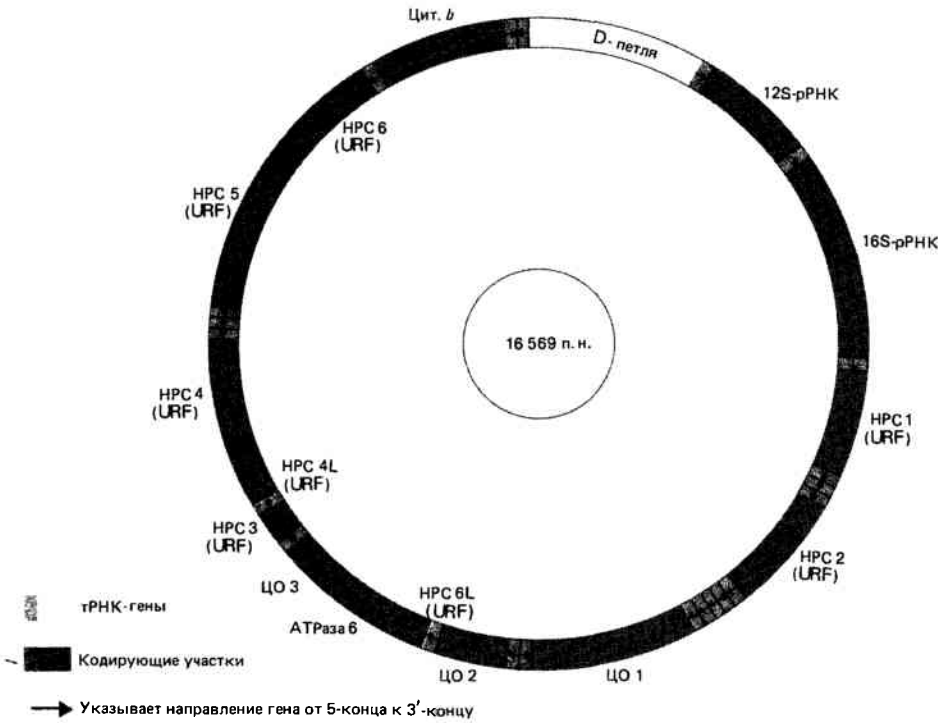


Рис. 22.3. В нуклеотидной последовательности митохондриальной ДНК человека обнаруживаются 22 гена тРНК, 2 гена рРНК, 13 участков, возможно, кодирующих белки (некоторые из них соответствуют известным белкам). ЦО – цитохромоксидаза.

оснований принадлежит какому-либо гену. За исключением D-петли, области, отвечающей за инициацию репликации ДНК, не более 87 из 16 569 п. н. митохондриального генома человека можно отнести к междистронным участкам.

Обнаружена значительная гомология в строении полных нуклеотидных последовательностей митохондриальных геномов человека и мыши. Карта митохондриального генома человека приведена на рис. 22.3. На ней имеется 13 областей, которые потенциально могут кодировать белки. К ним относятся области, кодирующие цитохром *b*, три обычные субъединицы цитохром-оксидазы и одну из субъединиц АТРазы. Остальные рамки считывания обозначены как URF [от англ. *unidentified reading frames* – неидентифицированные рамки считывания (HPC)]; функции кодируемых ими белков пока не известны.

Организация генов имеет ряд поразительных черт, отражающих особенности экспрессии генов в митохондриях млекопитающих. Большинство генов экспрессируется в одном и том же направлении; гены для тРНК располагаются между генами, кодирующими рРНК или белок.

Во многих случаях гены вовсе не отделены друг от друга. Последнее основание одного гена соседствует с первым основанием следующего за ним гена. В некоторых случаях гены перекрываются, обычно в пределах одного основания, так что последнее основание одного гена оказывается первым основанием следующего за ним гена. В пяти рамках считывания отсутствует терминирующий кодон; они оканчиваются на U или UA, и поэтому при полиаденилировании транскрипта образуется охрана-кодон (к 3'-концу митохондриальной мРНК присоединяется короткая poly(A)-последовательность). В трех случаях терминирующие кодоны – это AGA или AGG, обычно соответствующие аргинину (гл. 4).

Почти все гены экспрессируются в направлении по часовой стрелке, как показано стрелками на карте. Из 15 областей, кодирующих белок или рРНК, 14 транскрибируются в одном и том же направлении. (Исключительны-

ми свойствами обладают три URF-гена. URF4L и URFA6L необычайно сильно перекрываются с соседними генами. URF6 отличается от основной части генов только направлением его экспрессии.) Из 22 генов тРНК 14 экспрессируются в направлении по часовой стрелке и 8 – против часовой стрелки.

Был обнаружен только один случай, когда два гена, экспрессирующиеся в направлении по часовой стрелке, оказались расположенными рядом (гены для АТРазы 6 и ЦО3). Во всех других случаях соседние кодирующие участки разделены по меньшей мере одним геном тРНК. Были обнаружены мРНК, соответствующие всем белок-кодирующим участкам, и во всех случаях инициаторный кодон располагался на расстоянии в пределах шести оснований от начала молекулы. Иницирующими кодонами, по-видимому, могут являться кодоны AUG, AUA или AUU. Последние нуклеотиды мРНК – это нонсенс-кодоны или неполные нонсенс-кодоны. Поэтому у митохондриальной мРНК человека фактически отсутствуют не-транслируемые 5'- и 3'-участки.

Прерывание участков, кодирующих рРНК и белки, генами тРНК не оставляет места для промоторов, подобных обнаруживаемым в эукариотическом ядре или бактериях. По-видимому, для всех участков, транскрибируемых в направлении по часовой стрелке, имеется общий промотор, расположенный в области D-петли. На основе этого создана модель транскрипции митохондриальной ДНК, приведенная на рис. 22.4. Транскрипция начинается непосредственно перед геном тРНК, расположенным перед геном 12S-рРНК, проходит почти по всему кругу и терминируется в D-петле. Образующаяся при этом цепь называется **Н-цепью** (поскольку митохондриальную ДНК можно разделить на две одинаковые цепи – тяжелую (*heavy*) и легкую (*light*), различающиеся по своей плавучей плотности).

Значение чередования тРНК-генов с генами, кодирующими рРНК и белки, заключается в том, что гены тРНК обозначают сайты расщепления. При расщеплении первичного продукта транскрипции по сайтам, располо-

женным по обе стороны каждого гена тРНК, на всех генах кроме генов АТФазы 6 и цитохромоксидазы 3 (ЦОЗ), синтезируются моноцистронные продукты. Молекулы рРНК, по-видимому, синтезируются в больших количествах, чем молекулы мРНК. Это может быть вызвано преждевременной терминацией синтеза части транскриптов в некоторой точке, после того как транскрибировались два гена рРНК.

Таким образом, митохондриальная ДНК млекопитающих представляет собой ближайший аналог бактериального оперона, обнаруженный у эукариот. Ее транскрипция осуществляется при участии одного промотора, а индивидуальные тРНК и мРНК выделяются уже из транскрипта. Процессинг транскрипта поэтому является основным этапом экспрессии генов. (Имеются также и бактериальные опероны, в которых разделение участков, соответствующих тРНК, рРНК и белкам, осуществляется посттранскрипционно; см. гл. 23.)

Как же экспрессируются гены тРНК, кодируемые L-цепью? По-видимому, в этом случае действует сходный механизм, когда на L-цепи синтезируется гигантский транскрипт, от которого остаются только некоторые последовательности тРНК и, возможно, некоторые последовательности мРНК, тогда как все остальные разрушаются. Точка начала транскрипции L-цепи (происходящей против часовой стрелки) также располагается в области D-петли. Выделяемые транскрипты L-цепи могут достигать очень больших размеров, так что их процессинг происходит, по-видимому, некоторое время спустя после транскрипции. Больших транскриптов H-цепи не обнаружено, поэтому можно предположить, что их процессинг происходит сразу же или почти сразу после завершения транскрипции каждой промежуточной тРНК.

ДНК некоторых органелл участвует в процессе рекомбинации

Предпосылкой для рекомбинации служит одновременное присутствие геномов двух родителей – условие, не выполняющееся при преобладании однородительского наследования генетического материала. Но, поскольку существуют примеры, когда потомству передаются геномы обеих родительских органелл, сразу возникает вопрос: как они могут взаимодействовать между собой? Возможна ли комплементация различных мутаций; может ли происходить рекомбинация между геномами? Для комплементации необходимо перемешивание продуктов экспрессии генов; для рекомбинации требуется расположение двух геномов физически рядом друг с другом. Оба требования несколько противоречат устоявшемуся взгляду на индивидуальную органеллу, который мог возникнуть на основе представления об эндогенной экспрессии генов органеллы как о процессе, строго ограниченном ее пределами. Однако (по меньшей мере у одного из видов организмов) может происходить рекомбинация хлоропластных ДНК; а у другого вида между митохондриальными геномами может происходить и комплементация, и рекомбинация.

У водорослей *C. reinhardtii* обычно наследуется хлоропластная ДНК одного из родителей. Однако при необычных условиях, когда гены одного из родителей, обычно доминирующего, облучают ультрафиолетом, хлоропластные гены другой родительской особи также наследуются. В образующихся в этом случае зиготах, включающих генотипы обоих родителей, между хлоро-

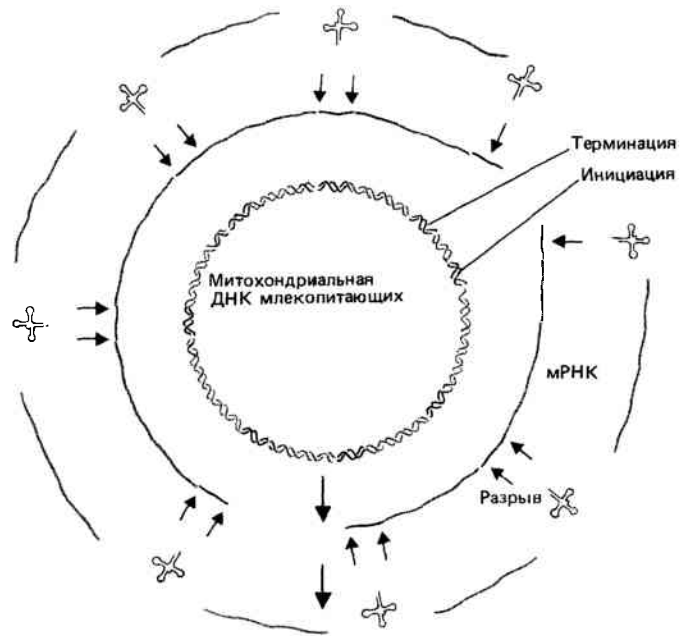


Рис. 22.4. При транскрипции митохондриальной ДНК человека образуется единый транскрипт, из которого вырезаются молекулы тРНК и освобождаются молекулы рРНК и мРНК.

пластными геномами происходит рекомбинация, и ее результаты могут быть использованы для построения генетической карты.

У дрожжей *S. cerevisiae* рекомбинация между митохондриальными маркерными генами – обычное явление. Природа такой рекомбинации не известна, но кроссинговер между двумя родительскими митохондриальными ДНК, имеющими значительные различия, приводит к появлению потомства, обладающего рекомбинантными ДНК. Этот факт непосредственно свидетельствует о том, что митохондриальные ДНК обеих родительских особей вступают в контакт.

Каким образом ДНК родительских органелл вступают в контакт друг с другом? Наиболее правдоподобное объяснение состоит в том, что органеллы могут сливаться, образуя единую структуру, содержащую оба родительских генома. Эта точка зрения получила дальнейшее развитие в случае дрожжей, где, возможно, вместо множества отдельных митохондрий имеется большая разветвленная структура (или структуры), соответствующая многим митохондриям. Возможно, она находится в состоянии непрерывного изменения, сопровождающегося образованием и разрушением мембран. Это увеличивает вероятность того, что геном связан с некоторой структурной единицей меньшего размера, чем сама органелла; число геномов, приходящихся на органеллу, тогда определялось бы числом таких субъединиц.

У дрожжей проходит несколько часов от момента образования зиготы до осуществления рекомбинации. Однако во время этого промежуточного периода можно обнаружить комплементацию между различными мутациями. Это предоставляет уникальную возможность определения генетической организации генома органеллы, поскольку мутации могут быть картированы и в то же время отнесены к определенным группам комплементации. (Аналогичный подход также оказался успешным в случае хлоропластов *C. reinhardtii*.)

Перестройки митохондриальной ДНК дрожжей

Митохондриальный геном дрожжей имеет удивительно непостоянное строение. Мы уже упоминали о том, что у разных штаммов *S. cerevisiae* обнаружены различные формы некоторых генов. Общий размер митохондриального генома у разных штаммов может значительно различаться (на величину около 10 000 п.н.). Значительная часть митохондриальной ДНК может утрачиваться в результате делеции, как в случае **петит** (*petite*)-мутантов *S. cerevisiae* (и некоторых других штаммов дрожжей). Делеции могут возникать случайно или могут быть индуцированы определенными воздействиями, такими, как обработка бромистым этидием.

Все петит-мутации нарушают функционирование митохондрий. Они не становятся летальными, поскольку дрожжи могут существовать и в аэробных условиях (когда необходимо дыхание) и в анаэробных (когда без него можно обойтись). Следовательно, мутации митохондриального генома являются условно-летальными, вызывающими гибель дрожжей лишь в аэробных условиях; в анаэробных же условиях они выживают. Таким образом, нарушение функций митохондрий приводит к переходу на анаэробный способ существования. Несомненно, такой переход невозможен в случае, например, клеток животных, для которых прекращение функционирования митохондрий оказывается летальным. (Аналогичная генетическая ситуация наблюдается в случае хлоропластов *S. reinhardtii*, которые не нуждаются в фотосинтезе в присутствии ацетата.)

Петит-мутации *S. cerevisiae* распадаются на три типа. **Ядерные петиты**—это менделевские (т.е. ядерные) мутации, полностью нарушающие деятельность митохондрий. **Нейтральные петиты** (ρ^0) представляют собой крайний случай, когда отсутствует вся митохондриальная ДНК. Это рецессивный генотип. В митохондриальной ДНК **супрессируемых петитов** (ρ^-) имеются значительные аномалии. Митохондрии ρ^- содержат кольцевые молекулы ДНК, размеры которых значительно меньше размеров обычного генома, а сложность составляет только небольшую часть ее обычной величины, находясь в пределах от 0,2 до 36%. Это означает, что большая часть митохондриального генома делетирована.

Последовательности, сохраняющиеся в случае петит-мутаций ρ^- , часто амплифицируются с образованием большого числа копий. Амплификация осуществляется в результате либо увеличения ploидности (числа молекул ДНК), либо образования мультимерных молекул ДНК, каждая из которых состоит из большого числа копий определенной последовательности. Сохраняющаяся последовательность не обязательно соответствует исходно непрерывному участку генома и может состоять из соединенных вместе последовательностей его разных областей. Иногда могут происходить дальнейшие перестройки этой последовательности (в этом случае петит нестабилен).

ДНК, сохранившаяся в петитном штамме, может рекомбинировать с ДНК других петит-мутантов или штаммов дикого типа (называемых по контрасту *гранд-штаммами*). Используя рекомбинацию, можно определить положение генетических маркеров, присутствующих

в каждом петит-штамме, в сохранившейся последовательности ДНК. Фактически именно использование петит-мутаций для картирования делеций позволило построить карту митохондриального генома, показанную на рис. 22.2.

У петит-мутанта может сохраниться *любая* последовательность митохондриальной ДНК. Это означает, что все участки ДНК могут реплицироваться независимо. Мы не знаем, какой механизм обеспечивает такую способность, и связана ли она с наличием структуры, аналогичной точке начала репликации (гл. 31).

При скрещивании петит-штамма со штаммом дикого типа у определенной части потомства обнаруживается только петит-генотип. Митохондриальная ДНК штамма дикого типа исчезает. Это явление называется **супрессией**, а доля потомства, в которой она осуществляется, характеризует степень **супрессируемости**.

Причина возникновения супрессии у петит-мутантов с высокой степенью супрессируемости состоит в преимущественной репликации петитной митохондриальной ДНК. Это, по-видимому, происходит потому, что в данной ДНК имеется большая концентрация последовательностей, аналогичных последовательностям участков начала репликации, и называемых областями *rep*, по сравнению с последовательностями ДНК дикого типа. Таким образом, мы видим, что дрожжевая митохондриальная ДНК содержит специфическую последовательность (специфические последовательности), которая может быть использована для инициирования репликации. Но в отличие от ее аналогов, присутствующих в других геномах, эта последовательность не является обязательной для выживания, поскольку петитная ДНК, в которой она отсутствует, может сохраняться (хотя она супрессируется в меньшей степени). Поэтому в случае петит-мутантов с меньшей степенью супрессируемости причина супрессии не обязательно совпадает с причиной ее возникновения у высоко-супрессируемых штаммов.

Общий вывод, который можно сделать на основе изучения митохондриальной ДНК петит-штаммов и штаммов дикого типа,—это значительная гибкость организации генома в отношении процессов его экспрессии и репликации. Возможно, что ее отличие от других митохондриальных ДНК обуславливает способность дрожжей существовать при полном отсутствии функционирования митохондрий. Было бы интересно выяснить, какие ферментные системы работают на этой ДНК.

Рекомендуемая литература

Общий обзор, касающийся передачи генетической информации хлоропластов и митохондрий, написан Джиллхэмом (*Gillham, Organelle Heredity*, Raven Press, 1978). Вопросы, рассматриваемые в этой главе, освещены в книге Льюина (*Lewine, Gene expression*, 2, Eucaryotic Chromosomes Wiley, New York, 1980, 583–637). Нуклеотидная последовательность митохондриального генома человека была определена Андерсоном и др. (*Anderson et al., Nature*, 290, 457–465, 1981). Транскрипции митохондриальной ДНК млекопитающих посвящен обзор Клейтона (*Clayton, Ann. Rev. Biochem.* 53, 573–594, 1984).

Глава 23

СХОДСТВО И РАЗЛИЧИЯ
КЛАСТЕРОВ ТАНДЕМНЫХ ГЕНОВ

Рассматривая геном с точки зрения индивидуальных генов, можно обнаружить самые разные варианты его организации. Ген может занимать такое положение, где его нуклеотидная последовательность будет единственной в своем роде, хотя в других участках генома могут находиться сходные последовательности. Он может входить в состав небольшого кластера генов, происшедших от общего гена-предка и выполняющих сходные функции (как в случае систем глобиновых генов). Другие кластеры могут иметь большие размеры и состоять из многократно повторяющихся генов, идентичных или обладающих очень большим сходством. Про гены (или другие последовательности), которые обнаруживаются в виде множества последовательно расположенных копий, говорят, что они **тандемно** повторяются. Один из видов кластеров тандемно повторяющихся генов кодирует гистоны, другой — рибосомные РНК. В обоих случаях многократная повторяемость генов, по-видимому, свидетельствует о потребности в больших количествах кодируемого ими продукта.

Компактное группирование генов в виде тандема в определенном участке генома, единственное назначение которого — кодирование определенной функции или определенных функций, определяет способ описания организации таких последовательностей. Каждый тандемный кластер может быть описан как **повторяющаяся единица**. Это основной структурный элемент, многочисленные повторы которого, расположенные конец в конец, и формируют кластер. Копии повторяющейся единицы могут быть идентичными или различаться.

Обычно такая повторяющаяся единица имеет большую длину, чем транскрибирующаяся на ней последовательность; следовательно, ее можно подразделить на две области. **Единица транскрипции** — это участок между первым и последним основаниями, транскрибирующимися с образованием РНК. **Нетранскрибирующийся спейсер** составляет остальную часть повторяющейся единицы и занимает область, расположенную между соседними транскрипционными единицами. Такой способ организации генетического материала показан на **рис. 23.1**.

Рестрикционная карта кластера тандемных генов обладает характерной особенностью. Она имеет вид кольца. Когда повторяющиеся единицы идентичны, расположение сайтов рестрикции в них совершенно одинаково. Соседние повторы имеют общие концевые фрагменты. В приведенном на рисунке примере фрагмент *A* находится по соседству с фрагментом *B*, соседствующим с фрагментом *C*, который в свою очередь находится рядом с фрагментом *A*; таким образом и получается кольцевая карта.

Такое чередование последовательностей нарушается только на двух концах кластера в целом, обозначенных на рисунке как фрагменты *X* и *Y*. Однако обычно эти фрагменты не обнаруживаются, поскольку каждый из них является уникальным в отличие от множества повторяющихся внутренних фрагментов. К примеру, при наличии 100 копий повторяющейся единицы фрагменты *X* и *Y* со-

ставят только 1% от материала, представленного фрагментами *A*, *B* и *C*.

Нетранскрибирующийся спейсер обычно короче, чем единица транскрипции. Его можно обнаружить по четкости расположения участков внутри кластера тандемных генов. Спейсер — это просто часть повторяющейся единицы, которая не экспрессируется с образованием РНК, но функции которой (если таковые имеются) остаются связанными с единственной функцией кластера: синтезом РНК, кодируемой единицей транскрипции.

Такая точка зрения может повлечь за собой следующий вопрос: следует ли рассматривать в таком же свете несколько более длинные последовательности, находящиеся между индивидуальными генами, например в кластере глобиновых генов? Однако между этими двумя случаями имеется важное различие. Все межгенные промежутки в кластере генов глубина различаются, тогда как нетранскрибирующиеся спейсеры кластера тандемных генов повторяются вместе с генами.

Гены гистонов образуют повторы

Гистоны — основные структурные белки ядра. В хромосомах всех (по-видимому) эукариот имеется пять типов гистонов: гистон **H1**, составляющий в молярном выражении половину каждого гистона, и гистоны **H2A**, **H2B**, **H3** и **H4**, присутствующие в эквимольных количествах (гл. 29). Суммарная масса гистонов примерно равна массе самой ДНК. В делящихся соматических клетках синтез гистонов происходит одновременно с репликацией ДНК; сразу же после окончания репликации цепи ДНК гистоны уже могут связываться с ней. Это означает, что за короткий период синтеза ДНК должно синтезироваться большое количество гистонов. Необходимость синтеза соматическими клетками всей массы гистонов за относительно короткое время, по-видимому, является основной причиной повторяемости гистоновых генов.

Между размерами генома и числом гистоновых генов нет определенной зависимости. Следовательно, можно думать, что у разных видов гены гистонов должны экспрессироваться с разной эффективностью. Однако повторяемость гистоновых генов — общее свойство, существенное для образования гистонов. Обычно в геноме имеется одинаковое число копий каждого гистонного гена. В геноме цыпленка частота их повторяемости равна примерно 10, у млекопитающих — примерно 20. Эта величина возрастает примерно до 40 у *X. laevis* и примерно до 100 у *D. melanogaster*. У нескольких видов морских ежей каждый гистоновый ген имеет 300–600 копий.

У этих видов морских ежей первые деления ядер в эмбриогенезе происходят очень быстро. Большое число гистоновых генов у них, по-видимому, необходимо для экспрессии гистоновых генов *de novo*, чтобы синтез гистонов происходил с той же скоростью, что и синтез ДНК. У амфибий экспрессия гистоновых генов предшествует

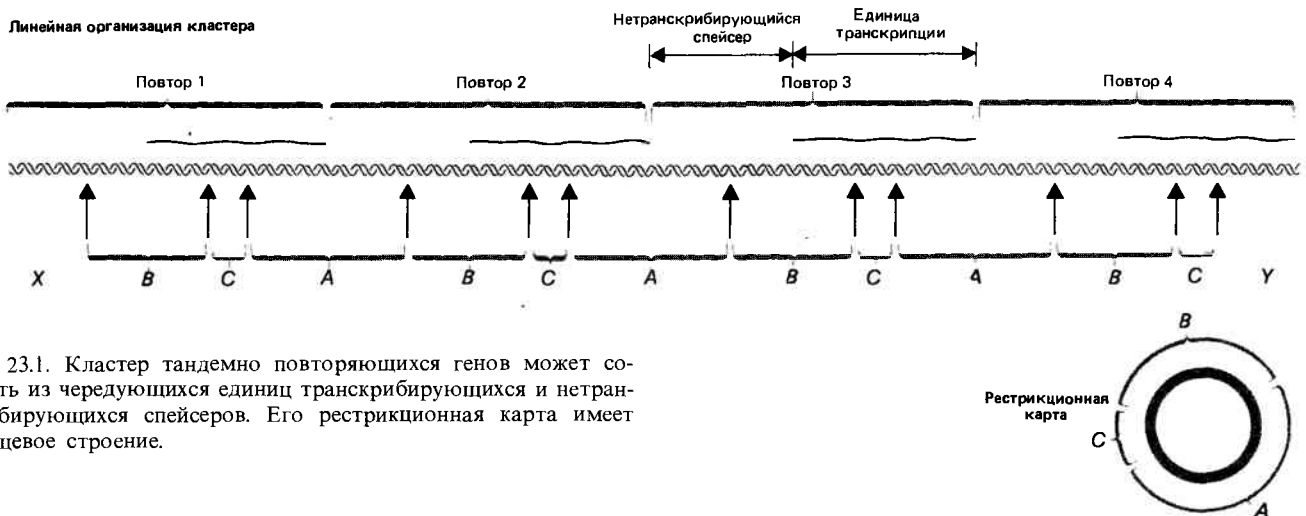


Рис. 23.1. Кластер тандемно повторяющихся генов может состоять из чередующихся единиц транскрибирующихся и нетранскрибирующихся спейсеров. Его рестрикционная карта имеет кольцевое строение.

началу эмбриогенеза и в ооците накапливается большое количество гистоновой мРНК. Могут также накапливаться и некоторые гистоны. Это исключение из того правила, что экспрессия гистоновых генов происходит только во время фазы синтеза ДНК.

На основе имеющихся в настоящее время данных об организации генов гистонов у охарактеризованных в этом плане видов организмов можно увидеть разницу в способе организации этих генов в геномах с их низкой повторяемостью (менее 50) и геномах с высокой повторяемостью (более 100). В случае низкой повторяемости генов гистонов встречаются различные варианты их организации. В случае высокой повторяемости они организованы по более жесткому принципу, когда все гистоновые гены расположены по единому типу, в виде кластера тандемных повторов. Во всех этих случаях гены гистонов имеют одинаковую структуру; в настоящее время общим правилом представляется отсутствие интронов у всех функционально активных генов гистонов.

Разнообразие кластеров тандемных генов гистона

Первыми охарактеризованными кластерами гистоновых генов были кластеры морского ежа. Три вида морских ежей (не являющихся близкими родственниками) — *P. miliaris*, *S. purpuratus* и *L. pictus* — имеют одинаковый тип организации.

В каждом случае имеются два обширных класса гистоновых генов. Ранние гены кодируют гистоны, интенсивно синтезирующиеся на ранних стадиях эмбриогенеза; к ним относится большинство гистоновых генов (более 300 копий на геном). Поздние гены кодируют несколько иные варианты гистонов и экспрессируются на более поздних стадиях эмбрионального развития; на геном приходится около 10 копий каждого гена.

Все пять ранних гистоновых генов входят в состав одной и той же единицы, в результате повторения которой в геноме образуется один кластер тандемных генов. Каждый ген отделен от следующего за ним гена нетранскрибируемым спейсером; все пять спейсеров различаются. Многочисленные копии повторяющейся единицы практически идентичны. Все они транскрибируются в одном и том же направлении, хотя на каждом из них независимо синтезируется своя собственная мРНК. В общих чертах организация кластеров гистоновых генов

у морских ежей приведена на рис. 23.2, откуда видно, что основные различия между видами заключаются в нетранскрибирующихся спейсерах. Спейсеры различаются не только по своей длине, но и по нуклеотидным последовательностям. Между соответствующими друг другу спейсерами различных видов отсутствует перекрестная гибридизация, хотя гены могут обладать значительным сходством.

Это парадоксальная ситуация. Предположим, что эти кластеры эволюционировали путем дупликации всей повторяющейся единицы, общая организация которой позволяет предположить, что она существовала еще до того, как произошло разделение морских ежей на виды. Естественный отбор, по-видимому, действовал в направлении сохранения функций этих генов, в то время как спейсеры разных видов дивергировали, сохраняя, однако, свое постоянство *внутри* каждого вида. Из этого, очевидно, следует, что на кластер генов каждого вида морских ежей должен воздействовать некий механизм коррекции.

В действительности же каждый из пяти спейсеров не остается абсолютно неизменным. Выявляются некоторые различия в длинах соответствующих спейсеров отдельных повторяющихся единиц, о чем свидетельствует обнаруженная гетерогенность (небольшая) размеров соответствующих фрагментов рестрикции. Такие различия между повторяющимися единицами называют **микрөгетерогенностью**.

Имеются также и различия между нуклеотидными последовательностями генов. У морских ежей синтезируются разные наборы гистонов в раннем эмбриогенезе (во время фазы быстрого деления ядра) и на более поздних стадиях развития.

У одного вида морских ежей были охарактеризованы два типа повторяющихся единиц, соответствующих ранним эмбриональным генам. Основное различие этих единиц заключается в различиях их спейсеров: они тандемно повторяются и при этом не перемежаются между собой. Эти данные указывают на то, что, какие бы механизмы ни были ответственны за поддержание гомогенности, они должны воздействовать на внутренние области кластера, и их действие не должно распространяться по всей его длине.

Поздние гены гистонов организованы иначе. Эти гены можно обнаружить в отдельных фрагментах ДНК; иногда встречаются пары тесно сцепленных генов (находящихся на расстоянии > 10 т. п. н. друг от друга). Постоянной формой организации таких генов не обнаружено.

Принцип организации всех пяти гистоновых генов в виде целой повторяющейся единицы характерен и для *D. melanogaster*, хотя в этом случае повторяющаяся единица имеет иное строение. Гены располагаются в другом порядке и транскрибируются в разных направлениях. Для преобразования кластеров морского ежа в кластер плодовой мушки потребовались бы две транслокации. В кластере *D. melanogaster* также имеются различные варианты спейсеров.

Необычная форма организации кластера гистоновых генов обнаружена у тритона *N. viridescens*, у которого имеется повторяющаяся единица единственного типа (отличающаяся от повторяющейся единицы и морского ежа, и плодовой мушки). Однако отдельные повторяющиеся единицы располагаются не рядом друг с другом, а разделены участками высокоповторяющейся ДНК длиной до 10–50 т.п.н.

В некоторых случаях обнаружена менее жесткая форма организации гистоновых генов. Гистоновые гены у *X. laevis* объединены в кластер, но их организация имеет гетерогенный характер, поскольку гены одного типа могут иметь различных соседей. При наличии тандемных повторов, по-видимому, имеется более одного класса повторов с разным порядком расположения генов. У цыплят имеется кластер гистоновых генов, но порядок расположения генов варьирует, а тандемные повторы отсутствуют. Переходя к млекопитающим, мы снова обнаруживаем, что гены могут не иметь упорядоченной организации в виде повторяющихся единиц, но могут располагаться небольшими группами или даже в виде индивидуальных генов. Несмотря на то что необходимо иметь намного больше данных, чтобы оценить значение таких способов организации гистоновых генов, в настоящее время их можно считать промежуточными вариантами между небольшим кластером генов (таким, как глобиновый) и кластером тандемно повторяющихся генов.

Между структурой отдельных гистоновых генов и их организацией в целом имеется интересное отличие. Все гистоновые гены имеют одинаковую общую непрерывную структуру, и соответствующие друг другу гены у разных видов организмов кодируют белки, обладающие большим сходством. Однако имеется значительное разнообразие в способах организации различных классов генов, которые варьируют от образования жестких тандемных кластеров до беспорядочного расположения генов. Все это служит подтверждением общих выводов, сделанных на основе примера кластера глобиновых генов. Индивидуальные гены могут оказаться приспособленными для выполнения своей функции в результате специфических мутаций в сохранившейся в других отношениях последовательности, но при этом происходит постоянное образование новых копий генов с последующей реорганизацией всего кластера.

рРНК и тРНК кодируются повторяющимися генами

Каждый вид РНК, входящий в состав аппарата белкового синтеза, кодируется большим числом генов. В некотором смысле структурные гены, кодирующие мРНК, составляют самый большой класс генов, хотя, поскольку имеется так много различных видов мРНК, каждый ген представлен только относительно небольшим числом молекул мРНК. Чтобы обеспечить трансляцию большого числа молекул мРНК, клетка содержит множество рибосом и молекул тРНК.

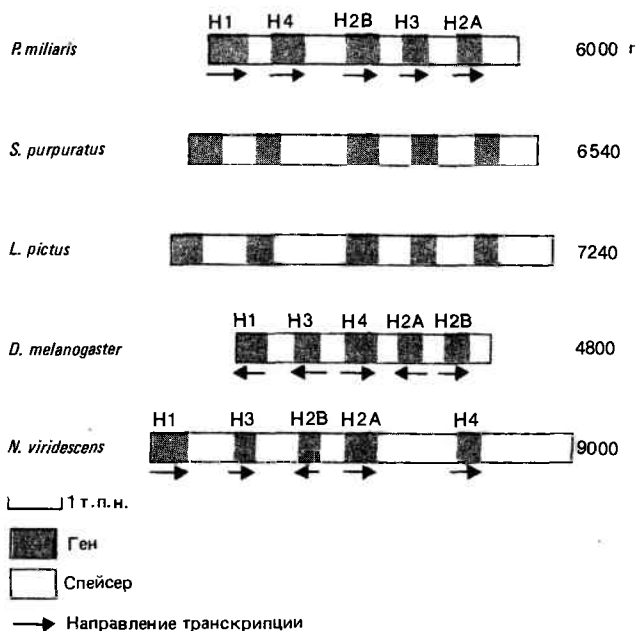


Рис. 23.2. Повторяющиеся единицы кластеров гистоновых генов организованы по-разному.

Рибосомная РНК, несомненно, является основным продуктом транскрипции, составляя примерно 80–90% всей РНК как эукариотической, так и прокариотической клетки. Такое высокое содержание рРНК обусловлено ее размером, поскольку на одну рибосому приходится несколько молекул тРНК. И рРНК, и тРНК кодируются множественными генами; в табл. 23.1 приведены данные о числе генов, приходящихся на гаплоидный геном, для наиболее изученных видов организмов.

Число генов, кодирующих высокомолекулярные рРНК, колеблется от 7 у *E. coli* до 100–200 у низших эукариот и до нескольких сотен у высших эукариот. Практически во всех случаях гены, кодирующие высоко- и низкомолекулярные рРНК, образуют тандемные пары. (Единственное исключение составляют рибосомные гены митохондрий дрожжей.) У бактерий и некоторых низших эукариот гены, кодирующие 5S-РНК, входят в состав той же транскрипционной единицы; поэтому общее число 5S-генов равно числу генов, кодирующих высокомолекулярные рРНК. У бактерий 5S-гены транскрибируются вместе с генами для высокомолекулярных рРНК;

Таблица 23.1
Во всех геномах содержится большое число генов, кодирующих рРНК и тРНК

Вид	Число генов, кодирующих 18S- и 28S-рРНК	Доля рДНК от всей ДНК клетки, %	Число генов, кодирующих 5S-рРНК	Число генов, кодирующих тРНК
<i>E. coli</i>	7	1,0	7	60
<i>S. cerevisiae</i>	140	5,5	140	250
<i>D. discoideum</i>	180	17,0	180	?
<i>D. melanogaster</i>				
(X)	250	1,3	165	850
(Y)	150			
Человек	280	0,4	2 000	1300
<i>X. laevis</i>	450	0,18	24 000	1150



Рис. 23.3. Одна единица транскрипции включает последовательности, кодирующие синтез большой и малой молекул рРНК.

у эукариот они транскрибируются независимо. У высших эукариот гены, кодирующие 5S-РНК, организованы в самостоятельные кластеры, и их число превышает число генов для высокомолекулярных рРНК.

Отсутствие сколько-нибудь заметных различий в нуклеотидных последовательностях молекул рРНК одного вида указывает на то, что все копии каждого гена должны быть идентичными или по крайней мере число различающихся копий настолько мало (менее 1%), что не обнаруживается при анализе. У бактерий многочисленны пары генов 16S–23S-рРНК рассеяны по геному. У эукариот ядерные гены для рРНК, как правило, организованы в виде tandemного кластера или просто кластеров. Такие области иногда называют рДНК. (В некоторых случаях доля рДНК в общем содержании ДНК достаточно велика, что в сочетании с ее необычным нуклеотидным составом дает возможность выделять ее в виде отдельной фракции непосредственно из фрагментированной ДНК генома.)

В случае тРНК точное число кодирующих ее генов определить трудно из-за того, что сильно выраженная вторичная структура молекулы делает проведение реакции гибридизации технически трудным. Поэтому имеющиеся оценки числа генов, по-видимому, ниже, чем действительное число генов.

Тандемно повторяющаяся единица включает оба гена рРНК

Две высокомолекулярные рРНК транскрибируются в виде единого предшественника и в бактериальной клетке, и в ядре эукариот. После завершения транскрипции предшественник расщепляется с образованием отдельных молекул рРНК; этот процесс более подробно обсуждается в гл. 25. Наиболее характерные особенности организации соответствующих генов показаны на рис. 23.3. Транскрипция начинается с 5'-лидерной последовательности, далее транскрибируется последовательность, кодирующая малую рРНК, затем область, называемая **транскрибирующимся спейсером**, и, наконец, у 3'-конца транскрипта располагается большая рРНК.

Таблица 23.2

Предшественник рРНК длиннее двух зрелых рРНК

Вид	Длина транскрипта	Длина рРНК		Относительное содержание зрелых рРНК в первичном транскрипте, %
		большая	меньшая	
<i>E. coli</i>	5 600	2904	1542	80
<i>S. cerevisiae</i>	7 200	3750	2000	80
<i>D. melanogaster</i>	7 750	4100	2000	78
<i>X. laevis</i>	7 875	4475	1925	79
<i>N. tabacum</i>	7 900	3700	1900	71
<i>G. domesticus</i>	11 250	4626	1800	57
<i>M. musculus</i>	13 400	5000	1950	52

Таким образом, транскрипционная единица имеет большую длину, чем общая длина зрелых рРНК. В табл. 23.2 приведены данные о соотношениях между первичными продуктами транскрипции и зрелыми рРНК у некоторых видов организмов.

Наименьшую длину имеет транскрипционная единица бактерий, у которых последовательности рРНК составляют 80% при общей длине транскрипционной единицы 6000 п. н. У эукариот, где длина транскрипта варьирует от 7000 до 8000 п. н. и 70–80% последовательностей представлено в рРНК, какой бы то ни было систематичности для разных видов не обнаружено. Самый длинный предшественник рРНК найден у млекопитающих (где по результатам определения скорости седиментации это — 45S-РНК). В данном случае последовательности зрелой рРНК составляют только несколько более 50% от всей длины первичного транскрипта, так что лидерная последовательность и транскрибирующийся спейсер здесь должны иметь большие размеры.

Что происходит с нерибосомными участками РНК-предшественника? Лидер и транскрибирующийся спейсер отбрасываются при созревании рРНК и, по-видимому, затем распадаются до нуклеотидов, которые в дальнейшем повторно используются. Однако в состав транскрибирующегося спейсера входят последовательности, кодирующие низкомолекулярные РНК, освобождающиеся при его расщеплении. У млекопитающих и у *Xenopus* имеется короткая последовательность, соответствующая 5,8S-РНК, молекуле небольшого размера, которая при помощи водородных связей связывается с 28S-рРНК в рибосоме. У бактерий, как мы увидим позже, транскрибирующимся спейсером могут кодироваться такие последовательности тРНК, которые могут также (иногда) располагаться у 3'-конца первичного транскрипта.

Чередование транскрипционных единиц и нетранскрибирующихся спейсеров в tandemных повторах ядер эукариот можно непосредственно увидеть на электронных микрофотографиях. На рис. 23.4, где в качестве примера показана транскрипция рДНК тритона *N. viridescens*, видно, что каждая транскрипционная единица интенсивно экспрессируется, так что в транскрипции каждой повторяющейся единицы одновременно принимает участие большое число РНК-полимераз. Эти полимеразы располагаются настолько близко друг к другу, что при увеличении длины их продуктов образуется характерная структура, движущаяся вдоль транскрипционной единицы.

Длина нетранскрибирующегося спейсера варьирует в широких пределах у разных видов и (иногда) внутри одного вида организмов. В табл. 23.3 приведены соответствующие данные.

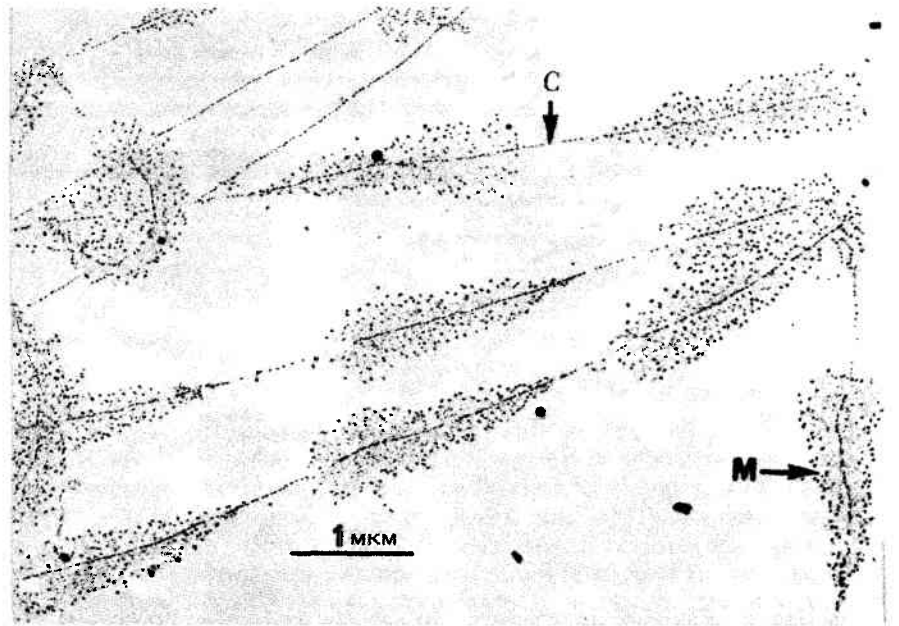
Таблица 23.3

Тандемно повторяющиеся единицы кластеров рДНК имеют разные размеры¹

Вид	Длина повторяющейся единицы	Длина нетранскрибирующегося спейсера	Длина транскрипта
<i>S. cerevisiae</i>	8950	1750	7 200
<i>D. melanogaster</i>	11 500–14 200	3750–6 450	7 750
<i>X. laevis</i>	10 500–13 500	2300–5 300	7 875
<i>M. musculus</i>	44 000	30 000	13 400

¹ Эти данные относятся к непрерывным генам. Диапазон длин был бы больше, если бы в таблицу были включены данные для прерывистых 28S-генов.

Рис. 23.4. При транскрипции кластеров рДНК образуется множество матриц (М), каждая из которых соответствует одной транскрипционной единице и отделяется от другой нетранскрибирующимся спейсером (С). (Фотография любезно предоставлена Oscar Miller.)



У дрожжей присутствует короткий нетранскрибирующийся спейсер, имеющий относительно постоянную длину. В разных копиях повторяющейся единицы *D. melanogaster* длина нетранскрибирующегося спейсера может отличаться почти в два раза. Аналогичная картина наблюдается и у *X. laevis*. Во всех случаях каждая повторяющаяся единица имеет вид единого tandemного кластера, располагающегося в одной определенной хромосоме. (В случае *D. melanogaster* это половая хромосома. Кластер X-хромосомы имеет большие размеры, чем кластер Y-хромосомы. Поэтому у самок имеется большее число копий генов для рРНК, чем у самцов. Однако, как мы увидим далее, не все эти гены активны.)

У млекопитающих повторяющаяся единица имеет существенно больший размер и включает транскрипционную единицу размером около 13 000 п. н. и нетранскрибирующийся спейсер размером около 30 000 п. н. Обычно гены располагаются в нескольких разбросанных кластерах, находящихся у человека и мыши в пяти и шести хромосомах соответственно. Возникает интересный вопрос: каким образом механизмы коррекции, по-видимому функционирующие в пределах одного кластера и поддерживающие постоянство последовательностей рРНК, могут работать при наличии нескольких кластеров.

Область ядра, где происходит синтез рРНК, имеет характерное строение с фибриллярной сердцевиной, окруженной гранулярной оболочкой. Фибриллярная сердцевина — это область, где на ДНК-матрице транскрибируется рРНК, а гранулярная оболочка образована рибонуклеопротеиновыми частицами, в которые встроена рРНК. Вся область называется **ядрышком**.

Можно видеть, что с ядрышком связаны определенные участки хромосом. Они называются **ядрышковыми организаторами**. Каждый ядрышковый организатор соответствует кластеру tandemно повторяющихся генов рРНК, поэтому в диплоидном ядре имеется в два раза больше ядрышковых организаторов, чем tandemных кластеров в гаплоидном геноме. Концентрация tandemно повторяющихся рРНК-генов вместе с их очень интенсивной транскрипцией и определяют характерное морфологическое строение ядрышка.

Некоторые гены рРНК располагаются не в хромосомах

В двух случаях транскрипционные единицы для рРНК располагаются не в хромосоме, а присутствуют в клетке в виде отдельных **внехромосомных** молекул.

У ряда низших эукариот наблюдается стабильная картина, когда гены, ответственные за синтез рРНК, располагаются во внехромосомных молекулах ДНК, причем на одно ядро приходится много копий такой ДНК. Пример такого рода, обнаруженный у *D. discoideum*, показан на рис. 23.5. В этом случае внехромосомная молекула представляет собой большой палиндромный димер, включающий две транскрипционные единицы с противоположной ориентацией, разделенные отрезком ДНК размером около 20 000 п. н. Каждый внехромосомный димер также включает два 5S-гена, хотя они и транскрибируются независимо от основной транскрипционной единицы. О механизмах наследования описанной молекулы ДНК известно очень немного.

При развитии некоторых животных возникает следующая ситуация, носящая временный характер. У многих видов животных во время оогенеза внезапно образуется большое число дополнительных рРНК-генов. Наиболее полно этот процесс был охарактеризован у *X. laevis*. Путем амплификации его хромосомных повторяющихся единиц образуются кольцевые внехромосомные молекулы ДНК, каждая из которых содержит много повторяющихся единиц. В отличие от хромосомных генов, где соседние повторяющиеся единицы могут иметь спейсеры разной длины, соседние единицы внехромосомных кольцевых молекул всегда имеют спейсеры одинаковой

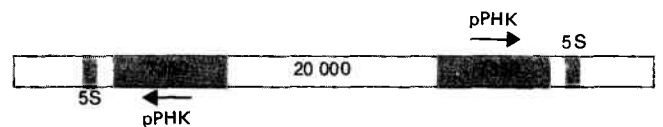


Рис. 23.5. Гены рРНК *D. discoideum* располагаются во внехромосомном палиндроме.

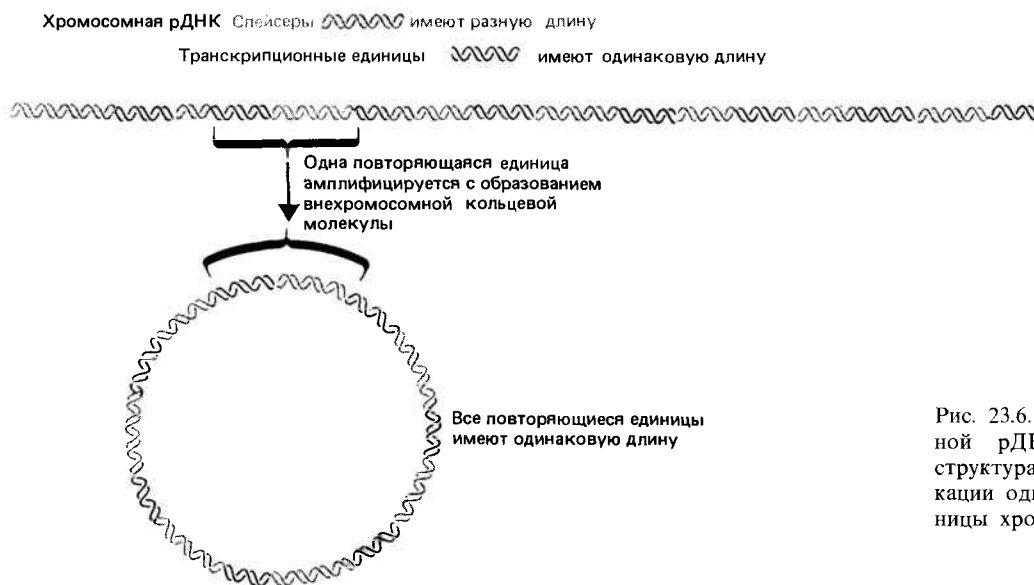


Рис. 23.6. Каждая молекула внехромосомной рДНК ооцита *X. laevis* – кольцевая структура, образовавшаяся путем амплификации одной и той же повторяющейся единицы хромосомной рДНК.

длины. Из этого, очевидно, следует, что каждая из кольцевых молекул образуется путем синтеза множества расположенных конец в конец копий одной хромосомной повторяющейся единицы, как это показано на рис. 23.6. Дополнительные гены нужны для образования необходимой для ооцита рРНК; их наличие не характерно для соматических клеток, потребность которых в рРНК обеспечивается хромосомными генами. [По своему механизму такая амплификация – разновидность репликации, происходящей по так называемому типу катящегося кольца (разматывающегося рулона), которая будет обсуждаться в гл. 31 и проиллюстрирована на рис. 31.14.]

О нетранскрибирующихся спейсерах и промоторах

В отличие от постоянства длины нуклеотидной последовательности транскрипционной единицы в пределах одного кластера генов длины нетранскрибирующихся спейсеров различаются. Несмотря на это различие, последовательности длинных нетранскрибирующихся спейсеров сохраняют гомологию с более короткими нетранскрибируемыми спейсерами. Следовательно, можно предположить, что каждый нетранскрибирующийся спейсер имеет внутренние повторы, так что различия в длине являются следствием различий в числе повторов некоторой субъединицы.

В общих чертах строение нетранскрибирующегося спейсера приведено на рис. 23.7, где в качестве примера показана его организация у *X. laevis*. Участки с постоянной длиной чередуются с участками, длина которых может варьировать. В состав каждого из трех повторяющихся участков входит разное число повторов довольно короткой последовательности (длиной менее 100 п.н.). Этим объясняются общие различия в длине таких участков.

Один из имеющих постоянную длину участков (А) уникален по нуклеотидной последовательности и длине. Другие представляют собой короткие последовательности одинаковой длины, называемые *Ват-островками*. (Такое название возникло, поскольку эти участки были выделены с помощью рестриктазы *Vat*.) Они имеют удивительное строение, так как вторая половина нуклеотидной

последовательности *Ват-островка* обладает большим сходством с последовательностью, непосредственно предшествующей началу транскрипционной единицы и, по-видимому, содержащей промотор. Наличие такого сходства вызывает вопрос: почему транскрипция не начинается на *Ват-островках* так же, как она начинается с настоящего промотора?

Другой случай – наличие идентичных на вид промоторов, часть которых инициирует транскрипцию, а другие нет, – обнаружен у *D. melanogaster*. Здесь около двух третей генов, кодирующих высокомолекулярную рРНК, содержат вставки. Имеются два типа вставок, длина которых может быть различной, но местоположение постоянно (гл. 37). Ни одну из вставок нельзя отнести к интронам, поскольку сами гены, в состав которых они входят, не экспрессируются. Однако их 5'-фланкирующие участки не отличаются от соответствующих участков активных непрерывных генов. Некоторым образом (каким, до сих пор не ясно) наличие вставки связано с нарушением инициации транскрипции в начальном участке гена, находящемся от нее на расстоянии 6800 п.н.

5S-гены и псевдогены перемежаются

У большинства эукариот 5S-гены по своей организации отличаются от генов, кодирующих высокомолекулярные виды рРНК. Общее число 5S-генов может быть выше; они отличаются по расположению (расположениям) в хромосоме и транскрибируются РНК-полимеразой III вместо РНК-полимеразы I (фермент ядрышка). Мы имеем весьма слабое представление о том, как координируется транскрипция 5S-генов и генов высокомолекулярных рРНК, и координируется ли она вообще.

5S-гены обычно организованы в виде кластера или кластеров тандемных повторов. Они были детально исследованы у *X. laevis*, где в большинстве или во всех хромосомах имеется кластер 5S-генов, расположенный очень близко к теломере или непосредственно в области теломеры (конец хромосомы).

У *X. laevis* обнаружено несколько типов последовательностей 5S-РНК, которые можно разделить на два основных типа. Последовательности одного из них встречаются исключительно в ооцитах, последовательности

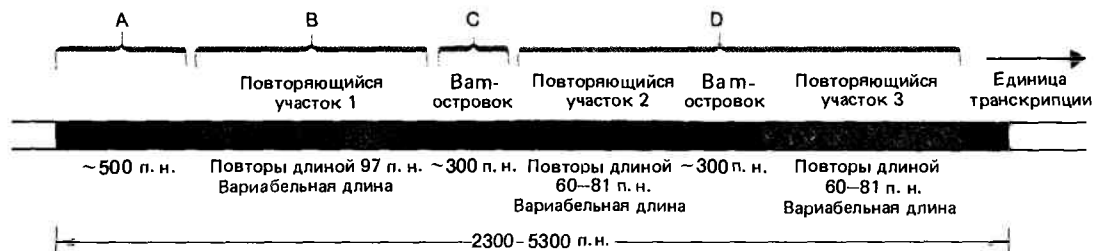


Рис. 23.7. В составе нетранскрибирующегося спейсера рДНК *X. laevis* имеются внутренние повторы, что объясняет различия в длинах спейсеров.

другого синтезируются в соматических клетках. Аналогичные типы последовательностей обнаружены у близкой к *X. laevis* лягушке *X. borealis*. Фактически сходство между соответствующими типами молекул 5S-РНК у этих двух видов больше, чем между молекулами двух типов в пределах одного вида. У *X. laevis* 5S-РНК ооцита и соматических клеток различаются шестью основаниями.

5S-РНК *X. laevis* кодируется около 20 000 генов, однако подавляющее большинство этих генов кодирует РНК, обнаруживаемую в ооцитах. Такое большое число генов, по-видимому, требуется для поддержания их экспрессии на уровне экспрессии амплифицированных генов для высокомолекулярных рРНК во время оогенеза. Гены, кодирующие 5S-рРНК, не амплифицируются.

Основной кластер 5S-генов ооцита имеет уже знакомую нам организацию в виде чередующихся генов и нетранскрибирующихся спейсеров. Однако повторяющаяся единица в свою очередь имеет внутренние повторы, возникшие, по-видимому, в результате ранее происшедшей дупликации. Структура этого кластера изображена на рис. 23.8.

В составе спейсера имеется (А—Т)-богатая последовательность, основу которой составляет повторяющаяся разное число раз последовательность длиной 15 п. н. и некоторые ее разновидности, обладающие большим сходством с ней. Длина каждого нетранскрибирующегося спейсера колеблется в соответствии с числом tandemных копий этого повтора. Средняя длина (А—Т)-богатой последовательности составляет около 400 п. н.

Остальные 300 п. н. повторяющейся единицы обогащены основаниями G—C. Этот участок начинается с последовательности длиной 49 п. н., за которой следует последовательность собственно 5S-гена длиной 120 п. н.

Далее следует около $1\frac{1}{2}$ (А—Т)-богатой повторяющейся единицы, за которой находится еще одна копия (G—C)-богатой спейсерной последовательности длиной 49 п. н. Затем снова повторяются первые 101 п. н. гена; поскольку эта неполная последовательность не транскрибируется в РНК, она является псевдогеном. Область, включающая псевдоген, — это повтор последовательности, окружающей начало активного гена, начиная от 73 п. н. слева от него и кончая 101 п. н. справа (от -73 до +101 п. н.). Удивительная особенность кластера в целом заключается в том, что он таким образом должен содержать одинаковое число генов и псевдогенов.

Промотор транскрипции 5S-РНК целиком расположен внутри гена (гл. 11). Псевдоген приобрел отличия от гена в результате всего лишь девяти точковых мутаций. Таким образом, в состав псевдогена входит участок, со-

державший слегка измененный внутренний промотор гена. Необходимо понять, являются ли эти изменения достаточными для того, чтобы псевдоген утратил способность к транскрипции с образованием РНК-продукта *in vivo*.

Не все 5S-гены идентичны. Нуклеотидная последовательность минорного компонента 5S-РНК ооцита (trace 5S-RNA) имеет некоторые отличия. Она также кодируется набором tandemно организованных генов и спейсеров, но длина повтора составляет только 350 п. н., и спейсер не имеет сходства со спейсером большого кластера. И снова мы видим, что нуклеотидная последовательность гена может претерпевать небольшие изменения, тогда как спейсер меняется полностью. Аналогичным образом нуклеотидная последовательность спейсера кластера 5S-генов *X. borealis* не обладает сходством ни с одной из охарактеризованных нуклеотидных последовательностей спейсеров *X. laevis*.

Эволюционная дилемма

При изучении всех кластеров идентичных генов возникает один и тот же вопрос: каким образом при наличии такого большого числа копий каждого гена поддерживается постоянство их нуклеотидных последовательностей? Говоря точнее, нас интересует, каким образом естественный отбор может препятствовать накоплению вредных мутаций.

При наличии большого числа копий гена непосредственное последствие мутации в какой-либо копии гена может быть лишь весьма незначительным. Эффект одиночной мутации разбавляется в связи с наличием большого числа копий гена, сохраняющих нуклеотидные последовательности дикого типа. Из этого можно сделать вывод, что к тому моменту, когда эффект мутаций станет достаточно явным, чтобы быть отброшенным в процессе эволюции, в геноме уже может накопиться достаточное количество мутантных копий. Летальность приобретает количественный характер — утверждение, согласующееся

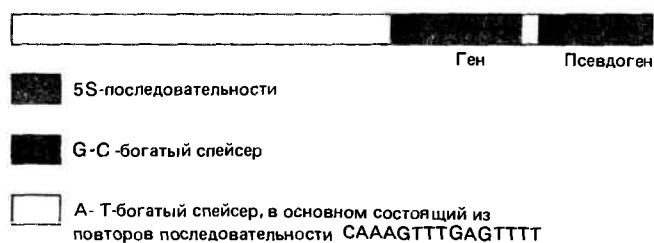


Рис. 23.8. В состав повторяющейся единицы большого кластера 5S-генов ооцита *X. laevis* входят ген и псевдоген.

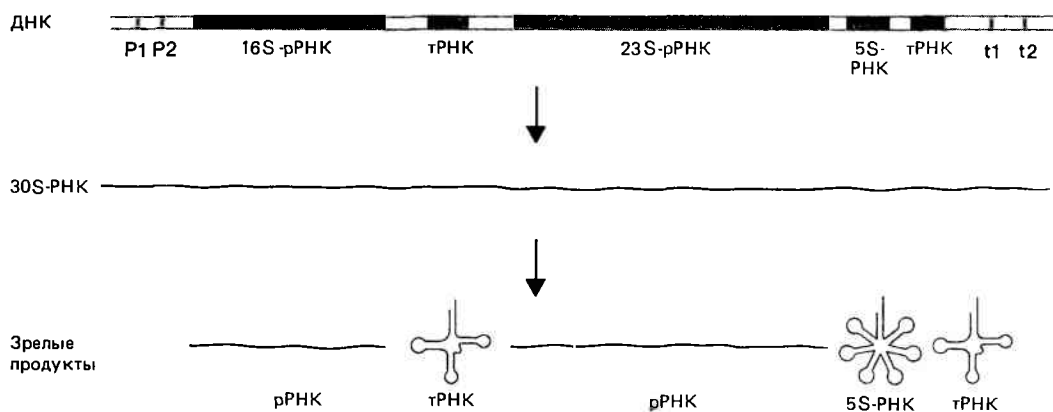


Рис. 23.9. В состав оперонов *rrr* входят гены, кодирующие рРНК, и тРНК.

Каждая функционально активная последовательность отделяется от следующей участком транскрибируемого спейсера; длины лидерной и кон-

цевой последовательностей зависят от того, какие промоторы и терминаторы используются. Каждый РНК-продукт образуется путем расщепления первичного транскрипта с двух сторон от освобождаемой РНК.

с тем наблюдением, что утрата половины единиц, входящих в состав кластеров рДНК *X. laevis* и *D. melanogaster*, может не привести к губительному эффекту. Каким же образом в этих единицах не происходит постоянного накопления вредных мутаций? И какова вероятность проявления полезных свойств редкой благоприятной мутации в кластере?

Кажется очевидным, что должен существовать некий механизм эволюционного отбора отдельных изменений. Были предложены два возможных типа механизмов.

В соответствии с моделью **внезапной коррекции** периодически весь кластер генов заменяется новым набором копий, происшедшим от одной или в крайнем случае от очень небольшого числа копий из генома предыдущего поколения. Чтобы отбор был достаточно эффективным, «периодически» должно означать каждые несколько поколений. Однако на практике любой механизм должен быть основан на регулярной повторяемости, поэтому «периодически» означает каждое поколение. Небольшое число копий, дающих начало новому кластеру, — это исходный набор, на который воздействует естественный отбор; отбрасываемые копии не подвергаются отбору. Исходные копии могут составлять определенный набор или могут быть отобраны случайным образом.

Недостаток этой модели заключается в том, что она предполагает регенерацию кластера из нескольких *повторяющихся единиц*. Это означает, что должны регенерировать и транскрипционные единицы, и спейсеры, а, значит, вся повторяющаяся единица должна быть представлена одинаковыми последовательностями или по крайней мере различающимися настолько же, насколько различается небольшое число исходных копий. В действительности же только транскрипционные единицы постоянны, а в спейсерах наблюдаются значительные вариации.

Модель **фиксации при кроссинговере** предполагает, что весь кластер подвергается постоянным перестройкам, осуществляемым путем неравного кроссинговера. По существу, это означает, что на тандемный кластер должны действовать в широком масштабе все те механизмы, которые мы уже обсуждали для случая глобиновых генов, где они носят более ограниченный и случайный характер. Более подробно значение таких механизмов будет обсуждено в гл. 24, но уже теперь ясно, что наличие тандемных кластеров может быть причиной частых неправильных спариваний генов, нуклеотидные последовательности ко-

торых практически одинаковы, но располагаются в разных участках кластеров.

В результате постоянного увеличения и уменьшения числа единиц при неравном кроссинговере могло получиться так, что все повторяющиеся единицы данного кластера произошли из относительно небольшого их числа в исходном кластере. Различие спейсеров по длине согласуется с предположением о том, что в неравном кроссинговере участвуют спейсеры, содержащие внутренние «неправильно» спаривающиеся участки. Это может объяснить постоянство нуклеотидных последовательностей генов по сравнению с их вариабельностью в спейсерах. Гены подвергаются воздействию естественного отбора, когда происходит амплификация отдельных повторяющихся единиц кластера; однако спейсеры отбору не подвергаются и могут накапливать изменения.

Бактериальные рРНК-гены и тРНК-гены входят в состав одних и тех же оперонов

Семь оперонов, кодирующих рРНК *E. coli*, называются *rrnA-G*. Тесная связь их в хромосоме отсутствует, и не известно, как поддерживается одинаковый нуклеотидный состав их последовательностей. Поскольку число генов невелико, можно определить их точные нуклеотидные последовательности и таким образом непосредственно увидеть, являются ли все рРНК-гены идентичными или имеют небольшие различия, не обнаруживаемые в популяции молекул рРНК.

Каждый такой оперон транскрибируется с образованием единого РНК-предшественника, при расщеплении которого образуются молекулы зрелых рРНК. У бактерий, у которых реакция расщепления блокирована в результате мутации гена, кодирующего соответствующий фермент, предшественник накапливается в виде РНК с константой седиментации примерно 30S. В целом все *rrn*-опероны организованы одинаково и содержат гены, кодирующие синтез трех молекул рРНК и расположенные в последовательности 16S-23S-5S, как показано на рис. 23.9. Гены, кодирующие высокомолекулярные рРНК, располагаются в том же порядке, что и в эукариотической транскрипционной единице; наличие в той же самой

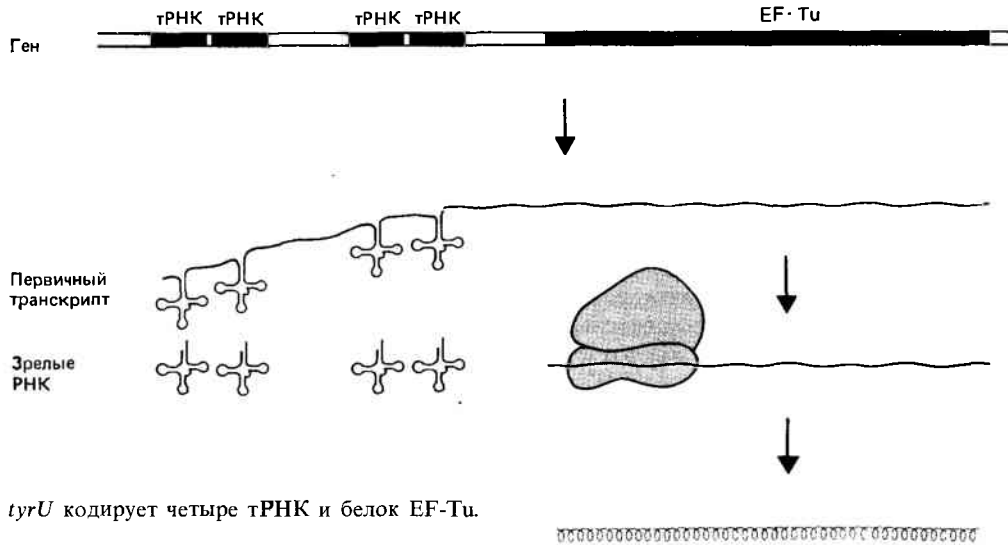


Рис. 23.10. Оперон *tyrU* кодирует четыре тРНК и белок EF-Tu.

транскрипционной единице генов для 5S-РНК присуще только прокариотам.

Все опероны *rrn* имеют двойной промотор. Первый промотор, *P1*, расположен на 300 п. н. левее точки начала синтеза 16S-рРНК. По-видимому, это основной промотор; для инициирования транскрипции он использует нуклеотиды АТР и GTP и имеет обычные универсальные последовательности. На протяжении первых 150 п. н. транскрипционные единицы могут различаться в разных *rrn*-оперонах. Внутри этого участка, на расстоянии около 110 п. н. от *P1*, располагается второй промотор, *P2*, который для инициирования транскрипции использует менее распространенный начальный нуклеотид СТР и универсальные последовательности которого могут быть не столь явно выражены. 5'-Конец индивидуального предшественника 16S-рРНК (размер которого слегка превышает размер зрелой рРНК) располагается после начала последовательности, общей для всех *rrn*-оперонов.

Между последовательностями, кодирующими 16S-рРНК и 23S-рРНК, находится транскрибирующийся спейсер размером 400–500 п. н. Однако внутри его располагается последовательность или последовательности, кодирующие тРНК. Как показано на рис. 23.9, в четырех *rrn*-оперонах в транскрибирующийся спейсер входит одна и та же последовательность, кодирующая тРНК (тРНК₂^{Gly}). В трех других *rrn*-оперонах в транскрибирующийся спейсер содержится последовательности, кодирующие две вида тРНК (тРНК₁^{Ile} и тРНК_{1B}^{Ala}). Таким образом, взяв за основу последовательности «спейсерной тРНК», можно выделить два типа оперонов, содержащих гены рРНК. тРНК образуются при расщеплении РНК-предшественника; остальные участки спейсера, по-видимому, деградируют.

Встречается и другая разновидность *rrn*-оперонов. В некоторых оперонах последняя кодируемая последовательность, входящая в состав предшественника РНК, — это последовательность, соответствующая 5S-РНК. В других могут присутствовать некоторые дополнительные последовательности, кодирующие тРНК. Например, у 3'-конца *rrnC*-оперона имеются два гена, кодирующие тРНК^{Trp} и тРНК^{Asp}. (Наличие гена для тРНК^{Trp} делает этот оперон необходимым для *E. coli*, поскольку данный ген присутствует в виде только одной копии и поэтому обеспечивает единственную возможность использования триптофана в синтезе белка.) Гены

тРНК, расположенные в спейсере и у 3'-конца оперона, не связаны между собой.

тРНК-гены могут быть организованы в виде кластеров

Мы имеем только лишь неполное представление об организации генов для тРНК и у бактерий, и у эукариот, но, по-видимому, тРНК-гены имеют тенденцию располагаться в виде кластеров.

Скудная информация об организации эукариотических тРНК-генов получена при исследовании клонированных фрагментов генома, комплементарных определенным тРНК. Часто оказывается, что в состав этих фрагментов входят сходные последовательности, кодирующие другие тРНК. В отличие от РНК, кодируемых тандемными генными кластерами, все тРНК, кодируемые генами, находящимися в одном и том же участке генома, могут быть различными. Нам не известно, транскрибируются ли организованные в виде кластера тРНК-гены независимо друг от друга или образуют общую транскрипционную единицу.

Аналогичная ситуация наблюдается в клетках *E. coli*, где тРНК-гены образуют кластеры, каждый из которых кодирует несколько акцепторов аминокислот. По крайней мере некоторые кластеры образуют опероны, имеющие один промотор и транскрибирующиеся с образованием одного большого РНК-предшественника. Фаг Т4 содержит кластер, включающий восемь генов для тРНК, экспрессирующихся аналогичным образом. В отличие от такого способа организации тРНК-генов в митохондриях млекопитающих они располагаются между структурными генами (гл. 22).

Во всех известных случаях образование тРНК происходит по общему принципу. тРНК входит в состав более длинного предшественника. В результате расщепления этого предшественника и других реакций процессинга образуется зрелая тРНК. Эти процессы рассматриваются в гл. 25.

Были детально исследованы два кластера тРНК-генов *E. coli*. Оказалось, что они содержат разные нуклеотидные последовательности, кодирующие разные тРНК. Эти кластеры включают гены, кодирующие тРНК для

тирозина, соответственно чему они и получили свои названия. Кластер *tyrU* включает в себя гены, кодирующие четыре тРНК: тРНК₄^{Thr}, тРНК₂^{Tyr}, тРНК₃^{Thr}, тРНК₂^{Gly}. В состав кластера *tyrT* входят два идентичных гена, кодирующие тРНК₁^{Tyr}.

В каждом кластере его единственный промотор расположен перед первым тРНК-геном, как показано в приведенном на рис. 23.10 примере. За последним тРНК-геном находится нуклеотидная последовательность, кодирующая белок. В кластере *tyrU* это последовательность гена *ufvB* (одного из двух генов, кодирующих фактор синтеза белка EF-TU). В кластере *tyrT* такая последовательность кодирует белок Р, подобный протамину полипептид (протамины – сильноосновные белки, связанные с ДНК сперматозоидов).

мРНК, кодирующие каждый такой белок, по-видимому, образуются при расщеплении первичного транскрипта на участке между последним тРНК-геном и началом структурного гена. Местом расщепления, очевидно, является общая для двух оперонов короткая нуклеотид-

ная последовательность. В отличие от других бактериальных оперонов, где рибосомы, по всей вероятности, связываются с мРНК практически сразу же после освобождения РНК-полимеразы из комплекса с промотором, в этом случае, по-видимому, вначале должно произойти образование мРНК путем ее «отрезания» от расположенной левее последовательности тРНК. Спейсеры в рДНК исследовал Ридер (*Reeder, Cell, 38, 349–351, 1984*).

Рекомендуемая литература

Анализ структуры гистоновых генов и меняющиеся представления о степени ее консервативности приведены в ряде обзоров, начиная с обзора Кидса (*Kedes, Ann. Rev. Biochem., 48, 837–870, 1979*) и кончая обзором Хентшела и Бернстила (*Hentschel, Birmstiel, Cell, 25, 301–313, 1981*). Структура кластеров эукариотической рДНК рассматривается в книге Б. Льюина (*Lewin, Gene Expression, 2, Eucaryotic Chromosomes Wiley, New York, 875–906, 1980*).

Глава 24

ОРГАНИЗАЦИЯ ПРОСТЫХ ПОСЛЕДОВАТЕЛЬНОСТЕЙ ДНК

В геноме эукариот имеется большое число нуклеотидных последовательностей, не кодирующих белки (или РНК). Составляя, по-видимому, основную часть последовательностей ДНК, они находятся под влиянием других факторов эволюционного отбора, чем те, которые обуславливают необходимость кодирования определенных аминокислот. Некоторые из них входят в состав транскрипционных единиц, например нетранслируемые фланкирующие последовательности мРНК или интроны, удаляемые при созревании мРНК.

Другие могут играть роль сигнальных последовательностей, узнаваемых белками; к ним относятся промоторы транскрипции, точки начала репликации ДНК, сайты скручивания хромосом, точки прикрепления кинетохоры и другие элементы, необходимые для осуществления клеточных функций. Очень немногие из этих последовательностей охарактеризованы до такой степени, чтобы определенной последовательности могла быть приписана определенная функция. В самом деле, единственная такая последовательность в эукариотическом геноме, некоторые свойства которой известны, – это промотор транскрипции. Если говорить и о вирусных геномах, то имеются также данные о точках начала репликации ДНК. На основе небольшого количества имеющихся данных можно предположить, что сигнальные последовательности такого типа, по-видимому, имеют небольшую длину.

Рассматривая эту проблему с другой стороны, можно задать вопрос о том, что мы знаем о коротких последовательностях, повторяющихся в геноме в виде идентичных или сходных копий. Конечно, это приведет нас к определению повторяющейся ДНК по характеру кинетики ее реассоциации, как описано в гл. 17. Резюмируя все сказанное, можно утверждать, что фракция умеренных повторов состоит из множества семейств после-

довательностей, характеризующихся разной степенью повторяемости и разной внутренней организацией. Расположению в геноме членов таких семейств последовательностей часто свойственно более или менее регулярное чередование с более длинными участками неповторяющейся ДНК. Фракция высокоповторяющейся ДНК, напротив, состоит из очень коротких последовательностей, многократно и тандемно повторяющихся в составе больших кластеров. Из-за короткой длины повторяющихся единиц их иногда называют «простыми последовательностями ДНК».

Ни для одного семейства последовательностей не имеется данных о его функции. В этой главе мы опишем известные свойства различных типов повторяющейся ДНК; при изучении некоторых из них получена интересная информация о механизмах их эволюционирования.

Семейство Alu

По крайней мере в геномах некоторых млекопитающих значительная часть рассеянных умеренно повторяющихся последовательностей входит в состав одного семейства. Небольшая длина и высокая степень повторяемости членов семейства делают его похожим на простые последовательности ДНК, за исключением того, что отдельные члены семейства рассеяны по геному, а не организованы в виде тандемных кластеров.

В геноме человека значительная часть умеренно повторяющейся ДНК представлена последовательностями длиной около 300 п.н., перемежающимися с неповторяющейся ДНК. Двухцепочечная ДНК, соответствующая фракции умеренно повторяющихся последовательностей, может быть выделена путем ренатурации при средних

значениях C_{0t} с последующим разрушением ренатурировавшихся соседних участков неповторяющейся ДНК. По меньшей мере половина ренатурировавшегося двухцепочечного материала расщепляется рестриктазой AluI в единственном сайте. Этот сайт расположен на расстоянии 170 п.н. от начала последовательности.

Все последовательности, расщепляемые рестриктазой AluI, по-видимому, являются членами одного семейства, названного (в соответствии со способом его обнаружения) Alu-**семейством**. На гаплоидный геном приходится примерно 300 000 членов, входящих в его состав (это эквивалентно тому, что на каждые 6 т.п.н. ДНК приходится один его член). Отдельные Alu-последовательности широко распространены в геноме; по меньшей мере одна копия Alu-последовательности обнаруживается практически в каждом клоне библиотеки фрагментов геномной ДНК со средней длиной 18 т.п.н. Семейство аналогичных последовательностей обнаружено у мыши (оно называется семейством B1; и в его состав входит 50 000 членов), у китайского хомячка (где оно называется Alu-подобным семейством) и у других млекопитающих.

Отдельные члены Alu-семейств не идентичны, но сходны друг с другом. По-видимому, в семействе каждого вида организмов существует общая для всех членов семейства каноническая (усредненная, универсальная, обобщенная) последовательность. У десяти членов Alu-семейства человека имеется каноническая последовательность, состоящая из тандемно дублировавшего участка размером 130 п.н. и последовательности размером 31 п.н., встроившей в правую часть димера. На каждом конце отдельных членов семейства находятся короткие, отличающиеся друг от друга повторы, напоминающие концевые участки транспозонов (гл. 38). Средняя степень гомологии отдельных последовательностей семейства с канонической последовательностью составляет 87%. Повторяющаяся последовательность B1 у мыши имеет размер 130 п.н., что соответствует размеру мономера Alu-семейства человека. Степень гомологии B1 с Alu-последовательностями человека составляет 70–80%.

Были выявлены многие свойства Alu-семейства, а его всеобщность породила множество предположений о его функции, но истинную роль этого семейства до сих пор установить не удалось.

В состав Alu-последовательности входит участок размером 14 п.н., практически идентичный последовательности, присутствующей в точке начала репликации паповавирусов (таких, как SV-40), и вируса гепатита В. Это, возможно, указывает на связь Alu-семейства с точками начала репликации генома эукариот, хотя этому противоречит количество членов семейства, поскольку их число примерно в десять раз превышает число возможных точек начала репликации (гл. 31).

Было установлено также, что по крайней мере некоторые члены семейства могут транскрибироваться *in vitro* с образованием небольших РНК с помощью РНК-полимеразы III (ответственной за синтез малых ядерных РНК, тРНК и 5S-РНК). Некоторые (хотя и не все) члены Alu-подобного семейства китайского хомячка, по-видимому, транскрибируются *in vivo*. Такого рода транскрипционные единицы обнаруживаются поблизости от других транскрипционных единиц. Например, две такие транскрипционные единицы расположены в области кластера β -глобиновых генов человека.

В эукариотических клетках обнаруживается несколько видов малых РНК (гл. 27). 7S-РНК, по-видимому, имеет-

ся в клетках высших эукариот и, кроме того, входит в состав вирионов некоторых ретровирусов. Была установлена нуклеотидная последовательность 7S-РНК клеток человека и крысы, где она имеет сходное строение и соответствует мономеру Alu-семейства со вставкой размером 140 нуклеотидов в точке, находящейся на расстоянии 100 нуклеотидов от 5'-конца. Таким образом, эта РНК гомологична первым 100 и последним 40 нуклеотидам Alu-последовательности. Центральная вставка одинакова у человека и крысы. Из этого следует, что формирование кодирующей последовательности из Alu-последовательности предшествовало разделению млекопитающих на виды.

Члены Alu-семейства могут входить в состав транскрипционных единиц структурных генов, что следует из их присутствия в гигантской ядерной РНК. Наличие множества копий Alu-последовательности в одной молекуле ядерной РНК может способствовать образованию вторичной структуры. Более того, присутствие членов Alu-семейства в виде обращенных повторов, описанное в следующем разделе, в основном и обуславливает вторичную структуру ядерной РНК, обнаруженную у млекопитающих (гл. 27).

Обращенные повторы мгновенно ренатурируют

Ренатурация основной части ДНК любого генома представляет собой бимолекулярную реакцию, протекающую в соответствии с кинетикой реакции второго порядка, описанной в гл. 17. Это относится к ренатурации неповторяющейся, умеренно повторяющейся и высокоповторяющейся фракций ДНК. Но обычно имеется также и другая фракция, реассоциирующая значительно быстрее, чем высокоповторяющаяся ДНК. Эту фракцию получают в виде двухцепочечных структур даже при самых коротких временах реассоциации. Будет, конечно, преувеличением сказать, что она реассоциирует «мгновенно», но реакция действительно протекает чрезвычайно быстро при значениях C_{0t} менее 10^{-4} или 10^{-5} .

Эту фракцию иногда называют **ДНК, ренатурирующей в нулевой момент времени**, поскольку в реакции реассоциации она обнаруживается в виде двухцепочечных молекул, по существу, в момент времени, равного нулю. Обычно ее удаляют до определения кинетики реассоциации других фракций ДНК. (Это означает, что ДНК денатурируют, а затем двухцепочечные молекулы удаляют хроматографией на гидроксипатите, после которой инкубируют одноцепочечную ДНК для осуществления реакции реассоциации второго порядка.)

Реассоциацию фракции, ренатурирующей в нулевой момент времени, называют **мономолекулярной реакцией**, так как в реакции участвует только одна молекула ДНК. Это означает, что два реассоциирующих компонента должны располагаться в одной цепи ДНК. Поэтому такая реакция является внутримолекулярной реассоциацией последовательностей, противоположно ориентированных и расположенных близко друг к другу. Эти последовательности — **обращенные повторы** или **палиндромы**, уже рассмотренные нами в гл. 13. Такую фракцию иногда называют «схлопнувшейся» ДНК.

Можно выделить два типа палиндромов, как показано на **рис. 24.1**. Когда обращенные повторы непрерывны, при их реассоциации образуется структура типа шпильки. Когда они разделены какой-либо другой последователь-

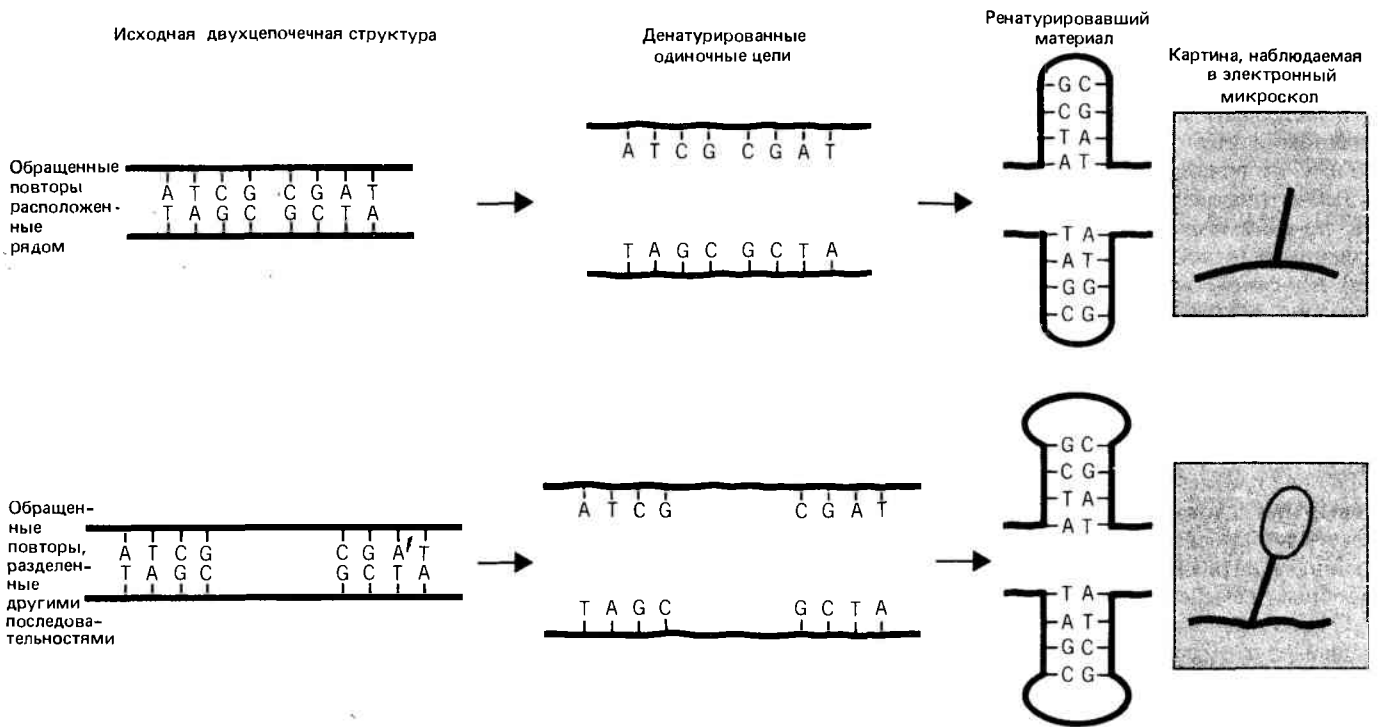


Рис. 24.1. При денатурации и ренатурации из расположенных рядом обращенных повторов образуются шпильки, а из обра-

щенных повторов, разделенных другими последовательностями, образуются стебли и петли.

ностью, при их реассоциации образуется двухцепочечный стебель с одноцепочечной петлей.

По-видимому, обращенные повторы рассеяны по всему геному. К их числу относятся последовательности всех классов, как неповторяющиеся, так и повторяющиеся. Об их функции известно немного.

Они присутствуют в больших транскрипционных единицах, кодирующих синтез ядерной РНК, и в этом случае в основном и обуславливают наличие у нее характерной вторичной структуры. Многие обращенные повторы, обнаруживаемые в ядерной РНК человека, являются членами Alu-семейства, описанного в предыдущем разделе. Действительно, если рассматривать Alu-семейство в целом, то около трети его членов — это расположенные непосредственно по соседству друг с другом обращенные повторы. Остальные представляют собой либо отдельные копии либо группу обращенных повторов, расположенных более отдаленно друг от друга. Всего к Alu-семейству относится половина или более обращенных повторов генома человека.

Высокоповторяющаяся ДНК образует сателлитную ДНК

Высокоповторяющаяся ДНК организована в виде коротких последовательностей, многократно повторяющихся в геноме в виде тандема. Таким образом, эта фракция образует блоки, каждый из которых включает множество копий высокоповторяющихся последовательностей. В некоторых случаях эти копии идентичны, в других они близки по строению.

Тандемно повторяющиеся короткие последовательности часто образуют фракцию с определенными физическими свойствами. Эти свойства могут быть использованы для ее выделения. В некоторых случаях состав

азотистых оснований повторяющихся последовательностей отличается от их состава в геноме в целом; в результате они могут образовать отдельную фракцию, характеризующуюся определенной плавучей плотностью.

Зависимость плавучей плотности двухцепочечной ДНК от содержания в ней пар G—C выражается эмпирической формулой

$$\rho = 1,660 + 0,00098 (\% \text{ G—C}) \text{ г} \cdot \text{см}^{-3}. \quad (12)$$

Плавучую плотность обычно определяют при центрифугировании ДНК в градиенте плотности CsCl. ДНК образует полосу в участке градиента, соответствующем ее собственной плотности. В градиенте плотности CsCl можно разделять фракции ДНК, плавучая плотность которых различается более чем примерно на $0,005 \text{ г} \cdot \text{см}^{-3}$. Эта разница соответствует различию в содержании (G—C)-пар, равному 5%.

При центрифугировании эукариотической ДНК в градиенте плотности основная часть генома представлена множеством фрагментов, образующих довольно широкий пик, вершина которого совпадает с плавучей плотностью, соответствующей среднему содержанию в геноме (G—C)-пар. Это основной пик ДНК. Иногда при другом значении плавучей плотности выявляется дополнительный пик (или пики) меньшего размера. Материал, соответствующий этому пику, называют сателлитной ДНК.

Хорошим примером такого рода служит ДНК мышцы; это показано на рис. 24.2. Приведенный график — результат сканирования пиков, образовавшихся при центрифугировании ДНК мышцы в градиенте плотности CsCl. Основной пик содержит 92% материала генома, а его вершина соответствует плавучей плотности $1,701 \text{ г} \cdot \text{см}^{-3}$ (при этом среднее содержание (G—C)-пар равно 42%, что типично для ДНК млекопитающих). Пик меньшего размера включает 8% материала генома и имеет плавучую плотность, равную $1,690 \text{ г} \cdot \text{см}^{-3}$. Это сателлитная ДНК

мышь, где содержание (G—C)-пар гораздо ниже, чем в любой другой части генома (примерно 30%).

Сателлитные ДНК присутствуют в геноме многих эукариот. В градиенте плотности CsCl они могут быть тяжелее или легче основной фракции ДНК; обычно они составляют не более 5% всей ДНК. Разрешающую способность при обнаружении сателлитной ДНК часто можно значительно повысить при центрифугировании в градиентах плотности Cs₂O₄, содержащих ионы серебра или другие химические соединения, включая различные красители. В некоторых случаях от основного пика четко отделяется один пик сателлитной ДНК; иногда выявляется несколько последовательностей сателлитной ДНК.

Сателлитная ДНК при центрифугировании в градиентах плотности чаще всего ведет себя аномально. Поэтому, когда определяют ее подлинный нуклеотидный состав, он часто отличается от состава, который можно было бы предсказать на основе ее плавучей плотности. Объясняется это тем, что величина ρ зависит не только от состава оснований, но и от того, какие пары являются ближайшими соседями. В случае простых последовательностей значение ρ , по-видимому, отличается от значения, характерного для случайного расположения пар и подчиняющегося уравнению 12. Сателлитная ДНК может быть метилирована, что также меняет ее плотность.

Основная часть высокоповторяющейся ДНК генома часто может быть выделена как сателлитная ДНК. Когда имеется некоторая фракция высокоповторяющейся ДНК, которая в градиенте плотности не ведет себя как сателлитная, после выделения ее свойства часто оказываются такими же, как свойства сателлитной ДНК. Это означает, что она состоит из множества tandemных повторов, аномально ведущих себя при центрифугировании. Выделяемый таким способом материал иногда называют **скрытой (криптической) сателлитной ДНК**. Скрытая и явная сателлитные ДНК вместе обычно включают в себя все tandemно повторяющиеся блоки высокоповторяющейся ДНК. Когда в геноме имеется более одного типа высокоповторяющейся ДНК, каждому из них соответствует свой блок сателлита (хотя иногда различные блоки расположены рядом).

Сателлитная ДНК часто располагается в области гетерохроматина

Где располагаются блоки сателлитной ДНК? Развитие техники гибридизации нуклеиновых кислот позволило определить расположение последовательностей сателлитной ДНК непосредственно в хромосоме. При гибридизации *in situ* или **цитологической гибридизации** хромосомную ДНК денатурируют, обрабатывая клетки, распластанные на покровном стекле. Затем к препарату добавляют радиоактивный ДНК- или РНК-зонд. Зонд гибридизуется с комплементарными ему участками денатурированного генома. Расположение участков гибридизации можно определить с помощью радиоавтографии.

Меченая сателлитная ДНК часто располагается в области гетерохроматина, обнаруживаемого вокруг центромер митотических хромосом. **Гетерохроматин** — термин, используемый для обозначения участков хромосом, постоянно находящихся в компактной спирализованной форме и инертных, в отличие от **эухроматина**, составляющего основную часть генома (гл. 28).

На рис. 24.3 приведен пример локализации сателлит-

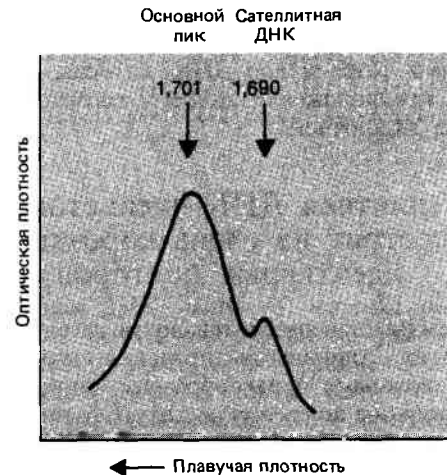


Рис. 24.2. При центрифугировании в градиенте плотности CsCl ДНК мыши разделяется на основную фракцию и фракцию сателлитной ДНК.

ной ДНК в хромосомах клеток мыши. В этом случае мечеными оказываются концы хромосом, поскольку именно там расположены центромеры хромосом *M. musculus*.

Отождествление сателлитной ДНК с гетерохроматином означает, что высокоповторяющаяся ДНК не экспрессируется с образованием РНК или белка (последнее следует также из простоты составляющих ее нуклеотидных последовательностей). Попытки обнаружить РНК-продукты, соответствующие сателлитной ДНК, подтвердили предположение о том, что высокоповторяющиеся последовательности обычно не транскрибируются.

Локализация сателлитной ДНК в области центромеры говорит о том, что она может выполнять некоторую структурную функцию в хромосоме. Поскольку центромеры — это области, где при митозе и мейозе образуются кинетохоры, обеспечивающие движение хромо-



Рис. 24.3. При цитологической гибридизации обнаружено, что сателлитная ДНК мыши расположена в области центромер. (Фотография любезно предоставлена Mary Lou Pardue и Joe Gall.)

сом, функция сателлитной ДНК может быть связана с процессом расхождения хромосом. Но это все, что мы знаем о роли сателлитной ДНК, которая, если не считать этих предположений общего характера, остается совершенно загадочной.

Сателлитная ДНК членистоногих состоит из очень коротких идентичных повторов

Нуклеотидные последовательности сателлитной ДНК членистоногих, типичными представителями которых являются насекомые и ракообразные, по-видимому, довольно гомогенны по своему составу. Обычно одна очень короткая повторяющаяся единица составляет более 90% всех последовательностей сателлитной ДНК. Это облегчает определение ее нуклеотидной последовательности, что и было сделано в ряде случаев.

У *Drosophila virilis* имеются три основных типа сателлитной ДНК, а также скрытый сателлит, которые вместе составляют значительную часть генома, около 40%. Нуклеотидные последовательности сателлитной ДНК приведены в табл. 24.1. Три основных пика сателлитной ДНК состоят из обладающих большим сходством последовательностей. Замена одного нуклеотидного основания достаточна для образования из фракции I сателлитной ДНК фракций II или III. Скрытая сателлитная ДНК может быть образована из фракции II путем замены двух оснований.

Последовательность фракции I сателлитной ДНК имеется и у других видов *Drosophila*, родственных *virilis*; следовательно, ее появление могло предшествовать разделению дрозофил на виды. Последовательности фракций II и III, по-видимому, специфичны для *D. virilis*, и поэтому их образование из фракции I в процессе эволюции могло произойти после разделения видов.

Основное свойство этих сателлитных ДНК — очень маленькая длина повторяющейся единицы: только 7 п.н. Аналогичные сателлитные ДНК обнаружены у других видов организмов. Например, в геноме *D. melanogaster* имеется множество фракций сателлитной ДНК, причем повторяющиеся единицы некоторых имеют небольшую

длину (5, 7, 10 или 12 п.н.). Сходные сателлитные ДНК обнаружены у ракообразных.

Большое сходство последовательностей сателлитных ДНК, обнаруживаемое у *D. virilis*, не является непременным свойством геномов других организмов, сателлитные ДНК которых могут различаться. Очевидно, каждая сателлитная ДНК возникла в результате амплификации очень короткой последовательности. Эта последовательность, вероятно, представляет собой разновидность ранее существовавшей сателлитной ДНК или имеет какое-то иное происхождение. По-видимому, во всех случаях сателлитные ДНК постоянно образуются в геноме и исчезают из него. Поэтому трудно установить, как эволюционировали сателлиты, поскольку существующий в настоящее время сателлит мог произойти от некоторого ранее существовавшего сателлита, позже исчезнувшего из генома. Важная особенность таких сателлитных ДНК заключается в том, что они представляют собой очень длинные участки ДНК с очень низкой генетической сложностью, внутри которых может поддерживаться постоянство последовательностей нуклеотидных оснований.

Одно из свойств многих сателлитных ДНК — выраженная асимметрия в ориентации нуклеотидных пар двух цепей ДНК. В случае *D. virilis* (см. табл. 24.1) в каждой из основных фракций сателлитной ДНК одна цепь существенно богаче основаниями Т и Г. Это увеличивает ее плавучую плотность, поэтому после денатурации **тяжелая цепь (H)** может быть отделена от **легкой цепи (L)**. Это может оказаться полезным при определении нуклеотидной последовательности сателлитной ДНК.

Сателлитная ДНК млекопитающих состоит из иерархически организованных повторов

У млекопитающих, как это показано на примере грызунов, нуклеотидные последовательности, входящие в состав тандемных повторов каждой сателлитной ДНК, сильно различаются. Общие для всех тандемов короткие последовательности могут быть обнаружены по их преобладанию среди олигонуклеотидных фрагментов, полученных при химической или ферментативной обработке ДНК. Однако преобладающая короткая последовательность обычно составляет только небольшую часть копий. Другие короткие последовательности отличаются от нее множеством замен оснований, делеций и вставок.

Ряд таких вариантов короткой повторяющейся единицы может сформировать более длинную, которая сама по себе тандемно повторяется с некоторыми изменениями. Таким образом, сателлитные ДНК млекопитающих состоят из иерархически организованных повторяющихся единиц. Более длинные повторяющиеся единицы входят в состав последовательностей, ренатурирующих при проведении реассоциации. Их также можно обнаружить при расщеплении ДНК рестриктазами.

Некоторая неточность при оценке сложности повторяющейся ДНК обусловлена неправильным спариванием при реассоциации сходных, но не идентичных повторяющихся единиц. Этот эффект особенно выражен в случае сателлитной ДНК, и это означает, что при оценке длины реассоциирующей единицы можно ошибиться примерно в два раза.

Когда сателлитную ДНК расщепляют ферментом, сайт узнавания которого находится в повторяющейся единице ДНК, из каждой повторяющейся единицы будет

Таблица 24.1
Сателлитные ДНК¹ *D. virilis*

Фракция сателлитной ДНК	Основная последовательность	Число копий	Относительное содержание ДНК, %
I	A C A A A C T T G T T T G A	1,1 · 10 ⁷	25
II	A T A A A C T T A T T T G A	3,6 · 10 ⁶	8
III	A C A A A T T T G T T T A A	3,6 · 10 ⁶	8
Скрытая	A A T A T A G T T A T A T C		

¹ Более 95% каждой фракции сателлитной ДНК состоит из тандемно повторяющейся основной последовательности. Обратите внимание на то, что фракции II и III имеют в точности совпадающий состав нуклеотидных оснований (1 G—C-пара из 7, т.е. 14% G—C), однако значения их плавучих плотностей равняются соответственно 1,688 и 1,671 г·см⁻³.

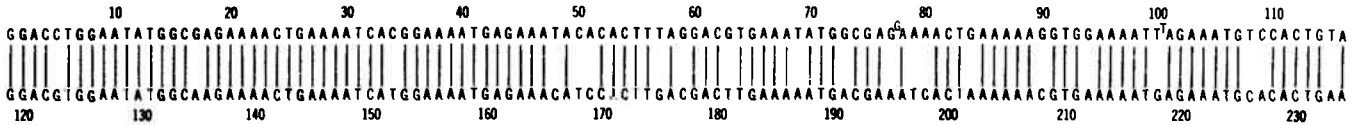


Рис. 24.4. Повторяющаяся единица сателлитной ДНК мыши содержит два полуповтора, гомологичные участки которых обозначены вертикальными линиями.

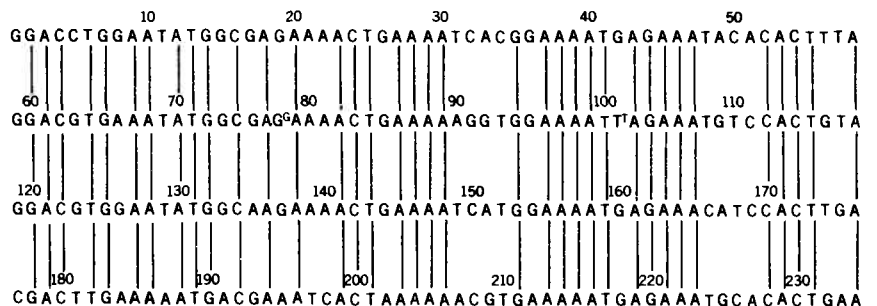
получен один фрагмент, в котором находится этот сайт. Действительно, при обработке эукариотической ДНК рестриктазами основная ее часть при электрофорезе имеет вид пятна вследствие случайного распределения сайтов рестрикции. Однако сателлитная ДНК образует четкие полосы, поскольку при расщеплении ДНК по сайтам рестрикции, расположенным через регулярные интервалы, образуется множество одинаковых или почти одинаковых фрагментов.

Прямое определение нуклеотидной последовательности сателлитной ДНК может оказаться сложным. Можно попытаться сделать это, используя дискретный набор последовательностей, получаемых при расщеплении ДНК рестриктазами. Однако при наличии существенных различий между нуклеотидными последовательностями отдельных повторяющихся единиц в одном и том же положении разных повторов будут находиться разные нуклеотиды, поэтому при электрофорезе получится нечеткая картина. Если различия не слишком велики, скажем находятся в пределах 20%, то можно непосредственно выявить повторяющуюся единицу усредненного состава.

Отдельные участки сателлитной ДНК могут быть встроены в плазмиды с целью клонирования. Трудность в этом случае состоит в том, что последовательности сателлитной ДНК иногда имеют тенденцию исчезать из химерных плазмид при рекомбинации в бактерии-хозяине. Однако при успешном клонировании можно однозначно определить нуклеотидную последовательность клонированного фрагмента. Таким способом можно установить реальную нуклеотидную последовательность повторяющейся единицы или единиц, но при этом для определения характера различий, типичных для сателлитной ДНК в целом, необходимо располагать данными о большом числе таких индивидуальных последовательностей.

При любом методе определения последовательности ДНК можно получить информацию об участке ДНК, который может быть исследован в пределах одного секвенирующего геля. Поскольку повторяются различающиеся тандемные копии, невозможно реконструировать более длинные последовательности путем получения перекрывающихся индивидуальных рестриктов.

Рис. 24.5. При расположении четвертьповторов один под другим обнаруживается гомология между первой и второй половинами каждого полуповтора.



Сателлитная ДНК мыши *M. musculus* расщепляется рестриктазами EcoRII или Sau96I на ряд фрагментов, включая основной мономерный фрагмент, электрофоретическая подвижность которого соответствует длине 230–240 п.н. При исследовании этого фрагмента обнаруживается одна нуклеотидная последовательность длиной 234 п.н. Эта последовательность, по-видимому, повторяется с небольшими вариациями и составляет 60–70% сателлитной ДНК, расщепляемой с образованием мономерных фрагментов. Проанализируем эту последовательность с точки зрения составляющих ее следующих друг за другом повторяющихся единиц меньшего размера.

На рис. 24.4 показана последовательность, содержащая два полуповтора. Изобразив последовательность длиной 234 п.н. так, чтобы первые 117 п.н. располагались напротив вторых 117 п.н., мы увидим, что две ее половины обладают большим сходством. Они различаются 22 нуклеотидными основаниями, что соответствует степени дивергенции этих последовательностей, равной 19%. Это означает, что существующая в настоящее время повторяющаяся единица размером 234 п.н., по-видимому, когда-то в прошлом возникла в результате дубликации повторяющейся единицы размером 117 п.н., после чего произошло накопление различий между двумя копиями.

Внутри последовательности размером 117 п.н. можно обнаружить две другие субединицы. Каждая из них – это четвертая часть повторяющейся единицы, входящей в состав сателлитной ДНК. На рис. 24.5 показаны четыре фрагмента, составляющие четвертьповтора. Две верхние строки соответствуют первому повтору, изображенному на рис. 24.4, две нижние – второму. Видно, что дивергенция между четырьмя четвертьповторами возросла до различий в 23 положениях из 58, или 40%. При этом первые три четвертьповтора обладают несколько большим сходством, а значительная степень дивергенции обусловлена изменениями в четвертом четвертьповторе.

Рассматривая фрагменты, составляющие четвертую часть повторяющейся единицы, можно обнаружить, что каждый из них состоит из двух сходных субединиц (составляющих каждая одну восьмую часть повтора), обозначенных на рис. 24.6 как α - и β -последовательности. Во

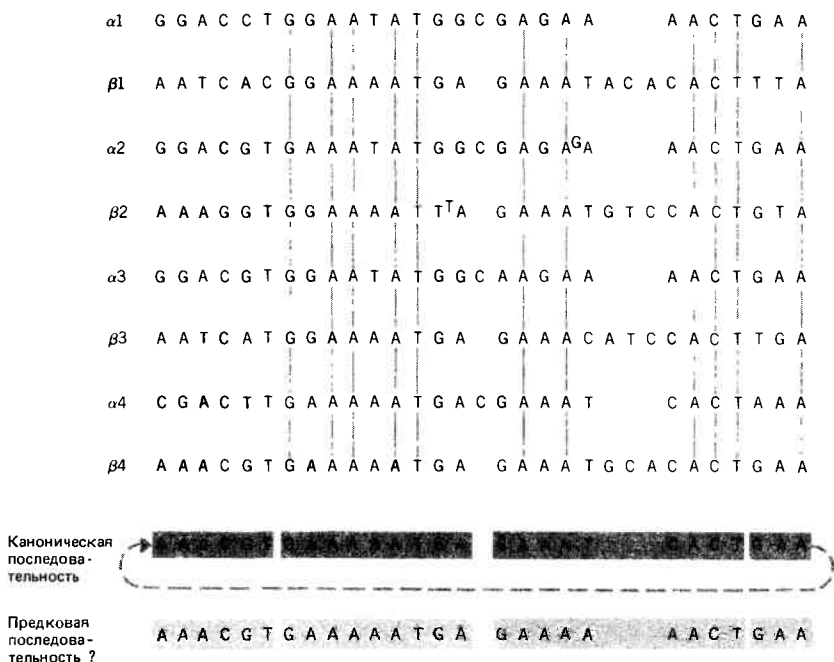


Рис. 24.6. При расположении одного под другим фрагментов, составляющих одну восьмую часть повтора, обнаруживается, что каждый четвертьповтор состоит из α - и β -половин. В каждом положении канонической последовательности расположено наиболее часто встречающееся в этом месте нуклеотидное основание. В составе «предковой» последовательности имеются три фрагмента размером 9 п.н., обладающие большим сходством с канонической последовательностью, которая могла быть предшественником α - и β -единиц.

всех α -последовательностях имеется вставка С, а во всех β -последовательностях – вставка тринуклеотида, сходная с канонической последовательностью. Это позволяет предположить, что четвертьповторы образовались в результате дупликации последовательности, сходной с канонической, после чего в них произошли изменения, которые привели к образованию компонентов, обозначенных нами как α и β . В дальнейшем изменения между тандемно повторяющимися последовательностями $\alpha\beta$ накапливались, и образовывались отдельные четверть- и полуповторы, существующие в настоящее время. Степень дивергенции между существующими сейчас участками, составляющими одну восьмую часть повтора, равняется $(\frac{19}{31}) \cdot 100 = 61\%$.

В канонической последовательности можно выделить три сходные последовательности размером по 9 п.н. (Не следует забывать, что нуклеотидная последовательность сателлитной ДНК непрерывна; поэтому, чтобы установить нуклеотидный состав канонической последовательности, можно рассматривать ее как структуру с циклическими перестановками, как это показано на рисунке присоединением последнего триплета GAA к первым шести нуклеотидным парам.)

Если в одном из повторов наиболее часто встречающееся основание заменить в двух положениях на следующее по частоте встречаемости, то мы получим три близкие последовательности длиной 9 п.н.

```

G A A A A A C G T
G A A A A A T G A
G A A A A A A C T

```

Это предположение, которое можно сделать относительно исходного состава сателлитной ДНК.

На рис. 24.7 приведены данные по непосредственному определению канонической последовательности, при котором обнаруживается, что существующая в настоящее время нуклеотидная последовательность сателлитной ДНК, по-видимому, произошла от рассмотренной после-

довательности длиной 9 п.н. Каноническая последовательность целиком может быть записана как $G A A A A A T G A$, и она фактически представляет собой мозаику трех повторов размером 9 п.н., приведенных на рис. 24.6.

Реконструкция этапов эволюции сателлитной ДНК мыши

По нуклеотидной последовательности сателлитной ДНК мы можем воссоздать картину ее эволюции и объяснить присущие ей в настоящее время свойства. Модель, демонстрирующая возможные этапы такой эволюции, приведена на рис. 24.8.

Эта модель основана на том, что в разные моменты может внезапно происходить латеральная амплификация группы повторяющихся единиц с образованием большого числа идентичных тандемных копий. Такое событие называется **скачкообразной репликацией**. Далее в результате последующего накопления мутаций копии перестают быть идентичными. Позднее в группе таких копий снова может произойти скачкообразная репликация. Степень различия между копиями, амплифицировавшимися в каждый момент времени, будет зависеть от времени, прошедшего с момента последней скачкообразной репликации, когда образовался набор идентичных сателлитных ДНК. Она может эволюционировать путем ряда таких скачкообразных репликаций, чередующихся с накоплением мутаций.

Предположим, что существующая в настоящее время сателлитная ДНК произошла от тандемно повторяющейся последовательности с нуклеотидным составом $G A A A A A T G T$ или близким к нему. (Эта последовательность могла входить в состав сателлитной ДНК или иметь совершенно иное происхождение; она существует столько же времени, сколько и сам сателлит.) Все исходные фрагменты размером 9 п.н. были идентичны, но со временем в результате накопления мутаций между ними появились различия. Потом произошла амплифика-

ция трех расположенных рядом фрагментов вместе с исходной гипотетической нуклеотидной последовательностью, приведенной на рис. 24.6, и образовался tandemный повтор размером 27 п. н.

В этой структуре происходили мутации, в том числе и такие, когда в один из фрагментов встраивался дополнительный остаток С, тогда как в соседнем происходила вставка триплета (неизвестного нуклеотидного состава). Такая пара повторов, составляющих теперь вместе в длину 58 п. н., подвергалась скачкообразной репликации, в результате чего образовался сателлит, состав которого можно обозначить как $(\alpha\beta)_n$.

И снова происходит накопление точковых мутаций, делеций и вставок в сателлитной ДНК, приводящее к возникновению различий между входящими в ее состав повторяющимися единицами. Две расположенные рядом пары участвуют в следующей скачкообразной репликации с образованием повторяющейся единицы $(\alpha_c \beta_c \alpha_d \beta_d)$ размером 116 п. н. В результате дальнейших мутаций пара расположенных рядом единиц амплифицировалась с образованием существующей ныне сателлитной ДНК.

Различия в существующей в настоящее время повторяющейся единице сателлитной ДНК

Свойства сателлитной ДНК объясняются строением типичной последовательности мономерного фрагмента, получаемого при расщеплении сателлита мыши рестриктазой *Sau*. При расщеплении рестриктазой обнаруживается самая длинная повторяющаяся единица размером 234 п. н. Единица реассоциации отдельных цепей денатурированной сателлитной ДНК – это, по-видимому, полуповтор размером 117 п. н., поскольку фрагменты размером 234 п. н. могут реассоциировать как полностью, так и наполовину (в последнем случае первый полуповтор одной цепи ренатурирует со вторым полуповтором другой цепи). Наиболее распространенные фрагменты олигонуклеотидного гидролизата, составляющие 4% от общего количества ДНК, – это GA_5TGA , GA_4TGA и GA_4CTGA , каждый из которых может быть обнаружен во фрагменте размером 234 п. н. и обладает сходством с предполагаемой родоначальной единицей.

До сих пор мы рассматривали существующую в настоящее время сателлитную ДНК так, как будто она состоит из идентичных копий повторяющейся единицы размером 234 п. н. Хотя эта единица и составляет основную часть сателлитной ДНК, обнаруживаются также и ее разновидности. Некоторые из них распределены в сателлитной ДНК случайно, другие организованы в виде кластеров.

Существование разновидностей повторяющейся единицы подразумевалось, когда мы называли исходный материал для анализа нуклеотидной последовательности сателлитной ДНК «мономерным» фрагментом. При расщеплении сателлитной ДНК ферментом, для которого имеется один сайт рестрикции в последовательности размером 234 п. н., образуются также димеры, тримеры и тетрамеры, обладающие сходством с последовательностью размером 234 п. н. Это происходит, когда повторяющаяся единица в результате мутации утрачивает сайт рестрикции.

Мономерная единица размером 234 п. н. образуется в том случае, когда каждый из двух соседних повторов

1	2	3	4	5	6	7	8	9	
G G A C C T									
G	G	A	A	T	A	T	G	G	C
G	A	G	A	A	A	A	C	T	
G	A	A	A	A	T	C	A	C	
G	G	A	A	A	A	T	G	A	
G	A	A	A	T*	C	A	C	T	
T	T	A	G	G	A	C	G	T	
G	A	A	A	T	A	T	G	G	C
G	A	G	G _A	A	A	A	C	T	
G	A	A	A	A	A	G	G	T	
G	G	A	A	A	A	T	T	A	
G	A	A	A	T*	C	A	C	T	
G	T	A	G	G	A	C	G	T	
G	G	A	A	T	A	T	G	G	C
A	A	G	A	A	A	A	C	T	
G	A	A	A	A	T	C	A	T	
G	G	A	A	A	A	T	G	A	
G	A	A	A	C*	C	A	C	T	
T	G	A	C	G	A	C	T	T	
G	A	A	A	A	A	T	G	A	C
G	A	A	A	T	C	A	C	T	
A	A	A	A	A	A	C	G	T	
G	A	A	A	A	A	T	G	A	
G	A	A	A	T*	C	A	C	T	
G	A	A							
G ₂₀	A ₁₆	A ₂₁	A ₂₀	A ₁₂	A ₁₇	T ₈	G ₁₁	A ₅	
				T ₇	C ₅	A ₈	C ₉	T ₁₅	
						C ₇			

* указывает на включение триплета нуклеотидов в β-последовательность

Основание С в положении 10 – дополнительное основание в α-последовательности

Рис. 24.7. Наличие полной канонической (обобщенной) последовательности можно показать, записав нуклеотидный состав сателлитной ДНК в виде повторяющихся фрагментов размером 9 п. н.

имеет сайт узнавания фермента. Димер возникает, когда одна единица утратила сайт узнавания, тример – когда две соседние единицы утратили этот сайт, и т. д. При обработке рестриктазами *EcoRII* или *Sau96I* основная часть сателлитной ДНК расщепляется на ряд групп таких повторяющихся единиц, как показано для ДНК, изображенной на рис. 24.9. Уменьшающееся число димеров, тримеров и т. д. указывает на то, что повторы, из которых в результате мутации исчез сайт узнавания фермента, распределены в сателлитной ДНК случайно.

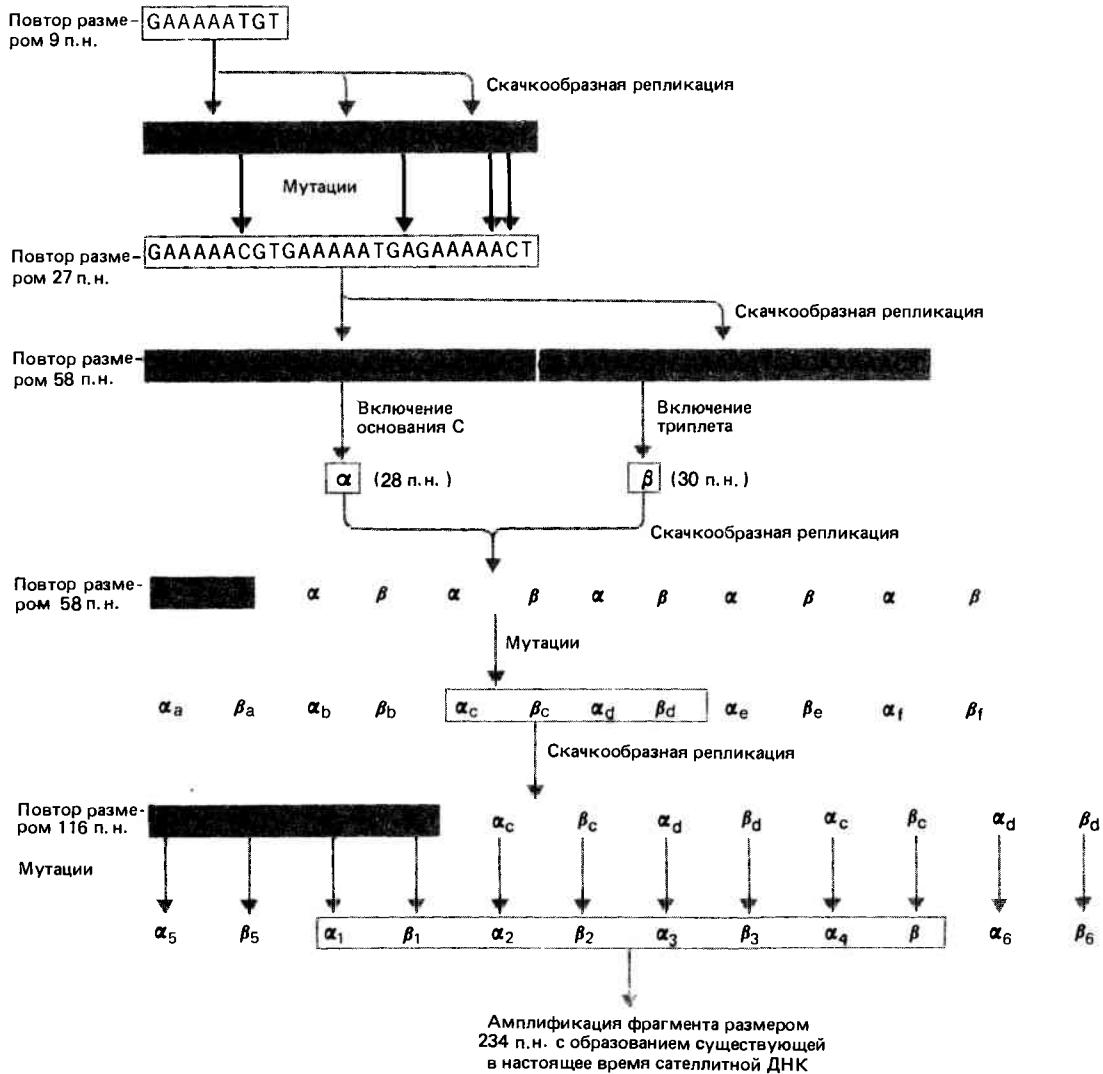


Рис. 24.8. Эволюция сателлитной ДНК мыши могла осуществляться путем чередования скачкообразной репликации и накопления мутаций.

Некоторые рестриктазы по отношению к сателлитной ДНК ведут себя иначе. При обработке этими ферментами образуются последовательности одного и того же размера. Но они расщепляют только небольшую часть ДНК, скажем 5–10%. Из этого следует, что повторяющиеся единицы, имеющие такой специфический сайт рестрикции, сосредоточены в определенном участке сателлитной ДНК. По-видимому, все группы повторов в этом домене произошли от родоначальной последовательности (утратили его в результате мутаций).

Роль неравного кроссинговера

В случае неповторяющейся ДНК рекомбинация осуществляется между точно соответствующими друг другу точками двух гомологичных хромосом; при этом образуются реципрокные рекомбинантные хромосомы; в общих чертах мы обсудили это в гл. 1 и детально рассмотрим на молекулярном уровне в гл. 35. Точность спаривания обеспечивается тем, что две последовательности двухцепочечной ДНК выстраиваются точно одна против другой. Мы знаем, что неравная рекомбинация

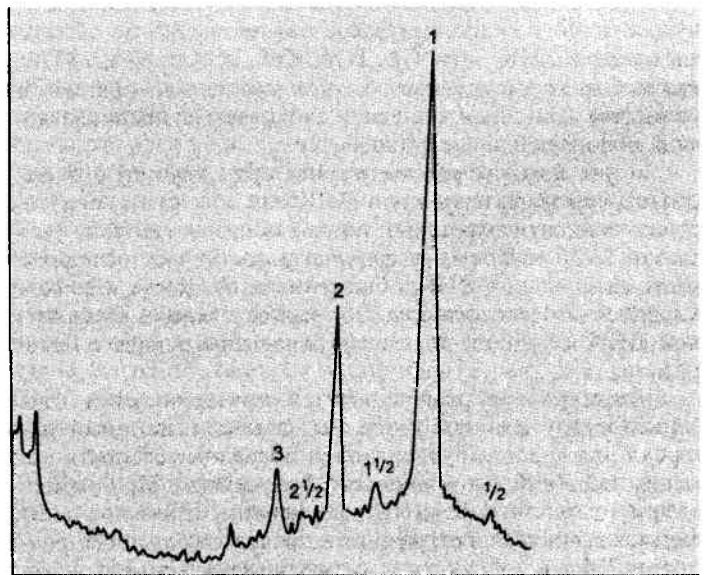


Рис. 24.9. При расщеплении сателлитной ДНК мыши рестриктазой EcoRII обнаруживаются группы повторяющихся единиц (1, 2, 3), представляющие собой мультимеры фрагмента размером 234 п. н., а также небольшие группы (1/2, 1 1/2, 2 1/2), включающие полуповторы (см. ниже). Пик, находящийся далеко слева, соответствует фракции, устойчивой к действию рестриктаз.

может осуществиться при наличии множества копий генов со сходными экзонами, даже в том случае, если их фланкирующие последовательности и интроны различаются (гл. 21). Это происходит вследствие неправильного спаривания соответствующих экзонов *неаллельных* генов. Можно представить себе, насколько чаще наблюдается неточное расположение цепей ДНК друг против друга в тандемных кластерах идентичных или почти идентичных повторяющихся последовательностей. За исключением концов кластера, наличие взаимосвязи между следующими друг за другом повторами может даже сделать невозможным выявление точно соответствующих повторов.

Рассмотрим последовательность, состоящую из повторяющихся единиц «аб» с концами «х» и «у». Если обозначить одну хромосому черным цветом, а другую красным, то при точном расположении «аллельных» последовательностей одна под другой они будут иметь следующий вид:

хабабабабабабабабабабабабабабабабабу
 хабабабабабабабабабабабабабабабабабу

Но предположим, что *любая* последовательность *аб* одной хромосомы может спариваться с *любой* последовательностью *аб* другой хромосомы. В случае неправильного расположения цепей ДНК, такого, как

хабабабабабабабабабабабабабабабабабу
 хабабабабабабабабабабабабабабабабу

образовавшийся двухцепочечный участок будет не менее стабилен, чем при расположении спаривающихся последовательностей одна строго против другой, хотя, конечно, он будет короче. Немногое известно о том, как иницируется спаривание оснований, предшествующее рекомбинации, но весьма вероятно, что оно начинается со спаривания коротких соответствующих друг другу участков, а затем захватывает и другие участки ДНК. Если этот процесс начинается с сателлитной ДНК, то в него скорее всего будут вовлечены повторяющиеся единицы, локализация которых в соответствующих кластерах точно не совпадает.

Теперь предположим, что рекомбинация происходит на участке, где основания спарены со смещением. В одном случае кластер станет длиннее, а в другом короче:

хабабабабабабабабабабабабабабабабабу
 ×
 хабабабабабабабабабабабабабабабабу
 ↓
 хабабабабабабабабабабабабабабабабабу
 +
 хабабабабабабабабабабабабабабабабу

где «х» обозначает место кроссинговера.

Если такой процесс носит общий характер, то размер кластеров тандемных повторов будет постоянно увеличиваться или уменьшаться. К сожалению, мы не располагаем достаточным количеством данных о степени разли-

чий размеров кластеров сателлитной ДНК индивидуальных геномов одного вида организмов.

Неравная рекомбинация приводит к другому результату при наличии в повторяющейся единице внутренних повторов. В только что приведенном примере два кластера располагаются друг под другом со сдвигом по отношению к расположению повторяющихся единиц внутри каждого кластера, но **совпадают во всех точках**, что и видно по расположению отдельных повторов *аб* и *аб*. Но предположим, что и сами компоненты *а* и *б* повторяющейся единицы обладают достаточно большим сходством, чтобы они могли спариваться. Тогда два кластера могут располагаться с таким **сдвигом**, что последовательность *а* одного кластера будет располагаться под последовательностью *б* другого. Частота такого события будет определяться степенью сходства между двумя половинами повторяющейся единицы. (Мы уже упоминали о том, что *in vitro* реассоциация между цепями денатурированной сателлитной ДНК часто происходит по второй схеме.)

В результате рекомбинации происходит изменение длины повторяющихся единиц, участвующих в этом процессе.

хабабабабабабабабабабабабабабабабу
 ×
 хабабабабабабабабабабабабабабабабу
 ↓
 хабабабабабабабабабабабабабабабабу
 +
 хабабабабабабабабабабабабабабабабу

В верхнем кластере, участвующем в рекомбинации, произошла замена участка *а* на участок *aab*. В нижнем кластере участок *аб* заменен на участок *б*.

Такая рекомбинация объясняет картину, получаемую при обработке сателлитной ДНК мышцы рестриктазами. На рис. 24.9 показано, что помимо групп целых повторяющихся единиц в меньшем количестве представлены $\frac{1}{2}$,

$1\frac{1}{2}$, $2\frac{1}{2}$ и $3\frac{1}{2}$ от длины повторяющейся единицы. Пред-

положим, что в предыдущем примере *аб* – повторяющаяся единица сателлитной ДНК мышцы размером 234 п.н., образовавшаяся при расщеплении по сайту, находящемуся во фрагменте *б*. Фрагменты *а* и *б* соответствуют полуповторам размером 117 п.н.

Тогда в верхнем рекомбинантном кластере участок *aab* соответствует фрагменту, длина которого в $1\frac{1}{2}$ раза

больше размера обычной повторяющейся единицы. А в нижнем рекомбинантном кластере участок *б* соответствует фрагменту, длина которого составляет $\frac{1}{2}$ длины

повторяющейся единицы. (Более длинные фрагменты, составленные из полуповторов, образуются так же, как и более длинные фрагменты, составленные из целых повторов. Это происходит в результате утраты при мутации некоторыми повторяющимися единицами сайтов рестрикции.)

И наоборот, можно утверждать, что анализ фракций, составленных из полуповторов, при гель-электрофорезе

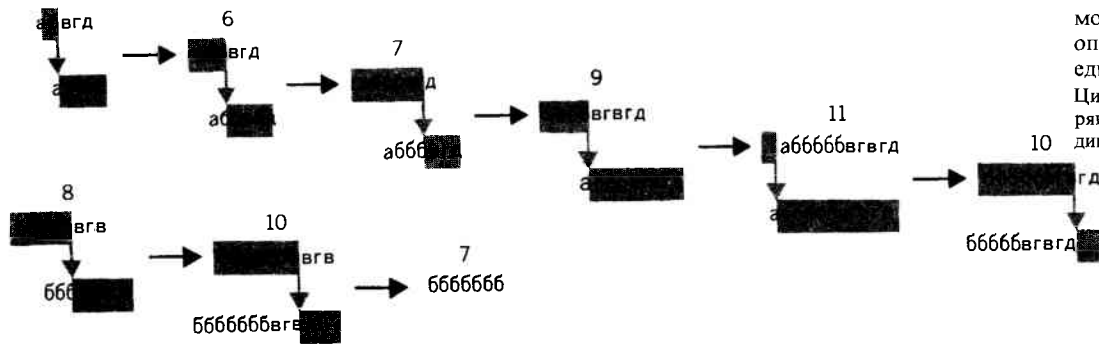


Рис. 24.10. В результате неравной рекомбинации весь кластер может образоваться из одной определенной повторяющейся единицы. Цифрами показан размер повторяющейся единицы на каждой стадии.

показал, что повторяющаяся единица сателлитной ДНК размером 234 п.н. состоит из двух полуповторов, обладающих достаточным сходством, чтобы иногда спариваться и участвовать в рекомбинации. На рис. 24.9 видны также несколько слабо выраженных пиков, соответствующих фрагментам, размер которых составляет $\frac{1}{4}$ и $\frac{3}{4}$ от размера основной повторяющейся единицы. Такие фрагменты образуются так же, как и фрагменты, составленные из полуповторов при рекомбинации кластеров со сдвигом между повторяющимися единицами на четверть их длины. Уменьшение частоты пиков, соответствующих фрагментам размером $\frac{1}{4}$ и $\frac{3}{4}$, по сравнению с пиками полуповторов объясняется уменьшением степени сходства между четвертьповторами по сравнению с полуповторами.

Фиксация при кроссинговере может обеспечивать существование идентичных повторов

Основное предположение, касающееся сателлитной ДНК, состоит в том, что ее последовательность не находится под интенсивным давлением естественного отбора (если она, разумеется, вообще находится под каким-либо давлением). В отличие от последовательностей, кодирующих белок, мутации в которых могут привести к инактивации продукта, сателлитной ДНК, для выполнения ее функций (какие бы они ни были), по-видимому, достаточно одного лишь ее присутствия, а не каких-то конкретных особенностей нуклеотидной последовательности. Это предположение хорошо согласуется со структурой тех сателлитных ДНК, повторяющиеся единицы которых похожи, а не идентичны. С того последнего момента, когда сателлитная ДНК состояла из идентичных последовательностей, в ней произошло накопление мутаций. Но как объяснить наличие у членистоногих сателлитных ДНК, основная часть повторяющихся единиц которых остается идентичной? Даже если имеет значение последовательность нуклеотидов сателлитной ДНК, все-таки трудно объяснить, как может действовать отбор на такое большое число копий.

Для объяснения эволюции сателлитной ДНК кроме модели скачкообразной репликации была предложена модель неравной рекомбинации. В ее основе лежит предположение о том, что неравный кроссинговер часто происходит в случайных точках. В результате ряда таких случайных кроссинговеров одна повторяющаяся единица

распространяется на всю сателлитную ДНК. Этот процесс называют **фиксацией при кроссинговере**.

На рис. 24.10 показано, как определенная повторяющаяся единица распространяется на всю сателлитную ДНК. Предположим, что первоначально сателлит состоял из последовательности *авгд*, где каждая буква соответствует одной повторяющейся единице. Разные повторяющиеся единицы обладают достаточно большим сходством для того, чтобы при рекомбинации происходило смещенное спаривание оснований. Затем в результате ряда неравных рекомбинаций размер повторяющегося участка может увеличиться или уменьшиться и, кроме того, одна единица может распространиться на всю сателлитную ДНК и вытеснить все другие.

Фактически из этой модели следует, что *любая* последовательность ДНК, не подверженная давлению естественного отбора, будет вытеснена множеством идентичных тандемных повторов, образовавшихся подобным образом. Важное допущение этой модели состоит в том, что процесс фиксации при кроссинговере должен происходить быстрее, чем мутационный процесс, так, чтобы возникающие мутации либо отбрасывались (в этом случае соответствующие повторы утрачиваются), либо распространялись на весь кластер (как это показано на примере распространения фрагмента *б* на рис. 24.10).

Интересно было бы знать, могли ли области ДНК, содержащие определенные повторяющиеся единицы (как в случае сателлитной ДНК мыши), образоваться путем частичного распространения одной повторяющейся единицы исходной последовательности. Например, можно увидеть, что на промежуточных стадиях распространения определенной последовательности в сателлитной ДНК, что показано на рис. 24.10, обнаруживаются кластеры, в состав которых входят участки, включающие группы фрагментов *б*.

Рекомендуемая литература

Идею о существовании скачкообразной репликации выдвинул Саузерн (*Southern, J. Mol. Biol.*, **94**, 51–70, 1975). Представления о роли неравного кроссинговера были развиты Смитом (*Smith, Science*, **191**, 528–535, 1976). Организация обращенных и тандемных повторов рассматривается в книге Льюина (*Lewin, Gene Expression*, **2**, Eucaryotic Chromosomes, Wiley, New York, 531–569, 1980). Данные об *Alu*-последовательностях рассматриваются Шмидом и Джлинеком (*Schmid, Jelinek, Science*, **216**, 1065–1070, 1982).

Часть VII

СОЗРЕВАНИЕ РНК: ПРОЦЕССИНГ

Если молекулы гетерогенной ядерной РНК, значительно превышающие по размеру мРНК, — это предшественники мРНК, то напрашивается вывод, что, по меньшей мере некоторые повторяющиеся последовательности играют роль элементов, осуществляющих регуляцию процессинга... Один из подходов к решению важнейшей проблемы выяснения взаимоотношений между мРНК и геномом, а именно вопроса о том, каким образом экспрессируются структурные гены, может состоять в определении структуры гетерогенных ядерных РНК-предшественников, в состав которых входят первичные транскрипты, с тем чтобы сопоставить единицу транскрипции как с единицей трансляции, так и с организацией генома.

Бенджамин Льюин, 1975

ОБРАЗОВАНИЕ СТАБИЛЬНОЙ РНК ПУТЁМ РАЗРЕЗАНИЯ И ПОДРАВНИВАНИЯ ПРЕДШЕСТВЕННИКА

Процесс формирования зрелых молекул РНК из предшественника (процессинг) – важнейший этап образования многих видов мРНК. Все стабильные РНК бактерий синтезируются в виде первичных транскриптов, из которых затем должны выделиться зрелые рРНК или тРНК. Аналогичным путем происходит образование и эукариотических рРНК и тРНК. В отличие от этого большинство бактериальных мРНК сами представляют собой первичные транскрипты, так что только в исключительных случаях их трансляция должен предшествовать процессинг. У эукариот, напротив, образование мРНК, кодируемой прерывистым геном, – наиболее сложный процесс из всех видов процессинга РНК, иногда включающий очень большое число реакций сплайсинга, обеспечивающих соединение экзонов путем удаления интронов (гл. 26).

Все реакции процессинга высокоспецифичны; в результате этих реакций образуются зрелые молекулы РНК с уникальными 5'- и 3'-концами. Поэтому возникает вопрос: какие же признаки РНК-субстратов должны узнаваться ферментами процессинга, чтобы процесс созревания происходил с такой точностью?

Реакции процессинга осуществляются специфическими рибонуклеазами, поскольку их протекание зависит от способности фермента расщеплять фосфодиэфирные связи в молекуле РНК. Как и все нуклеазы, рибонуклеазы можно разделить на два класса: экзонуклеазы и эндонуклеазы.

Эндонуклеазы расщепляют определенные связи *внутри* молекул РНК, что приводит к образованию дискретных фрагментов. Эти ферменты участвуют в реакциях разрезания, в которых нуклеотидные последовательности зрелой РНК отделяются от фланкирующих последовательностей.

Экзонуклеазы удаляют остатки с конца молекулы по одному, в результате чего образуются мононуклеотиды. По способу воздействия на субстрат они могут быть разделены на ферменты, действующие **случайно** или **последовательно**. Фермент, действующий на субстрат случайным образом, удаляет азотистое основание из одной молекулы РНК и затем диссоциирует из комплекса с РНК; при осуществлении следующей каталитической реакции фермент может атаковать либо ту же самую, либо другую молекулу РНК-субстрата. Последовательное воздействие на субстрат означает, что фермент остается связанным с одной и той же его молекулой, удаляя основания одно за другим до полного осуществления своей функции.

Экзонуклеазы участвуют в реакциях подравнивания молекул РНК. Ненужные основания отщепляются от молекулы РНК-субстрата по одному. Все охарактеризованные к настоящему времени экзонуклеазы движутся вдоль цепи нуклеиновой кислоты начиная с 3'-конца. Это означает, что таким способом может быть удален избыточный материал, находящийся на 3'-конце, но материал на 5'-конце может быть отделен только путем разрезания.

Помимо участия в специфических реакциях созревания рибонуклеазы осуществляют деградацию «избыточной» РНК. К «избыточным» молекулам относятся зрелые РНК, которые уже выполнили свои функции, а также материал, который образовался при разрезании предшественников. В частности, должны быть разрушены мРНК, время жизни которых истекло (время жизни бактериальных мРНК короче, у эукариот оно длиннее, но в очень редких случаях мРНК достаточно стабильна, чтобы существовать в течение длительного периода времени). Можно предположить, что рибонуклеазы, участвующие в таком превращении, менее специфичны, чем рибонуклеазы, участвующие в определенных реакциях созревания, поскольку их функция заключается в полном разрушении РНК, а не в образовании специфических концов. К этому классу рибонуклеаз относятся и эндорибонуклеазы, и экзорибонуклеазы. В некотором смысле, конечно, деградация РНК до мононуклеотидов представляет собой завершающий этап ее созревания, приводящий к повторной утилизации ее компонентов в соответствующих метаболических путях.

На самом деле рибонуклеазы сильно различаются по своей специфичности. Не известно ни одной рибонуклеазы, которая бы воздействовала на последовательность из нескольких нуклеотидных оснований субстрата. Поэтому их нельзя считать аналогами рестриктирующих нуклеаз (рестриктаз), узнающих специфические последовательности ДНК. Многие рибонуклеазы обладают некоторой небольшой специфичностью в том смысле, что они расщепляют фосфодиэфирные связи в местах расположения определенных оснований. Однако они, по-видимому, значительно более специфичны по отношению к конформации РНК-субстрата.

Узнавание субстрата может состоять в обнаружении специфических особенностей вторичной структуры (таких, как наличие шпилек определенного размера) или в реакции на вторичную или третичную структуру в целом (что, по-видимому, характерно для процессинга тРНК).

E. coli – единственный объект, для которого есть надежда получить достаточное количество данных для определения всего набора клеточных рибонуклеаз. В этом случае не только можно изучать ферменты биохимически, но также можно определить их действие *in vivo* путем селекции бактерий с мутацией по синтезу определенного фермента. В действительности часто бывает необходимо использовать оба подхода, чтобы установить различие между ферментами или показать, что два совершенно разных вида активности присущи одному и тому же ферменту. (Возможно, со временем аналогичные данные будут получены и для дрожжей.)

Мы пока еще точно не знаем, каково число рибонуклеаз в клетках *E. coli*; к настоящему времени известно менее 10 ферментов. По-видимому, путем сочетания активностей небольшого числа рибонуклеаз обеспечивается необходимый набор реакций при процессинге рРНК и тРНК. Хотя мРНК подвергается активной деградации,

этот процесс до сих пор представляется трудным для изучения, и мы все еще не имеем подробной информации об участвующих в нем ферментах.

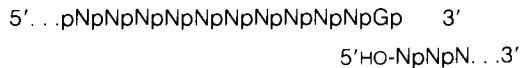
Фосфодиэфирные связи могут расщепляться с обеих сторон

Нуклеазы различаются как по своей ферментативной активности, так и по субстратной специфичности. Фосфодиэфирная связь, связывающая два нуклеотида, в принципе может быть разорвана с любой стороны от фосфатной группы (рис. 25.1). Внесение разрыва по одну сторону от фосфатной группы приводит к образованию 3'-гидроксильных и 5'-фосфатных концов. Разрыв связи с другой стороны приводит к образованию 3'-фосфатных и 5'-гидроксильных концов.

Две рибонуклеазы, Т1 и панкреатическая рибонуклеаза, имеют практическое значение, хотя их физиологическая роль и не изучена. Оба фермента относятся к эндорибонуклеазам, расщепляющим довольно простые последовательности. Рибонуклеазу Т1 получают из *Aspergillus oryzae* (но она имеется и у других грибов и бактерий). Эта рибонуклеаза гидролизует связь у 3'-конца остатка G. Таким образом, она расщепляет следующую последовательность:



с образованием продуктов



Обратите внимание, что в результате образуются новые концы: 3'-фосфатный и 5'-гидроксильный. Такие же концы образуются при действии панкреатической рибонуклеазы, расщепляющей фосфодиэфирную связь у 3'-конца пиримидиновых оснований.

До того как появились современные методы определения последовательностей нуклеиновых кислот, эти ферменты использовали для расщепления РНК до олигонуклеотидов, анализом которых можно было определить нуклеотидную последовательность всей молекулы. В силу своей низкой специфичности они, по-видимому, ответственны за деградацию РНК *in vivo*. Этим может объясняться и то, какие типы концов они образуют.

В естественном состоянии все молекулы РНК полярны; при этом на 3'-конце находится гидроксильная группа, а 5'-конец фосфорилирован (на нем может находиться трифосфат, монофосфат или даже «кэп», в зависимости от того, был ли изменен первоначальный 5'-конец, и если был, то как). Ферменты, участвующие в процессинге РНК, образуют такие же концы. Таким образом, воздействие на РНК всех ферментов, рассмотренных в этой главе, как разрезающих молекулу изнутри, так и отщепляющих концевые нуклеотиды, приводит к образованию 5'-фосфатных и 3'-гидроксильных концов.

РНКаза III «вырезает» ранние мРНК фага Т7 из полицистронного продукта транскрипции

Почти все бактериальные мРНК транслируются в том же виде, в котором они образуются при транскрипции. Большинство из них полицистронны, некоторые моноци-

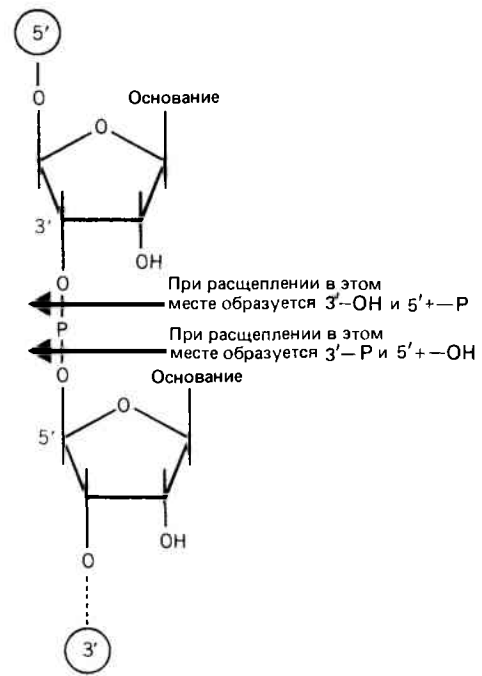


Рис. 25.1. Природа новых концевых участков определяется тем, с какой стороны была расщеплена фосфодиэфирная связь.

стронны. Исключительная ситуация имеет место в случае ранних генов фага Т7. В этом случае при транскрипции образуется полицистронная мРНК. Но затем молекула расщепляется по нескольким сайтам с образованием мРНК, соответствующих индивидуальным генам. (Аналогичное расщепление, возможно, происходит и в случае поздней РНК, но этот процесс менее изучен.)

Такой вид экспрессии помимо фага Т7 обнаружен и у фага Т3. Поскольку нуклеотидные последовательности ДНК двух данных фагов совершенно различны, наличие общего пути созревания мРНК указывает на то, что он обеспечивает определенные преимущества при отборе. Но нам не известно, почему у этих двух фагов существует такой уникальный вариант процессинга.

В гл. 9 мы упоминали об образовании индивидуальных мРНК путем разрезания полицистронных РНК, образующихся в процессе транскрипции вирусов эукариот. В этом случае нам ясна цель: получение моноцистронных мРНК, которые могут быть транслированы. Однако нам ничего не известно о механизме этого процесса.

Транскрипция области ранних генов фага Т7 проиллюстрирована на рис. 25.2. Иницирование транскрипции происходит на одной из трех групп промоторов (A1, A2, A3) и продолжается до терминатора (*t*), расположенного на расстоянии 7000 н.п. Эта область охватывает шесть генов, и для образования пяти индивидуальных мРНК (четырех моноцистронных и одной двухцистронной) молекула разрезается по границам между цистронами.

Разрезание молекулы осуществляется РНКазой III. При транскрипции ДНК фага Т7 *in vitro* РНК-полимеразой *E. coli* образуется одна молекула РНК большого размера, соответствующая всей области ранних генов. Она может быть расщеплена очищенной РНКазой III на индивидуальные мРНК, обнаруживаемые *in vivo*. Мутации у бактерий (называемые *ms⁻*), инактивирующие РНКазу III, *in vivo* блокируют процессинг. Поэтому в мутантных клетках *ms⁻* ДНК фага Т7 транскрибируется с образова-

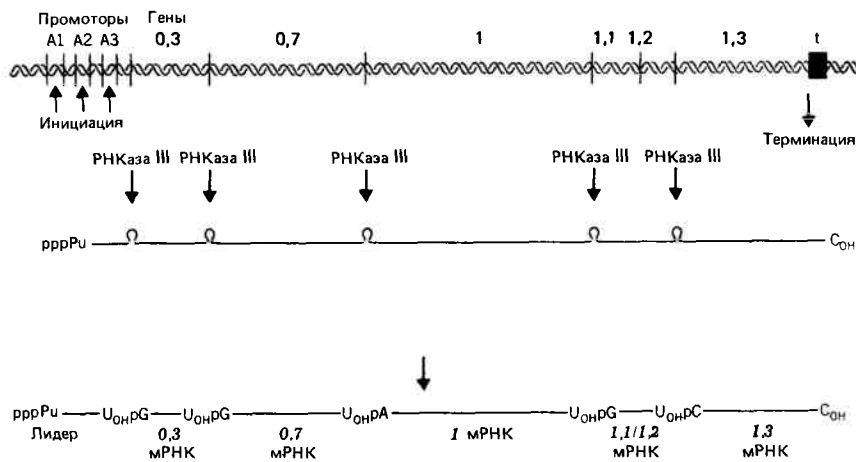


Рис. 25.2. Область ранних генов фага T7 транскрибируется с образованием одной молекулы РНК, расщепляемой РНКазой III на пять индивидуальных мРНК.

нием такой же большой молекулы предшественника, как и *in vitro*. Несмотря на то что в мутантной бактериальной клетке-хозяине предшественник остается нерасщепленным, *in vivo* он может быть транслирован и в виде полицистронной мРНК; в результате образуются обычные ранние белки и таким образом осуществляется процесс инфекции.

Успешная экспрессия ранних генов фага T4 в *ms⁻*-мутантах усиливает наше сомнение в необходимости расщепления. Однако известно, что оно необходимо в случае трансляции двухцистронной 1,1/1,2 мРНК. У мутантных фагов с нарушенным процессом расщепления, когда в состав 1,1/1,2 мРНК входит 29 дополнительных оснований, кодирующие белок участки РНК не транслируются. Причина заключается в том, что более длинная мРНК приобретает вторичную структуру, при которой не может осуществиться инициация транскрипции кодирующей области 1,1 (а нарушение трансляции кодирующей последовательности 1,1 в свою очередь нарушает трансляцию области 1,2). Этот факт был выявлен при попытках выращивания фага T7 на мутантных штаммах *E. coli*, у которых для роста фага необходима экспрессия гена 1,2 (обычно это свойство не является жизненно необходимым). Поэтому в таких условиях, а возможно, и в каких-то других природных условиях расщепление молекулы-предшественника, по-видимому, оказывается необходимой реакцией.

Другой пример влияния РНКазы III на трансляцию можно наблюдать при трансляции РНК фага лямбда, несущей последовательность гена *int*. Эта мРНК обычно не транслируется (белок *int* синтезируется на матрице другой мРНК). Мутации по сайту *sib*, расположенному правее кодирующей последовательности *int*, приводят к трансляции белка *int* на неактивной в обычных условиях мРНК. Ингибиторный эффект последовательности *sib* дикого типа называют ретрорегуляцией. Последовательность *sib* утрачивает способность нарушать трансляцию в клетках-хозяевах *ms⁻*. Это указывает на то, что она может служить мишенью для РНКазы III; при расщеплении РНК в этом месте она либо приобретает конформацию, которая не дает осуществиться иницированию трансляции, либо разрушается. Таким образом, обе мутации *sib* и *ms*, препятствуют расщеплению РНК и способствуют осуществлению трансляции.

В обоих случаях возможный механизм связан с

влиянием разрыва на вторичную структуру РНК. При отделении части молекулы в результате внесения разрыва изменяется структура других частей молекулы, в том числе сайтов, имеющих важное значение для функционирования РНК.

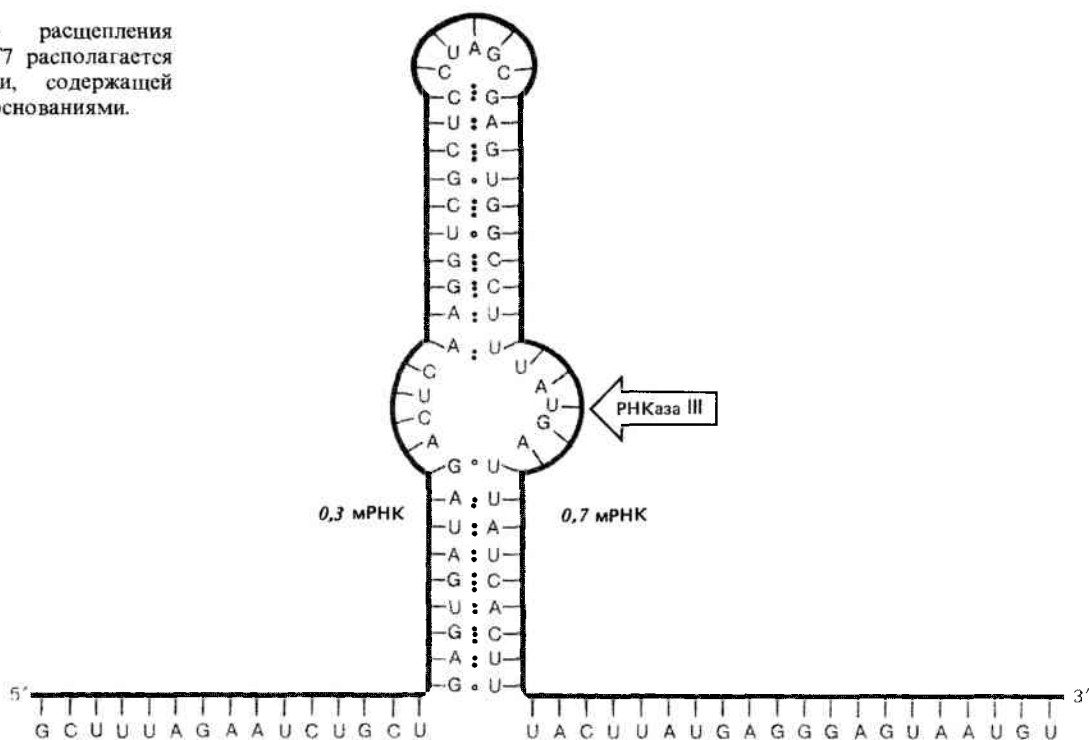
К числу затрагиваемых функций могут относиться иницирование трансляции, чувствительность к деградации, способность к сплайсингу (у эукариот; см. сл. главу) и, без сомнения, другие события. Мы видели в гл. 15, что изменение вторичной структуры приводит к терминированию транскрипции, а в результате и к затуханию следующих за транскрипцией процессов. Вопрос о связи между вторичной (а, возможно, также и третичной) структурой и функцией РНК приобретает все большее значение по мере расширения наших представлений о диапазоне функций, осуществляемых РНК (см. сл. главу).

Какова природа сайтов, узнаваемых РНКазой III? Последовательности-мишени для фермента были обнаружены путем определения нуклеотидных последовательностей 5'- и 3'-концов молекул мРНК, полученных при расщеплении. Эти последовательности можно сопоставить с известной последовательностью ДНК фага T7 и выяснить полное строение участка РНК вокруг сайта расщепления.

Фермент РНКазы III первоначально был обнаружен как эндорибонуклеаза (тогда ее функция была неизвестна), активная в отношении двухцепочечной РНК. Наличие двухцепочечной структуры — важное свойство сайтов узнавания фермента.

Три сайта расщепления в области ранних генов фага T7 обнаруживают высокую степень гомологии, но два других различаются. Наибольшая степень гомологии в случае всех пяти сайтов присуща триплету нуклеотидов, расположенному около сайта расщепления. Все сайты, однако, имеют одну и ту же общую структуру двухцепочечной шпильки, содержащей неспаренное «вздутие». Сайт такого рода изображен на рис. 25.3. РНКазы III вносит разрывы в одну точку каждого из четырех сайтов и в две точки пятого сайта, причем все точки находятся внутри «вздутия» или около него. Единственная информация, которой мы располагаем относительно зависимости процесса расщепления РНК от ее структуры, заключается в том, что мутации, уменьшающие стабильность шпильки, могут уменьшать или нарушать «разрезание» молекулы ферментом.

Рис. 25.3. Типичный сайт расщепления РНКазой III в РНК фага Т7 располагается внутри короткой шпильки, содержащей «вздутие» с неспаренными основаниями.



Рибосомные рРНК образуются из своих предшественников под действием РНКазы III

В клетках бактерии-хозяина РНКазы III принимает участие в процессинге рибосомных рРНК. 16S-, 23S- и 5S-рРНК синтезируются в составе общего первичного транскрипта (см. рис. 23.9). В клетках *trc*⁻ не обнаруживается зрелых рРНК; вместо этого там накапливается 30S-предшественник. Этот предшественник можно расщепить РНКазой III *in vitro* с образованием молекул, являющихся индивидуальными предшественниками зрелых рРНК. В случае 16S- и 23S-рРНК эти предшественники известны под названиями p16 и p23. Каждый предшественник несколько длиннее, чем рРНК, обнаруживаемая в рибосоме, вследствие наличия дополнительных последовательностей как на 5'-, так и на 3'-концах.

Концы p16- и p23-рРНК образуются посредством общего механизма. В каждом случае участки первичного транскрипта, содержащие концы, комплементарны друг другу. Поэтому они могут спариваться с образованием двухцепочечной структуры, содержащей 5'- и 3'-концевые участки продукта. Концевые участки последовательности p16 могут образовать двухцепочечную структуру типа стебля размером 26 п.н. Последовательности, расположенные на концах области p23, могут спариваться с образованием более длинного (хотя и несовершенного) дуплекса. Необычное свойство этого процесса заключается в том, что комплементарные последовательности разделены большим участком размером около 1600 нуклеотидов в p16 и около 2900 нуклеотидов в p23.

На рис. 25.4 показано, что в каждом случае РНКазы III вносит разрывы в обе цепи двухцепочечного стебля, вызывая таким образом одновременное образование 5'- и 3'-концов молекул p16 и p23. Такое действие РНКазы III в двух отношениях отличается от наиболее часто наблюдаемого ее воздействия на РНК ранних генов фага Т7.

В данном случае узнаваемая последовательность

представляет собой непрерывный стебель без какого-либо заметного «вздутия». Кроме того, чаще вносятся два разрыва, а не один (что также происходит и в случае атипичного сайта расщепления фага Т7).

Между сайтами расщепления РНКазы III в молекулах p16 и p23 или между этими сайтами и сайтами расщепления в РНК фага Т7 сколько-нибудь значительная гомология отсутствует. Поэтому вопрос о том, какие свойства молекулы РНК необходимы для ее специфического расщепления РНКазой III, остается загадкой, за исключением общего требования к наличию двухцепочечной структуры вокруг сайта расщепления или в самом сайте. Функционирование фермента может находиться в зависимости от сочетания определенной первичной и вторичной структуры РНК.

Образование сайта узнавания для РНКазы III путем спаривания оснований между 5'- и 3'-концами нуклеотидной последовательности каждого предшественника означает, что каждая реакция процессинга может осуществляться только после того, как синтезировались два конца последовательности-мишени (p16 или p23). *In vivo*, по всей вероятности, РНК связывается с некоторыми рибосомными белками до осуществления процессинга. Однако известно, что это не является необходимым условием для его протекания, поскольку выделенная из клеток 30S-рРНК может быть расщеплена РНКазой III *in vitro*. Таким образом, способность 5'- и 3'-концевых участков к спариванию сохраняется как в отсутствие, так и в присутствии рибосомных белков.

Поскольку расщепление РНКазой III не приводит к образованию истинных 5'- или 3'-концов 16S- или 23S-рРНК, необходимо осуществление дальнейших реакций процессинга. За исключением того, что их протекание возможно после только предварительного воздействия на РНК РНКазы III, об этих реакциях известно немного. Процессинг молекул тРНК, последовательности которых также кодируются участками транскрипционных единиц *rrn*, происходит при участии РНКазы Р и других ферментов процессинга тРНК (об этом см. ниже).

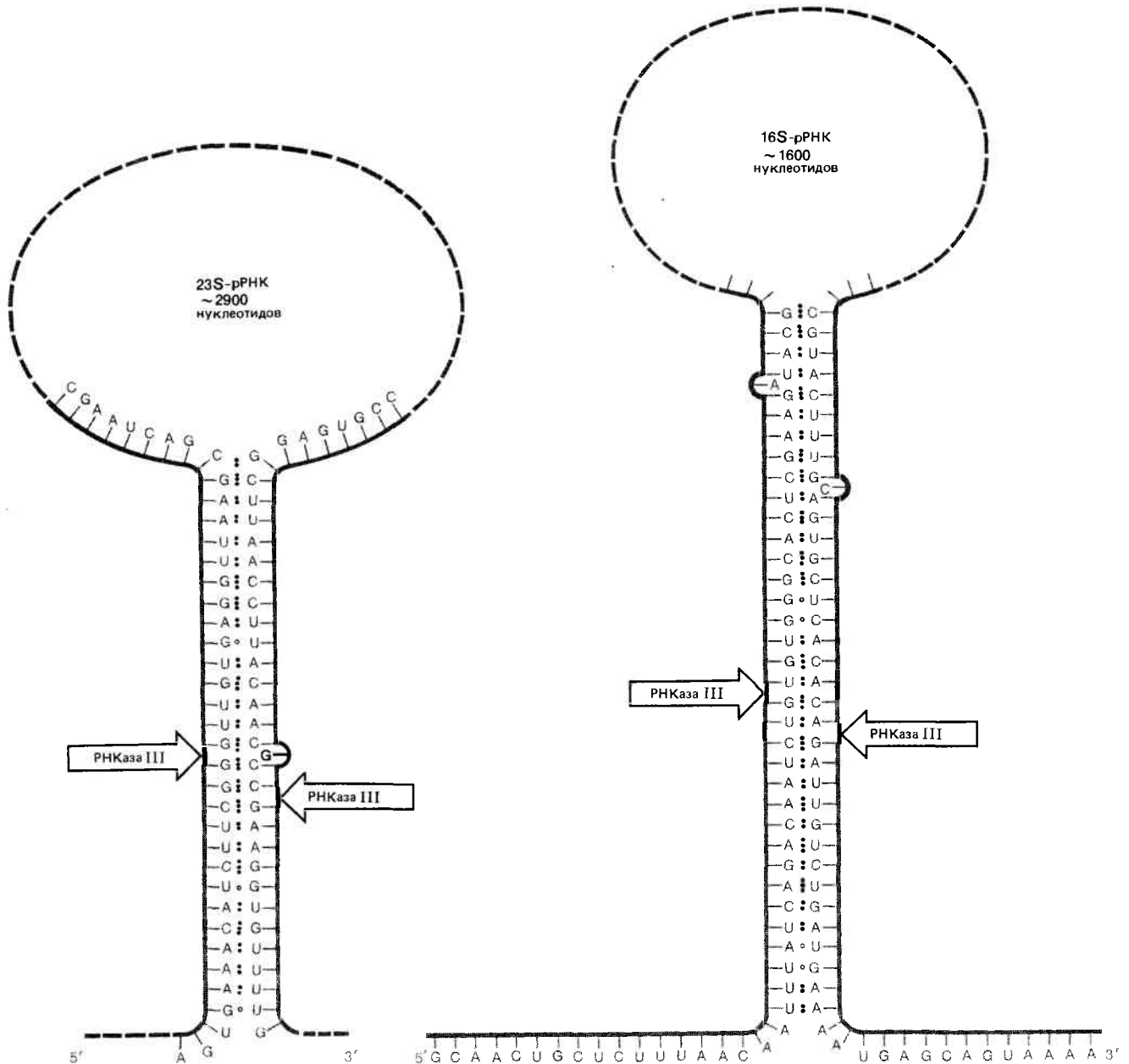


Рис. 25.4. Области, содержащие 5'- и 3'-концевые участки п16- и п23-рРНК, спариваются с образованием двухцепочечных стеб-

лей, каждый из которых расщепляется РНКазой III в двух точках.

Сайты расщепления при созревании эукариотической рРНК

Процесс образования зрелой эукариотической рРНК может быть детально изучен вследствие относительно невысокой скорости созревания первичного транскрипта, включающего ряд промежуточных стадий, с образованием зрелых молекул рРНК. Основные промежуточные продукты процессинга могут быть выделены из многих клеток, в частности из клеток млекопитающих, где процессинг рРНК хорошо охарактеризован.

Первичный продукт транскрипции рибосомных генов млекопитающих — это 45S-рРНК, содержащая последовательности как 18S-, так и 28S-рРНК, но по длине превышающая их суммарную длину почти в два раза (см. табл. 23.2). В клетках HeLa (клетки человека) предше-

ственный рРНК содержит около 110 метильных групп, присоединяющихся во время транскрипции или сразу же после нее. Почти все метильные группы связаны с остатками рибозы в разных олигонуклеотидных последовательностях. В ходе процессинга предшественника с образованием зрелых рРНК метильные группы *сохраняются*. В зрелой 18S-рРНК содержится около 39 из исходно присоединенных к первичному транскрипту групп и, кроме того, еще 4 группы, добавленные позже в цитоплазме. В 28S-рРНК содержится около 74 метильных групп (все они присутствовали еще в первичном транскрипте). Из этого можно сделать вывод, что с помощью метилирования «метятся» участки первичного транскрипта, которые войдут в состав зрелой рРНК.

Обнаружено, что существует не один путь созревания рРНК. Об этом свидетельствуют данные о размерах мо-

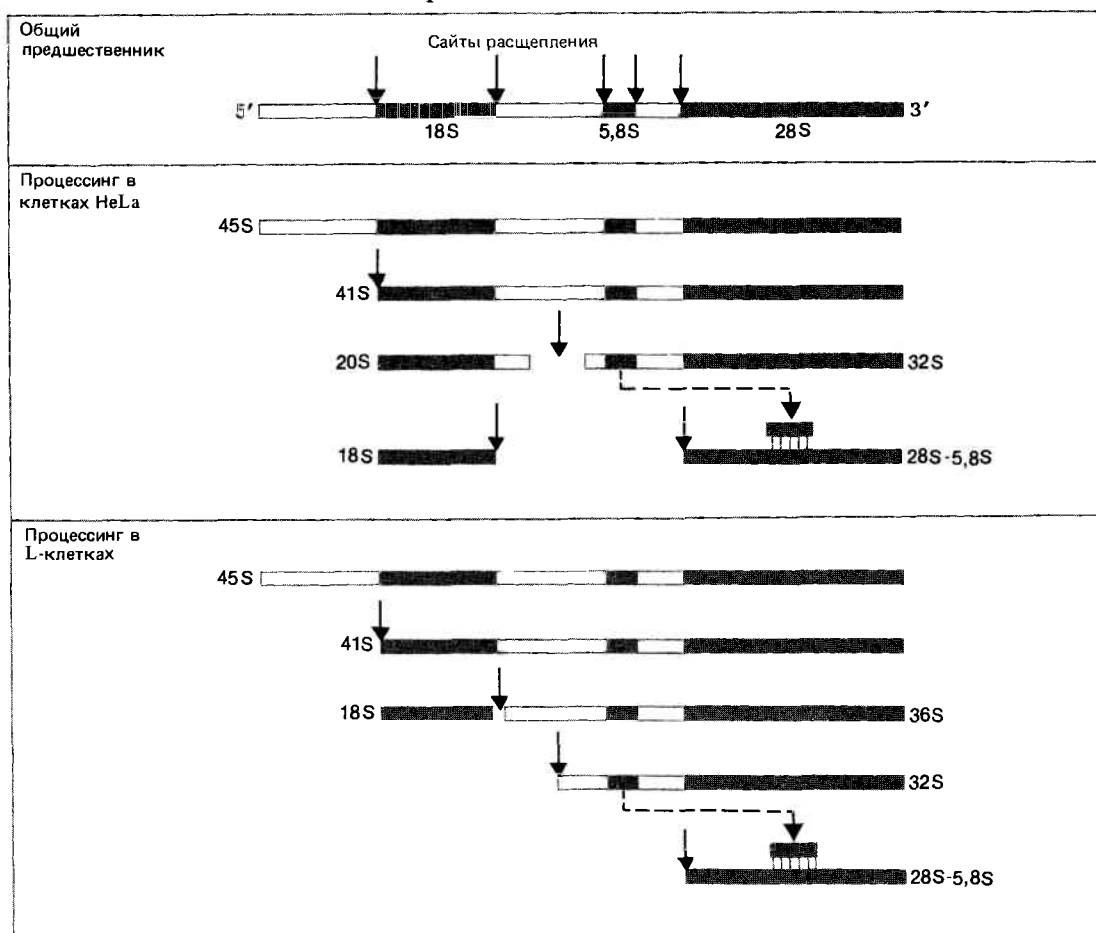


Рис. 25.5. Зрелые рНК млекопитающих образуются из 45S-предшественника в результате внесения разрывов в каждый спейсерный участок.

лекул, соответствующих промежуточным стадиям созревания. Однако все пути созревания можно свести к одной схеме, если предположить, что расщепление происходит по небольшому числу одних и тех же сайтов, но в разном порядке. Два способа осуществления процессинга у млекопитающих, изученные на примере клеток HeLa (человека) и L-клеток (мыши), показаны на рис. 25.5. В каждом случае имеются сайты расщепления, расположенные у 5'-конца гена 18S-рНК, внутри спейсера, разделяющего последовательности 18S- и 5,8S-рНК, и спейсера, разделяющего 5,8S- и 28S-рНК. (Из 5,8S-последовательности образуется низкомолекулярная рНК, связанная с 28S-рНК в результате спаривания оснований.) Различия между двумя способами созревания рНК, показанными на рисунке, заключаются только в порядке использования сайтов расщепления. Возможны также и другие варианты последовательности расщепления. В некоторых случаях в одном и том же типе клеток обнаруживается несколько путей созревания рНК.

На рисунке показано минимальное количество сайтов, при расщеплении которых могут образоваться все 5'- и 3'-концы. Мы не знаем, приводят ли такие расщепления в действительности к образованию концов зрелых рНК или при расщеплении образуются индивидуальные предшественники, которые затем разрезаются или подравниваются. Не имеется также и какой-либо информации об активности рибонуклеаз, ответственных за осуществление этого процесса. Однако известно, что 45S-рНК связы-

вается с белками непосредственно после своего синтеза. Таким образом, процессингу подвергается рибонуклеопротеин, а не свободная рНК.

тРНК разрезаются и подравниваются несколькими ферментами

Процессинг — необходимый этап образования всех известных бактериальных и эукариотических цитоплазматических тРНК. В клетках эукариот новосинтезированные тРНК имеют константу седиментации, равную или превышающую 4,5S, что соответствует молекулам, состоящим примерно из 100 нуклеотидов; зрелая же 4S-тРНК состоит из 70–80 нуклеотидов. У бактерий все тРНК синтезируются в виде предшественников, имеющих дополнительные последовательности на 5'- и 3'-концах. В некоторых случаях один первичный транскрипт может содержать последовательности нескольких тРНК (до шести последовательностей), которые должны быть отделены друг от друга. К настоящему времени известно, что часть молекулы предшественника, соответствующая тРНК, обычно имеет такую же вторичную структуру, как и зрелая тРНК, и не исключено, что это играет важную роль при ее узнавании рибонуклеазами во время созревания. Мутации в участках, соответствующих тРНК, могут нарушать процессинг; в результате дефектная рНК в клетке не будет накапливаться.

Образование 5'-концов всех молекул тРНК в клетках

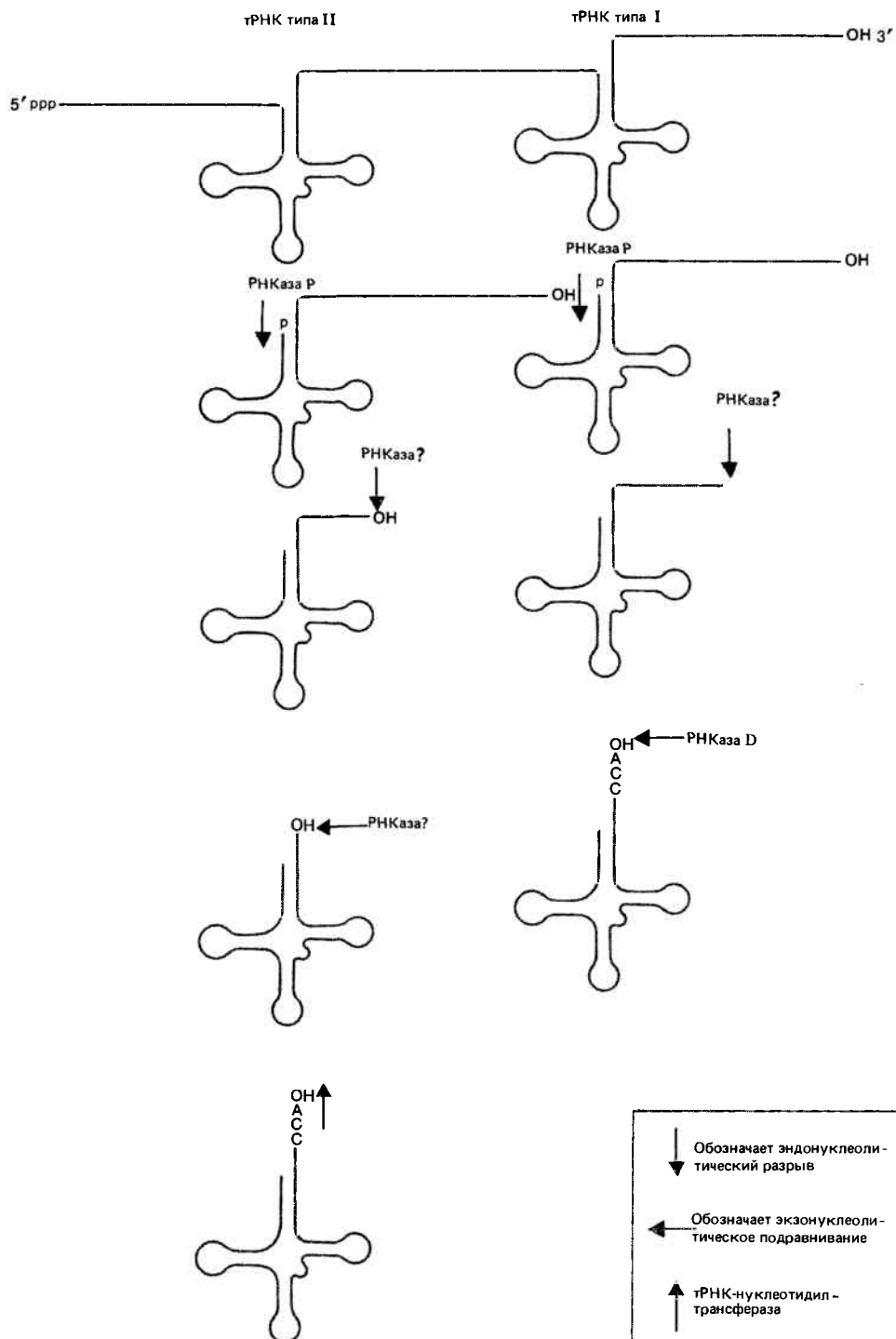


Рис. 25.6. Процессинг (гипотетического) предшественника-димера, содержащего тРНК типа I и типа II, требует присутствия нескольких ферментов.

E. coli, по-видимому, происходит при участии одного-единственного фермента. Это **РНКаза Р**, эндорибонуклеаза, расщепляющая молекулу либо по границе с 5'-лидерной последовательностью, либо по межцистронной последовательности (в том случае, когда тРНК представляет собой часть полицистронного транскрипта). Расщепляемые последовательности первичного транскрипта различаются во всех молекулах-предшественниках. Мутации, нарушающие процессинг тРНК, осуществляемый РНКазой Р, затрагивают несколько участков молекулы.

Из этого можно сделать вывод, что РНКазы Р должны узнавать общие черты третичной структуры тРНК, присутствующие предшественнику и определяемые различными участками тРНК.

Рибонуклеаза Р — необычный фермент. В его состав входит и белок, и РНК.

РНК, входящая в состав РНКазы Р, необходима для проявления каталитической активности. Направляется предположение о том, что ее роль заключается в узнавании тРНК при помощи механизма, включающего спаривание

оснований. Такое взаимодействие может обеспечивать правильную ориентацию белкового компонента для осуществления специфического расщепления субстрата. Потенциальные аналоги РНКазы Р были обнаружены в клетках эукариот.

В клетках *E. coli* имеются два гена, мутации по которым нарушают активность РНКазы Р: *mpA* и *mpB*. Об этих генах известно только то, что они располагаются довольно далеко друг от друга. Может быть, один из них кодирует белковый компонент, а другой – РНК?

У таких мутантов накапливается большое количество молекул-предшественников тРНК. Некоторые полицистронные предшественники расщепляются с образованием индивидуальных моноцистронных предшественников, хотя ни у одного из них не имеется настоящих 5'-концов. Характер расщепления указывает на то, что существует другой фермент, способный вносить разрывы в межцистронные участки. Таким ферментом может быть РНКазы III.

Почему в отсутствие РНКазы Р одни предшественники расщепляются с образованием моноцистронных предшественников, а другие нет? Даже если в расщеплении предшественников всех тРНК участвует один и тот же набор ферментов, порядок, в котором они воздействуют на разные предшественники, может быть различным. Это подтвердилось в ходе изучения индивидуальных предшественников, у которых аналогичные реакции могут осуществляться на разных стадиях процессинга. Необходимость осуществления реакций в определенном порядке может быть обусловлена тем, что в исходном транскрипте расщеплению доступен только один сайт, причем происшедшее расщепление приводит к изменению структуры транскрипта и появляется возможность расщепления по сайту следующего фермента.

Другие этапы созревания тРНК мало изучены. Совокупность основных этапов созревания приведена на рис. 25.6.

Эндонуклеолитическое... расщепление, по-видимому, происходит также и на 3'-конце нуклеотидных последовательностей тРНК. К настоящему времени не известно, какой фермент (ферменты) ответствен за этот процесс.

3'-конец зрелой тРНК образуется при действии экзонуклеазы. Фермент, называемый РНКазой D, удаляет основания по одному с 3'-конца (возможно, с 3'-конца исходного предшественника, а возможно, с 3'-конца молекул, образованных при расщеплении предшественника), пока не достигнет концевой ССА-последовательности, общей для всех зрелых тРНК. По-видимому, сигналом для остановки фермента служит структура тРНК в целом, а не последовательность ССА, поскольку он может удалять дополнительную ССА-последовательность, присоединенную к 3'-концу исходной молекулы. РНКазы D от-

носится к экзонуклеазам, действующим случайным образом.

Другой фермент, обладающий экзонуклеолитической активностью, – РНКазы II. Вначале считали, что РНКазы II принимает участие в процессинге тРНК, но фермент может и полностью разрушать последовательности тРНК. Таким образом, он, по-видимому, осуществляет разрушение тРНК, а не созревание специфических тРНК. Действие фермента состоит в последовательном отщеплении нуклеотидов.

В бактериальных клетках имеются два типа предшественников тРНК, различаемых по их 3'-концевым последовательностям. В молекулах типа I имеется триплет ССА, отделяющий нуклеотидную последовательность зрелой тРНК от дополнительного материала, находящегося на 3'-конце. Молекулы же типа II (кодируемые генами некоторых фагов) не имеют ССА-последовательности. После удаления из молекулы предшественника дополнительных нуклеотидов, расположенных на 3'-конце, к оставшейся части молекулы должна быть присоединена последовательность ССА. Мы не знаем, как образуется 3'-конец, к которому присоединяется ССА, у тРНК типа II, – осуществляет ли этот процесс РНКазы D, как это происходит в случае тРНК типа I, или в нем участвует другой фермент? У эукариот, очевидно, все тРНК относятся к типу II.

Добавление ССА осуществляет фермент тРНК-нуклеотидилтрансфераза, присоединяющая триплет (или его часть) к любой последовательности тРНК, где он отсутствует. Мутанты *E. coli* со сниженным содержанием этого фермента (называемые *cca*⁻) растут медленно, что указывает на необходимость фермента для биосинтеза тРНК. Концевой остаток А у всех бактериальных тРНК подвержен постоянным превращениям; поэтому тРНК-нуклеотидилтрансфераза так же необходима для осуществления репарации, как и для образования тРНК типа II.

Рекомендуемая литература

Имеется несколько обзоров, посвященных анализу ферментов, действительно участвующих в процессинге тРНК и рРНК, а также отличиям этих ферментов от других ферментативных активностей, выявляемых при этом процессе. В работе Гегенхеймера и Апириона сделана попытка классифицировать генетические и биохимические свойства бактериальных ферментов (*Gegenheimer, Apirion, 45, 502–541, 1981*). Статья Брэма, Янга и Стейтца, посвященная РНКазе III, включает также и полученные ранее данные (*Bram, Young, Steitz, Cell, 19, 393–401, 1980*). Процессинг тРНК рассматривается Альтманом (*Altman, Transfer RNA, MIT Press, Cambridge Mass., 48–77, 1978*), современные данные приводятся в «Cell» (*23, 3–4, 1981*).

Глава 26

МЕХАНИЗМЫ СПЛАЙСИНГА РНК

Когда было обнаружено, что прерывистые гены транскрибируются с образованием предшественников РНК, из которых в результате сплайсинга выделяются экзоны, возник ряд вопросов, касающихся экспрессии генов. Рассматривая сплайсинг РНК просто как один из механизмов экспрессии генов, интересно выяснить, какие участки

РНК затрагивает этот процесс по сравнению с другими модификациями РНК («экспириванием» и полиаденилированием). Удаляются ли интроны из предшественника в определенном порядке? Чем объясняется необычайная точность осуществления этого процесса? Происходит ли он в определенном участке ядра, связан ли с другими

процессами, такими, как, например, транспорт из ядра в цитоплазму? Обусловливает ли связь с другими процессами различия в способах экспрессии прерывистых и непрерывных генов? Является ли процесс созревания РНК (путем сплайсинга или других механизмов) фактором регуляции экспрессии генов, осуществляющейся путем отбора имеющихся предшественников?

Сплайсинг РНК включает разрыв фосфодиэфирных связей на границах экзон—интрон и образование связи между концами экзонов. Это может произойти путем либо двух независимых последовательных реакций, либо одной согласованной реакции переноса. Разрыв фосфодиэфирных связей происходит и при других реакциях процессинга, но их образование в молекуле РНК—уникальная особенность сплайсинга. Каковы бы ни были ферментативные механизмы возникновения связей, основное условие этого процесса—расположение рядом групп, реагирующих с образованием новых связей.

По-видимому, не существует какого-либо уникального механизма сплайсинга РНК. Можно выделить по меньшей мере четыре четко различимых типа систем сплайсинга. Удаление интронов из ядерных предшественников тРНК у дрожжей происходит при участии ферментов, взаимодействие которых с субстратом напоминает действие ферментов процессинга тРНК, поскольку решающим фактором в обоих случаях служит конформация предшественника тРНК. Удаление интрона из предшественника рРНК у инфузорий, по-видимому, свойство самой РНК. При удалении интронов из РНК митохондрий у дрожжей единый механизм, вероятно, отсутствует, но известно, что по крайней мере в некоторых случаях необходимо наличие белка, кодируемого интроном, который должен быть удален (гл. 20). Удаление интронов из ядерных РНК высших эукариот происходит при участии ферментной системы, по-видимому, не узнающей никаких более специфических последовательностей, чем короткие канонические последовательности, расположенные на границах экзон—интрон у большого числа видов организмов. Таким образом, если гены с прерывистым строением и имели единое эволюционное происхождение, то с тех пор они сильно дивергировали в плане реализации их генетической информации в виде РНК.

Сплайсинг дрожжевой тРНК включает разрезание и сшивание

Около 40 из ~400 ядерных генов тРНК дрожжей имеют прерывистое строение. В каждом случае в состав гена входит один интрон, расположенный на расстоянии одного нуклеотида от 3'-конца антикодона. Длина интронов варьирует от 14 до 64 п. н. Нуклеотидные последовательности интронов генов, кодирующих тРНК для одной аминокислоты, сходны между собой, но у интронов генов, кодирующих тРНК для разных аминокислот, сходство отсутствует. *Канонической последовательности, которая могла бы узнаваться ферментами сплайсинга, не имеется.* Однако все интроны содержат последовательность, комплементарную антикодону тРНК. Это приводит к изменению конформации антикодонной ветви, при которой основания антикодона спариваются, и в результате ее длина увеличивается. Пример такого рода приведен на рис. 26.1. Меняется только ветвь антикодона, остальная часть молекулы сохраняет свою обычную структуру. Таким образом, в основе сплайсинга ле-

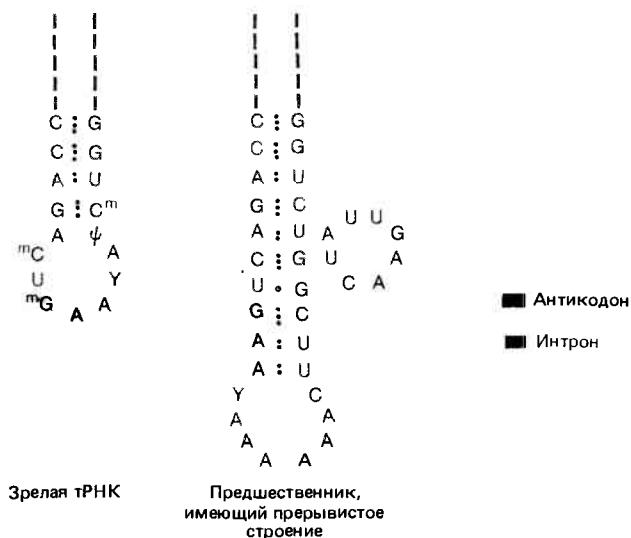


Рис. 26.1. Спаривание интрона дрожжевой тРНК^{Phe} с антикодонном приводит к изменению структуры антикодонной ветви.

жит, по-видимому, узнавание вторичной структуры определенного типа, а не определенной нуклеотидной последовательности.

Удобный объект для изучения сплайсинга—температурочувствительные мутанты дрожжей, у которых не удаляются интроны, и в ядре накапливаются предшественники тРНК, имеющие прерывистое строение. Процессинг фланкирующих последовательностей может происходить обычным способом. Молекулярный механизм возникновения этого дефекта неясен, но известно, что он специфичен для прерывистых генов, поскольку тРНК, кодируемые непрерывными генами, могут созревать и транспортироваться в цитоплазму.

Это свойство оказывается очень полезным, поскольку позволяет выделять предшественники тРНК в такой форме, когда единственное отличие их нуклеотидной последовательности от последовательностей зрелых тРНК состоит в наличии интронов. (У предшественников могут также отсутствовать и некоторые модифицированные основания.) Затем такие молекулы используются в качестве субстратов в бесклеточной системе, полученной из клеток дикого типа. За сплайсингом предшественника можно легко проследить очень просто—определяя степень уменьшения размеров образовавшегося продукта. Об этом уменьшении свидетельствует изменение положения полосы РНК при гель-электрофорезе, как показано на рис. 26.2. Такое уменьшение размеров, по-видимому, объясняется появлением полосы, соответствующей интрону.

Для изучения сплайсинга тРНК можно использовать бесклеточную систему. Реакция *in vitro* требует присутствия АТР, и ее можно разбить на два этапа, осуществляя сплайсинг в отсутствие АТР. Первый этап включает расщепление фосфодиэфирной связи, протекающее как необычная нуклеазная реакция. На втором этапе происходит образование связи; эта реакция называется «сшиванием» (лигированием), а осуществляющий ее фермент называется РНК-лигазой. До сих пор не известно, присуши ли нуклеазная и лигазная активности одному ферменту или разным ферментам. Две стадии реакции сплайсинга изображены на рис. 26.3.

В отсутствие АТР нуклеазная активность приводит к расщеплению предшественника по двум концам интро-

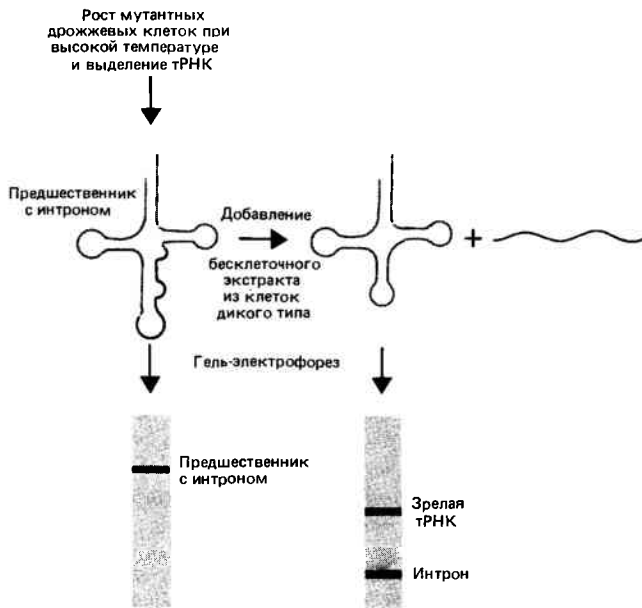


Рис. 26.2. Сплайсинг дрожжевой тРНК *in vitro* можно изучать с помощью гель-электрофореза предшественника тРНК и продуктов сплайсинга.

на. В результате образуются две «половинки» молекулы тРНК и линейная промежуточная последовательность. Эти промежуточные продукты имеют необычные концы: на 5'-конце имеется гидроксильная группа, а на 3'-конце — фосфат. (Как упоминалось в гл. 25, все другие известные ферменты процессинга РНК расщепляют фосфатную связь с другой стороны.)

При добавлении АТФ осуществляется вторая реакция. Основания двух половинок тРНК спариваются с образованием тРНК-подобной структуры. Под действием РНК-лигазной активности происходит ковалентное «сшивание» двух половинок путем образования фосфодиэфирной связи. Для осуществления лигазной реакции необходимо наличие 3'-концевого фосфата.

Циклическая фосфатная группа раскрывается и образуется 2'-фосфатный конец. Эта реакция, по-видимому, требует присутствия циклофосфат-диэстеразы. В состав продукта входит 2'-фосфатная группа и 3'-ОН-группа.

5'-ОН-группа, образовавшаяся под действием нуклеазы, должна превратиться в 5'-фосфатную. Обычно это происходит путем обычной киназной реакции, в которой донорной группой служит γ -фосфат АТФ.

Первый этап лигазной реакции — активация фермента АТФ с образованием аденилированного белка. Фосфатная группа АМР ковалентно связана с аминогруппой белка.

Освобождение АТФ из комплекса с ферментом сопряжено с самой реакцией лигирования. Если эта реакция протекает по тому же пути, что и ранее описанная реакция РНК-лигазы Т4, она будет включать два этапа.

Сначала осуществится перенос АМР с фермента на 5'-фосфатную группу с образованием 5'-5'-фосфатной связи. Затем в результате атаки со стороны 3'-ОН-группы другой половины тРНК происходит вытеснение АМР. Обратите внимание, что фосфатная группа, связывающая два экзона, исходно не входила в состав РНК, а во время реакции была перенесена с АТФ.

Молекула после сплайсинга становится ковалентно-непрерывной с наличием 5'-3'-фосфатной связи в сайте сплайсинга, но она также содержит 2'-фосфатную груп-

пу — маркер осуществления сплайсинга. Эта лишняя группа, по всей вероятности, удаляется фосфатазой. (На промежуточных этапах сплайсинга ее присутствие может быть полезным для обозначения сайта сшивания.)

Образование 2'-3'-циклического фосфата не является уникальным свойством дрожжей. Такая циклическая группа возникает при сплайсинге тРНК у растений и млекопитающих. Реакции ее образования у растений и дрожжей, по-видимому, не различаются.

Однако у млекопитающих осуществляется другая реакция. В клетках человека (HeLa) лигаза непосредственно присоединяет конец РНК с 2'-3'-циклофосфатной группой к концу РНК, несущему 5'-гидроксильную группу.

Сплайсинг предшественников дрожжевой тРНК может осуществляться также и в экстракте, полученном из зародышевых пузырьков (ядер) ооцитов *Xenopus*. Из этого следует, что реакция не является видоспецифической. У *Xenopus*, по всей вероятности, имеются ферменты, способные узнавать интроны дрожжевых тРНК.

Особенности строения предшественника, узнаваемые ферментами сплайсинга, можно выявить, внося изменения в молекулу РНК либо путем манипуляций с клонированным геном *in vitro*, либо с помощью мутаций *in vivo*. На отсутствие необходимости в наличии интрона со строго определенными последовательностью нуклеотидов и размером указывает существование природных различий между генами. Это подтверждается экспериментами, при которых в интрон встраивали дополнительный фрагмент размером 21 п.н., что приводило к увеличению длины дополнительной ветви предшественника, но не влияло на сплайсинг.

Системы для изучения сплайсинга тРНК отличаются тем преимуществом, что легко могут быть получены соответствующие мутанты. Например, существуют дрожжевые клетки, в которых прерывистый ген для тРНК^{Tyr} мутировал с возникновением охра-супрессорной мутации. Из мутантных клеток можно отобрать такие, в которых супрессия не осуществляется. В некоторых из них отсутствует супрессорная тРНК, поскольку мутации нарушили образование предшественника. На рис. 26.4 приведены три типа структур тРНК: предшественника тРНК из клеток дикого типа, предшественника тРНК из клеток с супрессорной мутацией (который подвергается нормальному сплайсингу) и предшественника из мутантных клеток, у которых сплайсинг протекает в семь раз менее эффективно.

Вторичная структура интронов во всех случаях разная. Предшественник из клеток с супрессорной мутацией имеет укороченный двухцепочечный стебель, и один конец интрона входит в состав двухцепочечного участка, а не одноцепочечного. Мутация *al22* затрагивает область разрезания при сплайсинге, приводя к образованию большого неспаренного участка в двухцепочечном стебле; другой конец интрона вследствие спаривания оснований входит в состав двухцепочечного участка, а не одноцепочечной петли. Таким образом, сайты сплайсинга узнаются ферментами независимо от того, располагаются ли они в области спаренных оснований или одноцепочечной области, и даже изменения на границе экзон—интрон приводят всего лишь к снижению эффективности сплайсинга и не приводят к замене сайтов сплайсинга.

Это напоминает выводы, сделанные нами в гл. 7 и 25 относительно структуры тРНК и ее экспрессии. Соответственно в основе реакций сплайсинга лежит узнавание не специфических оснований, а конформации молекулы в це-

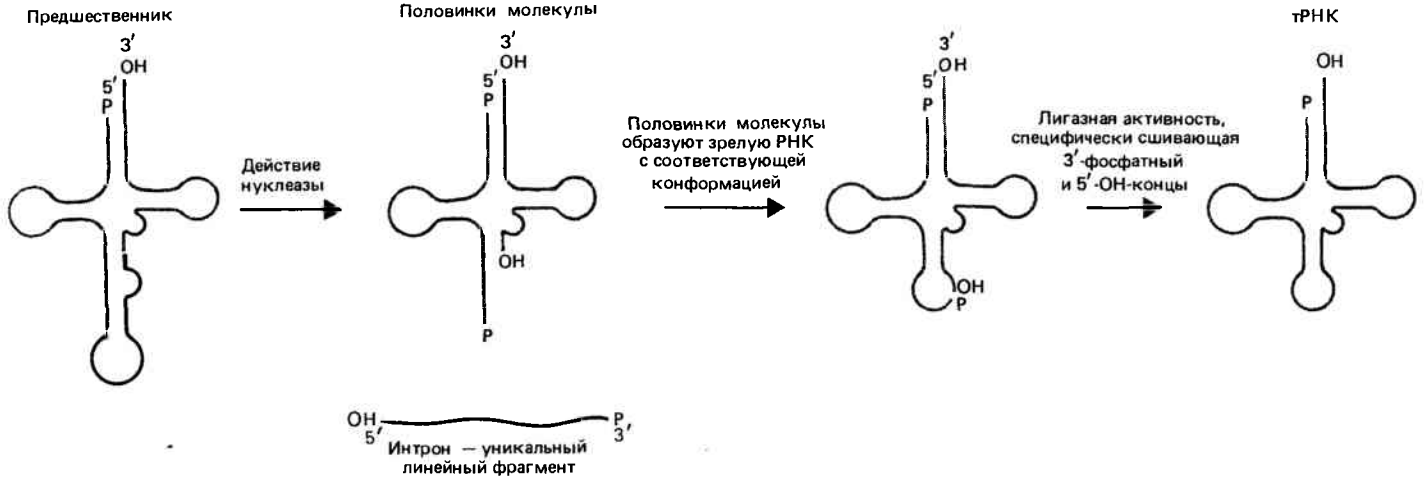


Рис. 26.3. Сплайсинг осуществляется путем нуклеазного расщепления, приводящего к образованию интрона и двух поло-

винок молекулы tРНК, за которым следует ковалентное сшивание половин молекулы с помощью лигазы.

лом. В частности, важную роль, по-видимому, играет узнавание таких структурных особенностей предшественника, которые сохраняются и в зрелой tРНК. Это согласуется с тем наблюдением, что зрелая tРНК конкурентно ингибирует сплайсинг предшественника, по-видимому, тем, что узнается ферментом (ферментами) сплайсинга.

в результате амплификации образуется большое число внехромосомных линейных молекул, каждая из которых представляет собой палиндромный димер рДНК. Две транскрипционные единицы идентичны, и при экспрессии каждой из них образуется 35S-предшественник РНК, на 3'-конце которого находится последовательность 26S-рРНК.

Необычный сплайсинг рРНК *Tetrahymena*

Ген, кодирующий 26S-рРНК у некоторых штаммов *Tetrahymena thermophila*, прерывается одним интроном. Ситуация с этим геном необычная. Диплоидный геном *Tetrahymena* сосредоточен в микронуклеусе, в то время как в макронуклеусе происходит амплификация генов. В ре-

При инкубации изолированных макронуклеусов интрон может удаляться из предшественника. Он накапливается в виде линейных фрагментов размером 400 нуклеотидов. При более длительной инкубации эти фрагменты превращаются в кольцевые молекулы РНК. Описанные события проиллюстрированы на рис. 26.5. Могут ли они быть воспроизведены при инкубации выделенного предшественника РНК в системе in vitro?

Поразительный результат был получен при выделении

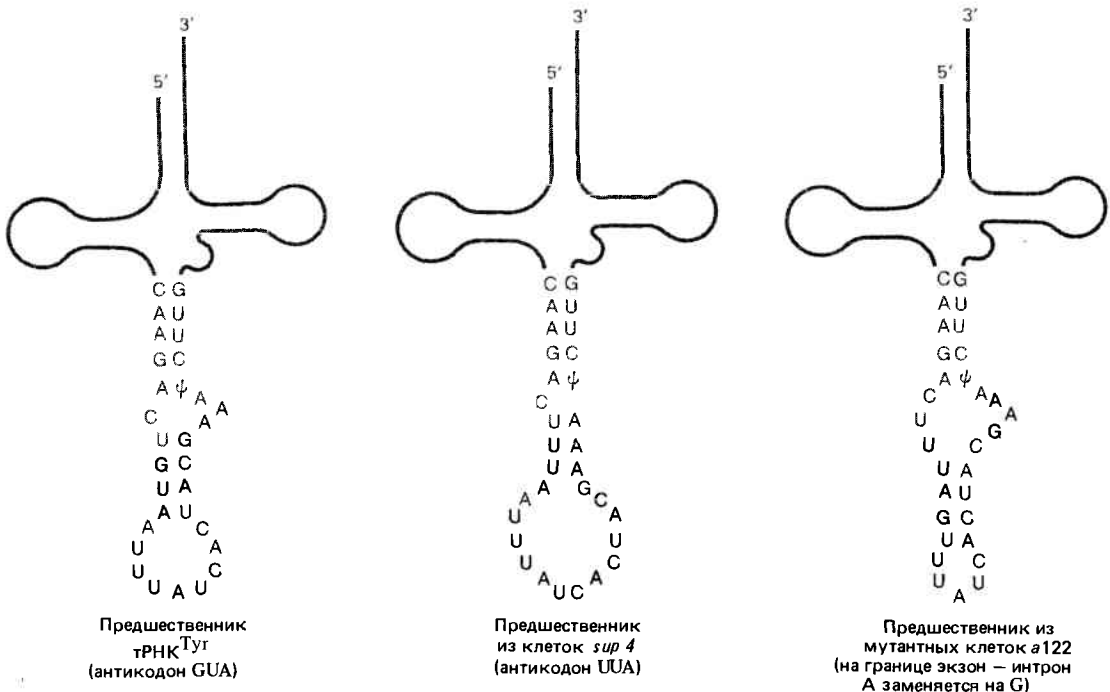


Рис. 26.4. Три варианта дрожжевой tРНК^{Tyr} имеют разную конформацию области антикодона предшественника с прерывистым строением.

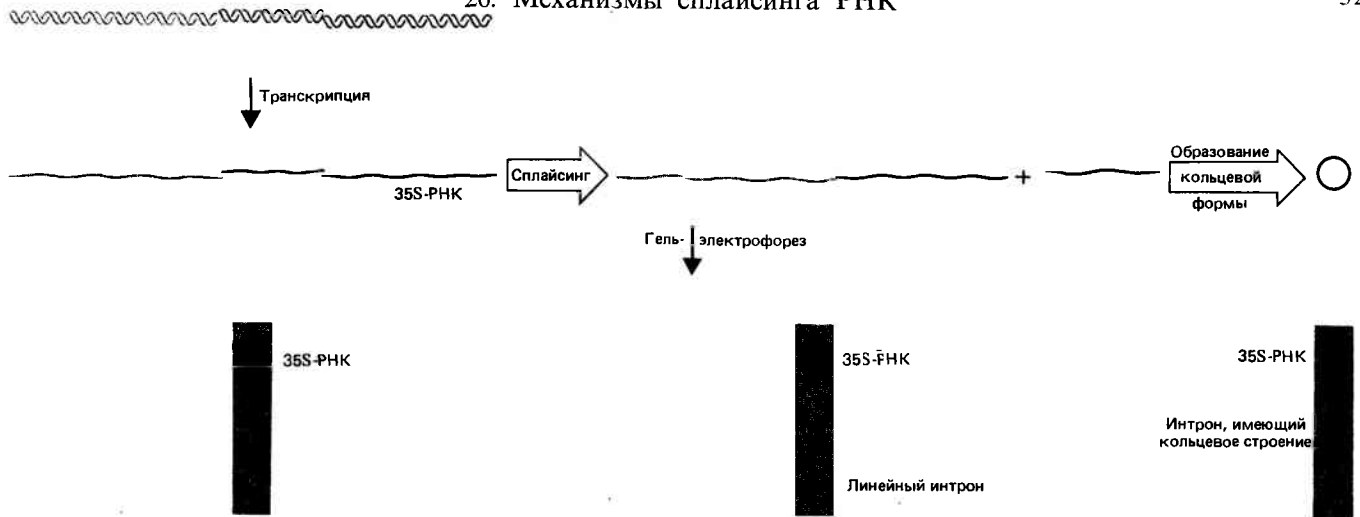


Рис. 26.5. Сплайсинг 35S-предшественника рРНК *Tetrahymena* можно изучать с помощью гель-электрофореза.

35S-предшественник рРНК образует в геле довольно широкую полосу. Удаление интрона обнаруживается по появлению узкой полосы с большой подвижностью. (В полосе 35S-рРНК изменений не обнаруживается

вследствие незначительного уменьшения ее размеров по сравнению с исходной шириной. Свободных экзонов не наблюдается.) Когда интрон приобретает кольцевую форму, его электрофоретическая подвижность уменьшается, как видно из более высокого положения соответствующей полосы.

35S-предшественника РНК. Его сплайсинг происходил в присутствии только моновалентного и дивалентного катионов и GTP. Это означает, что вместе с предшественником РНК должен был выделиться некий фермент, осуществляющий сплайсинг. Но метод выделения предшественника включал использование фенола, обработку протеазами (разрушающими любые белки) и применение такой химической обработки, которая также должна была разрушить все присутствующие белки. Принято считать, что при таком методе выделения не может сохраниться ни один белок.

В результате остаются две возможности, каждая из которых а priori представляется достаточно маловероятной. Ферментативной активностью, необходимой для протекания сплайсинга, может обладать фермент, который действительно очень прочно связан с РНК (возможно, с помощью ковалентных связей), сохраняя свою активность и после такой жесткой обработки. Или же никакого фермента сплайсинга вообще не существует, а возможность протекания реакции обеспечивается свойствами самой молекулы РНК. В таком случае молекула РНК должна иметь специфическую конформацию, при которой соответствующие связи располагаются непосредственно по соседству, делая возможным автономное осуществление реакций разрыва связей и «сшивания». Наличие способности к такому осуществлению сплайсинга предшественником РНК, транскрибированным на копиях гена, клонированных в клетках *E. coli*, указывает на то, что действительно эта способность — свойство самой РНК (хотя *in vivo* в этой реакции может принимать участие и белок).

Самой реакции присущи некоторые необычные свойства. Уникальное свойство — потребность в нуклеотиде, содержащем гуанозин; он не может быть заменен никаким другим основанием. Но присутствие трифосфата не обязательно; в реакции могут быть использованы GTP, GDP, GMP и сам гуанозин. Следовательно, хотя в реакции участвует остаток G, это не обусловлено энергетическими потребностями. Гуаниновый нуклеотид должен содержать 3'-гидроксильную группу.

В том случае, когда гуаниновый нуклеотид несет радиоактивную метку, она включается в вырезанный ли-

нейный интрон. Оказалось, что остаток G связан с 5'-концом интрона при помощи обычной фосфодиэфирной связи. Это свойственно каждому фрагменту, освобождающемуся при сплайсинге. Когда реакция сплайсинга продолжается *in vitro*, линейный интрон приобретает кольцевую форму. Реакция включает присоединение 3'-конца интрона к сайту, расположенному на расстоянии 15 нуклеотидов от его 5'-конца (концевой участок размером 15 нуклеотидов, включающий дополнительный остаток G, удаляется).

Формально эту реакцию сплайсинга можно считать эквивалентной реакции переноса фосфоэфирных связей, как показано на рис. 26.6. Гуаниновый нуклеотид служит источником свободной 3'-гидроксильной группы, на которую переносится 5'-конец интрона. За этой реакцией может последовать вторая аналогичная реакция, в которой 3'-гидроксильная группа, образовавшаяся на конце первого экзона, используется для его присоединения ко второму экзону. Эти два переноса, по-видимому, сопряжены. Поскольку в реакции не обнаруживается свободных экзонов, их сшивание — скорее всего составная часть той же самой реакции, при которой удаляется интрон. Образование интрона с кольцевым строением можно объяснить другим переносом фосфоэфирной связи, при котором разрывается связь, находящаяся на расстоянии 15 нуклеотидов от 5'-конца, а освободившийся 5'-конец переносится на 3'-гидроксильную группу молекулы интрона. (Можно провести параллель между этой реакцией и катализируемой топоизомеразой реакцией переноса 3'-ОН-группы на 5'-Р-конец, не сопряженной с расходом энергии (гл. 32).

РНК как катализатор: расширение понятия биохимического катализа

В биохимии глубоко укоренилось представление о том, что только белки обладают ферментативной активностью. (Однако следует вспомнить и о том, что приверженцы представления о белках как молекулах, осуществляющих все важные функции, считали, что только белки могут быть генетическим материалом!) В основе

разумного объяснения идентичности ферментов с белками лежит убежденность в том, что только белки с многообразным их трехмерной структуры и множеством белковых групп обладают гибкостью, позволяющей создавать активные сайты, необходимые для осуществления катализа биохимических реакций. Однако при изучении двух систем процессинга РНК обнаружилось, что это слишком упрощенное представление.

В состав описанной выше системы сплайсинга рРНК *Tetrahymena* *in vitro* не входит никакого добавленного извне белка, и приведенные свойства системы наглядно демонстрируют, что способность к сплайсингу — *внутреннее свойство самой РНК*.

Такое свойство было названо **автокатализом**, поскольку РНК сама обеспечивает свою перестройку. Она не катализирует реакции с участием других молекул. Мы делаем заключение о том, что РНК, по-видимому, может образовывать специфическую вторичную (третичную) структуру, при которой соответствующие группы располагаются друг рядом с другом таким образом, что реакции, показанные на рис. 26.6, становятся возможными. Хотя реакционная способность — несомненно, свойство самой РНК, протеканию реакции *in vivo* могут способствовать белки, функция которых — стабилизация структуры РНК.

Еще более веское доказательство в пользу способности РНК к автономному сплайсингу было получено при изучении рибонуклеазы Р, эндонуклеазы, участвующей в процессинге тРНК *E. coli* (рассмотренном в гл. 25). Рибонуклеаза Р может диссоциировать на два компонента: РНК размером 375 оснований и полипептид с молекулярной массой 20 000 дальтон. В условиях, обычно используемых для исследования активности фермента *in vitro*, оба компонента были необходимы для расщепления тРНК-субстрата.

Но изменение ионного состава среды — увеличение концентрации ионов Mg^{2+} — делает ненужным присутствие белкового компонента. Катализировать реакцию может и одна РНК! Если интерпретировать эти результаты, считая, что функцию фермента выполняла РНК, то получится, что каждый такой «фермент» катализирует расщепление по меньшей мере четырех субстратов. В самом деле, оказалось, что активность РНК не намного меньше активности неочищенных препаратов рибонуклеазы Р.

Поскольку мутации, затрагивающие как ген, кодирующий РНК, так и ген, кодирующий белок, могут приводить к инактивации РНКазы Р *in vivo*, то очевидно, что для обеспечения ферментативной активности природной РНКазы Р необходимы оба компонента. Как и следовало ожидать, было выдвинуто предположение о том, что каталитической активностью обладает белок, в то время как РНК выполняет некоторую вспомогательную роль, например роль помощника при связывании с субстратом (в ее состав входит ряд коротких последовательностей, комплементарных участкам тРНК, с которыми взаимодействует фермент). Но оказалось, что дело обстоит наоборот!

Каким образом РНК обеспечивает создание каталитического центра? Вполне возможно, что она способна образовывать активный центр, если представить его себе как поверхность, на которой, занимая определенное положение друг относительно друга, находится ряд реакционноспособных групп. В случае белка такие группы образуются боковыми цепями аминокислот, отличающимися заметным разнообразием и включающими в себя как

группы, несущие положительный и отрицательный заряд, так и гидрофобные группы. У РНК количество доступных участков меньше, и они в основном представляют собой группы азотистых оснований, находящиеся на поверхности молекулы. Можно предположить, что при создании определенной вторичной (третичной) конформации в молекуле имеются некие короткие участки, образующие поверхность с активными группами. Это создает условия для осуществления разрыва и возникновения связей в другой молекуле. Кажется несомненным, что в основе взаимодействия между РНК-катализатором и РНК-субстратом лежит спаривание оснований.

Очевидно эволюционное значение этих данных. Всегда казалась удивительной «расщепленность» генетического аппарата, во всех компонентах которого присутствует РНК, а каталитические функции которого выполняют белки. Маловероятно, что в состав самых первых систем репликации могли входить и нуклеиновая кислота, и белок.

Однако представим себе, что первые такие системы содержали только автономно реплицирующуюся нуклеиновую кислоту, обладающую примитивными каталитическими свойствами, необходимыми только лишь для создания и расщепления фосфодиэфирных связей. Если предположить, что участие 2'-связей в обнаруживаемых в настоящее время реакциях сплайсинга имеет своим истоком такую примитивную каталитическую активность, то можно утверждать, что исходной нуклеиновой кислотой была РНК, поскольку в молекуле ДНК нет 2'-ОН-группы, и она поэтому не может вступать в такие реакции.

Но можно представить себе, что белки могли быть привнесены в эту систему благодаря их свойству стабилизировать структуру РНК, которая, по-видимому, не была надежно стабильной. Позже вследствие большего разнообразия свойств белков они могли приобрести функцию катализаторов, что в конце концов привело к созданию изоцирено сложного современного аппарата экспрессии генов.

Следы ранее существовавшей системы можно еще обнаружить. Наиболее явно они видны в случае РНКазы Р и рРНК *Tetrahymena*. Менее очевидно их наличие в случае таких органелл как рибосомы. Если считать, что рибосомы вначале состояли из РНК, наделенной каталитическими свойствами, которая затем медленно «обрастала» белками, постепенно принимающими на себя каталитические функции, и в конце концов стала играть только структурную роль, то очевидно, что два исходных компонента рибосом полностью поменялись ролями.

Реакция сплайсинга РНК осуществляется в определенной предпочтительной последовательности

Ряд интересных проблем, касающихся экспрессии генов, связан с существованием генов с высокой степенью мозаичности. При их экспрессии должно быть удалено большое число интронов, а экзоны должны быть соединены в правильной последовательности. То, что реакция не распространяется последовательно вдоль предшественника, означает, что внутри предшественника РНК должны распознаваться определенные пары концов интронов.

До появления метода блоттинга было практически невозможно установить соответствие ядерных предшествен-

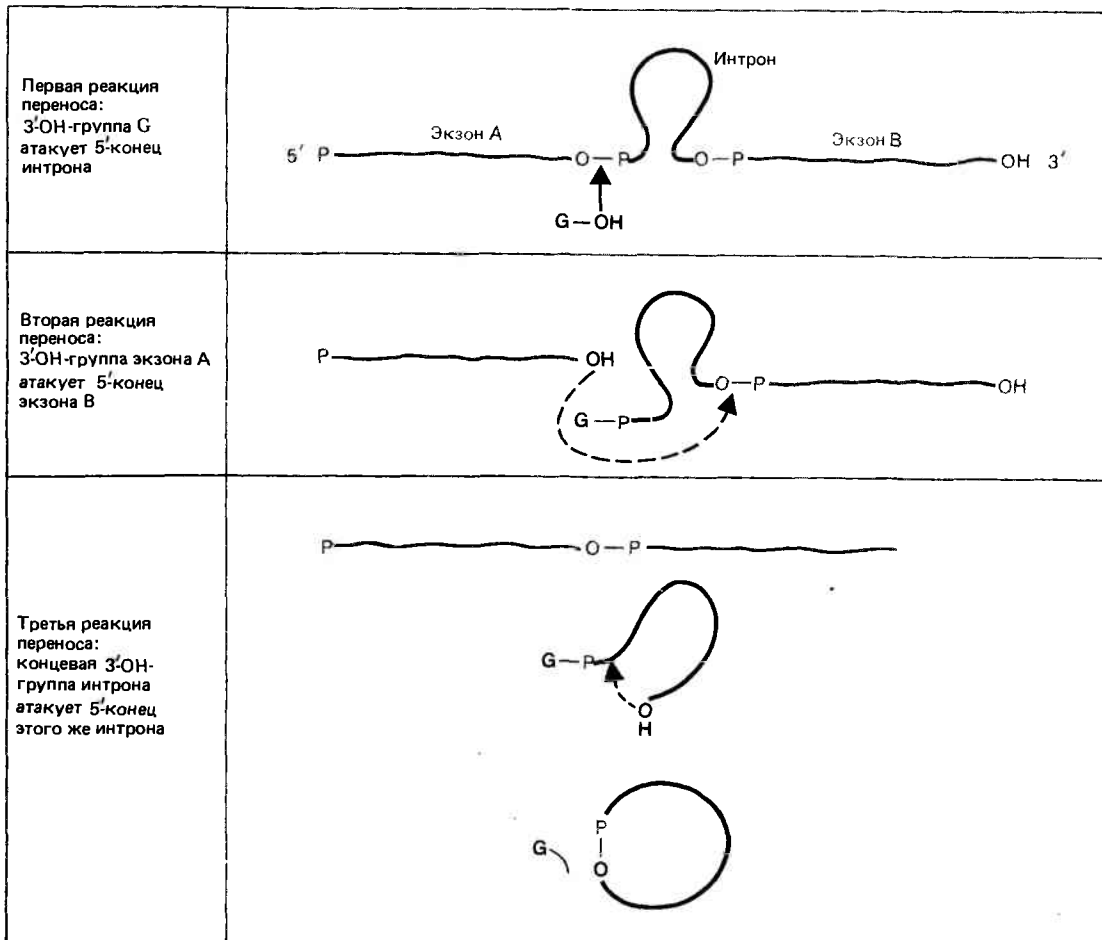


Рис. 26.6. Молекулы РНК, образующиеся при сплайсинге у *Tetrahymena*, могут возникать в результате реакций переноса фосфоэфирных связей.

ников индивидуальным мРНК вследствие незначительного содержания в клетке ядерной РНК и ее неизбежного загрязнения цитоплазматической мРНК. Использование метода «Нозерн-блоттинга» позволяет фракционировать ядерную РНК в геле, переносить материал на специальную бумагу и гибридизовать радиоактивный зонд со специфической последовательностью. Затем с помощью радиоавтографии выявляют полосы, положение которых позволяет идентифицировать молекулы РНК, содержащие гибридизующуюся последовательность. (Аналогичный метод блоттинга по Саузерну показан на рис. 19.6.) Зондом, используемым для выявления полос РНК, может служить кДНК, соответствующая мРНК, или клонированная последовательность, соответствующая определенному участку гена (например, экзону или интрону).

При анализе ядерной РНК из куриного яйцевода с помощью зонда для мРНК овомукоида или овальбумина в каждом случае получают ряд дискретных полос. Это само по себе указывает на то, что сплайсинг может происходить в определенной последовательности. (Если бы семь интронов каждого из этих генов удалялись в абсолютно случайном порядке, имелось бы более 300 предшественников с различными комбинациями интронов и никаких дискретных полос не наблюдалось.)

Полоса мРНК самого большого размера соответ-

ствует размеру гена и, по-видимому, представляет собой первичный транскрипт. Полоса РНК наименьшего размера — это мРНК. Каждая полоса, расположенная в промежутке между ними, соответствует предшественнику (предшественникам), из которого удалены некоторые (но не все) интроны. На рис. 26.7 приведен результат анализа предшественников мРНК овомукоида с помощью Нозерн-блоттинга с основными промежуточными продуктами сплайсинга б (отсутствуют интроны Е и F), з-д (отсутствуют интроны Е, F, D и G) и ж (содержит только интрон С).

Дальнейшая информация о протекании сплайсинга может быть получена при помощи электронно-микроскопического анализа молекул индивидуальных ядерных РНК. Каждую молекулу гибридизуют с ДНК интактного гена, а затем определяют, какие интроны удалены. Любой удаленный интрон выглядит как петля ДНК. Результаты такого анализа приведены в табл. 26.1. Они свидетельствуют о том, что какого-либо *обязательного* порядка сплайсинга, по-видимому, нет, поскольку при анализе могут быть обнаружены промежуточные продукты, у которых интроны удалены в разных комбинациях. Однако имеются основания предполагать наличие *предпочтительного* порядка протекания сплайсинга.

Когда удаляется один интрон, это практически всегда интрон Е или F. Но каждый из них может быть удален



Рис. 26.7. При анализе ядерной РНК с помощью блоттинга по Нозерну и гибридизации с зондом, содержащим овомукоидные последовательности, обнаруживается ряд дискретных предшественников мРНК.

первым. При удалении двух интронов это опять-таки чаще всего интроны Е и F, но возможны и другие комбинации.

Интрон С на первых трех стадиях сплайсинга никогда не удаляется или удаляется очень редко. Отсюда следует, что имеется предпочтительный порядок сплайсинга, при котором первыми удаляются интроны F и E, затем удаляются интроны G и D; далее, по-видимому, удаляется интрон В, и последними вырезаются интроны А и С. Но очевидно, что сплайсинг может происходить и в другой последовательности, поскольку, например, в некоторых молекулах последним утрачивается интрон D или G.

Интерпретировать такие результаты необходимо с большой осторожностью, так как у нас нет никаких данных, доказывающих, что все эти промежуточные продукты действительно превращаются в зрелую мРНК; однако выдвинутые предположения вполне разумны.

Общий вывод, который можно сделать на основе такого анализа, состоит в том, что конформация РНК может влиять на доступность точек разрезания при сплайсинге. При удалении определенных интронов конформация молекулы меняется, и становятся доступными новые пары сайтов сплайсинга. Но способность предшественника утрачивать интроны более чем в одном порядке указывает на то, что либо молекула может принимать разные конформации, либо при каждой конкретной конформации более чем одна пара сайтов может быть доступной разрезанию.

Способность одного интрона приобретать свойственное ему строение может зависеть от событий, происхо-

дящих с другими интронами. С увеличением числа интронов число таких вариантов должно увеличиваться, и при рассмотрении генов большего размера становится труднее понять, чем обеспечивается создание определенной вторичной структуры.

Один из выводов, который следует из такого анализа, заключается в том, что все промежуточные продукты сплайсинга полиаденилированы. Это означает, что обычно после завершения транскрипции и до начала сплайсинга по крайней мере к значительной части молекул добавляется фрагмент poly(A). В этом случае сплайсинг следует за транскрипцией, а не сопутствует ей. Однако такая последовательность событий не всегда имеет место, поскольку сплайсинг может происходить и при ингибировании полиаденилирования (гл. 27).

До сих пор мы предполагали, что удаление каждого интрона осуществляется на определенном этапе сплайсинга. Но это не обязательно так. В случае гена β -глобина мыши обнаруживается 15S-предшественник, содержащий только большой интрон. Это, возможно, означает, что первым удаляется малый интрон. Но затем большой интрон удаляется, возможно, в два этапа, поскольку имеется промежуточный продукт, содержащий только часть этого интрона. Таким образом, пытаюсь обнаружить сайты, по которым происходит сплайсинг, не следует забывать, что, хотя они и должны содержать границы экзон—интрон, другие сайты, расположенные между ними, также могут быть использованы в качестве промежуточных.

Что происходит с удаленными интронными участками? Эксперименты по гибридизации с популяцией ядерных РНК свидетельствуют о том, что концентрация овалбуминовых интронных последовательностей примерно в 10 раз ниже, чем концентрация последовательностей экзонов. Возможно, это означает, что удаленные из молекул интроны довольно быстро разрушаются, а предшественники, идентифицируемые при Нозерн-блоттинге или при электронно-микроскопическом анализе, представляют собой популяцию более стабильных ядерных молекул. Все промежуточные продукты, которые подвергаются более быстрому процессингу, будут присутствовать в относительно небольших количествах, так что обнаруживаемые предшественники—это промежуточные продукты, процессинг которых происходит медленно и которые поэтому присутствуют в более высоких концентрациях.

Для полного выяснения вопроса о путях осуществления сплайсинга необходимо иметь данные опытов по вытеснению импульсной метки, в которых в течение короткого времени осуществляется включение радиоактивной метки в первичный предшественник, а затем прослеживают ее судьбу на разных стадиях созревания РНК.

Таблица 26.1

Интроны удаляются из ядерной РНК овомукоида в предпочтительной последовательности

Общее число удаленных интронов	Частота удаления определенных интронов						
	А	В	С	Д	Е	F	G
1	5	0	0	0	30	60	5
2	20	20	0	2	60	60	25
3	5	5	5	30	100	95	60
4	10	25	35	95	90	90	55
5	40	75	65	85	100	75	60
6	55	100	80	90	100	100	80

Границы сплайсинга могут быть взаимозаменяемыми

Для точного выяснения молекулярных механизмов сплайсинга необходимо исследовать природу границ сплайсинга (границ экзон—интрон), последовательностей, непосредственно окружающих сайты разрезания и сшивания. Как уже отмечалось в гл. 20, границы сплайсинга называются соответственно их расположению в интроне. Левая граница (иногда называемая донорной) находится

на левом конце интрона; правая граница (иногда называемая акцепторной) располагается на другом его конце.

Зависит ли узнавание правильной пары левой и правой границ сплайсинга от индивидуальных свойств последовательности, или оно обусловлено каким-то внешним фактором ее взаимодействия с аппаратом сплайсинга? Иными словами, присущи ли каждой паре границ сплайсинга уникальные особенности нуклеотидной последовательности или структуры, или же все левые границы функционально эквивалентны и все правые границы также неразличимы?

Из нуклеотидных последовательностей границ сплайсинга и окружающих их участков очевидно, что никакой комплементарности между соответствующими левой и правой границами нет; это исключает возможность их непосредственного соединения друг с другом путем спаривания оснований. Другая возможность состоит в специфическом попарном узнавании последовательностей на границах экзон—интрон белками или РНК. Модели такого рода могут быть применены к митохондриям дрожжей, где сплайсинг осуществляется только по нескольким парам границ (как описано в гл. 20). В случае гораздо большего числа границ в ядерных молекулах труднее проверить справедливость таких моделей.

Возникает связанный с этим вопрос о том, существует ли какая-нибудь тканеспецифичность сплайсинга РНК? Имеется ли для транскриптов всех генов единый механизм сплайсинга, не зависящий от конкретных обстоятельств, так что узнавание нужных границ сплайсинга определяется только РНК и общим для всех транскриптов аппаратом сплайсинга? Или же сплайсинг специфической ядерной РНК происходит только в тех клетках, где обычно экспрессируется соответствующий ген? Ответы на эти вопросы могут быть получены при изучении сплайсинга предшественников РНК в различных системах клеток, а также путем конструирования искусственных генов с новыми комбинациями границ сплайсинга.

Описанный в гл. 19 метод клонирования чужеродных последовательностей в бактериофагах или плаزمиде может быть использован и для встраивания генов в эукариотические вирусы. Путем внесения необходимых разрывов рестриктазами и последующего объединения полученных фрагментов глобиновый, инсулиновый и другие гены из множества источников были встроены в эукариотические вирусные векторы. Один из наиболее распространенных векторов—обезьяний вирус SV-40, который может размножаться в клетках ряда млекопитающих. Новые гены могут быть встроены в область ранних или поздних генов транскрипционной единицы вируса. Их экспрессия включает стадию образования РНК-предшественника, на 5'-конце которого обычно имеются вирусные последовательности, за которыми следует последовательность встроеного гена. Такие РНК подвергаются нормальному сплайсингу, как показано на **рис. 26.8**.

При использовании такой системы получают важную информацию двух видов. Во-первых, правильное протекание сплайсинга не зависит от целостности природного РНК-предшественника, поскольку транскрипт встроеного гена (или иногда только его часть) может подвергаться нормальному процессингу, находясь в составе вирусных последовательностей. Во-вторых, информация, заключенная в вирусной последовательности, способна обеспечить протекание сплайсинга начиная с обычного места в клетках, полученных из другой ткани и (или) другого вида организмов. Таким образом, сигнальные после-

довательности, участвующие в сплайсинге, должны быть эволюционно высококонсервативными.

При использовании подхода реконструкции генов был сделан дальнейший шаг: получены синтетические «гены», в которых экзон одного природного гена соединился с экзонем другого гена. Это схематично показано на **рис. 26.9**. В проведенном эксперименте первый экзон области ранних генов транскрипционной единицы вируса SV-40 был сшит с третьим экзонем β -глобинового гена мыши. Интрон такой гибридной молекулы успешно удалялся при сплайсинге. Таким образом, левая граница интрона SV-40 (на рисунке это *l1*) может быть при сплайсинге присоединена к правой границе интрона β -глобинового гена мыши (область *r2* на рисунке). Из этого следует, что в принципе любая левая граница сплайсинга может взаимодействовать с любой правой границей.

Парадоксальность такого заключения по сравнению с обычной ситуацией, когда сплайсинг осуществляется только по левой и правой границам *одного и того же* интрона, наглядно проиллюстрирована на **рис. 26.10**. На рисунке сравнивается экспрессия иммуноглобулинового гена дикого типа с мутантным геном, у которого в результате делеции была удалена область, включающая границу между экзонем *VI* и интроном 2. (Структура и экспрессия иммуноглобулиновых генов детально обсуждаются в гл. 39.)

В случае гена дикого типа левая граница каждого интрона при сплайсинге соединяется только вместе с правой границей этого же интрона. Взаимодействия между левой границей интрона 1 и правой границей интрона 2 не происходит. В случае же делеционного мутанта правая граница интрона 1 не участвует в сплайсинге и левая граница этого интрона присоединяется к правой границе интрона 2. Это означает, что двум таким границам присуще свойство узнавать друг друга при сплайсинге, несмотря на то что обычно при наличии левой границы интрона 2 взаимодействия между ними не происходит.

Вывод, который можно сделать, исходя из этих экспериментов, состоит в том, что, по-видимому, все левые границы сплайсинга, так же как и все правые границы, воспринимаются аппаратом сплайсинга как одинаковые. Можно также считать, что основные свойства границ сплайсинга каждого типа относятся к общим свойствам, присущим всем последовательностям на границах экзон—интрон.

Мутации в канонических последовательностях могут влиять на сплайсинг

Мы уже убедились в том, что все интроны ядерных структурных генов, кодирующих белки, можно сопоставить на основе правила GT...AG так, что обнаружится гомология коротких универсальных последовательностей на границах экзон—интрон. Как уже мы рассматривали в гл. 20, каноническая, или усредненная, последовательность имеет вид



Стрелками обозначены связи, сплайсинг которых приводит к превращению предшественника в зрелую РНК, но, являются ли они действительно сайтами сплайсинга,

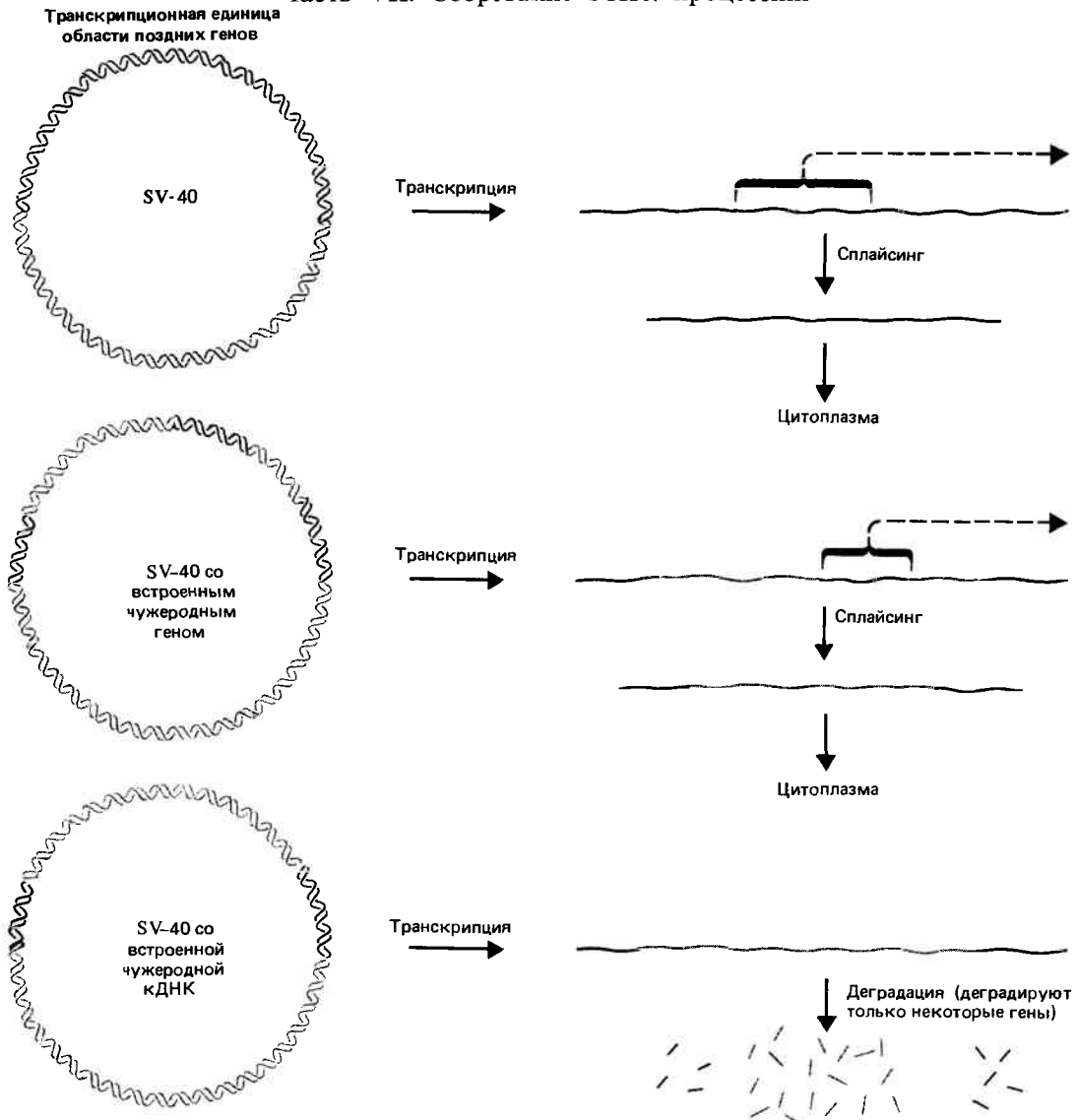


Рис. 26.8. Последовательности вектора SV-40, в единицах транскрипции которых на месте поздних генов находятся прерывистые гены другого организма, транскрибируются, подвергаются нормальному сплайсингу, и конечные продукты экспор-

тируются в цитоплазму. Когда встроенный фрагмент представляет собой непрерывную кДНК, представляющую ген, транскрипция происходит, но в некоторых случаях образующийся продукт расщепляется в ядре (см. также текст).

до сих пор не установлено, поскольку не изучались концевые последовательности промежуточных продуктов. (Во многих случаях наличие повторяющихся последовательностей приводит к тому, что сайты сплайсинга не могут быть однозначно установлены на основе только последовательности нуклеотидов; в случае самой канонической последовательности стрелки можно сдвинуть на одно или два основания влево от каждой границы.)

Тщательное изучение нуклеотидных последовательностей границ сплайсинга показывает, что единственное их общее свойство — наличие таких коротких канонических последовательностей. Они находятся на границах экзон—интрон чрезвычайно большого числа видов организмов, от животных до дрожжей. Но никакого другого существенного сходства не удается выявить даже при рассмотрении отдельных групп интронов (таких, как интроны определенного вида животных, интроны семейства генов и т. д.).

Могут ли такие короткие последовательности обуславливать специфичность протекания сплайсинга? Мы знаем, что по крайней мере в некоторых случаях на

границе экзон—интрон не требуется присутствия нуклеотидной последовательности значительного размера. Например, делеция нуклеотидной последовательности вируса SV-40, при которой удаляются почти все нуклеотиды левее левой универсальной последовательности, не мешает ее использованию при сплайсинге. Для того чтобы установить, является ли термин «граница сплайсинга» синонимом термина «каноническая, или усредненная, последовательность», необходимо проводить дальнейшие исследования такого рода. Известно, что границы экзон—интрон не всегда функционируют независимо от присутствия других последовательностей. Делеция небольшого участка нуклеотидной последовательности вируса SV-40, захватывающая участок, находящийся на расстоянии 200 п.н. от левой границы сплайсинга, может препятствовать ее использованию при сплайсинге (возможно, в результате изменения конформации РНК). Аналогичным образом *цис*-мутации, нарушающие сплайсинг интрона гена *box* дрожжевой митохондриальной ДНК, могут быть локализованы на некотором расстоянии от границы сплайсинга (гл. 20).

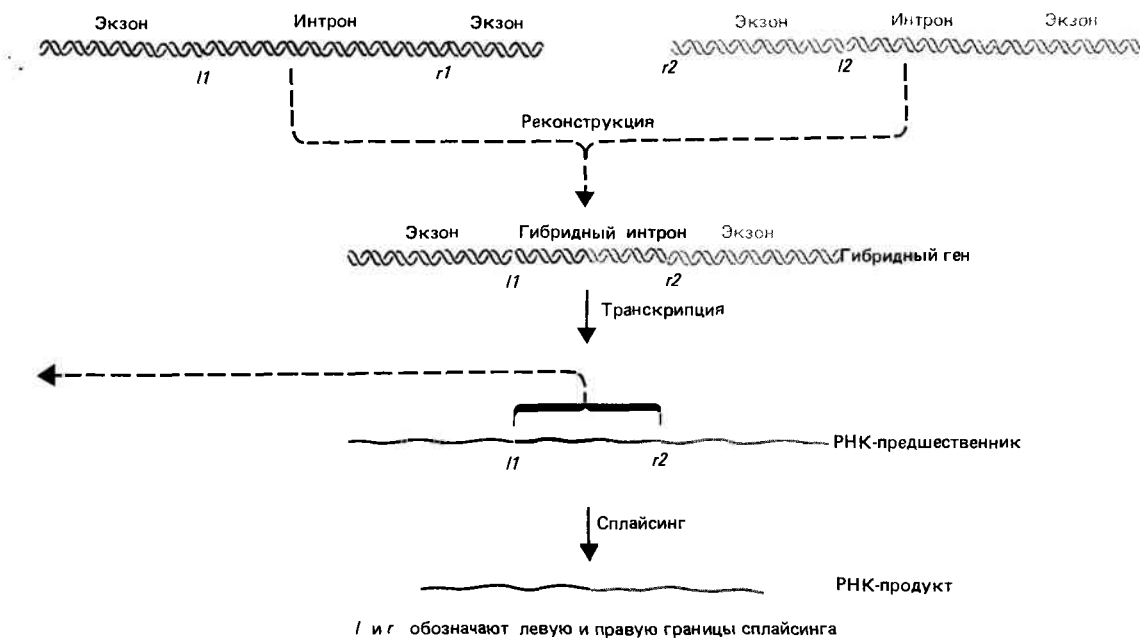


Рис. 26.9. Сплайсинг транскриптов, полученных на последовательности, сконструированной путем сшивания экзонов двух разных генов, протекает нормально.

Предположение о том, что канонические последовательности –этой сайты, узнаваемые при сплайсинге, было высказано лишь за неимением лучшего объяснения. Но важная роль этих последовательностей непосредственно подтверждается существованием некоторых точковых мутаций, затрагивающих канонические последовательности и влияющих на сплайсинг. Такие мутации были обнаружены в случае талассемий человека, при которых нарушается образование либо α -, либо β -глобинов. (Имеется много причин возникновения талассемий; некоторые из них рассматривались в гл. 21.)

При α -талассемии образование мРНК, соответствующей $\alpha 2$ -гену, нарушается в результате делеции пентануклеотидной последовательности на левой границе интрона 1. В этом случае не обнаруживается стабильной РНК.

β^0 -Талассемия, при которой не происходит образования β -глобина, возникает в результате точковой мутации в канонической последовательности. В результате этой мутации динуклеотид GT на левой границе экзон—ин-

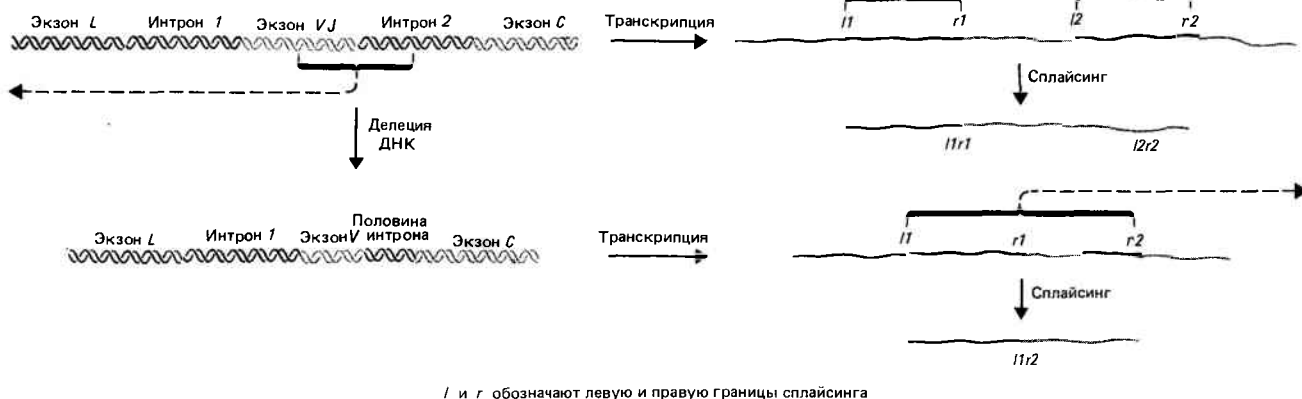
трон заменяется на AT, приводя, по-видимому, к нарушению образования мРНК.

При β^+ -талассемии наблюдается значительное снижение содержания β -глобиновой мРНК и синтеза β -глобина. Но образующиеся мРНК и белок, по-видимому, являются нормальными. Для изучения природы этого дефекта были клонированы некоторые гены, обуславливающие возникновение β^+ -талассемий.

В одном случае при β^+ -талассемии имеется единственная точковая мутация, находящаяся на расстоянии 21 п.н. от правого конца малого (первого) интрона. Она приводит к замене нуклеотидного основания G в последовательности дикого типа на A; в результате образуется мутантная последовательность



Рис. 26.10. Экзоны L и C объединяются при сплайсинге в случае делеции левой границы интрона 2, но не тогда, когда она интактна.



l и r обозначают левую и правую границы сплайсинга

Последовательность, содержащая мутировавший сайт, практически идентична последовательности, содержащей границы сплайсинга; мутация приводит к тому, что шесть из семи оснований оказываются гомологичными. (А различающееся основание занимает положение, в котором и в нормальных сайтах сплайсинга других β -подобных генов глобина находятся разные нуклеотидные основания.) Обратите внимание на то, что последовательность, состоящая из семи нуклеотидов, захватывает и область справа от сайта сплайсинга, не входящую в состав участка универсальной последовательности правой границы экзон—интрон.

Из-за такого сходства сплайсинг осуществляется по мутантному сайту вместо нормального. Это имеет место у больных талассемией и при экспрессии клонированного гена в культуре клеток обезьяны или человека. Вплоть до 90% молекул предшественника РНК подвергается сплайсингу по мутантному сайту вместо нормального. В тех немногих случаях, когда сплайсинг происходит по нормальному сайту, это приводит к образованию активной β -глобиновой мРНК, обнаруживаемой у больных. Мы не знаем, почему мутантный сайт используется с большей эффективностью; возможно, это обусловлено тем, что он располагается ближе к левой границе интрона или вследствие некоторых особенностей конформации РНК.

При образовании мутантной РНК снижается общая эффективность сплайсинга; кроме того, в ядре накапливается предшественник РНК, вообще не подвергавшийся сплайсингу. Отсюда следует вывод, что нарушение сплайсинга интрона 1 может привести к нарушению сплайсинга интрона 2 (это согласуется с данными о том, что интрон 1 удаляется первым; см. ранее). Возможно, такой эффект обусловлен особенностями вторичной структуры РНК.

Что происходит с РНК, образовавшейся в результате неправильного сплайсинга? Это до конца еще не установлено, и не исключено, что все зависит от конкретных обстоятельств. У большого талассемией такая РНК не накапливается в цитоплазме ретикулоцитов. Из этого следует, что либо предшественник РНК или продукт неправильного сплайсинга нестабилен либо нарушен ее транспорт из ядра в цитоплазму. Однако в клетках культуры ткани РНК, образовавшаяся в результате неправильного сплайсинга, транспортируется в цитоплазму столь же эффективно, как и нормальная мРНК. Не известно, присоединяется ли она к полисомам, но если и присоединяется, то на ней транскрибируется аномальный белок, синтез которого терминируется преждевременно. (Осуществление сплайсинга по мутантному сайту приводит к появлению в данной рамке считывания бессмысленного кодона.)

Необходим ли сплайсинг для образования мРНК, кодируемых прерывистыми генами? (Известно, что в случае мРНК, кодируемой непрерывными генами, нужды в нем нет.) Использование в качестве вектора вируса SV-40 позволяет непосредственно сравнить экспрессию прерывистого и непрерывного вариантов одного и того же гена. Прерывистый ген в этом случае представлен фрагментом генома; непрерывный ген представлен кДНК, соответствующей мРНК. В тех случаях, когда транскрипционная единица не содержит границ сплайсинга, она нормально транскрибируется. Но в некоторых случаях (например, в случае глобиновой мРНК) образующаяся мРНК оказывается нестабильной, как было показано ранее на рис. 26.8. Следовательно, можно предположить, что судьба транскриптов определенных ядерных генов с прерыв-

истым строением может быть двоякой: они могут подвергнуться сплайсингу и затем транспортироваться в цитоплазму или же деградировать в ядре. Для других генов этого не наблюдается, и они могут в равной степени успешно экспрессироваться, имея как прерывистое, так и непрерывное строение.

Участвуют ли в сплайсинге малые ядерные РНК?

Изменения сплайсинга, вызванные мутациями, приводящими к появлению или разрушению канонических последовательностей, указывают на то, что эти последовательности существенны для узнавания границ экзон—интрон. Но до сих пор все еще не ясно, как именно узнаются такие последовательности. Можно предложить две общие схемы этого процесса. Либо фермент узнает специфически пару канонических последовательностей, либо они могут спариваться с основаниями РНК, что приведет к созданию вторичной структуры, которая и будет узнаваться ферментом. При этом конформация РНК играет такую же роль, как и в случае других ферментативных реакций в процессинге РНК (хотя здесь спаривание оснований не обязательно происходит автономно, на него, возможно, оказывают влияние и белки).

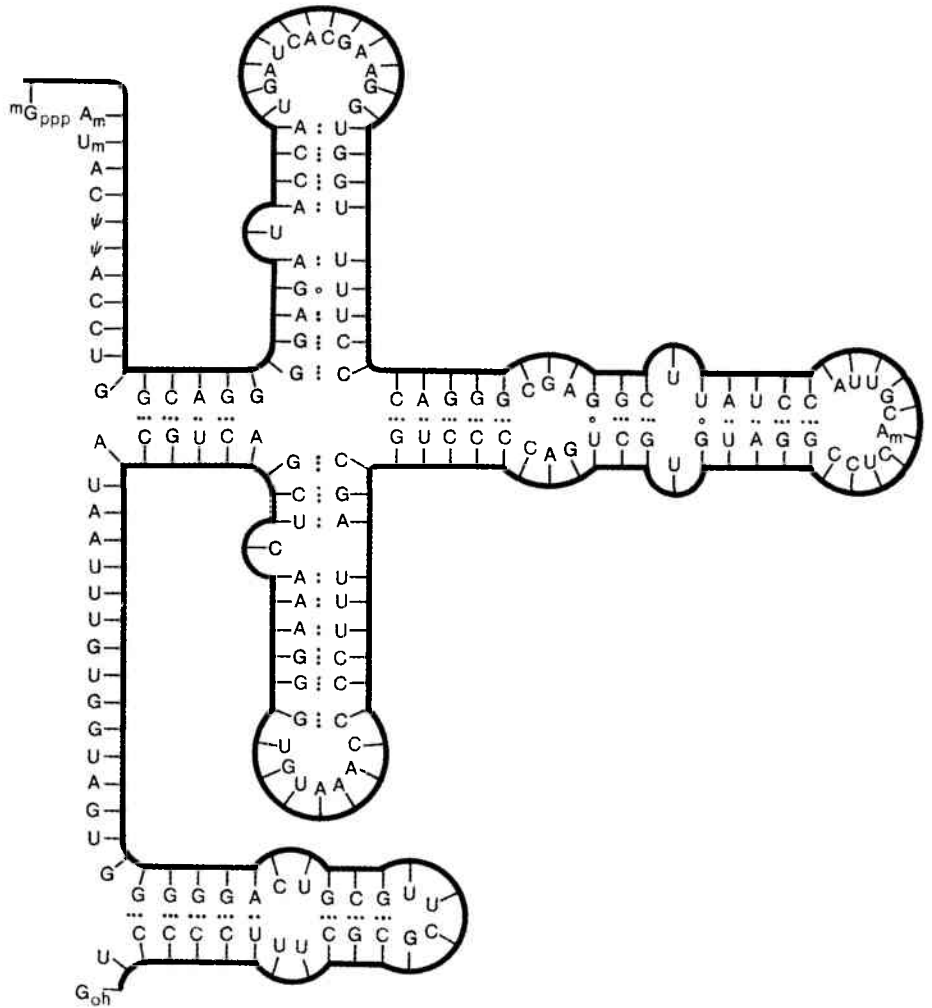
Каково происхождение такой РНК? Ее источником может служить последовательность, расположенная в каком-либо другом участке того же транскрипта, что, однако, представляется маловероятным, поскольку ни в одной из ядерных последовательностей этого не наблюдалось. Или может существовать самостоятельная РНК, которая для соединения концов интрона спаривается с предшественником, не подвергнувшимся сплайсингу.

И ядро, и цитоплазма эукариотических клеток содержат довольно большое число видов различных малых РНК. Их размер варьирует от 100 до 300 нуклеотидов, а число молекул, приходящихся на одну клетку, составляет примерно от 10^5 до 10^6 . Некоторые малые РНК синтезируются с помощью РНК-полимеразы III, другие — РНК-полимеразы II; среди таких транскриптов встречаются и некоторые имеющие экз, как и молекулы мРНК.

Такие молекулы, встречающиеся в ядре, называются **малыми ядерными РНК** (мяРНК, или snRNA — от small nuclear RNAs); аналогичные РНК, обнаруживаемые в цитоплазме, называются **малыми цитоплазматическими** (мцРНК, или scRNA — от small cytoplasmic RNAs). Природные малые РНК существуют в виде рибонуклеопротеиновых частиц (мяРНП и мцРНП, или snRNP и scRNP соответственно). В обиходе такие РНК иногда называют *snurps* и *scyrps*. Обнаружено некоторое сходство последовательностей различных мяРНК; взаимодействие мяРНП с антителами позволяет предположить, что в состав некоторых из них могут входить одни и те же белковые субъединицы.

5'-концевая последовательность одного из видов таких РНК, U1-мяРНК, комплементарна последовательностям, находящимся на границах экзон—интрон. U1-мяРНК, обладающие лишь очень небольшими различиями нуклеотидных последовательностей, обнаружены в клетках животных, птиц и насекомых. Предположительная вторичная структура U1-мяРНК человека приведена на рис. 26.11. Обратите внимание на то, что одиннадцать 5'-концевых нуклеотидов находятся в одноцепочечной форме и, следовательно, потенциально реакционноспособными.

Рис. 26.11. U1-мяРНК человека имеет вторичную структуру, при которой ее 5'-конец является одноцепочечным. Другие виды мяРНК имеют аналогичные нуклеотидные последовательности и структуру.



Модель, объясняющая, как происходит спаривание этого участка с границами экзон—интрон, приведена на рис. 26.12. В верхней части рисунка изображен интрон, расположенный напротив U1-мяРНК. В нижней части рисунка показано, как осуществляется спаривание оснований U1-мяРНК и последовательностей, расположенных на концах типичного интрона. Как в точности происходит спаривание оснований, предсказать трудно, поскольку последовательность UCC U1-мяРНК может спариться с AGG, находящейся на любой из границ экзон—интрон.

Спаривание происходит на участке небольшого размера, и мы не знаем, достаточен ли его размер для того, чтобы границы сплайсинга соединились вместе. На рисунке приведен лучший из возможных вариантов этого процесса. На самом деле в области левой границы экзон—интрон, по-видимому, всегда имеется 4–6 п. н., способных к образованию пар, но в области правой границы могут оказаться только две такие пары, вместо трех.

Интактная частица U1 мяРНК *in vitro* может связываться с левой границей экзон—интрон. Способностью к связыванию обладает частица в целом; очищенная U1 РНК связываться с границей сплайсинга не может. Пока еще не имеется данных о том, может ли U1 мяРНК связываться с правой границей экзон—интрон.

Говоря о роли, которую такое спаривание оснований, вероятно, играет при сплайсинге, следует помнить, что специфичность этой реакции может зависеть от внешних

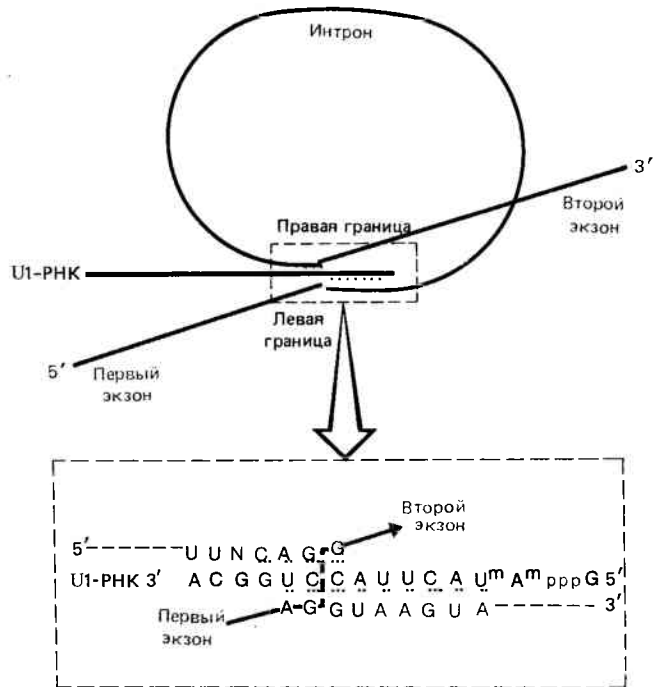


Рис. 26.12. 5'-конец U1-мяРНК может спариваться с левой, а также, возможно, и с правой границами сплайсинга.

условий. Так, мы не стали бы а priori утверждать, что взаимодействие триплетов кодон—антикодон обеспечивало бы достаточную специфичность трансляции, но в тех условиях, которые имеют место в рибосоме, это именно так. Возможно ли, чтобы специфичность взаимодействия U1-мяРНК·ядерная РНК подобным образом зависела от внешних факторов, обусловленных наличием белковых компонентов?

Каким образом, не имея системы сплайсинга ядерных РНК *in vitro*, из которой компоненты аппарата сплайсинга могут быть выделены в чистом виде, выяснить, действительно ли U1-мяРНК участвует в сплайсинге? Некоторым косвенным подтверждением этого служат свойства антител, полученных от больных с аутоиммунным заболеванием **красной волчанкой**. Антитела таких больных вступают в самые разные реакции с нормальными компонентами клеток человека, и среди них можно обнаружить антитела к нуклеиновым кислотам. В некоторых случаях мишенью для них оказывается один (или несколько) тип **мяРНК**. Антитела взаимодействуют с рибонуклеопротеиновой частицей и, вероятно, узнают белковую субъединицу или субъединицы.

В ядрах, выделенных из клеток HeLa, зараженных аденовирусом, может продолжаться сплайсинг вирусных РНК. При инкубации этих ядер с антителами к мяРНК сплайсинг прекращается. По-видимому, в этом случае происходит специфическое ингибирование сплайсинга, а не общее нарушение экспрессии генов, поскольку другие ее стадии (такие, как транскрипция или полиаденилирование) протекают как обычно. Эффективно подавляют сплайсинг только антитела, взаимодействующие с U1-мяРНК; другие мяРНК в сплайсинге не участвуют.

Дополнительное свидетельство возможного существования связи между U1-мяРНК и ядерной РНК получено в экспериментах с псораленами. Это химические соединения, вызывающие образование сшивок и специфически воздействующие на двухпочечные участки нуклеиновых кислот. Сшивки могут быть разрушены при облучении ультрафиолетом. При обработке препаратов ядерных РНК псораленом, а затем ультрафиолетом обнаруживается некоторое количество свободной U1-мяРНК. Это означает, что она была связана с ядерной РНК спариванием их оснований. Только очень небольшая часть находящейся в ядре U1-мяРНК участвует в этой реакции.

Только ли U1-мяРНК содействует сплайсингу РНК, или имеются и другие «помощники»? Геном некоторых вирусов кодирует синтез молекул малых РНК. Один из таких вирусов—аденовирус, продуцирующий VAI-РНК. Эта РНК комплементарна последовательностям на границах экзон—интрон некоторых генов аденовируса; она может участвовать в их узнавании. Если в создании структур, узнаваемых при сплайсинге, участвуют малые подвижные РНК, то для осуществления определенных этапов сплайсинга может потребоваться синтез специфических РНК-помощников.

Окончательное выяснение вопроса о том, какие компоненты участвуют в сплайсинге, невозможно до тех пор, пока в нашем распоряжении не будет таких систем сплайсинга *in vitro*, в которых интроны могут быть удалены из РНК-предшественников. В последнее время в получении таких систем были достигнуты значительные успехи. Первые результаты указывают на то, что между этими и другими системами сплайсинга имеется определенное формальное сходство, что скорее всего подтверждает наши предположения об участии в сплайсинге U1 РНК.

Для изучения сплайсинга РНК-предшественников аденовируса были использованы ядерные экстракты из клеток HeLa. В одной из таких систем, требующей присутствия ионов Mg^{2+} и АТР, сплайсинг стимулируется неочищенным белковым экстрактом, активный компонент которого при очистке выделяется в виде небольших рибонуклеопротеиновых частиц. Способность этой системы функционировать с очищенным РНК-предшественником указывает на то, что осуществление сплайсинга не требует одновременного осуществления транскрипции.

Не обнаружено жесткой связи и между сплайсингом и процессом модификации РНК. Сплайсинг глобиновых РНК, у которых отсутствуют концевые poly(A)-последовательности, протекает совершенно нормально, и хотя у копированных РНК сплайсинг значительно более эффективен, чем у неэкспонированных, каких-либо других различий в осуществлении этого процесса не наблюдается. И *in vivo*, и *in vitro* сплайсинг РНК-предшественников больных талассемией происходит по одним и тем же аномальным сайтам; это открывает дорогу детальному исследованию последовательностей, участвующих в сплайсинге. Данные по изучению сплайсинга у больных талассемией говорят о том, что для выяснения роли определенных сайтов сплайсинга в предшественник РНК можно вносить и другие изменения.

Сплайсинг *in vitro* включает различные реакции разрезания и лигирования. При использовании в качестве субстратов сплайсинга аденовирусной или глобиновой РНК был обнаружен промежуточный продукт с необычной структурой, что, возможно, свидетельствует о наличии общей для разных субстратов последовательности реакций сплайсинга в ядре.

Для осуществления сплайсинга *in vitro* необходимо внести разрыв в левую границу экзон—интрон. В результате внесения разрыва в левый конец интрона образуются две молекулы РНК, одна из которых включает левый экзон, а вторая—правый экзон и интрон. По-видимому, оба вида молекул РНК объединяются вместе под действием белков, участвующих в сплайсинге.

Левый экзон представляет собой линейную молекулу, тогда как молекула экзон—интрон не линейна. 5'-конец, образовавшийся на левой границе экзон—интрон, присоединяется с помощью 5'-2'-связи к основанию А-последовательности CTGAC, расположенной на расстоянии около 30 оснований левее правого конца интрона. Эта последовательность-мишень имеет сходство с пятью последними основаниями ТАСТААС-блока интронов дрожжевой ядерной РНК. Ее участие в сплайсинге—возможно, центральное звено механизма интрона. В результате реакции образуется молекула, имеющая форму петли или ручки кастрюли.

Далее при внесении разрыва в правую границу экзон—интрон образуются петлеобразная молекула—интрон и линейная молекула—правый экзон. Затем петля распрямляется, образуя линейную молекулу вырезанного интрона. Левый и правый экзоны сшиваются.

Почему в процессе сплайсинга должна образовываться структура, включающая конец интрона, который будет удален, и сайт внутри интрона? Возможно, это необходимо для отделения свободного 5'-конца интрона таким образом, чтобы он не препятствовал протеканию реакции. Возможно также, что такая структура служит маркером, указывающим на интрон, участвующий в сплайсинге.

Эта реакция интересна наличием явного сходства

с процессом сплайсинга тРНК, происходящим по 2'-положению, и отщеплением 5'-конца предшественника рРНК *Tetrahymena*. Возможно, окажется, что конформация РНК играет важную роль при сплайсинге в ядре, особенно если вспомнить о наличии каталитической активности у других РНК. Основное значение приведенных данных состоит в том, что они свидетельствуют о возможности разделения сплайсинга на несколько отдельных стадий.

Активность систем сплайсинга *in vitro* специфически ингибируется антителами к белковым компонентам U1 мяРНК. Другие данные указывают на то, что удаление 5'-концевых нуклеотидов U1 мяРНК нарушает протекание сплайсинга *in vitro*. Поскольку наши представления, основанные на использовании этих двух систем, носят предположительный характер, мы многого ожидаем от получения систем, сплайсинг в которых полностью зависит от добавления U1 РНК, что послужило бы подтверждением его участия в сплайсинге и позволило бы исследовать механизм его действия.

Возможное участие U1 РНК в сплайсинге может быть только одним из примеров зависимости реакций процессинга РНК от малых РНК. МяРНК, по-видимому, опять-таки U1, участвует в полиаденилировании (гл. 27). МяРНК участвует в образовании 3'-концов природной гистоновой мРНК *Xenopus* (гл. 27). В свете автономного характера сплайсинга предшественника рРНК *Tetrahymena* и наличия каталитической активности у РНК рибонуклеазы Р кажется можно предположить, что реакции взаимодействия РНК—РНК могут играть важную роль при осуществлении большинства или всех этапов процессинга РНК.

Такие реакции могут протекать по-разному, с образованием специфических конформаций внутри одной и той же молекулы (в этом случае взаимодействие сайтов происходит по *цис*-типу) или на основе *транс*-взаимодействия между субстратом и другой РНК. Такие взаимодействия—возможно, основной след, оставшийся от начальных этапов эволюции нуклеиновых кислот как молекул—носителей наследственности. Будет не удивительно, если окажется, что молекулы мяРНК обладают свойствами, подобными каталитическим, при сплайсинге и других реакциях процессинга.

Мы до сих пор не имеем ответа на главный вопрос. Имеющиеся в настоящее время данные не объясняют, каким образом процесс узнавания при сплайсинге ограничивается узнаванием только левой и правой границ одного и того же удаляемого интрона. Модели, в которых за основу взята конформация РНК, представляются малоубедительными в свете нормального протекания сплай-

синга продуктов гибридных генов. Возможно, что компоненты аппарата сплайсинга функционируют последовательно. После того как фермент узнал границу сплайсинга, он должен двигаться вдоль молекулы РНК в нужном направлении до тех пор, пока не дойдет до другой границы интрона. Это ограничит протекание сплайсинга с соседними границами экзон—интрон. Таким образом можно объяснить осуществление сплайсинга в тех случаях, когда границы экзон—интрон являются уникальными и непосредственно взаимодействуют друг с другом (хотя, разумеется, некоторые интроны имеют очень большую длину). Однако это не объясняет наличия нескольких стадий сплайсинга интрона, не допускает возможности протекания сплайсинга другим способом, как в случае вируса SV-40, когда одна и та же левая граница сплайсинга может быть присоединена более чем к одной правой границе. (Конечно, во всех этих случаях мы предполагаем, что сплайсинг протекает точно; однако нет никаких данных о том, как часто может в действительности осуществляться неправильный сплайсинг.)

Рекомендуемая литература

В журнале «Cell» было опубликовано несколько кратких обзоров, посвященных сплайсингу. Общие механизмы сплайсинга с участием канонических последовательностей, находящихся на границах экзон—интрон, рассматриваются Льюином (*Lewin, Cell, 22, 324–326, 1980*) и Шарпом (*Sharp, Cell, 23, 643–646, 1981*). Структура и функции мяРНК обсуждаются в обзоре Зива (*Zieve, Cell, 25, 296–297, 1981*) и Лернера и Стейтца (*Lerner, Steitz, Cell, 25, 298–300, 1981*). Особенности протекания сплайсинга в разных органеллах посвящен обзор Льюина (*Lewin, Cell, 22, 645–646, 1980*). Данные об автономном осуществлении сплайсинга в клетках *Tetrahymena* в системе *in vitro* были представлены Цехом и др. (*Cech et al., Cell, 27, 487–496, 1981*). Цехом опубликован также обзор о ферментативных механизмах сплайсинга (*Cech, Cell, 34, 713–716, 1983*). Имеется и более ранний обзор, касающийся общих механизмов сплайсинга, включая сплайсинг дрожжевой тРНК, написанный Абелсоном (*Abelson, Ann. Rev. Biochem., 48, 1035–1069, 1979*). Каталитическую активность РНК обнаружил Герье-Такада (*Guerrier-Takada, Cell, 35, 849–857, 1983*). Каталитической активности РНК посвящен также обзор Альтмана (*Altman, Cell, 36, 237–239, 1984*). Структуру «лассо» описали Раскин и др. (*Ruskin et al., Cell, 38, 317–331, 1984*).

Глава 27

РЕГУЛЯЦИЯ ПРОЦЕССИНГА РНК

В клетках высших эукариот транскрипция ядерных структурных генов, кодирующих белки, осуществляется в нуклеоплазме. Но нуклеоплазматическая РНК отличается от мРНК. Она значительно больше по среднему размеру, очень нестабильна и характеризуется гораздо более высокой сложностью входящих в ее состав нуклеотидных последовательностей. Поскольку размеры ядер-

ной РНК колеблются в очень широких пределах, ее называют **гетерогенной ядерной РНК** (гяРНК, или hnRNA).

До того как были открыты прерывистые гены, происходили горячие споры относительно того, образуется ли мРНК из гяРНК путем уменьшения размеров последней или гяРНК может выполнять и совершенно другие функции. В настоящее время имеются основания

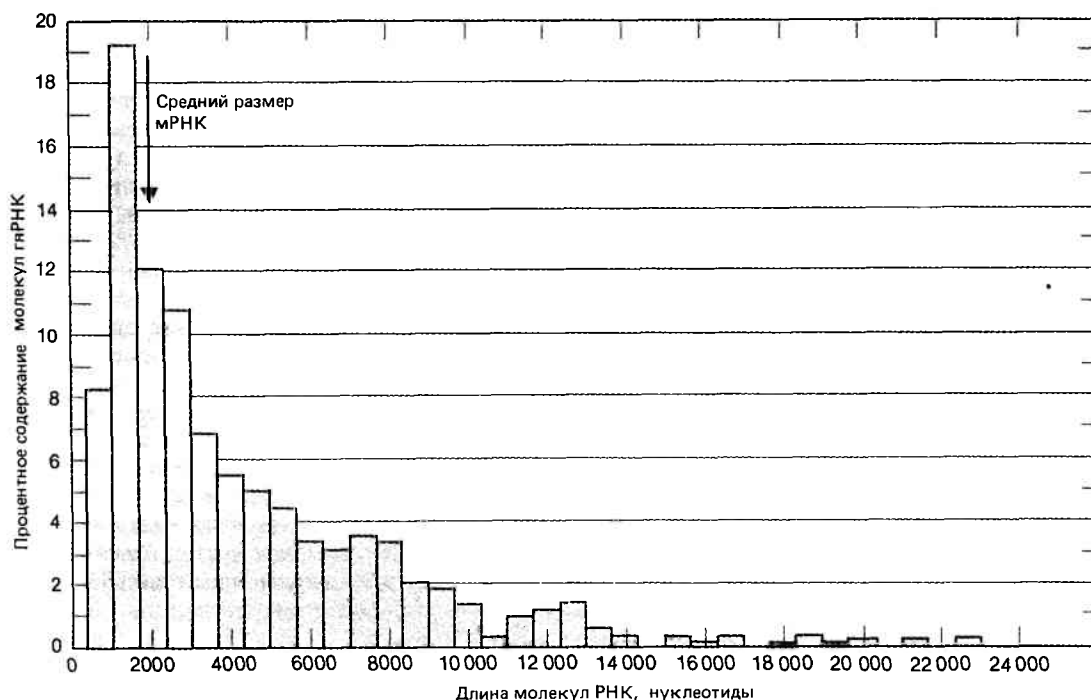


Рис. 27.1. Размеры гЯРНК варьируют в широких пределах, начиная примерно от размеров мРНК до более чем 14 000 п.н.

думать, что некоторые различия между гЯРНК и мРНК объясняются уменьшением размеров и сложности РНК, происходящим при ее сплайсинге. Многие структурные гены, кодирующие белки, имеют существенно большую длину, чем соответствующие мРНК, а первичные транскрипты должны иметь по меньшей мере такую же длину, как и ген (они могут быть и длиннее, если в их состав входят дополнительные фланкирующие последовательности).

Но мы увидим, что высокую нестабильность гЯРНК нельзя объяснить одним лишь включением в состав первичного транскрипта интронов и экзонов. В среднем только небольшая часть транскриптов каждого гена действительно преобразуется в мРНК-продукт. А в некоторых случаях имеются также транскрипционные единицы, продукты которых обнаруживаются в ядре, но из которых, по-видимому, вовсе не образуется цитоплазматическая мРНК. Таким образом, ответ на вопрос, осуществляется ли регуляция экспрессии некоторых генов на уровне гЯРНК, может быть получен, если будет установлено, образуется ли мРНК в результате процессинга специфической ядерной РНК.

гЯРНК имеет большие размеры и нестабильна

На долю транскрипции генов в нуклеоплазме приходится только небольшая часть общей РНК-синтезирующей активности клетки. Интенсивная транскрипция генов рРНК в ядрышке обеспечивает образование основной части (до 90%) клеточной РНК. Это затрудняет исследование синтеза гЯРНК в целом, поскольку, присутствуя в небольшом количестве, она трудно отделяется от общей массы рРНК. Однако эта задача может быть решена при тщательном изучении только части популяции гЯРНК или при использовании специальных условий.

Исследователь оказывается перед выбором, изучать ли ему фракции гЯРНК, имеющие самые большие размеры (и отличающиеся от предшественников рРНК, поскольку их константы седиментации $> 45S$), или избирательно ингибировать синтез рРНК актиномицином. При использовании другого метода в течение короткого времени наблюдают за судьбой импульсной (т.е. очень коротковременной) радиоактивной метки, которая включается в гЯРНК гораздо быстрее, чем в рРНК. Ни один из этих подходов не дает абсолютно убедительных результатов, но они вполне адекватно отражают общие свойства транскрипции, происходящей в нуклеоплазме.

Два основных свойства гЯРНК — это ее размер и нестабильность. Превращения основной части радиоактивной метки, включившейся в гЯРНК, практически полностью происходят в ядре. Она даже никогда не поступает в цитоплазму. Время жизни гЯРНК несколько различается в клетках разных типов. В соответствии со скоростью распада радиоактивной РНК могут быть выделены один или два компонента; но время ее полужизни обычно колеблется в пределах от нескольких минут до часа.

Превращения РНК в ядре свидетельствуют о том, что основная часть гЯРНК синтезируется и полностью распадается в ядре. При этом остается неясной судьба индивидуальных молекул или частей молекул; поэтому на основе таких данных нельзя судить о том, соответствует ли распадающийся в ядре материал только интронам или в его состав входят и целые молекулы гЯРНК.

Одна из сложностей при определении истинного размера гЯРНК связана с ее склонностью к образованию агрегатов. В гЯРНК имеются участки со вторичной структурой, большая часть которых образуется в результате реакций между обращенными повторами Alu-семейства или эквивалентного семейства (гл. 24). Кроме того, имеют место и межмолекулярные взаимодействия, которые, по-видимому, и приводят к образованию очень быстро седиментирующего материала.

На рис. 27.1 показано распределение по размерам гяРНК в сопоставлении со средним размером мРНК. На гистограмме приведены значения размеров молекул гяРНК, растянутых с целью их электронно-микроскопического исследования в денатурирующих условиях (препятствующих агрегации молекул). Возможно, что приведенные значения размеров первичных транскриптов занижены. Это может быть вызвано двумя причинами. Любой разрыв длинной молекулы при ее выделении приведет к образованию более коротких молекул. Кроме того, некоторые молекулы могут представлять собой промежуточные продукты процессинга (вспомните набор предшественников мРНК овомукоида, изображенных на рис. 26.7, каждый из которых присутствует в популяции стабильных гяРНК).

Но даже с учетом таких оговорок очевидно, что основная часть молекул гяРНК имеет гораздо большую длину, чем мРНК. В результате экспериментов с использованием импульсных меток, включающихся в новосинтезированные транскрипты, к настоящему времени сложилось представление, что средние размеры транскрипционных единиц (у млекопитающих) находятся в диапазоне от 8000 до 10 000 п.н. Весь диапазон размеров составляет от ~ 2000 до ~ 14 000 п.н. Таким образом, средний размер молекулы гяРНК в 4–5 раз больше среднего размера мРНК, составляющего 1800–2000 оснований.

На основе размеров гяРНК сразу же можно предположить, что молекула гяРНК имеет слишком большую длину, чтобы существовать в ядре в виде вытянутой цепи РНК. Действительно обнаружено, что она входит в состав рибонуклеопротеина (гяРНП), в котором гяРНК связана с несколькими белками. Эти белки, по-видимому, отличаются большим разнообразием, чем белки, связывающиеся с мРНК. Общая масса белка в гяРНП примерно в четыре раза больше массы гяРНК.

гяРНП связана с **ядерным матриксом**. Это плотная сеть фибрилл, расположенная с внутренней стороны от ядерной мембраны. Матрикс образует своего рода внутреннюю оболочку ядра, которая также распространяется и внутрь ядра. На электронной микрофотографии, приведенной на рис. 27.2, показан ядерный матрикс, находящийся среди остатков окружающей его цитоплазмы. Структура фибрилл до сих пор не установлена, хотя известно, что они состоят из большого числа белков.

Сам хроматин на протяжении многих участков связан с ядерным матриксом, и, по-видимому, индивидуальные первичные транскрипты присоединяются к матриксу вскоре после транскрипции или даже во время нее. В самом деле, их процессинг может происходить на матриксе ядра. Нам не известно, является ли такая локализация процессинга необходимой в том смысле, что он может осуществляться *только* здесь. Но, анализируя этапы созревания РНК, не следует забывать, что этот процесс может иметь сложную топографию и осуществляться не просто набором «растворимых» ферментов, передвигающихся по внутреннему содержанию ядра.

Мы практически не имеем никакой информации о том, как происходит транспорт зрелой мРНК в цитоплазму. Возможно, что он может быть пространственно связан с процессингом. Что касается пути транспорта РНК в цитоплазму, то в ядерной мембране имеется множество пор, через которые он мог бы осуществляться, однако не известно, действительно ли эти поры имеют именно такое назначение.

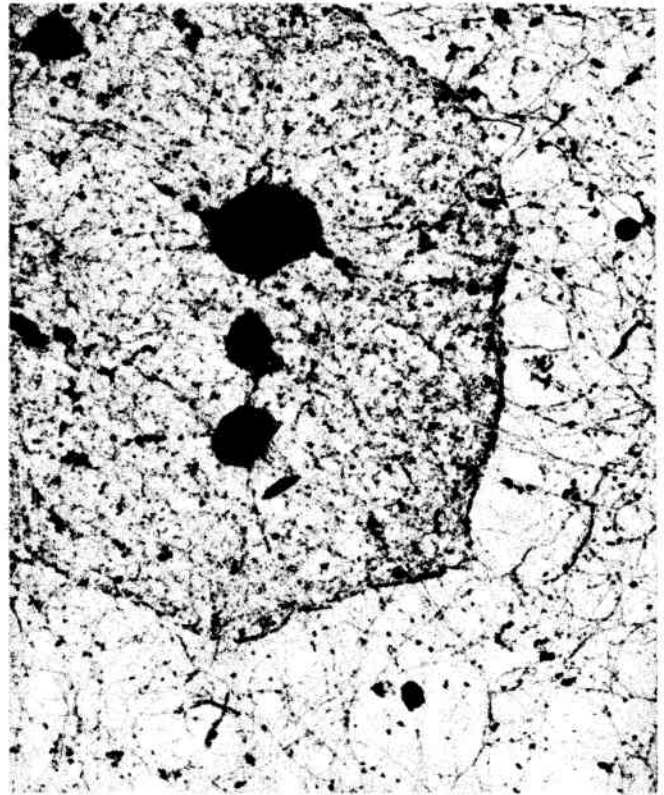


Рис. 27.2. Ядерный матрикс представляет собой сеть фибрилл. Сначала для удаления мембран проводили мягкую обработку клеток HeLa неионным детергентом, а затем для удаления хроматина и элементов цитоскелета (нитей окружающей цитоплазмы) материал обрабатывали ДНКазой и высококонцентрированным соевым раствором. Маленькие кольцевые образования на поверхности ядерного матрикса, возможно, представляют собой поры ядерной мембраны. Большие темные пятна — ядрышки. (Фотография любезно предоставлена David Carco, Sheldon Penman.)

мРНК образуется из гяРНК

На концах некоторых молекул гяРНК имеются такие же модификации, как и на концах мРНК. На 5'-конце находится метилированный (частично) кэп; 3'-конец полиаденилирован. Сходство таких концевых модификаций указывает на то, что между гяРНК и мРНК могут существовать взаимоотношения предшественник—продукт. (Кроме того, основная белковая фракция с мол. массой 74 000 дальтон, входящая в состав гяРНП и связанная с poly(A)-фрагментом,—это тот же белок, что и обнаруживаемый в мРНК.) Явное сходство концевых участков гяРНК и мРНК вначале представлялось парадоксальным, поскольку гяРНК имеет существенно большие размеры. Было не понятно, каким образом оба конца молекулы оказываются в мРНК. Теперь нам известно, что это происходит в результате уменьшения размеров молекулы при удалении интронов.

Для изучения судьбы модифицированных концов можно использовать эксперименты с вытеснением метки. Схематически такой эксперимент показан на рис. 27.3. На очень короткое время в систему вносится радиоактивный предшественник РНК. Это импульсное мечение РНК. Затем его заменяют на немеченый предшественник. Это и есть опыт с вытеснением метки, называемый также чейз-экспериментом. Далее через определенные интер-

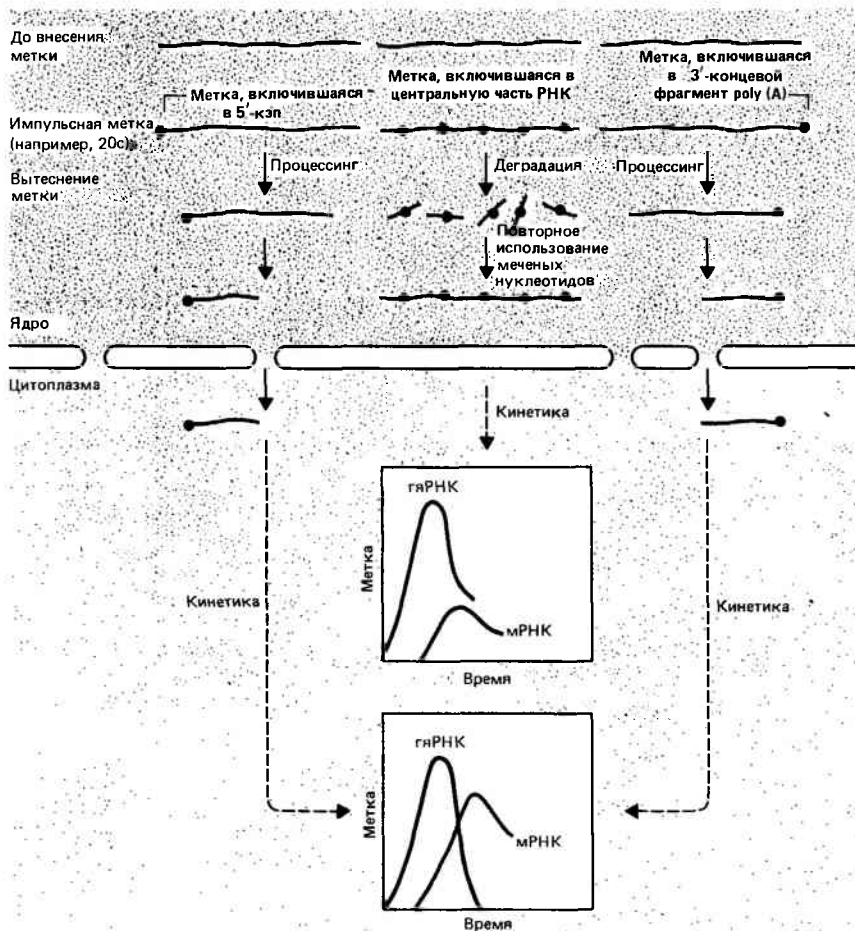


Рис. 27.3. Радиоактивная метка, включившаяся как в 5'-кэпы, так и в 3'-poly(A)-фрагменты, сначала обнаруживается в гяРНК, а затем переходит в мРНК, но основная часть метки, включившейся в центральную часть гяРНК, оказывается в продуктах ее деградации, а затем повторно используется при синтезе РНК.

валы времени определяют количество метки, включившейся в гяРНК и в мРНК.

В течение короткого времени после внесения импульсной метки она накапливается в гяРНК. Вследствие нестабильности молекул гяРНК количество метки в ее составе убывает. В зависимости от того, включилась ли метка в один из концевых участков РНК или в ее центральную часть, получаются различные результаты эксперимента.

Поскольку модифицированные концевые участки отличаются стабильностью, можно проследить их переход из гяРНК в мРНК. Действительно, основной способ, посредством которого меченые концы покидают популяцию молекул гяРНК,—это процессинг РНК, за которым следует ее транспорт в цитоплазму. Поэтому спад кривой включения метки в гяРНК сопровождается подъемом кривой включения метки в мРНК. Совпадение спада кривой включения метки в гяРНК и подъема кривой ее включения в мРНК свидетельствует о том, что концы мРНК образуются из концов гяРНК. Для этих участков молекулы между гяРНК и мРНК существуют обычные взаимоотношения предшественник—продукт.

По-видимому, после транскрипции около 20% гяРНК подвергается кэпированию. Можно обнаружить, что по меньшей мере значительная часть кэпов переходит из гяРНК в мРНК. Хотя полученные данные не свидетельствуют о том, что все 5'-кэпы гяРНК переходят в мРНК, наличие взаимоотношений предшественник—продукт в данном случае кажется правдоподобным. Кэпирование осуществляется сразу после начала транскрипции, о чем

говорит тот факт, что не удастся охарактеризовать исходные 5'-фосфатные концы первичного транскрипта.

Длина фрагмента poly(A), находящегося на другом конце молекулы ядерной РНК, слегка превышает длину poly(A)-фрагмента мРНК. Несколько оснований удаляются на стадии переноса молекулы из ядра в цитоплазму. Обсуждался вопрос о том, вся ли метка, включившаяся в poly(A), может перейти из ядра в цитоплазму. Однозначного ответа на этот вопрос пока еще не получено—количество метки, переходящей из ядра в цитоплазму, может быть различным в разных клетках и при разных обстоятельствах, но очевидно, что по меньшей мере основная часть poly(A)-фрагментов гяРНК переходит в состав мРНК.

Когда меченые нуклеотиды включаются во внутренние участки гяРНК, их основная участь—освободиться из молекулы при ее деградации. Затем они повторно используются при дальнейшем синтезе РНК, включаясь как в рРНК, так и в гяРНК. В результате наблюдается уменьшение общего количества метки в гяРНК. Небольшая часть метки включается в мРНК в цитоплазме, но, поскольку количество включившейся метки мало, ее появление в мРНК не совпадает с уменьшением ее количества в гяРНК и поэтому не может свидетельствовать о наличии взаимоотношений предшественник—продукт.

Результаты первых попыток проследить путь нуклеотидов из центральной части гяРНК в мРНК с помощью экспериментов с вытеснением метки, таким образом, не дали четкого ответа из-за тех превращений, которым

подвергается основная часть метки в ядре. Но позже усовершенствование метода позволило показать, что мРНК действительно образуется из гяРНК. Результаты таких экспериментов позволяют определить, какая часть общей массы гяРНК превращается в мРНК. В клетках млекопитающих она составляет около 5% от всей гяРНК, что с учетом пятикратных различий средних размеров гяРНК соответствует превращению (примерно) 25% молекул гяРНК.

Это означает, что уменьшением размеров транскрипта путем потери интронов можно объяснить судьбу только примерно четвертой части молекул гяРНК во всей их популяции. Остальная же часть представлена молекулами, превращения которых *полностью* происходят в ядре: из этих молекул цитоплазматическая РНК вообще не образуется. Единственный недостаток такого анализа состоит в том, что в его основе лежит определение количества poly(A)⁺-мРНК; из некоторой же части популяции молекул гяРНК может образовываться и poly(A)⁻-мРНК. Но, хотя это и приведет к увеличению количества гяРНК, приходящейся на долю предшественников мРНК, основной вывод останется прежним.

Значение полиаденилирования

Последовательность poly(A) присоединяется к 3'-концам молекул гяРНК уже после того, как закончится транскрипция. Полиаденилирование происходит под действием фермента poly(A)-полимеразы, использующего АТР в качестве субстрата. Присоединению poly(A) к гяРНК может помешать аналог АТР, 3'-дезоксиаденозин, известный также под названием **кордицепин**. Кордицепин — специфический ингибитор полиаденилирования. Он не препятствует транскрипции гяРНК, но его добавление нарушает выход мРНК в цитоплазму, что указывает на *необходимость* полиаденилирования для образования мРНК из гяРНК. [Конечно, это справедливо только в отношении poly(A)⁺-мРНК; для poly(A)⁻-мРНК полиаденилирование не требуется.]

Только часть молекул гяРНК полиаденилирована (около 30% в клетках млекопитающих по сравнению с 70% в случае мРНК). Часто наблюдается приблизительная корреляция между долей полиаденилированных молекул гяРНК и долей гяРНК, из которой образуются молекулы цитоплазматической poly(A)⁺-мРНК. Из этого можно сделать вывод, что полиаденилирование, возможно, служит сигналом для процессинга определенной молекулы гяРНК. Однако, хотя полиаденилирование и необходимо, самого по себе его, по-видимому, еще недостаточно, поскольку в противном случае соотношение poly(A) в ядре и цитоплазме было бы строго постоянно.

Один из важных нерешенных вопросов — происхождение 3'-концов, к которым присоединяется фрагмент poly(A). Мы знаем, что у некоторых вирусов эти концы образуются при расщеплении более длинных предшественников мРНК, как это показано на **рис. 27.4**. Общее свойство клеточных и вирусных мРНК — наличие последовательности AAUAAA, расположенной на участке между 11 и 30 нуклеотидами левее сайта полиаденилирования. При удалении этой последовательности из транскрипционной единицы области поздних генов вируса SV-40 расщепление РНК и полиаденилирование уже не осуществляются по обычным сайтам. Последовательность AAUAAA необходима, но недостаточна для осуществления этой реакции.

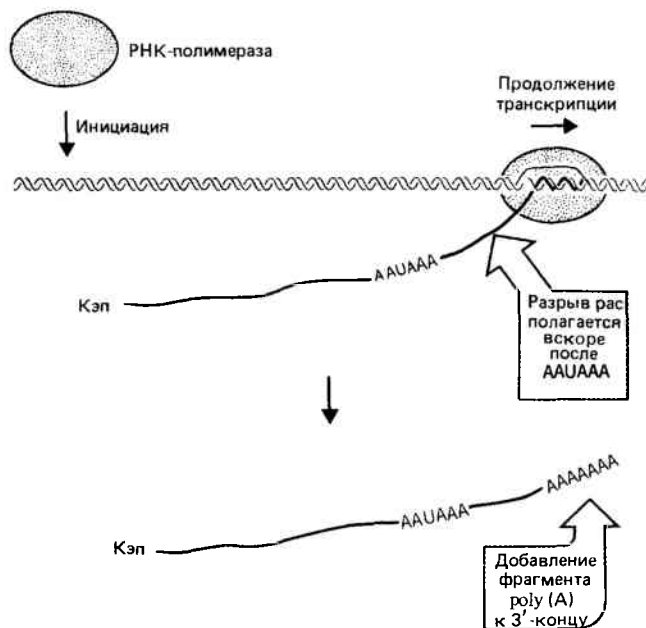


Рис. 27.4. Последовательность AAUAAA необходима для расщепления и полиаденилирования мРНК вируса SV-40.

Повсеместная распространенность сигнальной последовательности указывает на то, что тот же самый механизм — расщепление и полиаденилирование — может участвовать в экспрессии транскрипционных единиц клетки. Данных о механизме терминирования транскрипции РНК-полимеразой II практически нет; поэтому в настоящее время нам не известно, в какой доле случаев 3'-конец, используемый для полиаденилирования, образуется в результате терминирования, а не расщепления.

Полиаденилирование обычно предшествует сплайсингу РНК. При образовании клеточных транскриптов как молекулы предшественников полного размера, так и молекулы различных промежуточных размеров оказываются полиаденилированными. Изучение кинетики процессинга мРНК, транскрибируемых на поздних генах аденовируса, показало, что расщепление, приводящее к образованию 3'-конца, предназначенного для полиаденилирования, *предшествует* сплайсингу состоящей из трех частей лидерной последовательности, приводящему к образованию центральной части молекулы мРНК. (Этот процесс изображен на рис. 20.20.)

Порядок осуществления перечисленных этапов процессинга может означать, что сплайсинг — это более медленный процесс, чем транскрипция и полиаденилирование; возможно также, что этот порядок обусловлен в определенной мере зависимостью сплайсинга от полиаденилирования. Однако сплайсинг продолжается в ядрах, в которых синтез poly(A)-фрагментов подавлен добавлением кордицепина (3'-дезоксиаденозина). Это указывает на отсутствие механической причинно-следственной связи между этими событиями, хотя в тех случаях, когда для сплайсинга используются альтернативные границы экзон — интрон, расщепление с образованием 3'-конца может определять *выбор* границы сплайсинга (см., например, гл. 39).

Получение системы полиаденилирования *in vitro* облегчает изучение этих процессов. Система работает

только с РНК, синтезированной в экстракте, хотя процесс полиаденилирования отделен от завершения транскрипции значительным промежутком времени. Это, возможно, означает, что РНК может приобрести необходимую для процессинга и полиаденилирования структуру *только* во время транскрипции, даже если само присоединение poly(A) произойдет позже.

Реакция, протекающая *in vitro*, ингибируется антителами к малым ядерным РНП, в частности, антителами к мяРНП. Участвует ли U1 мяРНП в полиаденилировании так же, как и в сплайсинге? Возможно, что U1 мяРНП непосредственно не участвует в этом процессе, но входит в состав более сложной структуры, необходимой для протекания реакции.

Получена система *in vitro*, при использовании которой у гистоновых мРНК *X. laevis* (неполиаденилированных) образуются 3'-концы. При инъекции в ооциты *Xenopus* клонированного гена гистонов (ген H3 морского ежа) успешно проходят инициация его транскрипции и сама транскрипция, но терминация транскрипции осуществляется в разных сайтах. Однако при одновременной инъекции гена и ядерного экстракта из клеток морского ежа транскрибирующаяся РНК имеет надлежащий 3'-конец.

Активный компонент ядерного экстракта — 12S-фактор с мол. массой около $2,5 \times 10^5$. При фракционировании он ведет себя как комплекс белков. Однако малая РНК из клеток морского ежа, выделенная из 12S-фракции, также может вызывать образование 3'-концов.

Эта РНК имеет размер около 60 п.н. (20 000 дальтон). Она не имеет сходства ни с одной известной мяРНК. По-видимому ее функционирование в ооцитах *Xenopus* возможно вследствие наличия в них пула белков-мяРНП (snurp), с которыми РНК может взаимодействовать с образованием рибонуклеопротеиновой частицы.

Мы пока не знаем, на какой стадии функционирует такой мяРНП; это может быть стадия терминации или расщепления. Однако участие мяРНП в этом процессе служит веским доводом в пользу представления о том, что многие, а возможно и все, реакции процессинга зависят от взаимодействия с мяРНП.

гяРНК имеют более сложную организацию, чем мРНК

Исходя из данных о различиях в размерах гяРНК и мРНК, можно ожидать, что гяРНК имеет более сложные нуклеотидные последовательности. В случае тех транскрипционных единиц, сплайсинг первичных транскриптов которых приводит к образованию мРНК меньшего размера, сложность гяРНК определяется наличием как экзонов, так и интронов, тогда как сложность мРНК определяется наличием только экзонов. К сожалению, сложность гяРНК трудно определить столь же точно, как сложность мРНК. Во-первых, нет функционального теста для выделения гяРНК, свободной от других РНК. Во-вторых, возможность использования метода исследования кинетики гибридизации ограничивается наличием в гяРНК повторяющихся последовательностей, которые перемежаются с неповторяющимися последовательностями, образующими основную часть мРНК.

При определении способности гяРНК к насыщающей гибридизации с неповторяющейся ДНК ее сложность оказывается больше сложности мРНК минимально в 4 раза и максимально в 10 раз. Гибридизационные экспе-

рименты свидетельствуют о том, что популяция молекул гяРНК действительно включает все последовательности, присутствующие в цитоплазматической мРНК. Какова природа остальных последовательностей?

Этот вопрос может быть сформулирован и иначе. Представляют ли собой молекулы гяРНК, полностью превращающиеся в ядре, транскрипты с *тех же самых* генов, с которых транскрибируется мРНК, или же это транскрипты с *других генов*. Эти возможности рассматриваются на рис. 27.5.

В случае тех же самых генов в типичной клетке млекопитающих средний ген транскрибируется с образованием молекул гяРНК, 75% которых полностью деградируют, а у 25% молекул 20% их длины (экзоны) переходят в мРНК. Таким образом, процессинг оказывается чрезвычайно расточительным. В рассматриваемом случае различия в сложности гяРНК и мРНК полностью объясняются потерей интронов. В примере с клетками млекопитающих сложность гяРНК и мРНК различается в четыре раза, что очень близко к различиям в размерах молекул. Эти данные недостаточно точны, чтобы можно было утверждать, что *все* последовательности гяРНК, обнаруживаемые в мРНК, представляют собой интроны, но, по всей вероятности, основная причина состоит именно в этом.

Если образование молекул гяРНК, которые сохраняются или полностью деградируют, обеспечивается различными генами, то 75% транскрипционных единиц должны соответствовать молекулам гяРНК, все копии которых деградируют: ни из одной из них не образуется цитоплазматическая мРНК. (Строго говоря, ни из одной молекулы не образуется мРНК в клетках данного типа; в других клетках она может и образовываться.) На остальных 25% генов образуются транскрипты, все копии которых созревают с образованием мРНК (сохраняющей 20% их длины). В этом случае сложность гяРНК должна быть существенно выше, чем сложность мРНК, не только вследствие большей длины молекул гяРНК, но и из-за того, что в ее состав входит набор транскриптов, из которых мРНК не образуется. В некоторых тканях морского ежа, где сложность ядерной и цитоплазматической РНК различается в 10 раз, это различие, по-видимому, слишком велико, чтобы его можно было объяснить только потерей интронов.

Конечно, две приведенные модели представляют собой крайние случаи, и нет сомнения, что в реальном ядре имеет место промежуточная ситуация. Однако ясно, что либо процессинг транскриптов экспрессирующихся генов — процесс малоэффективный, либо некоторые транскрипционные единицы экспрессируются с образованием только ядерной РНК.

Осуществляется ли регуляция на посттранскрипционном уровне?

Тот факт, что только небольшая часть генома представлена в виде гяРНК (скажем, менее 5% в случае клеток млекопитающих), указывает на первостепенную роль регуляции транскрипции в экспрессии генов. Нам известно, что некоторые гены транскрибируются только в тех клетках, где они экспрессируются. Два примера такого рода — гены глобина (которые транскрибируются только на определенных стадиях развития клеток эритроидного ряда) и гены овальбумина (которые транскрибируются только при воздействии эстрогенов на яйцеводы кур). Ко-

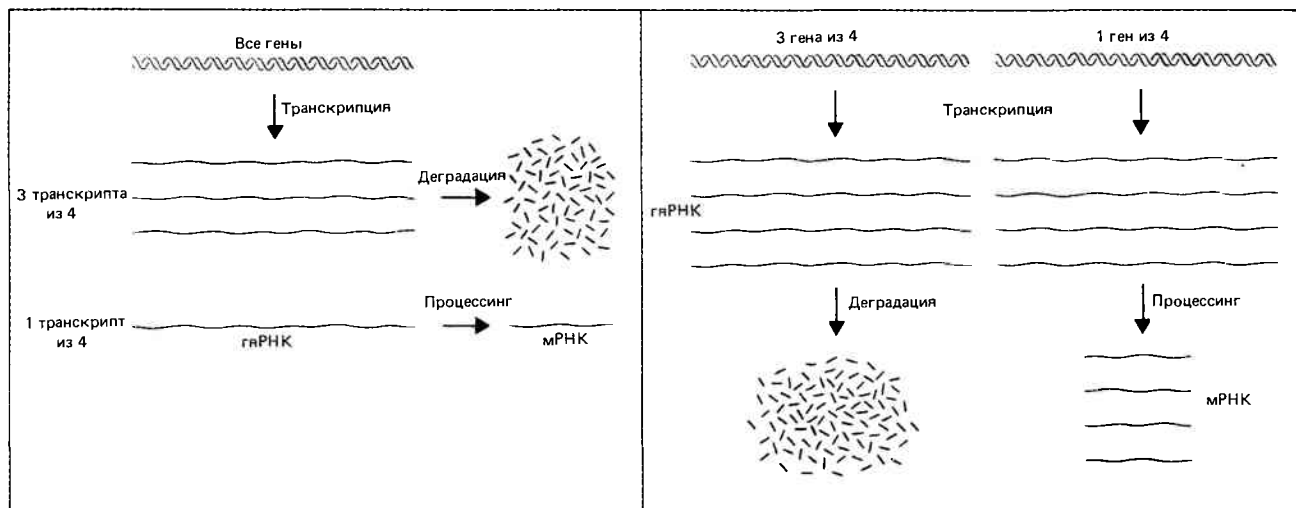


Рис. 27.5. Происходит ли потеря основной части всех гяРНК при их процессинге, как это показано в левой части рисунка, или он эффективно протекает в случае некоторых транскриптов, тог-

да как из других гяРНК вообще не образуется мРНК, как это показано в правой части рисунка?

нечно, такие гены экспрессируются очень активно, и специфический способ их регуляции может отличаться от способов регуляции других генов, экспрессирующихся менее активно.

Количественные различия в экспрессии генов могут проявляться на этапе между стадиями транскрипции и трансляции. В случае некоторых последовательностей обнаруживается разное число их копий в ядерной и в цитоплазматической РНК. Это означает, что интенсивность транскрипции — не единственный фактор регуляции содержания мРНК в клетке. Колебания стабильности РНК или эффективности процессинга могут иметь значение для изменения относительного количества последовательностей РНК при их транспорте из ядра в цитоплазму.

Общие количественные изменения обнаруживаются при переходе клеток из состояния покоя (т.е. когда клетки не делятся) в состояние роста (когда клетки активно проходят через стадии клеточного цикла).

В растущих клетках содержится больше РНК, чем в покоящихся клетках, и, в частности, в них содержится большее количество мРНК по сравнению с количеством рРНК. Однако в этих клетках не наблюдается изменений в скорости транскрипции гяРНК, в процентном содержании полиаденилированных молекул или в стабильности мРНК. Если что и происходит в растущих клетках, так это превращение большего количества poly(A)⁺-гяРНК в poly(A)⁺-мРНК.

Основная часть различий носит скорее количественный, а не качественный характер. Имеет место экспрессия некоторых новых последовательностей, но в большинстве случаев наблюдаемые изменения соответствуют эффективности, с которой один и тот же набор последовательностей гяРНК превращается в мРНК. Иными словами, в покоящихся клетках эффективность такого превращения уменьшается. В свете сказанного выше о природе последовательностей, встречающихся только в ядре, становится ясно, что, по крайней мере в покоящихся клетках, некоторые из этих последовательностей представляют собой избыточные копии экспрессирующихся последовательностей.

Обнаружено, что в клетках морского ежа регуляция отбора последовательностей может осуществляться на уровне ядерной РНК. У *S. purpuratus* значения суммарной сложности ядерных РНК из нескольких взрослых и эмбриональных типов клеток находятся в одних и тех же пределах (от 175 000 до 225 000 т.н. при средней длине молекул, равной 8800 нуклеотидам). Это указывает на отсутствие какого-либо сходства со значениями сложности молекул мРНК, составляющими от 6000 до 26 000 т.н. при среднем размере молекул 2000 оснований.

На рис. 27.6 сопоставляются последовательности РНК, присутствующие в двух тканях морского ежа. Сложность мРНК бластулы гораздо выше. Она содержит некоторые последовательности, входящие в состав и мРНК клеток кишечника, но основная часть мРНК бластулы обнаруживается только на эмбриональной стадии развития. Большая часть общих последовательностей (~80%) обнаруживается в составе ядерной РНК как клеток кишечника, так и бластулы.

Это означает, что в ядрах клеток кишечника имеются РНК-предшественники, содержание которых ограничено пределами ядра и из которых не образуется мРНК. Но в бластуле из тех же самых последовательностей образуются мРНК. Таким образом, в этом случае многие гены транскрибируются в тканях независимо от того, будут ли они в конечном счете экспрессироваться. Отбор транскриптов, которые подвергнутся процессингу и транспорту из ядра в цитоплазму, осуществляется только в определенных тканях. По-видимому, на уровне транскрипции осуществляется регуляция экспрессии только небольшого числа генов.

Модели регуляции экспрессии генов

Пока еще рано делать какие-либо общие выводы о том, насколько часто регуляция экспрессии генов осуществляется на том или ином уровне. Понятие «уровень регуляции» предполагает, что экспрессия гена, раз начавшись, не обязательно автоматически продолжается. В принципе специфическая для данного гена регуляция

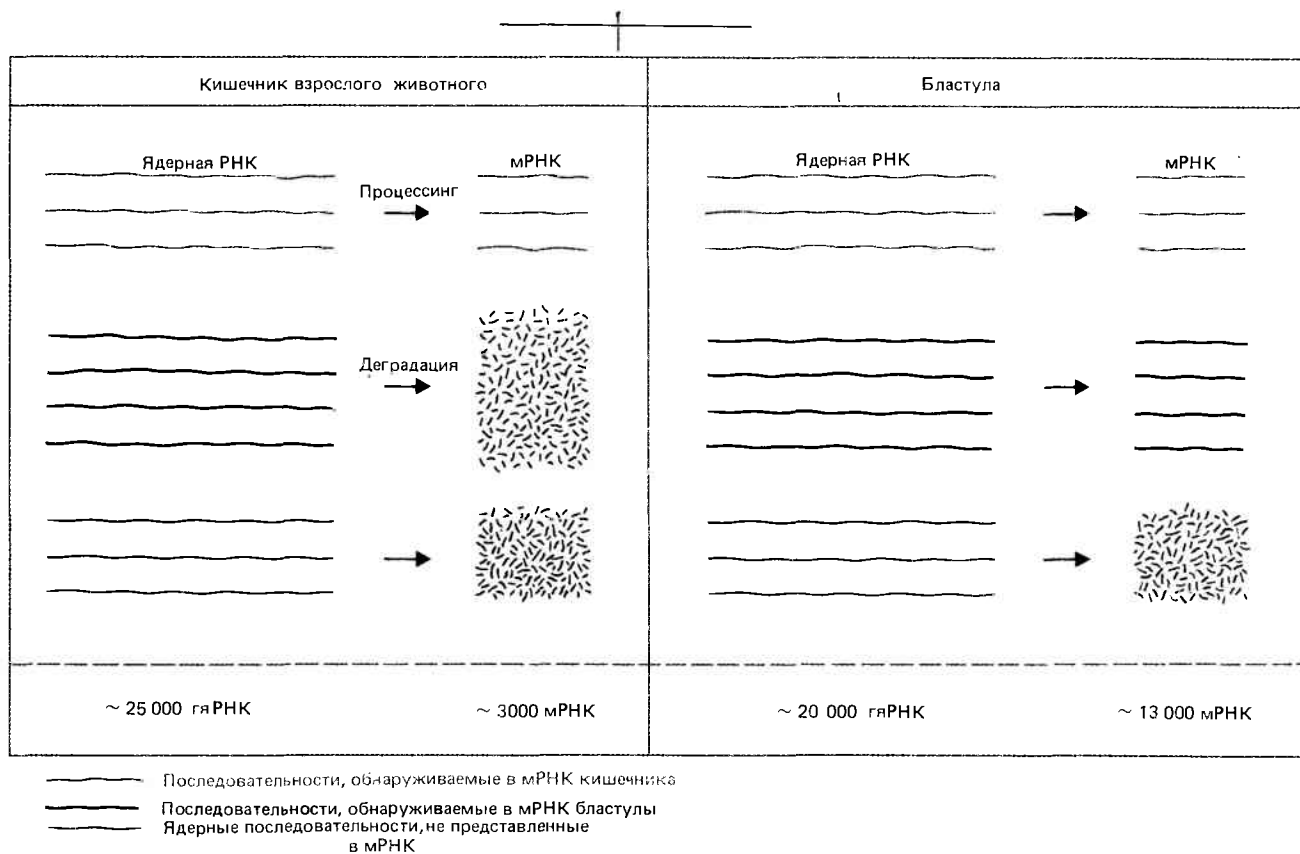
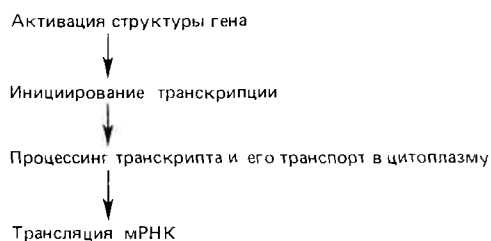


Рис. 27.6. мРНК бластулы морских ежей образуется из тех же самых ядерных транскриптов, которые деградируют в клетках кишечника взрослого организма.

его экспрессии может осуществляться на любом из нескольких последовательных этапов. Можно выделить по меньшей мере четыре этапа осуществления регуляции экспрессии гена:



На существование первого уровня регуляции указывает тот факт, что гены сами по себе могут находиться в одном из двух структурных состояний. Гены находятся в «активном» состоянии по отношению к основной части генома в тех клетках, в которых они транскрибируются. Природа такого состояния гена, о наличии которого судят по изменениям связей белков с ДНК, рассматривается в гл. 30. Изменение структуры гена предшествует транскрипции (возможно, их разделяет несколько делений клетки). Из этого можно сделать вывод, что возникновение «активной» структуры может быть первым этапом экспрессии генов.

Транскрипция гена, находящегося в активном состоянии, регулируется на стадии инициации, т.е. при взаимодействии РНК-полимеразы с промотором. На этом этапе регуляцию уже можно изучать в системе *in vitro* (гл. 11). В настоящее время пока еще не имеется данных о регуля-

ции на следующих стадиях транскрипции в эукариотических клетках, например с помощью механизмов, подавляющих терминацию транскрипции (антитерминация).

Первичный транскрипт во всех случаях должен быть модифицирован посредством каэпирования, а обычно также и полиаденилирования. Из транскриптов прерывистых генов должны быть удалены интроны. Зрелая РНК должна быть экспортирована из ядра в цитоплазму. Пока еще невозможно провести различие между этими этапами с точки зрения их роли в регуляции генной экспрессии. Следовательно, вполне возможно, что регуляция путем отбора последовательностей на уровне ядерной РНК может происходить на любом из этих этапов или же сразу на всех. Какие-либо предположения о молекулярных механизмах (как качественных, так и количественных) регуляции генной экспрессии на уровне ядерной мРНК отсутствуют.

Наконец, может осуществляться специфическая регуляция трансляции мРНК в цитоплазме. Данные об использовании этого механизма во взрослых соматических клетках немногочисленны. Однако в некоторых эмбриональных клетках, как описано в гл. 9, этот механизм обнаружен. Предполагают, что при помощи этого механизма блокируется инициирование трансляции некоторых мРНК специфическими белковыми факторами.

От вопроса, касающегося использования и типа механизма, работающего на каждом из потенциально возможных уровней регуляции экспрессии генов, мы должны перейти к общему вопросу о координации экспрессии генов между различными локусами. Гены, в результате экспрессии которых клетка приобретает специфический фенотип, могут занимать самые разные положения. Функционирование каждого гена, по-видимому, необходимо для создания набора клеток с определенным фенотипом. Каким образом в каждой клетке происходит активация

«правильной» комбинации генов, обеспечивающих ее специфический фенотип?

Принято считать, что в случае эукариотических генов имеет место положительная регуляция их активности: пока ген не включен, он неактивен. Это согласуется с предположением, что ген может экспрессироваться только после его перехода из обычного неактивного состояния в активное. На последующих этапах может осуществляться как отрицательная, так и положительная регуляция. Модели регуляции экспрессии эукариотических генов обычно предполагают наличие положительной регуляции, которая, однако, легко может быть заменена на отрицательную.

Имеются два общих пути для решения проблемы экспрессии перекрывающихся комбинаций генов в клетках с различным фенотипом. Либо гены повторяются, либо экспрессия отдельных генов регулируется множеством способов.

В предельном случае вместо повторяющихся генов в клетках каждого типа могла бы находиться одна копия гена, который должен экспрессироваться. Каждая копия при этом регулировалась бы своим способом, специфическим для данного типа клеток. Независимо от того, кодируют ли гены идентичные или сходные белки, не возникает проблем с осуществлением естественного отбора, если имеется потребность в каждом гене в определенном момент времени или в определенном месте. Конечно, некоторые гены (может быть, их и большинство) присутствуют в количестве большем, чем одна копия на геном. Нам до сих пор не известно, регулируется ли обычно экспрессия этих копий по-разному или они, как правило, представляют собой повторяющиеся последовательности, экспрессирующиеся одновременно и (или) в одном и том же месте. Начинает казаться, что обычно копии генов, по крайней мере слегка, различаются. На основе имеющегося к настоящему времени ограниченного числа данных можно предположить, что разные копии гена, по-видимому, экспрессируются в клетке в разных ситуациях.

Возникновение различающихся экзонов обусловлено не повторяемостью гена в целом, а наличием какого-то другого пути изменений его продукта. Так, изменения в сплайсинге могут привести к синтезу белков с различиями только в определенной части их аминокислотной последовательности. Такой же механизм может лежать в основе образования мРНК, нетранслируемые участки которых различаются, хотя белок и остается неизменным.

В тех случаях, когда имеется только одна копия гена, а его экспрессия бывает необходима в разных ситуациях, одна из возможностей реализации этого процесса состоит в снабжении гена множеством регуляторных элементов, каждый из которых может независимо обеспечивать его экспрессию; такие способы организации обнаружены в бактериальных геномах и их существование может быть постулировано и для эукариот.

Фенотип клетки есть результат координированной экспрессии определенного набора генов. Возможно, что в тех случаях, когда необходима координированная регуляция разбросанных по геному генов, в каждом локусе имеется один и тот же регуляторный элемент.

Если один и тот же ген экспрессируется при различных обстоятельствах в комбинации с различными другими генами, в его регуляции, по-видимому, участвуют разные элементы. Такую ситуацию обычно рассматривают по аналогии с моделью оперона. Довольно

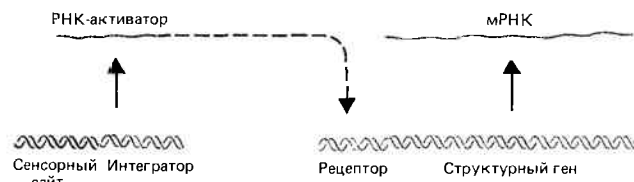


Рис. 27.7. К основным компонентам системы регуляции активности эукариотических генов могут относиться ген-интегратор с сенсорным сайтом и структурный ген с рецепторным сайтом, находящимся под контролем продукта гена-интегратора.

короткие последовательности ДНК могут служить сайтами узнавания при регуляции активности расположенных по соседству структурных генов. Нет необходимости объяснять, *каким* образом они регулируют активность генов: их действие может осуществляться путем структурных изменений, предшествующих транскрипции, оно может быть связано с ее инициацией в промоторном участке или влиянием на ядерный РНК-транскрипт.

Формальный анализ таких возможностей осуществили Бриттен и Дэвидсон (Britten, Davidson) на основе предложенной ими общей модели, показанной на рис. 27.7. Положительную регуляцию активности структурного гена обеспечивает прилегающий к нему **рецепторный сайт**. Его строение соответствует строению молекулы **активатора**, который в соответствии с данной моделью представляет собой РНК, но также может быть и белком. Активатор синтезируется в результате работы *гена-интегратора*, который, несомненно, является прямым аналогом бактериального гена-регулятора. Единственное различие между ними заключается в том, что ген-интегратор сам подвергается внешней регуляции. Участок, ответственный за регуляцию его активности, называется **сенсорным сайтом**. Предполагается, что сенсорный сайт взаимодействует с химическими соединениями, регулирующими экспрессию гена, например с гормонами. Предложенная модель объясняет, как происходит регуляция активности гена до транскрипции, но ее можно легко применить и к другим уровням регуляции.

Один активатор может контролировать работу многих генов в том случае, если каждый локус-мишень имеет копию соответствующего рецептора. Убедительным доводом в пользу существования таких повторяющихся участков служит наличие у бактерий, рассеянных по геному локусов, находящихся под общим контролем (гл. 15). Все гены, обладающие одним и тем же рецепторным участком, образуют **набор** генов, экспрессирующихся в результате взаимодействия с активатором. Такой набор в некотором смысле эквивалентен оперону бактерий, поскольку все относящиеся к нему гены всегда экспрессируются вместе. Из этого можно сделать вывод, что в его состав входят гены, кодирующие белки со сходными функциями, например ферменты одного метаболического пути.

Структурный ген может экспрессироваться в разных условиях, если он обладает более чем одним рецептором. Каждый рецептор может активировать ген путем взаимодействия с молекулой активатора. При наличии ряда рецепторов, расположенных по соседству со структурным геном, он может входить в состав нескольких разных наборов, как показано на рис. 27.8. В зависимости от степени повторяемости рецепторов каждый из трех струк-

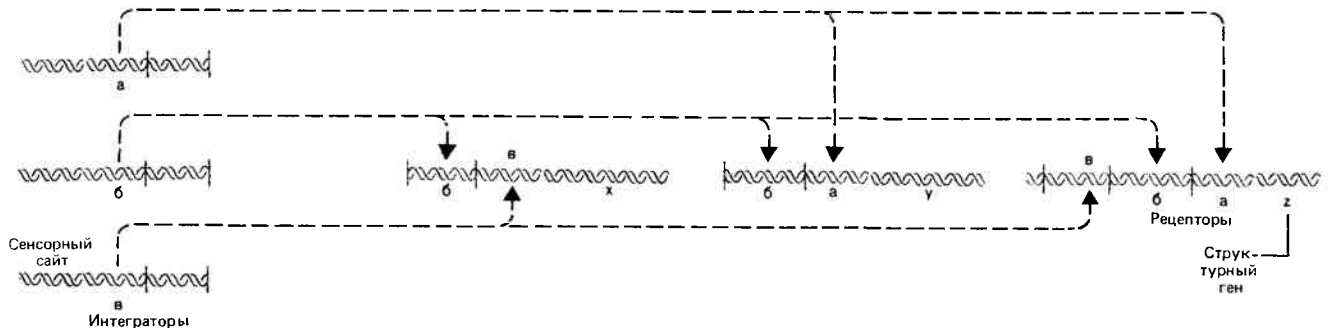


Рис. 27.8. Каждый структурный ген может быть активирован более чем одним рецептором, а потому и относиться к более чем одному набору генов.

турных генов (x, y, z) относится к двум из трех частично перекрывающихся наборов генов (a, б, в).

Регуляция активности различных наборов генов может происходить одновременно, если предположить, что сенсорный сайт способен регулировать активность нескольких интеграторов. Из этого следует, что одно стимулирующее воздействие на один сенсорный сайт приводит к активации нескольких наборов генов. Гены, находящиеся под контролем одного сенсорного сайта, называются **батареей**. Набор генов может входить в состав более чем одной батареи, если входящие в состав интегратора элементы повторяются, как это показано на рис. 27.9.

При такой иерархии элементов каждый индивидуальный ген способен экспрессироваться в сочетании с любым другим геном, что обеспечивает экономичность регуляции. В соответствии с рассмотренной моделью нуклеотидные последовательности интегратора и рецептора должны повторяться; поэтому их следует относить к повторяющейся ДНК. И действительно, основным толчком к созданию модели послужило стремление объяснить роль повторяющихся последовательностей в геноме.

К сожалению, однако, оказалось, что повторяющиеся последовательности, по крайней мере те, которые выявляются по кинетике их гибридизации, по-видимому, не

имеют никакого отношения к регуляции. Если регуляторные последовательности-повторы и имеются, то они, вероятно, представляют собой короткие канонические последовательности, слишком короткие для того, чтобы их можно было обнаружить с помощью гибридизационного анализа.

Независимо от того, участвует ли РНК в регуляции экспрессии генов в естественных условиях, но в искусственно созданных системах она может быть использована для этой цели. Если ген ориентирован противоположным образом относительно промотора, то может транскрибироваться имеющая противоположное направление цепь РНК. При введении в клетки имеющего противоположную ориентацию гена тимидин-киназы он ингибирует синтез этого фермента. Мы не знаем механизма ингибирования; противоположно направленная РНК может препятствовать транскрипции нормального гена, процессингу его РНК-продукта или трансляции мРНК.

Такой метод открывает большие возможности для выключения определенных генов; например, можно было бы исследовать функцию регуляторного гена, вводя в клетки его противоположно направленную (обращенную) копию. Если обращенный ген расположен так, что он будет находиться под контролем промотора, который сам является объектом регуляции, ген-мишень может

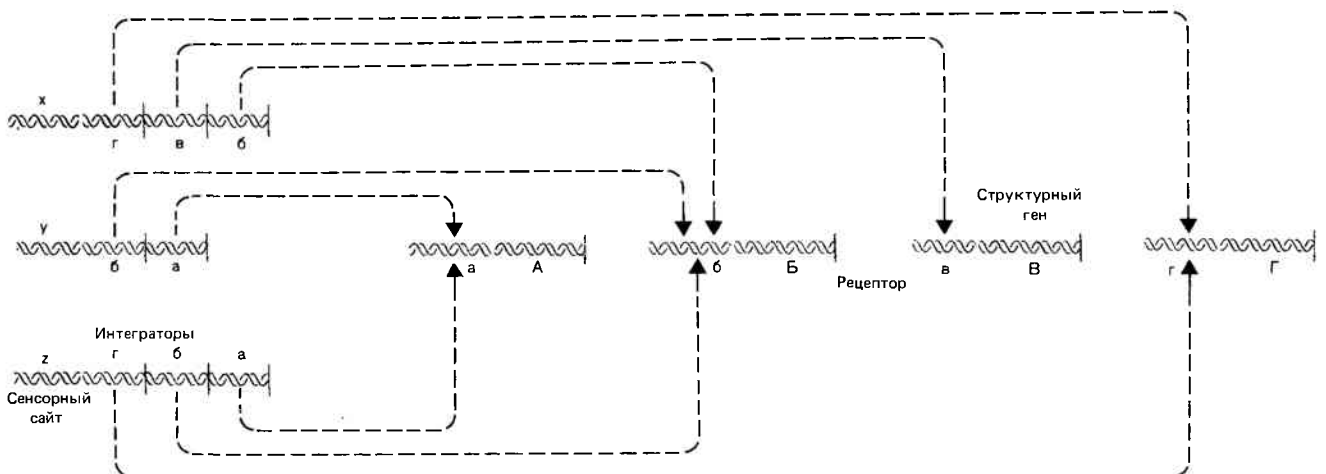


Рис. 27.9. Интеграторы могут повторяться, и каждый сенсорный сайт, вероятно, регулирует активность нескольких интеграторов, так что сигнал, поступающий на один сенсорный сайт,

способен активировать несколько наборов генов, а каждый набор может входить в состав более чем одной группы.

быть выключен или включен (по желанию) путем регуляции образования обращенной РНК. Этот метод, поможет вероятно, исследовать роль такого фактора, как время осуществления экспрессии определенного гена.

Рассмотренный подход эффективен в отношении как эукариотических клеток, так и бактерий. В случае эукариот обращенный ген соответствует нормальному гену, ориентированному противоположным образом. При некоторых способах регуляции экспрессии генов у бактерий РНК осуществляет регуляторную функцию благодаря своей комплементарности с последовательностью РНК-мишени. Регуляция может происходить на этапе репликации ДНК (гл. 33) или транскрипции (гл. 36). Направленную регуляцию можно осуществлять с помощью малой РНК, комплементарной 5'-концу мРНК-мишени.

Роль клеточных полипротеинов

В случае классического оперона координированная регуляция функционирования группы генов достигается тем, что они входят в состав одной транскрипционной единицы. При необходимости координированной регуляции активности рассеянных генов в каждом локусе должен присутствовать один и тот же регуляторный элемент. Этот механизм, значение которого доказано для бактериальных геномов (гл. 15), в случае эукариотических геномов является гипотетическим (см. предыдущий раздел). Когда необходимо, чтобы ген экспрессировался при разных обстоятельствах, регуляция его активности может осуществляться одним из двух способов: 1) использованием различных регуляторных элементов, каждый из которых способен независимо обеспечивать его экспрессию (и опять-таки это обнаружено в бактериальных геномах и может быть постулировано для геномов эукариот) или 2) повторяемостью самого структурного гена, обеспечивающей наличие в каждой его копии своего регуляторного элемента. (В определенном смысле это предположение противоположно модели Бриттена-Дэвидсона.) Активность некоторых членов семейств эукариотических генов может, по-видимому, регулироваться именно таким образом. (Повторяющиеся гены также, возможно, кодируют неидентичные, хотя и обладающие большим сходством, белки. Различия между белками могут отражать эволюцию различий в регуляции активности генов, так что каждый белок наилучшим образом соответствует условиям своего функционирования.)

Некоторые из рассмотренных вариантов наиболее ярко проявляются при синтезе **полипротеинов**. Этот термин впервые был использован для обозначения полифункциональных продуктов РНК-содержащих геномов ретровирусов. Несколько вирусных белков синтезируются в виде общего предшественника, расщепляемого в определенных участках с образованием индивидуальных функционально активных белков. Эта ситуация формально соответствует общему правилу, согласно которому эукариотические системы трансляции включают только моноцистронные мРНК, но при этом в результате одного акта трансляции может синтезироваться несколько белков.

К настоящему времени обнаружено несколько видов клеточных полипротеинов. В некоторых случаях в состав полипротеина входит ряд идентичных или сходных полипептидов. Так, в состав предшественника энкефалина входят шесть копий Мет-энкефалина (пептида Tyr-Gly-Gly-Phe-Met) и одна копия Лей-энкефалина (Tyr-Gly-Gly-Phe-

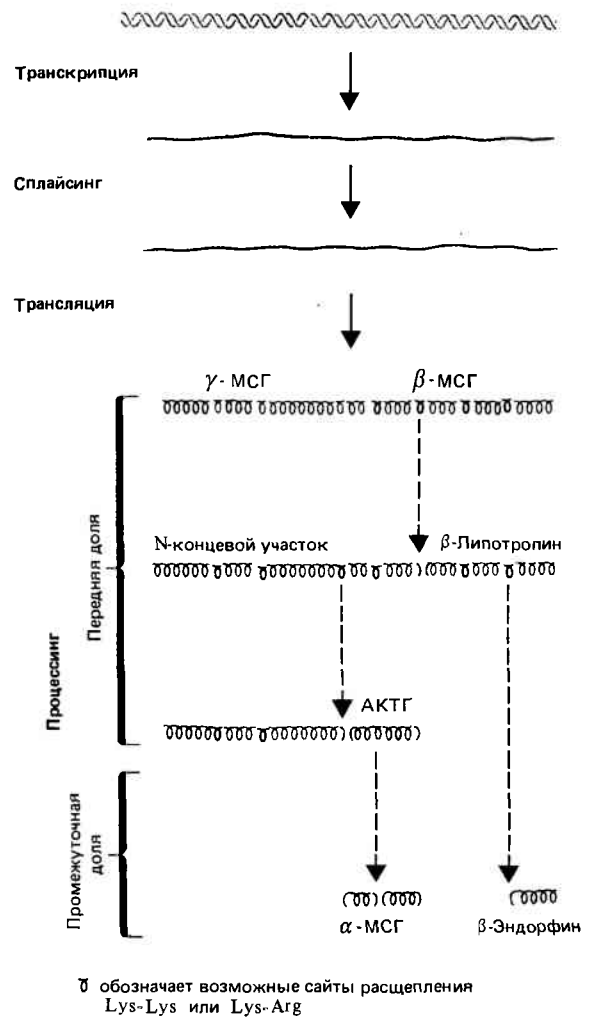


Рис. 27.10. Ген ПОМК экспрессируется с образованием полипротеина, продукты которого определяются тканеспецифичностью процессинга.

Leu). Полипротеин в этом случае образуется как бы в результате амплификации, поскольку при каждом цикле транскрипции и трансляции образуется несколько копий полипептидного продукта.

Более сложная ситуация, имеющая место в случае гена, кодирующего синтез проопиомеланокортина (ПОМК), изображена на **рис. 27.10**. Соответствующие гены ряда млекопитающих имеют сходное строение и содержат один интрон, находящийся в сигнальном участке кодирующей последовательности, и один экзон, кодирующий несколько белков. Для каждого вида выявлен только один такой ген, хотя в геноме могут присутствовать и другие (не обязательно идентичные) его копии. В состав единственного полипептидного продукта с мол. массой 32 000 дальтон входит несколько полипептидных гормонов, а также другие участки, аминокислотные последовательности которых говорят о том, что они, возможно, тоже имеют биологическое значение. Все реально или потенциально существующие индивидуальные продукты имеют сходные границы, которые представлены парами основных аминокислот, образующих участки, расщепляемые трипсиноподобными ферментами. В двух тканях, в которых экспрессируется ген ПОМК, расщепление продукта происходит по разным сайтам.

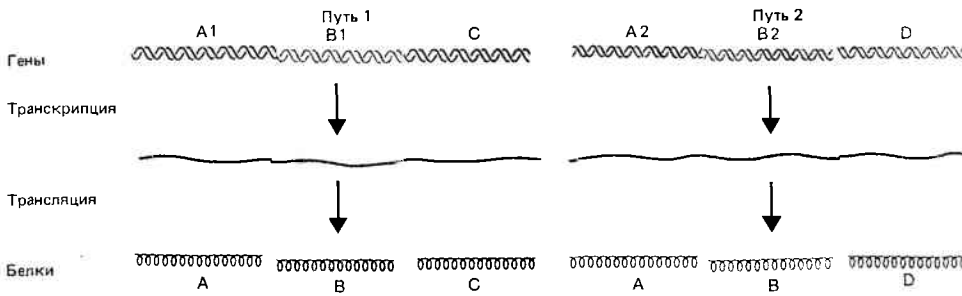


Рис. 27.11. Кодируют ли эти гены полипептиды, состоящие из перекрывающихся продуктов?

В передней доле гипофиза в белок ПОМК сначала вносится один разрыв, в результате которого образуется N-концевой фрагмент и β -липотропин (C-концевой участок, состоящий из 70–90 аминокислот). Затем N-концевой фрагмент расщепляется с образованием гормона АКТГ (длиной 39 аминокислот). В этой ткани реакция дальше не происходит.

В промежуточной доле гипофиза АКТГ расщепляется, образуя α -меланотропин (α -МСГ), состоящий из 13 N-концевых аминокислот. β -Липотропин также расщепляется еще в одном сайте, что приводит к образованию анальгетика β -эндорфина, включающего 31 C-концевую аминокислоту.

Кроме этих хорошо известных гормонов имеются еще две другие последовательности, которые потенциально могут обладать меланоцитстимулирующей активностью — их обозначают β -МСГ и γ -МСГ. Однако до сих пор не известно, когда они генерируются и генерируются ли вообще.

Для чего нужна такая форма организации? Нам до сих пор не ясны все функции, выполняемые этими гормонами, хотя известно, что они синтезируются в ответ на болевой стресс. Функциональное значение перекрывания последовательностей различных гормонов будет выяснено лишь тогда, когда мы лучше изучим их биологическую активность.

Приведенные данные указывают на то, что один ген способен кодировать различающиеся (хотя и перекрывающиеся) аминокислотные последовательности путем изменений в ходе процессинга белка (регуляция которого, по-видимому, осуществляется с помощью ферментов процессинга соответствующих тканей). Другая возможность заключается в том, что наличие повторяющихся копий генов, кодирующих полипептиды, может приводить к тому, что *один и тот же белок может синтезироваться в сочетании с различными другими белками*. Такой процесс в общих чертах изображен на рис. 27.11. Два гена для полипептидов содержат последовательности, соответственно кодирующие индивидуальные белки А, В, С и А, В, D. Активация каждого гена приводит к синтезу белков А и В, но в одном случае она сопровождается синтезом белка С, а в другом случае — синтезом белка D. Эта ситуация формально эквивалентна наличию повторяемости структурных генов в полицистронных транскрипционных единицах, несмотря на то что в данном случае белковые продукты разделяются на стадии процессинга белка, а не на стадии трансляции.

К этой категории, возможно, относится семейство генов, кодирующих ЭЛГ, нейропептид морского моллюска *Aplysia*. ДНК *Aplysia* содержит примерно 5 генов, кодирующих гормон ЭЛГ (полипептид, состоящий из 36 ами-

нокислот). Хотя полный набор генов еще не охарактеризован, было обнаружено несколько видов мРНК, кодирующих ЭЛГ. Эти мРНК транслируются с образованием разных полипептидов. В состав всех полипептидов входит последовательность ЭЛГ, но остальные последовательности соответствуют другим белковым продуктам, которые могут представлять собой иные полипептидные гормоны. В разных тканях возможен синтез разных мРНК, что позволяет в каждом случае синтезировать специфическую комбинацию гормонов. Таким образом, с помощью этого механизма путем активации определенных членов семейства повторяющихся генов могут синтезироваться перекрывающиеся наборы гормонов.

Видно, что в случае рассмотренных нами клеточных полипептидов образуются продукты небольшого размера (состоящие менее чем из 40 аминокислот). Отражает ли это тот факт, что сложности, возникающие при синтезе небольших белков *de novo*, приводят к необходимости их образования путем расщепления более длинных предшественников (и, таким образом, открывают путь для эволюции различающихся белковых продуктов)? Было бы интересно охарактеризовать гены, кодирующие множество сходных полипептидов. Действительно ли, что их сборка осуществляется путем «перетасовки» экзонов, аналогичной рассмотренной в гл. 20, или весь процесс в целом возник в результате эволюции одной из следующих стадий экспрессии повторяющихся генов, гены-предки которых не кодировали такие множественные функции?

Рекомендуемая литература

Свойства ядерной РНК детально рассматриваются в книге Льюина (*Lewin, Gene Expression, 2, Eucaryotic Genomes, Wiley, New York, 728–760, 1980*). Обзор, посвященный взаимосвязи гяРНК и мРНК, был опубликован ранее (*Cell, 4, 11–20 и 77–94, 1975*). Более современные данные, касающиеся сигналов для полиаденилирования и его взаимосвязи со сплайсингом, можно найти в оригинальных статьях Фитцджеральда и Шенка (*Fitzgerald, Shenk, Cell, 24, 251–260, 1981*) и Зиви, Невинса и Дарнелла (*Zeevi, Nevins, Darnell, Cell, 26, 39–46, 1981*). Модель регуляции активности генов была опубликована Бриттенем и Дэвидсоном (*Britten, Davidson, Science, 165, 349–357, 1969*). Полипептиды млекопитающих рассмотрены в обзоре Герберта и Ухлера (*Herbert, Uhler, Cell, 30, 1–2, 1982*). Полипептидам млекопитающих посвящен также обзор Дугласа, Чивелли и Херберта (*Douglas, Civelli, Herbert, Ann. Rev. Biochem., 53, 665–715, 1984*). Семейство генов, кодирующее нейропептид моллюска *Aplysia*, обнаружили Шеллер и др. (*Scheller et al., Cell, 28, 707–719, 1982*).

Часть VIII

УПАКОВКА ДНК

Есть ли ядро у бактерий? Этот вопрос долгое время привлекал внимание многих исследователей. Однако теперь в нем, кажется, не больше смысла, чем в вечном споре об авторстве пьес Шекспира – споре, разрешившемся тем, что автором был, вероятно, не Уильям Шекспир, о котором шла речь, а другой человек с тем же именем. В самом деле, главное свойство ядра растительной или животной клетки – не внешний вид или форма, а те упорядоченные изменения, которые соответствуют строго определенным биологическим функциям и которые можно непосредственно наблюдать.

Эрвин Чаргафф, 1947

Глава 28

О ГЕНОМАХ И ХРОМОСОМАХ

В общем виде рассматриваемая проблема сводится к тому, каким образом упакована ДНК в фагах и вирусах, в бактериальных клетках и эукариотических ядрах. Длина вытянутой молекулы ДНК во много раз превышает размеры содержащего ее компартмента. Поэтому ДНК (или РНК в случае некоторых вирусов) должна быть очень сильно сконденсирована, т.е. плотно упакована, чтобы уместиться в отведенном ей пространстве. Конденсация ДНК зависит от наличия белков, с которыми она связывается. Обычно это основные белки, положительные заряды которых нейтрализуют отрицательные заряды нуклеиновой кислоты. Интересно было бы узнать, насколько специфична эта упаковка: скручивается ли ДНК по какому-то *одному определенному* образцу или каждая отдельная копия укладывается по-своему?

Из примеров, суммированных в табл. 28.1, видно, насколько велико несоответствие между длиной нуклеиновой кислоты и размером ее вместилища. Бактериофаги и вирусы эукариот независимо от того, являются ли они длинными и тонкими (нитевидными) или имеют почти сферическую (икосаэдрическую) форму, содержат большое количество нуклеиновой кислоты (одноцепочечной или двухцепочечной ДНК или РНК), которая заполняет всю вмещающую ее оболочку.

В случае соответствующих компартментов у бактерий и эукариотических клеток это расхождение еще больше. ДНК содержится только в какой-то ограниченной области, а не заполняет всего компартмента. В бактериальной клетке она организована в **нуклеоид**, а в интерфазном ядре эукариот представлена массой **хроматина**. Среди примеров, приведенных в таблице, **кульминации достигает необходимость** разместить ДНК человека длиной 1,8 м в ядре, диаметр которого может составлять всего лишь 6 мкм (6×10^{-6} м). При делении эукариотической клетки генетический материал упаковывается еще более плотно и образует индивидуально различимые митотические **хромосомы**.

Степень конденсации ДНК может быть выражена как **плотность упаковки**, т.е. отношением длины ДНК к длине, содержащей ее хромосомы. Например, самая маленькая хромосома человека содержит ДНК, состоящую из $4,6 \cdot 10^7$ п.н. (примерно в 10 раз больше, чем размер генома *E. coli*). Длина такой ДНК составляет 14 000 мкм (1,4 см), а соответствующая хромосома, измеренная в сконденсированном состоянии в метафазе, имеет в длину всего лишь около 2 мкм. Таким образом, плотность упаковки этой хромосомы в метафазе равна примерно 7000.

Нельзя столь же уверенно установить плотность упаковки для бактериального нуклеоида или эукариотического хроматина, поскольку они имеют более аморфную структуру. Однако обычно считается, что митотические хромосомы упакованы примерно в 5–10 раз более плотно, чем интерфазный хроматин, плотность упаковки которого равна, следовательно, 1000–2000.

В организации всего генетического материала прослеживается общий принцип: генетический материал существует в виде компактной массы в строго ограниченных областях, и характерные для него процессы, такие, как репликация и транскрипция, должны происходить в пределах этих областей.

Упаковка вирусных геномов в оболочку

Характер упаковки *индивидуальной* последовательности ДНК существенно различается в зависимости от того, идет ли речь о геноме клетки или о вирусе. Размер клеточного генома не имеет строгих ограничений; число и расположение отдельных последовательностей в нем может изменяться в результате дупликации, делеции или перестройки. Таким образом, в этом случае требуется *общий* метод упаковки ДНК, нечувствительный к общему содержанию и распределению последовательностей. У ви-

Таблица 28.1

Длина нуклеиновых кислот намного больше, чем размеры окружающего компартмента

Пример	Форма	Размер, мкм	Нуклеиновая кислота	Длина
Вирус табачной мозаики	Нитевидная	$0,008 \times 0,3$	1 одноцепочечная РНК	2 мкм (6,4 т. п. н.)
Фаг fd	»	$0,006 \times 0,85$	1 одноцепочечная ДНК	2 мкм (6 т. п. н.)
Капсид аденовируса	Икосаэдрическая	Диаметр 0,07	1 двухцепочечная ДНК	11 мкм (35 т. п. н.)
Капсид фага T4	То же	$0,065 \times 0,10$	То же	55 мкм (170 т. п. н.)
Бактерия <i>E. coli</i>	Цилиндрическая	$1,7 \times 0,65$	»	1300 мкм (13 мм = $4,2 \cdot 10^3$ т. п. н.)
Клетки человека (HeLa), митохондрия	Сплюснутая сфероидальная	$3,0 \times 0,5$	~ 10 идентичных молекул двухцепочечной ДНК	50 мкм (5 мкм = 16 т. п. н. каждая)
Клетки HeLa, ядра	Сфероидальная	Диаметр 6	46 хромосом (двухцепочечная ДНК)	1,8 ($6 \cdot 10^6$ т. п. н.)

русов, напротив, есть два ограничения, обусловленные их природой. Во-первых, количество нуклеиновой кислоты, которую нужно упаковать, *предопределено* размером генома. И во-вторых, эта нуклеиновая кислота должна уместиться внутри оболочки, собранной из белка или белков, кодируемых вирусными генами.

При поверхностном рассмотрении вирусной частицы создается обманчивое представление, что она устроена очень просто. Геном, представленный нуклеиновой кислотой, находится внутри **капсида**. Это симметричная, или квазисимметричная, структура, построенная из одного или нескольких видов белков. К капсиду могут быть прикреплены или включены в него другие структуры, собранные из различных белков и необходимые для заражения клетки-хозяина.

С этой точки зрения вирусная частица сконструирована без больших допусков. Внутренний объем вирусной частицы в редких случаях намного превышает объем, необходимый для размещения нуклеиновой кислоты. Согласно простым расчетам, внутренний объем вириона превышает объем нуклеиновой кислоты обычно менее чем в два раза, а часто он лишь немного больше размеров нуклеиновой кислоты.

Ограничение, связанное с использованием для построения капсида исключительно белков, кодируемых вирусом, в крайней форме сводится к построению оболочки из белковых субъединиц только одного типа. Правила, по которым собираются плотные структуры из идентичных субъединиц, ограничивают форму образованного капсида всего двумя возможностями. Субъединицы белка могут примыкать друг к другу, располагаясь по спирали. При этом образуется **нитевидная**, или палочковидная, форма. Кроме того, возможно образование псевдосферической оболочки. Этот тип структуры в действительности является полиэдромом с **икосаэдрической симметрией**. Сейчас уже известно, что некоторые вирусные капсиды собраны из субъединиц более чем одного типа. Хотя это увеличивает возможности образования различных вариантов структур, все вирусные капсиды принадлежат к общим классам квазикристаллических нитей или икосаэдров.

Построение капсида, содержащего нуклеиновую кислоту, может осуществляться двумя способами. В первом случае белковая оболочка может собираться вокруг нуклеиновой кислоты; при этом конденсация ДНК или РНК будет происходить в результате взаимодействий белок—нуклеиновая кислота в процессе сборки. Во втором случае капсид может быть построен из его компонента (компонентов) в виде пустой оболочки и потом заполнен нуклеиновой кислотой, которая конденсируется по мере проникновения в капсид.

Сборка капсида вокруг генома происходит у нитевидных вирусов и фагов, содержащих одноцепочечную РНК. Наиболее хорошо охарактеризованным примером может служить вирус табачной мозаики (ВТМ). Его сборка начинается с двухцепочечной шпильки, находящейся в последовательности РНК. Из этого **центра нуклеации** процесс продолжается в двух направлениях вдоль РНК—до тех пор, пока не достигает концов. Одна структурная единица капсида представляет собой двухслойный диск: каждый слой содержит 17 идентичных субъединиц белка. Диск образует кольцевую структуру, которая при взаимодействии с РНК приобретает спиральную форму. Как видно на **рис. 28.1**, внутри белковой оболочки РНК скручена в виде спирали.

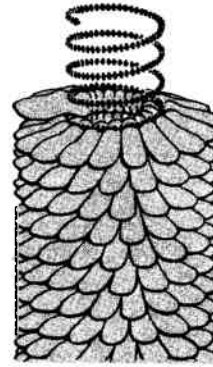


Рис. 28.1. Спиральное расположение РНК ВТМ обусловлено межплоскостными взаимодействиями белковых субъединиц вириона.

Характер структуры, в виде которой существует РНК-геном, определяется взаимодействием РНК с белковой оболочкой. Аналогичный в общих чертах тип организации генома характерен также для РНК-генома вируса желтой мозаики турнепса, имеющего капсид сферической формы. Здесь так же положение РНК внутри капсида прямо определяется ее связыванием с белками оболочки.

Последовательность событий при сборке сферических капсидов ДНК-содержащих вирусов лучше всего прослежена у фагов лямбда (λ) и T4. В обоих случаях пустая головка оболочки собирается путем взаимодействия небольшого набора белков. Затем двухцепочечный геном внедряется в головку, что достигается изменением структуры капсида.

На **рис. 28.2** схематически изображена сборка фага лямбда. Она начинается с маленькой оболочки головки, содержащей белковый слой. Затем эта структура превращается в пустую оболочку головки более отчетливой формы. Потом начинается упаковка ДНК, в процессе которой оболочка головки увеличивается в размере, сохраняя прежнюю форму. И наконец, заполненная головка соединяется с хвостовым отростком.

ДНК фагов лямбда и T4, заполняющая пустую головку, находится в виде **конкатемерных** молекул. Это множественные геномы, соединенные конец в конец, как показано на **рис. 28.3**. У каждого фага имеется свой собственный механизм, узнающий, какое именно количество ДНК должно быть включено в частицу.

Концы генома фага лямбда маркированы последовательностями, получившими название *cos*-сайтов. По левому (согласно обычной карте фага) *cos*-сайту происходит расщепление и образуется свободный конец, который внедряется в капсид. Внедрение ДНК продолжается вплоть до второго *cos*-сайта, который также отщепляется, образуя второй конец молекулы-генома. Тот конец, который при сборке входит в капсид последним, первым проникает в новую клетку-хозяина.

Взаимоотношения между ДНК и оболочкой головки исследовали прямо, определяя, какие именно участки ДНК химически пришиваются к белкам капсида. Ответ получился удивительный: оказалось, что все области ДНК в более или менее равной степени способны к образованию перекрестных швов. Вероятно, это означает, что укладка ДНК в головке происходит согласно общему правилу конденсации, но характер укладки не зависит от конкретных нуклеотидных последовательностей.

По-видимому, конденсация ДНК происходит по меха-

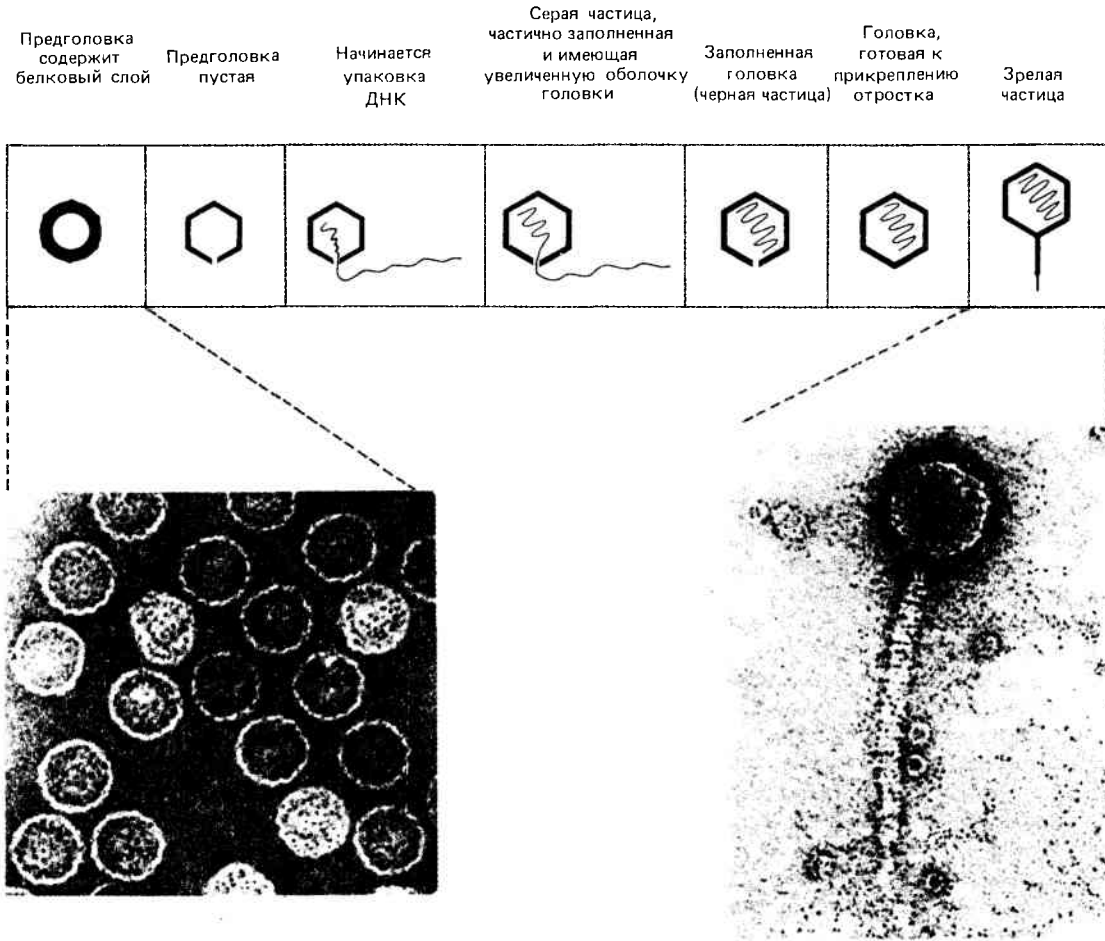


Рис. 28.2. Пустая головка фага лямбда при заполнении фаговой ДНК изменяет свою форму и увеличивается.

На электронной микрофотографии показаны частицы в начале и в конце цикла созревания. (Фотография любезно предоставлена А. Howatson.)

низму «наматывания на катушку», при котором она закручивается вокруг центральной оси (как нитка на катушке). Так как любой участок ДНК может находиться в контакте с белковой оболочкой, наматывание не может происходить таким способом, при котором сначала заполняется наружная часть оболочки, а потом внутренняя или наоборот. Была предложена модель, согласно которой конец ДНК, проникший в головку первым, фиксируется в точке, расположенной вблизи от места ее проникновения. ДНК продолжает входить вдоль центральной оси, закручиваясь вокруг «верхушки» головки. Витки на катушке располагаются равномерно, так что по мере того, как увеличивается количество вошедшей в головку ДНК, тем более тесным становится расположение витков. Реакция ДНК на уплотнение неодинакова в разных капсидах, так что точки контакта с белком оболочки различны у индивидуальных вирионов.

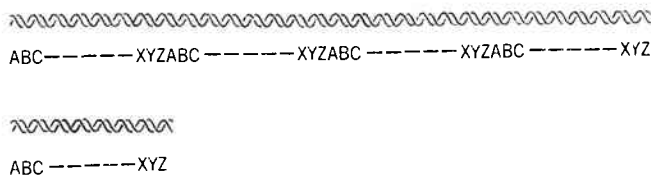


Рис. 28.3. Конкатемеры ДНК состоят из ряда фаговых геномов, расположенных тандемом. Верхняя полоска – конкатемер, нижняя полоска – индивидуальный геном.

Любая ДНК, находящаяся между двумя *cos*-сайтами, может быть упакована в оболочку фага. На этом основан метод клонирования с использованием «космид», описанный в гл. 19. Важно, однако, чтобы расстояние между двумя *cos*-сайтами не сильно отличалось от естественной длины фага лямбда. Если это расстояние будет слишком большим или слишком малым, то ДНК не будет упакована в оболочку. Этот факт свидетельствует о том, что для завершения процесса упаковки количество ДНК должно быть *достаточным* и что в головке фага может поместиться совсем немного избыточной ДНК (примерно 15%).

У фага Т4 упаковка ДНК начинается со *случайной* точки конкатемерного предшественника и продолжается до тех пор, пока головка не заполнится количеством ДНК, равным целому геному. Следовательно, можно предположить существование некоторого механизма, измеряющего длину включенной ДНК. На деле количество включенной ДНК несколько превышает размер одной геномной единицы; в результате создается концевая избыточность, соответствующая дополнительной длине. Согласно рис. 28.3, первый вирион может содержать ДНК от А до А, второй – от В до В и т. д. Таким образом, в каждом геноме есть (разные) буквы, повторяющиеся на обоих концах.

Итак, двухцепочечная ДНК, если рассматривать ее небольшие участки, представляет собой достаточно жест-

кую палочковидную структуру, которую нужно скрутить еще более компактно, для того чтобы она уместилась в капсиде. Каким образом это происходит, пока еще не ясно, но известно, что наряду с ДНК в капсиде существуют «внутренние белки». Возможно, они служат «остовом», на котором конденсируется ДНК. (Это было бы аналогично использованию белков оболочки вирусами растений, о котором мы упоминали ранее.) Какова специфичность этой упаковки? Во всяком случае она не зависит от определенной последовательности ДНК, поскольку делеции, вставки и замены не нарушают процесса сборки.

Все разнообразные способы сборки вирусов завершаются одинаково: упаковкой в капсид единичной молекулы ДНК или РНК. Однако существуют некоторые вирусы, геном которых состоит из многих молекул нуклеиновой кислоты. Реовирусы, например, содержат десять двухцепочечных фрагментов РНК, и все они должны быть упакованы в капсид. О том, каким образом в процессе сборки отбирается по одной копии каждой из разных молекул, для того чтобы собрать полный комплект генетической информации, ничего не известно.

Некоторые вирусы растений состоят из нескольких частей: их геномы разбиты на фрагменты, каждый из которых упаковывается в *отдельный* капсид. Примером может служить вирус мозаики люцерны, геном которого состоит из четырех различных одноцепочечных РНК. Каждая цепь упаковывается независимо в оболочку, состоящую из субъединиц одного и того же белка. Инфекционный процесс протекает успешно в том случае, когда в клетку попадает по одной из копий каждой цепи.

Четыре компонента вируса существуют в виде частиц различного размера. Это означает, что один и тот же капсидный белок может упаковать каждую РНК в характерную именно для нее частицу. Такой способ упаковки отличается от способа упаковки отрезков нуклеиновой кислоты одного размера в капсиды определенной формы. Однако и у вирусов, имеющих только одну правильную форму капсида, в процессе сборки могут возникать измененные формы капсида. Речь идет о частицах-монстрах, головка которых длиннее обычной. Таким образом, белку (или белкам) капсида присуща способность собираться в структуру определенного типа, но точный размер и форма этих структур не инвариантны. Во многих случаях существуют также **белки сборки**, которые не входят в состав оболочки головки, но способствуют правильной сборке частицы. Клеточные геномы также используют белки, функция которых — направлять сборку других белков (см. гл. 29).

Бактериальный геном свернут в нуклеоид

Хотя у бактерий не обнаружено структур, обладающих морфологическими признаками эукариотических хромосом, их геномы тем не менее организованы в виде определенных телец (нуклеоидов). Генетический материал имеет вид достаточно компактной группы телец или нескольких таких групп, занимающих (весьма приблизительно) около трети объема клетки. На **рис. 28.4** показан тонкий срез бактерии, на котором прекрасно виден этот **нуклеоид**.

У бактерий с частично реплицированной ДНК содержание ДНК в нуклеоиде превышает размер одного гено-



Рис. 28.4. На тонком срезе показан бактериальный нуклеоид в виде компактной массы в центре клетки. (Фотография любезно предоставлена J. Griffith.)

ма. Ко времени деления клетки генетический материал разделяется на два нуклеоида, которые распределяются между дочерними клетками. Механизм сегрегации, по-видимому, включает в себя этап прикрепления бактериального генома к мембране. Этот механизм аналогичен митотической сегрегации эукариотических хромосом. Однако вместо того, чтобы растаскиваться с помощью нитей веретена, специфические локусы каждого из дочерних геномов прикрепляются к определенному участку на мембране. По мере роста клетки эти два участка мембраны удаляются друг от друга, и после деления каждый оказывается в разном компартменте, захватив с собой свой нуклеоид (гл. 31).

Нуклеоиды из клеток *E. coli* были выделены непосредственно в виде очень быстро седиментирующего комплекса, который на 80% своей массы состоял из ДНК. (В аналогичных комплексах в клетках эукариот ДНК составляет около 50% всей массы; см. гл. 29.) У бактерий, содержащих более одной копии генома (в результате репликации ДНК), комплекс пропорционально увеличивается в размере. Этот комплекс можно перевести в менее компактное состояние, обработав его реагентами, действующими на РНК или белок. Из этого следует, что именно указанные компоненты способствуют стабилизации нуклеоида. Возможная роль белков очевидна. Роль, которую может играть РНК, неясна. Попытки выделить специфическую РНК, участвующую в образовании структуры, оказались безуспешными. Попытки изучения новосинтезированной РНК (т.е. незавершенных продуктов транскрипции) также не дали результата.

При лизисе клеток *E. coli* высвобождаются нити в виде петель, прикрепленных к остаткам клеточной оболочки. Как видно на **рис. 28.5**, ДНК этих петель находится не в виде свободно вытянутого дуплекса, а скручена в более компактные образования благодаря ее связи с белками. Быстро седиментирующий комплекс может состоять из таких петель, находящихся в компактном виде.

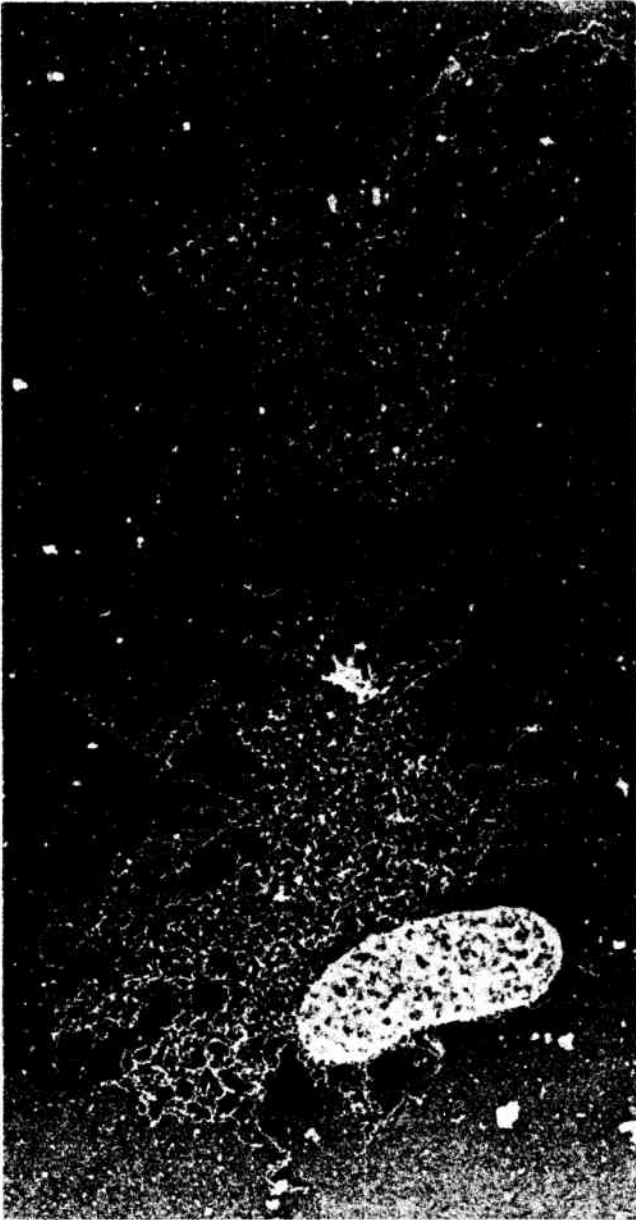


Рис. 28.5. Нуклеоид высвобождается из лизированной клетки *E. coli* в виде нитевидных петель.

(Фотография любезно предоставлена J. Griffith.)

Из *E. coli* было выделено несколько ДНК-связывающих белков, которые при поверхностном взгляде напоминают хромосомные белки эукариот. Свойства этих белков суммированы в табл. 28.2.

Два из них, HU и H находятся в клетке в достаточно большом количестве и по аминокислотному составу напоминают гистоны, связанные с эукариотической ДНК (гл. 29). Белки эти имеют небольшой размер, обладают сильноосновными свойствами и прочнее связываются с ДНК. Физиологическая роль этого связывания пока неясна.

Белок, потенциально связывающийся с ДНК, был обнаружен благодаря тому, что его ген расположен в опероне, содержащем тРНК (гл. 23). Это белок P. По своему аминокислотному составу он напоминает протамины, которые связываются с ДНК в некоторых сперматозоидах. Его количество и функции в клетке неизвестны.

Для одного из белков точно установлена принадлежность к генетическому аппарату. Это белок HLP1, кодируемый геном *firA*. Мутации в этом гене супрессируют устойчивость к рифампицину, обусловленную некоторыми мутациями в гене β -субъединицы РНК-полимеразы. Свойства этих мутаций не позволяют судить о том, взаимодействует ли этот белок с ДНК или с РНК-полимеразой. HLP1—это маленький основной белок, который, подобно другим белкам, может связываться с ДНК.

Нуклеоид содержит много суперспирализованных петель

Изолированная из нуклеоида ДНК, судя по ее реакции с бромистым этидием, ведет себя как замкнутая двухцепочечная структура. Маленькая молекула этого реагента встраивается (интеркалирует) между парами оснований ДНК и вызывает образование положительных супервитков в замкнутых кольцевых молекулах ДНК, у которых две цепи соединены ковалентно. В открытых кольцевых молекулах, содержащих одноцепочечный разрыв, или в линейных молекулах ДНК может свободно вращаться и таким образом освобождаться от дополнительного напряжения, создаваемого в результате включения бромистого этидия.

В природных кольцевых отрицательно суперспирализованных ДНК интеркаляция бромистого этидия сначала удаляет отрицательные супервитки и затем вводит положительные супервитки. По количеству бромистого этидия, необходимого для нулевой суперспирализации, измеряют исходную плотность отрицательных супервитков.

В процессе выделения компактного нуклеоида вводится некоторое количество одноцепочечных разрывов; такие же разрывы можно ввести в результате ограниченного гидролиза ДНКазой. Однако при этом не исчезает способность ДНК образовывать положительные супервитки в ответ на интеркаляцию бромистого этидия. Подобная реакция генома, содержащего одноцепочечные разрывы, в ответ на включение бромистого этидия говорит о том, что в нем должны существовать независимые

Таблица 28.2

У. E. coli существует несколько ДНК-связывающих белков

Белок	Строение	Содержание на 1 клетку	Родственный белок у эукариот	Локус
HU	2 неидентичные субъединицы по 9000 дальтон	40 000 димеров	Гистон H2B	Неизвестен
H	2 идентичные субъединицы по 28 000 дальтон	30 000 димеров	Гистон H2A	»
HLP1	Мономер, 17 000 дальтон	20 000 копий	Неизвестен	
P	Субъединица, 3000 дальтон	Неизвестно	Протамины	Неизвестен

домены. Суперспирализация в каждом домене не зависит от событий, происходящих в других доменах.

Для объяснения такой автономности предложена модель организации бактериальной хромосомы, схематически изображенная на **рис. 28.6**. Согласно этой модели, каждый домен состоит из одной петли ДНК, концы которой закреплены каким-то (неизвестным) образом так, что вращательное движение не передается из одного домена в другой. Существует около 100 таких доменов на 1 геном, так что каждый состоит примерно из 40 т.п.н. (13 мкм). При этом ДНК образует более компактную нитевидную структуру, которая еще детально не описана.

Благодаря существованию отдельных доменов в различных областях генома может сохраняться разный уровень суперспирализации. Этим фактором также можно объяснять различную чувствительность определенных бактериальных промоторов к суперскручиванию (гл. 11).

Суперспирализация ДНК может иметь два последствия. Если ДНК остается свободной, ее движения не сдерживаются и отрицательные супервитки вызывают напряжение скручивания, которое может быть снято раскручиванием двойной спирали, как это описано в гл. 2. ДНК может находиться в динамическом равновесии между состояниями напряжения и раскручивания (см. гл. 32). Однако суперспирализация может сдерживаться, если белки связываются с ДНК и поддерживают ее в определенной трехмерной конфигурации. В этом случае супервитки будут представлены по ходу ДНК, связанной с белками. Энергия взаимодействия между белками и суперскрученной ДНК влияет на стабильность двойной спирали. Например, если отрицательно суперспирализованная ДНК связывается с белком, относительно специфичным к одноцепочечной ДНК, в области связывания может перманентно происходить локальная денатурация ДНК. С другой стороны, связывание гистоновых белков с образованием нуклеосом стабилизирует двойную спираль отрицательно суперспирализованной ДНК (гл. 29).

Снимается ли большинство образовавшихся супервитков в ДНК *E. coli* *in vivo*, или двойная спираль испытывает напряжение скручивания, характерное для свободной от белков ДНК? Измерить степень суперспирализации *in vitro* трудно из-за того, что удерживающие белки могут теряться в процессе выделения. Однако ситуацию *in vivo* можно исследовать, используя реагент псорален, вызывающий перекрестные сшивки и более охотно связывающийся с ДНК в состоянии напряжения скручивания. По связыванию псоралена с ДНК *E. coli* *in vivo* средняя плотность суперспирализации соответствует одному отрицательному супервитку на каждые 200 п.н. Этот результат показывает, что супервитки действительно создают напряжение скручивания *in vivo*. Большинство естественных супервитков в ДНК *E. coli* может находиться в свободном состоянии.

Наиболее важные черты структуры компактного нуклеоида еще предстоит изучить. Какова специфичность образования доменов – всегда ли одни и те же последовательности лежат в тех же самых местах, или содержание отдельных доменов может изменяться? Как сохраняется целостность домена? Какую структурную роль играют ДНК-связывающие белки? Биохимический анализ сам по себе, вероятно, не может полностью ответить на эти вопросы. Однако молекулярный анализ структуры нуклеоида (при использовании соответствующих методов отбора) можно было бы выполнять с помощью структурных мутантов.

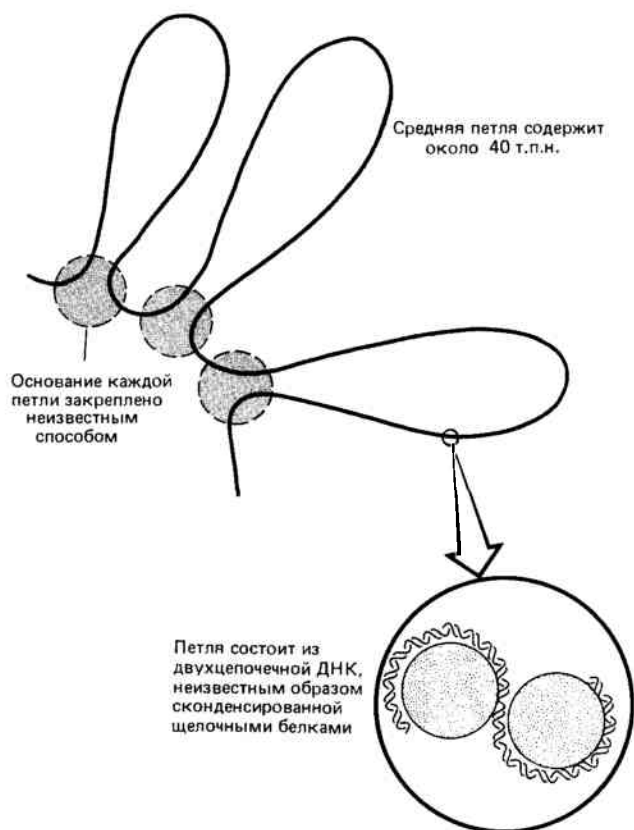


Рис. 28.6. Бактериальный геном состоит из большого числа петель двухцепочечной ДНК (уложенных в виде волокон). Каждая петля закреплена в основании и образует независимый структурный домен.

Различия между интерфазным хроматином и митотическими хромосомами

Индивидуальные хромосомы различимы только в течение короткого промежутка времени в процессе деления клетки. Лишь в это время каждая из них видна в виде компактного образования с плотностью упаковки 10 000. На **рис. 28.7** приведена электронная микрофотография пары сестринских хроматид, снятая в метафазе. (Сестринские хроматиды – это дочерние хромосомы, образованные во время предшествующей репликации, которые на этой стадии митоза еще соединены вместе; см. гл. 1.) Каждая хроматида состоит из узловатой нити диаметром около 30 нм, уложенной в хромосомную структуру.

В течение большей части жизненного цикла эукариотической клетки генетический материал находится в ядре в таком состоянии, что индивидуальные хромосомы нельзя различить. Структура интерфазного хроматина в период между делениями не претерпевает видимых изменений. Например, обычно не видно каких-либо перемен в период репликации, когда количество хроматина удваивается. Хроматин имеет нитевидную структуру, однако общую конфигурацию нити в пространстве трудно проследить в деталях. Сами по себе нити хроматина сходны, а может быть, и идентичны тем, которые наблюдаются в митотических хромосомах.

Различают два типа хроматина. Как это видно на **рис. 28.8**, в большей части нити упакованы менее плотно, чем в хромосоме. Этот материал называют эухромати-

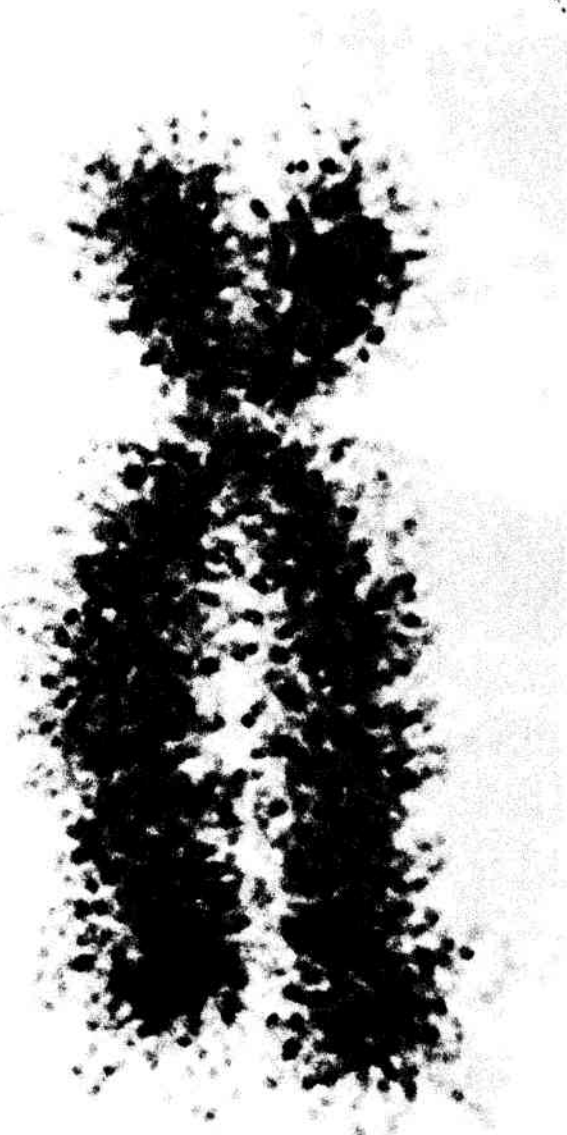


Рис. 28.7. Сестринские хроматиды митотической пары, каждая из которых состоит из нити (~30 нм в диаметре), компактно свернутой в хромосому.
(Фотография любезно предоставлена E. DuPraw.)

ном. Общая плотность его упаковки в интерфазе составляет, вероятно, около 1000–2000, а во время митоза он конденсируется еще в 5–10 раз.

В некоторых участках хроматина нити упакованы очень плотно и находятся в состоянии, напоминающем митотические хромосомы. Этот материал называют **гетерохроматином**. Степень конденсации таких участков почти не изменяется на протяжении всего клеточного цикла. Различные гетерохроматиновые участки часто агрегируют и образуют темноокрашенный **хромоцентр**.

Одни и те же фибриллы, не прерываясь, проходят через эу- и гетерохроматин; из этого следует, что эти состояния представляют собой различную степень конденсации генетического материала. Аналогично этому различна степень конденсации и эухроматиновых участков в интерфазе и во время митоза. Из этих наблюдений вытекает важный вывод о том, что генетический материал упакован таким образом, что в хроматине могут существовать рядом друг с другом оба альтернативных со-

стояния и возможны периодические изменения в упаковке эухроматина при переходе от интерфазы к делению.

Существует корреляция между структурой генетического материала и его транскрипционной активностью. Митотические хромосомы транскрипционно инертны; в период клеточного деления транскрипция прекращается. В гетерохроматине также не наблюдается транскрипционной активности. Существуют два класса гетерохроматина, каждый из которых содержит последовательно разного типа, но оба они в равной мере лишены транскрипционной активности.

Конститутивный гетерохроматин состоит из определенных областей, которые *никогда* не экспрессируются. Сюда относятся сателлитные последовательности ДНК. В виде **факультативного гетерохроматина** могут находиться целые хромосомы, которые оказываются неактивными в ряду клеточных поколений, хотя при иных обстоятельствах они способны к экспрессии. Отличным примером может служить X-хромосома млекопитающих, одна из копий которой (случайная) полностью инактивирована у каждой самки. (Этот механизм компенсирует наличие двух X-хромосом у самки по сравнению с одной у самцов.) Неактивная X-хромосома сохраняется в гетерохроматическом состоянии; активная X-хромосома является частью эухроматина. В данном случае можно видеть корреляцию между транскрипционной активностью и структурной организацией, поскольку в альтернативных состояниях находятся *идентичные* последовательности ДНК.

Отсутствие активности связано (и, вероятно, обусловлено) с чрезвычайно высокой степенью конденсации генетического материала. Заметим, однако, что обратное утверждение было бы неверно. Хотя активные гены действительно находятся в неконденсированном эухроматине, только незначительная часть входящих в него последовательностей когда-либо транскрибируется. Таким образом, необходимым условием экспрессии гена является его локализация в эухроматине, но одного этого **недостаточно**. Мы можем только гадать, используется ли тот же самый механизм, который вызывает резкие видимые различия между эу- и гетерохроматином, но только в меньшей степени для регуляции транскрипции в самом эухроматине. Иными словами, скрыты ли в структуре эухроматина менее выраженные участки с разной степенью конденсации, которые обуславливают различие между транскрибируемыми и нетранскрибируемыми районами.

В интерфазном хроматине и митотических хромосомах, выделенных *in vitro*, наблюдаются большие петли, состоящие из непрерывной нити, которая, очевидно, закреплена у основания и образует независимые домены, подобные тем, которые обнаруживают в бактериальном нуклеотиде. Эти структуры обсуждаются в гл. 29.

Из-за диффузного состояния хроматина мы в настоящее время не можем узнать детали его организации. Но мы можем поставить вопрос: насколько жестко организована структура хромосомы? Всегда ли определенные последовательности располагаются в одном и том же участке хромосомы, или скручивание нити в общую структуру является достаточно случайным? Несколько лет назад в результате разработки метода дифференциального окрашивания на **G-сегменты** было показано, что для каждой хромосомы характерна особая воспроизводимая при редупликации ультраструктура. Если хромосомы обработать специальным образом и затем окра-

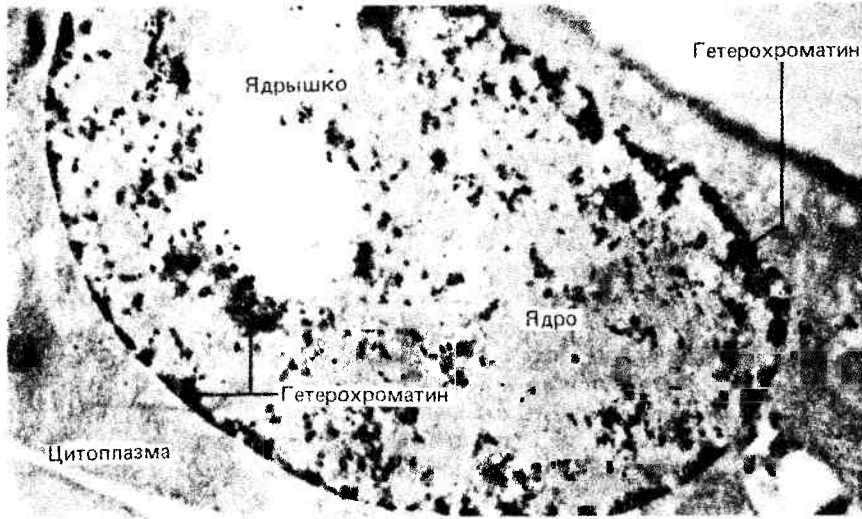


Рис. 28.8. Тонкий срез ядра окрашен по Фельгену. Виден гетерохроматин в виде компактных участков, сгруппированных около ядрышка и ядерной мембраны. (Фотография любезно предоставлена E. Puvion.)

сить краской Гимза, на них образуется серия специфических полос (G-полосы; G-сегменты). В качестве примера на рис. 28.9 показан хромосомный набор человека.

До того как появился этот метод окрашивания, хромосомы различали только по размеру и по относительной локализации центromеры (см. далее). Теперь же каждую хромосому можно идентифицировать по характерному расположению полос. Картина расположения полос воспроизводится настолько постоянно, что можно различить участок, транслоцированный с одной хромосомы на другую, сравнив его по структуре с исходным диплоидным набором хромосом.

Метод G-сегментации находит самое широкое практическое использование, однако природа дифференциального окрашивания до сих пор остается загадкой. Известно наверняка лишь то, что без предварительной обработки хромосомы окрашиваются краской Гимза более или менее равномерно. Таким образом, образование полос обусловлено тем, что различные виды предварительной обработки по-разному влияют на отдельные участки хромосомы. (Вероятно, в результате экстракции каких-то компонентов, связывающих краску, из участков, образующих промежутки между G-сегментами.) Однако разнообразие эффективных воздействий так велико, что не позволяет понять общей природы реакции. Можно только предполагать существование какой-то большой по длине структуры, но, что лежит в ее основе, не известно.

Наиболее простая модель хромосомы представляет собой одну протяженную двухцепочечную молекулу ДНК. Эта модель согласуется с фактом существования нити размером в 30 нм, которая обнаруживается на протяжении всей хромосомы. В ранних экспериментах было показано, что такая нить состоит из одной конденсированной двухцепочечной ДНК. Этот вывод нашел детальное подтверждение в последующих экспериментах, которые обсуждаются в гл. 29. Хромосома, так же как и индивидуальная молекула ДНК, реплицируется полуконсервативным путем (см. рис. 2.10 и гл. 31). Такой способ репликаций свидетельствует против модели хромосомы, состоящей из многих независимых молекул ДНК. Таким образом, можно объяснить структуру интерфазного хроматина и митотических хромосом, исходя из представлений об очень длинной молекуле ДНК. Эта молеку-



Рис. 28.9. При использовании метода C-окрашивания на каждом из членов хромосомного набора образуется характерная серия поперечных полос.

(Это изображение хромосомного набора человека предоставлено N. Davidson.)

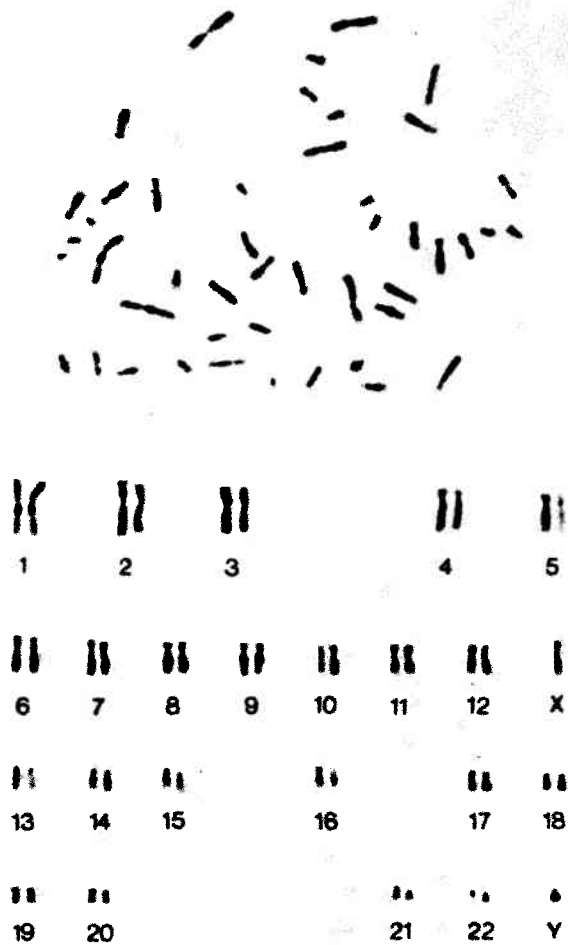


Рис. 28.10. При использовании метода С-окрашивания интенсивно окрашиваются центромеры всех хромосом.
(Фотография любезно предоставлена N. Davidson.)

ла уложена в форму, в которой она может периодически менять степень конденсации, что способствует осуществлению транскрипции и репликации.

Эукариотическая хромосома как единица сегрегации

В митозе сестринские хроматиды перемещаются к противоположным полюсам клетки (как показано на рис. 1.5). Это передвижение возможно благодаря прикреплению хромосомы к микротрубочкам, которые соединены с центриолями или закреплены рядом с ними у полюсов клетки. Место прикрепления микротрубочек к хромосоме ограничено определенным участком, который называют **центромерой**. Локализация центромеры хорошо видна на фотографии (рис. 28.7), где представлены сестринские хроматиды на стадии метафазы митоза. Хроматиды еще соединены друг с другом в области центромеры.

Исторически термин «центромера» использовали в структурном и в функциональном смысле для описания участка хромосомы, ответственного за ее передвижение. В митозе центромера тянется по направлению к полюсу

и тащит за собой прикрепленную к ней хромосому. С этой точки зрения хромосому можно рассматривать как специальное приспособление, используемое для прикрепления большого числа генов к аппарату клеточного деления.

Как показывают результаты разрывов хромосом, центромера служит для расхождения хромосомы по дочерним клеткам. В результате единичного разрыва образуется один фрагмент с центромерой и другой **ацентрический фрагмент**, т.е. лишенный центромеры. Ацентрический фрагмент не прикреплен к митотическому веретену и может не попасть ни в одно из ядер дочерних клеток. (Заметим, что каждая хромосома может иметь *только* одну центромеру.) Если в результате транслокации получатся хромосомы с более чем одной центромерой, в митозе образуются aberrантные структуры, поскольку две центромеры могут тащить *одну и ту же* сестринскую хроматиду к разным полюсам, что приведет к разрыву хромосомы.

Области, находящиеся по обеим сторонам от центромеры, часто богаты сателлитными последовательностями ДНК и могут содержать значительное количество конститутивного гетерохроматина. В отличие от интерфазного хроматина гетерохроматин не сразу заметен в митотических хромосомах, но его можно увидеть, используя метод С-окрашивания. На **рис. 28.10** видны темноокрашенные области в районе всех центромер. Это свойство (так же, как и характерные черты интерфазного гетерохроматина) связано не с какими-то особенностями сателлитной ДНК, а зависит от белков, специфически присутствующих в этом месте. Конститутивный гетерохроматин, вероятно, не связан непосредственно с механизмом деления, так как он не всегда обнаруживается рядом со *всеми* центромерами.

Какое свойство центромеры прямо связано с механизмом расхождения хромосом? Внутри центромеры можно увидеть темноокрашивающееся волокнистое образование с диаметром или длиной около 400 нм. Это вещество называют **кинетохором**. Кинетохор — это различимая структура, которая, по-видимому, непосредственно прикрепляется к микротрубочке. Обычно считают, что некая специфическая последовательность ДНК определяет место формирования кинетохора на хромосоме. У хромосом разных организмов структуры кинетохоров сильно варьируют, что затрудняет анализ их функции. Однако было показано, что кинетохоры содержат ДНК; поэтому можно предполагать, что кинетохор образуется *непосредственно* на соответствующей последовательности хромосомы.

Если за расхождение хромосом отвечает последовательность ДНК в центромере, то любая молекула ДНК, имеющая эту последовательность, будет надлежащим образом двигаться в процессе клеточного деления, тогда как ДНК без этой последовательности утратит такую способность. Это предсказание было использовано для выделения центромерной ДНК у дрожжей *S. cerevisiae*. У хромосом дрожжей нет заметно выраженных кинетохоров, как у хромосом высших эукариот. В то же время митотическое деление и расщепление в мейозе происходят с помощью аналогичных механизмов.

Недавно методами генной инженерии были получены плазмиды, которые реплицируются подобно хромосомам (см. гл. 31). Однако из-за неравномерного распределения они нестабильны в митозе и мейозе и исчезают из большинства клеток. Выделены фрагменты хромосомной

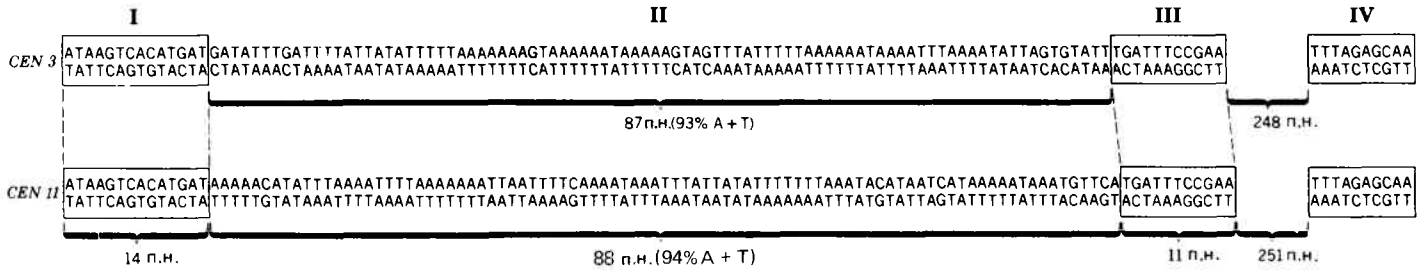


Рис. 28.11. При сравнении нуклеотидной последовательности локусов *CEN3* и *CEN11* дрожжей найдены четыре консервативных участка.

ДНК, которые придавали этим плазмидам митотическую стабильность. В одном из случаев в таком фрагменте оказался генетический маркер, расположенный рядом с центромой, что подтверждало происхождение фрагмента. Несколько *CEN*-фрагментов были охарактеризованы. У них отсутствуют протяженные области гомологии, но тем не менее им присущи общие свойства.

Элемент I состоит из последовательности в 14 п. н., которая с некоторыми вариациями встречается во всех центромерах. Элемент II представляет собой АТ-богатую (> 90%) последовательность из 82–89 п. н., которая обнаружена во всех центромерах. Вполне возможно, что ее функция зависит от размера, а не от точной последовательности. По своему составу этот фрагмент напоминает некоторые сателлитные ДНК высших эукариот. Элемент III состоит из 11 п. н. и почти полностью гомологичен во всех известных центромерах. Ни в одной из центромерных последовательностей нет открытых рамок считывания.

Для того, чтобы прямо определить, какие признаки необходимы для выполнения функции центромер, можно индуцировать делеции и другие изменения в этой области. Однако трудно исследовать функции центромер непосредственно в плазмидах. Хотя плазмиды с *CEN*-последовательностями стабильнее, чем плазмиды без них, они все-таки гораздо менее стабильны по сравнению с аутентичными центромерами. Отсутствие стабильности, по-видимому, связано с малым размером плазмиды, а не с каким-то повреждением в *CEN*-области.

Другим способом использования доступных центромерных последовательностей может быть их модификация *in vitro* с последующим введением в дрожжевые клетки, в которых они могут заменить соответствующие центромеры хромосом. Если хромосомная центромера заменена *CEN*-фрагментом, из которого удалены консервативные последовательности, то соответствующая хромосома ведет себя в митозе, как ацентрический фрагмент: она не может нормально сегрегировать. Это прямо показывает, что *CEN*-фрагмент действительно несет информацию, необходимую для функционирования центромеры.

Если *CEN*-фрагмент из одной хромосомы переносят вместо центромеры в другую хромосому, то не обнаруживают никаких изменений. Этот результат говорит о взаимозаменяемости центромер. Они просто используются для прикрепления хромосомы к веретену и не играют никакой роли в том, чтобы отличить одну хромосому от другой.

Первыми были обнаружены два таких фрагмента, *CEN3* длиной 627 п. н. и *CEN11* длиной 900 п. н. (Цифро-

вое обозначение локуса соответствует порядковому номеру хромосомы.) Интересно, что эти маленькие участки обеспечивают адекватное поведение в митозе, но не выполняют сегрегационной функции в мейозе. Возможно, для спаривания и сегрегации хромосом в первом мейотическом делении нужны более протяженные участки ДНК, чем для простого митотического деления.

Нуклеотидная последовательность обоих участков *CEN3* и *CEN11* определена. Эти участки не содержат протяженных областей гомологии. Существует несколько коротких гомологичных участков, а также (А–Т)-богатые области равной длины. На рис. 28.11 показаны консервативные последовательности этих двух участков. Важное значение имеет, очевидно, то обстоятельство, что в обеих центромерах короткие гомологичные участки расположены на одинаковом расстоянии друг от друга. Ни в одной из центромер не обнаружено открытых рамок считывания. Существует некоторое сходство между этой последовательностью и определенными последовательностями сателлитной ДНК у высших эукариот. Эти участки можно будет использовать для получения делеций и других изменений, необходимых для того, чтобы непосредственно установить, какие именно свойства центромеры обеспечивают выполнение ее функций.

Другим важным признаком всех хромосом является **теломера**. Мы до сих пор еще не понимаем до конца природу этих «печатей» на концах хромосом. Уже давно ясно, что концы должны иметь особую структуру, поскольку при разрыве хромосом образуются «липкие» концы, склонные к взаимодействию с другими хромосомами, тогда как естественные концы хромосом стабильны.

При идентификации теломерной последовательности следует использовать два критерия. Во-первых, такая последовательность должна лежать на конце хромосомы (или, по крайней мере, на конце аутентичной линейной молекулы ДНК). Во-вторых, она должна придавать линейной молекуле стабильность.

Используя критерий локализации, теломеры можно выделить из необычных компонентов геномов низших эукариот. Одним из источников являются линейные молекулы рДНК, другим – ДНК макронуклеуса некоторых инфузорий. Несмотря на значительную эволюционную отдаленность этих форм генетического материала, все теломеры имеют некоторые общие особенности.

Теломеры состоят из длинных серий коротких, tandemно повторяющихся последовательностей. В табл. 28.3 приведен перечень повторяющихся единиц, идентифицированных на концах линейных молекул ДНК. Их всех можно записать в общем виде как $C_n(A/T)_m$, где $n > 1$, а m равно 1–4.

Таблица 28.3
Теломеры содержат короткие тандемные повторы общего типа¹

Организм	Источник ДНК	Повторяющаяся единица (5-3)
Инфузории отряда Holotricha (например, <i>Tetrachymena</i> , <i>Paramecium</i>)	Макронуклеус	ССССАА
Инфузория отряда Hypotricha (например, <i>Stylonchia</i> , <i>Oxytricha</i>)	Макронуклеус	ССССАААА
Жгутиковые (например, <i>Trypanosoma</i> , <i>Zeichmania</i>)	Мини-хромосома	СССТА
Слизевики <i>Physarum</i>	рДНК	СССТА _n
<i>Dictyostelium</i>	рДНК	C ₁₋₈ T
Дрожжи (<i>Saccharomyces</i>)	Хромосома	C ₂₋₃ A(CA) ₁₋₃

¹ Повторяющиеся единицы даны в виде последовательности одной цепи в направлении от теломеры к центру

В пределах теломерной области существует специфический ряд одноцепочечных разрывов, структура которых препятствует их сшиванию лигазой, зашивающей обычно разрывы в одноцепочечной ДНК. Основания, расположенные на самом конце, оказываются каким-то образом заблокированными — они могут образовывать шпильку — так что нуклеазы их не узнают.

Для изучения данной функции на молекулярном уровне и на этот раз были использованы дрожжи. Все плазмиды, выживающие в дрожжевых клетках, являются кольцевыми молекулами ДНК. Линейные плазмиды нестабильны (поскольку они деградируют). Могут ли аутентичные теломерные последовательности ДНК придавать стабильность линейной плазмиде?

Оказалось, что концевой участок из естественной линейной последовательности ДНК — экстрахромосомной рДНК *Tetrahymena* — способен придать стабильность линейной дрожжевой плазмиде. Эта предполагаемая теломера состоит из повторяющейся последовательности, содержащей серию одноцепочечных разрывов. Эти разрывы располагаются у дрожжей в одних и тех же сайтах и проявляют явную межвидовую консервативность. Каждый конец линейной ДНК может иметь форму шпильки.

Последовательность ДНК *Tetrahymena* можно заменить некоторыми фрагментами дрожжевой ДНК, содержащими родственные последовательности. Пока что они выделены в виде фрагментов размером около 4 т.п.н. Предстоит доказать, что они являются естественными теломерами дрожжевых хромосом. Для того чтобы выяснить, как они функционируют, следует провести исследования такого же типа, как и при изучении центромерной ДНК — а именно использовать метод уменьшения размера и индуцировать мутации.

Некоторое представление о том, как функционирует теломера, дают необычные свойства концов линейных молекул ДНК. В популяции трипаносом длина концов молекул рДНК различна. Если проследить за судьбой одного клона, то оказывается, что длина теломеры увеличивается на 7–10 п.н. в расчете на одну генерацию. Еще более показательна судьба теломер инфузорий, введенных в дрожжевые клетки. После репликации в этих клетках к концам повторов *Tetrahymena* присоединяются теломерные последовательности дрожжевых клеток.

Похоже на то, что добавление теломерных повторов к концу хромосомы происходит в каждом цикле репликации (таким образом, вероятно, разрешается проблема репликации линейных молекул ДНК, обсуждающаяся в гл. 33). Некий механизм должен препятствовать чрезмерному росту концов, очевидно, удаляя часть повторов.

Если теломеры постоянно удлиняются, их последовательности могут не соответствовать точно друг другу. Все, что от них требуется, это только чтобы данный конец был узнан как подходящий субстрат для добавления. Это объясняет, каким образом теломера инфузорий функционирует в дрожжах. Однако мы еще не понимаем, как теломерная последовательность связана с устойчивостью хромосомных концов к повреждениям.

Минимум требований, необходимых для существования индивидуальной хромосомы, сводится к наличию 1) соответствующей целостности теломеры, 2) центромеры для поддержания сегрегации и 3) точки инициирования репликации (гл. 31). Все эти элементы были объединены для создания синтетической дрожжевой хромосомы. Оказалось, что синтетическая хромосома стабильна только в том случае, если ее длина превышает 20–50 т.п.н. Мы пока еще не знаем природы этого явления, но само создание синтетической хромосомы дает потенциальную возможность изучать природу сегрегации генетического материала в контролируемых условиях.

Деспирализованное состояние хромосом «ламповых щеток»

Для того чтобы установить, какие структурные изменения связаны с транскрипцией, было бы очень полезно пронаблюдать экспрессию генов в их естественном состоянии. Однако организация генетического материала такова, что подобный анализ возможен только в некоторых исключительных случаях. Из-за плотной упаковки ДНК в хромосоме и трудности идентификации местоположения отдельных генов визуализация транскрипции отдельных генов невозможна. (Однако, как это описано в гл. 30, активные гены проявляют некоторые характерные особенности, которые можно исследовать *in vitro*.)

Митотические хромосомы транскрипционно инертны. Во всяком случае они так компактны, что в них нельзя различить отдельных локусов. В отдельных участках, которые становятся видимыми при окрашивании методом G-сегментов (как это показано на рис. 28.9), содержится примерно 10⁷ п.н. ДНК, и в них может поместиться много сотен генов.

Когда хромосомы вступают в мейоз, у многих из них видна латеральная дифференцированность структуры. На этой стадии хромосомы похожи на нитку бус. Бусинки представляют собой темноокрашенные гранулы, известные под названием **хромомер**. В мейозе, однако, экспрессия генов незначительна, и этот материал практически не используют для идентификации индивидуальных генов.

Экспрессию генов можно непосредственно наблюдать в двух случаях, когда хромосомы находятся в очень вытянутой форме и можно различать индивидуальные локусы (или группы локусов). Это хромосомы типа ламповых щеток (лучше всего изученные у некоторых амфибий) и политенные хромосомы (лучше всего изученные у некоторых насекомых).

Хромосомы типа ламповых щеток образуются в про-

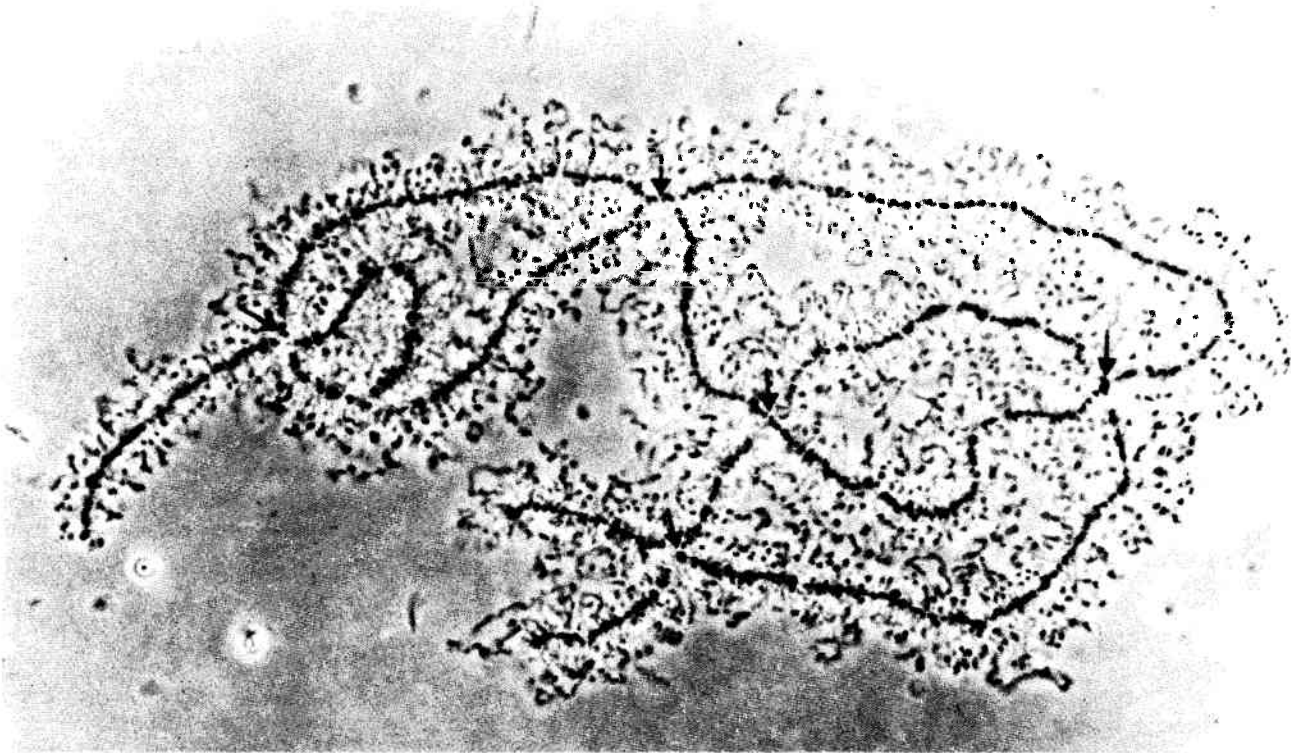


Рис. 28.12. Хромосома типа «ламповой щетки», представляющая собой мейотический бивалент, у которого две пары сестрин-

ских хроматид удерживаются вместе в местах хиазм (указано стрелками). (Фотография любезно предоставлена Joe Gall.)

цессе необычайно растянутого мейоза, который может длиться несколько месяцев! В течение этого времени хромосомы находятся в растянутом состоянии и их можно видеть под световым микроскопом. В позднем мейозе хромосомы снова возвращаются к своей обычной компактной форме. Таким образом, растянутые хромосомы являются раскрученным вариантом нормального состояния хромосом.

Хромосомы типа ламповых щеток являются мейотическими бивалентами; каждая такая хромосома состоит из двух пар сестринских хроматид. На рис. 28.12 изображен случай, когда пары сестринских хроматид максимально разделены и удерживаются вместе только хиазмами. Каждая пара сестринских хроматид образует серии эллипсоидных хромомер диаметром около 1–2 мкм, соединенных очень тонкой нитью. Эта нить состоит из двух сестринских двухцепочечных ДНК и тянется непрерывно вдоль хромосомы через все хромомеры. Размеры хромосом типа ламповых щеток у тритона *Notophthalmus viridescens* колеблются от 400 до 800 мкм, а их длина в позднем мейозе составляет 15–20 мкм. Следовательно, хромосомы типа ламповых щеток упакованы почти в 30 раз менее компактно. Общая длина полного набора хромосом типа ламповых щеток может составлять 5–6 мм и включает в себя около 5000 хромомер.

Хромосомы типа ламповых щеток получили свое название из-за наличия боковых петель, выступающих в определенных положениях из хромосом. (Они напоминают щетку для чистки керосиновой лампы, которая сейчас уже вышла из употребления.) Петли располагаются парами, по одной у каждой сестринской хроматиды. Петли являются продолжением осевой нити и представляют собой хромосомный материал, который выступает из более компактно уложенного материала хромомеры.

Петли окружены матриксом из рибонуклеопротеинов, содержащих новосинтезированные цепи РНК. Часто единицу транскрипции можно определить по увеличению длины РНП, движущегося вдоль петли. Пример этого показан на рис. 28.13.

Таким образом, петля — это активно транскрибируемый выступающий участок ДНК. В некоторых случаях петли соответствуют определенным генам, которые были идентифицированы. В этих случаях можно было изучить структуру транскрибированного гена и природу его продукта *in situ*.

Гигантские хромосомы образуются в результате политении

Интерфазные ядра в некоторых тканях личинок двукрылых мух содержат хромосомы во много раз большего размера, чем обычные хромосомы. У этих хромо-

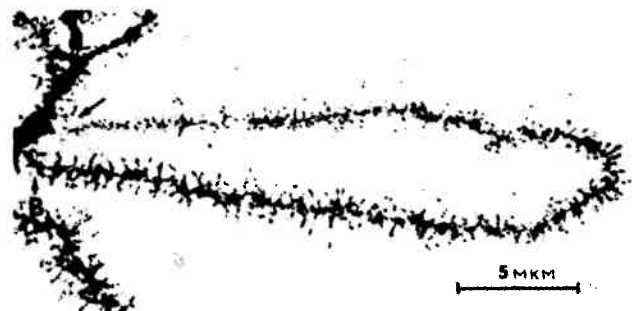


Рис. 28.13. Петля хромосомы «ламповая щетка», окруженная матриксом из рибонуклеопротеина. (Фотография любезно предоставлена O. Miller.)

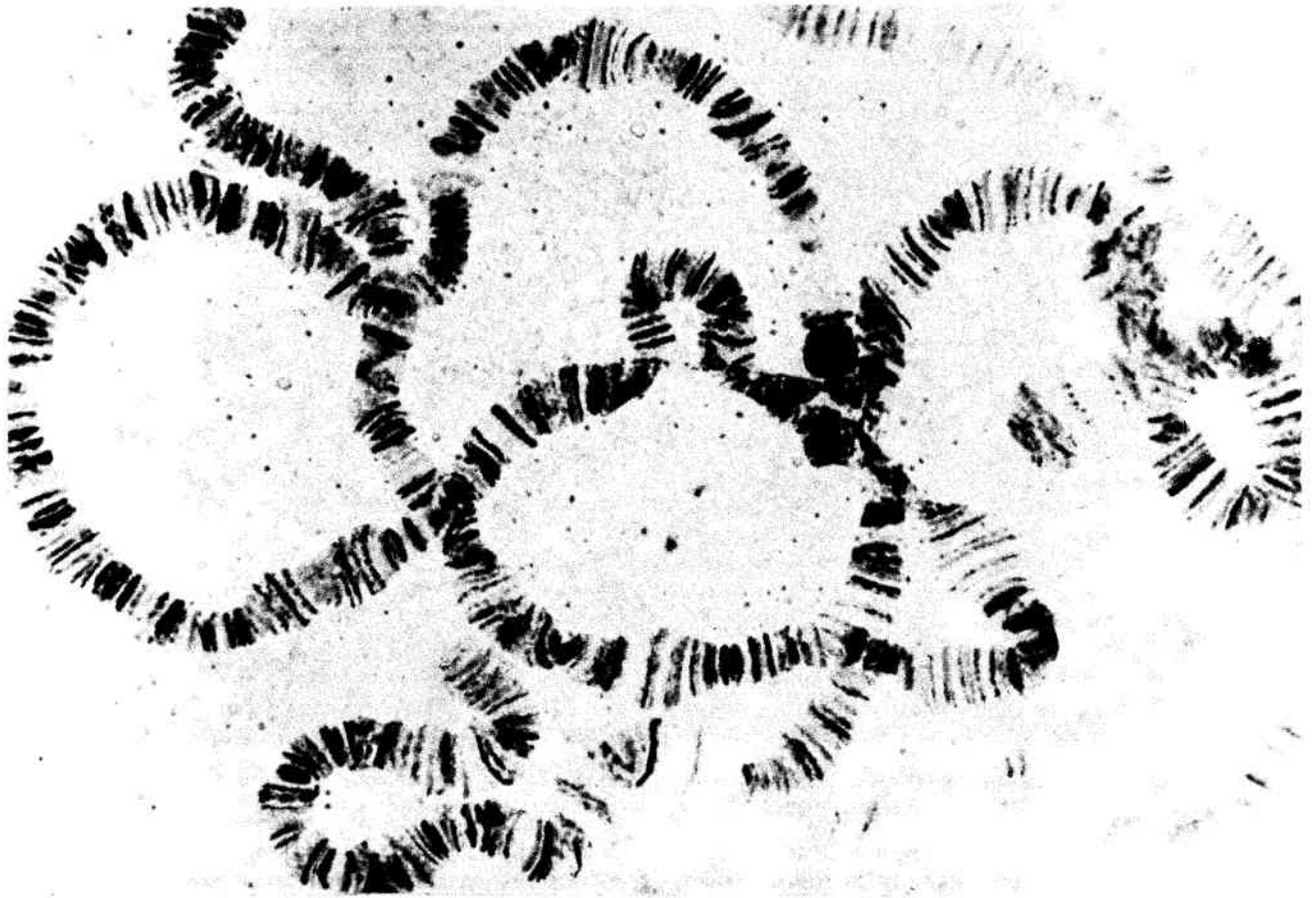


Рис. 28.14. Политенные хромосомы *D. melanogaster* образуют серии дисков и междисков.

(Фотография любезно предоставлена J. Bonner.)

сом увеличены диаметр и длина. На рис. 28.14 в качестве примера изображен хромосомный набор клеток слюнной железы *D. melanogaster*. Такие хромосомы называют **политенными**.

Каждая хромосома состоит из видимой серии полос-дисков (более правильно, но редко описываемых как хромомеры). Существует примерно десятикратное различие в размере дисков. Наибольшая ширина составляет примерно 0,5 мкм, а наименьшая — 0,05 мкм. (Наименьшие диски видны только под электронным микроскопом.) В дисках находится большая часть массы ДНК и они интенсивно окрашиваются соответствующими реагентами. Участки между отдельными дисками окрашены светлее; их называют междисками. В хромосомном наборе *D. melanogaster* содержится около 5000 дисков.

Центромеры всех четырех хромосом *D. melanogaster* агрегируют с образованием хромоцентра, который состоит в значительной степени из гетерохроматина (а у самцов включает целиком Y-хромосому). Учитывая это, можно считать, что около 75% гаплоидного набора ДНК организовано в диски и междиски. Общая длина хромосомного набора (измеренная в метафазе митоза) составляет около 2000 мкм, а эти 75% ДНК содержат $1,3 \cdot 10^8$ п. н., которые вытянуты на 40 000 мкм, так что средняя плотность упаковки равна ~ 20 . Это ясно показывает, в каком растянутом состоянии находится генетический материал в политенной хромосоме по сравнению с его обычными состояниями в интерфазном хроматине или в митотических хромосомах.

Какова структура этих гигантских хромосом? Каждая хромосома образуется в результате последовательных актов репликации конъюгировавшей диплоидной пары. Реплики не разделяются, а остаются прикрепленными друг к другу по всей длине. В начале процесса каждая конъюгировавшая пара содержит количество ДНК, равное $2C$ (где C — это содержание ДНК в одной индивидуальной хромосоме). Затем число удвоений происходит до 9 раз; в результате максимальное содержание ДНК может достигнуть $1024C$. Число удвоений варьирует в различных тканях личинки *D. melanogaster* и, конечно, у разных двукрылых мух тоже.

Каждая политенная хромосома видна в виде большого числа параллельных нитей, сильно сконденсированных в дисках и слабее в междисках. Вероятно, каждая такая нить представлена единичной копией (C) гаплоидной хромосомы. Отсюда и происходит термин «политения». Степень политении определяется числом гаплоидных хромосом, содержащихся в гигантской хромосоме.

Для каждой линии дрозофил характерен свой набор дисков. Постоянство числа и линейного расположения дисков впервые было отмечено в 30-х годах, когда исследователи поняли, что диски образуют **цитологическую карту** хромосом. В результате перестроек — таких, как делеции, инверсии или дупликации, — происходит изменение линейного порядка дисков. Бриджес (Bridges) первый осознал, что линейный порядок дисков соответствует линейному расположению генов. (Его исходная карта хромосом с некоторыми модификациями до сих пор служит

основой всех работ.) Таким образом, генетические перестройки, как это видно по карте сцепления, можно привести в соответствие со структурными перестройками на цитологической карте. В конечном итоге конкретную мутацию можно локализовать в определенном диске.

Положение определенных генов на цитологической карте можно определить непосредственно с помощью метода *in situ*, или **цитологической гибридизации**, рассмотренной в гл. 24 в связи с высокоповторяющимися последовательностями ДНК. Общая схема кратко повторена на **рис. 28.15**. Радиоактивная проба, представляющая некий ген (чаще всего меченый клон кДНК, полученный с мРНК), гибридизуется с денатурированной ДНК политенных хромосом *in situ*. Методом радиоавтографии можно определить местоположение одного или нескольких исследуемых генов, наблюдая за скоплением зерен серебра над одним или несколькими дисками, как это показано на **рис. 28.16**. Используя этот метод, можно прямо установить диск, в котором расположена данная последовательность ДНК.

Нарушение хромосомной структуры при транскрипции

Одна из наиболее удивительных особенностей политенных хромосом — это возможность непосредственного наблюдения их активных участков. Некоторые диски иногда «разбухают» или **образуют пuffs**; при этом наблюдается выпячивание хромосомного материала из осевой структуры. Пример очень больших пuffs (названных кольцами Бальбиани) изображен на **рис. 28.17**.

Какова природа пuffs? Пuff представляет собой участок, в котором хромосомные нити находятся в более раскрученном виде по сравнению с их обычным состоянием, когда они упакованы в диске. Вместе с тем эти нити остаются непрерывно связанными с нитями, находя-



Рис. 28.16. Под большим увеличением показана гибридизация *in situ* на дисках 87A и 87C с меченой РНК, выделенной из клеток после теплового шока.

(Фотография любезно предоставлена J. Bonner.)

щимися в осевой части хромосомы. Пuffs обычно образуются из материала отдельных дисков, хотя некоторые особенно крупные пuffs, такие, как кольца Бальбиани, разбухают так сильно, что заслоняют собой соседние диски.

Порядок образования пuffs соответствует транскрипции генов. В процессе развития личинки пuffs появляются и исчезают в определенном тканеспецифическом порядке. Каждая ткань в любой момент развития характеризуется определенной картиной расположения пuffs. Пuffs индуцируются гормоном экдизоном, который контролирует развитие дрозофилы. Некоторые пuffs ин-

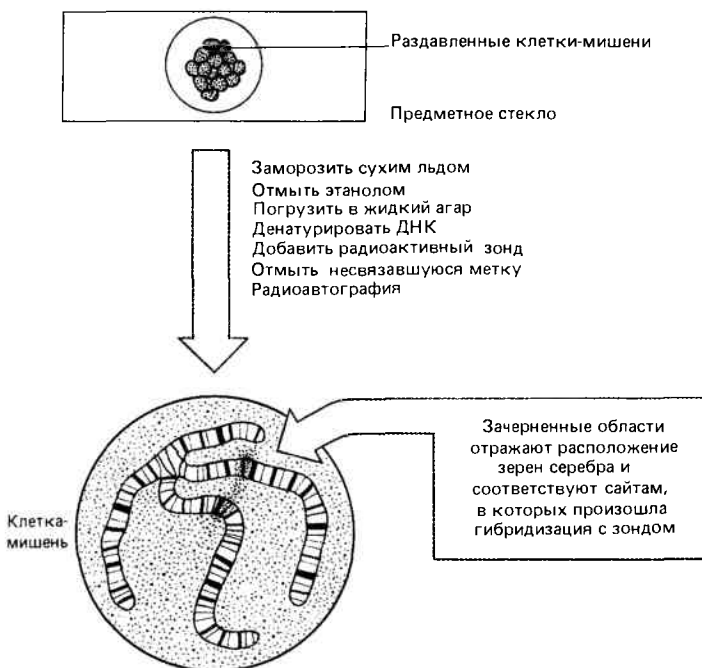


Рис. 28.15. Индивидуальные диски, содержащие определенные гены, можно идентифицировать методом гибридизации *in situ*.

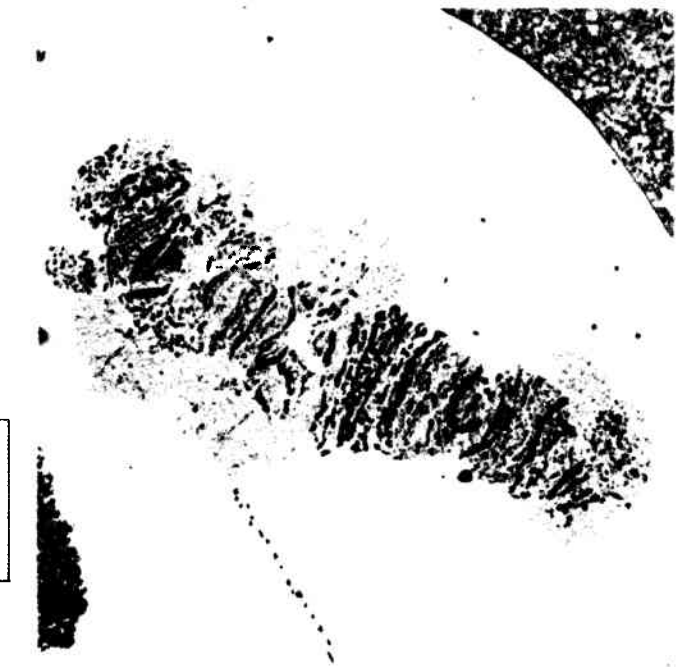


Рис. 28.17. На хромосоме IV из слюнной железы *C. tentans* видны три кольца Бальбиани.

(Фотография любезно предоставлена B. Daneholt.)

дуцируются непосредственно гормонами, тогда как индукция других опосредована продуктами более ранних пуфов.

Пуфы — это сайты, в которых происходит в данный момент синтез РНК. Согласно общепринятым представлениям, образование пуфов представляет собой увеличение размеров одного диска, связанное с необходимостью деконденсировать его структуру для синтеза РНК. Поэтому образование пуфов рассматривают как следствие транскрипции.

Однако есть некоторые основания считать, что образование пуфов, напротив, может быть условием транскрипции. В пользу такого предположения свидетельствует наличие мутанта, у которого в сайте, соответствующем синтезу белков глии, имеется пуф, даже несмотря на то, что мутация препятствует транскрипции генов, локализованных в данном диске. Вероятно, мутация останавливает выражение гена на какой-то стадии после образования пуфа, но до начала транскрипции. Идея о том, что образование пуфов может быть отделено от акта транскрипции, является новой, поскольку этот мутант остается пока единственным исключением из очевидного правила, что наличие пуфа непосредственно отражает протяженность области синтеза РНК.

От интенсивности синтеза зависит размер пуфа. Участки с пуфами отличаются от обычных дисков по накоплению дополнительных белков. В настоящее время эти белки охарактеризованы только в общем виде, но в свое время возможно будет установить, являются ли они причиной или следствием образования пуфов. И тогда можно будет выяснить природу молекулярных событий, вызывающих деконденсацию генетического материала.

Таким образом, пуф — это диск, содержащий активно транскрибируемый ген. Является ли диск единицей транскрипции? На этот несколько спорный вопрос сейчас еще нет окончательного ответа.

Генетические опыты были сконцентрированы в основном на изучении корреляции летальных локусов с определенными дисками. Летальный локус определяют по мутациям, нарушающим жизнеспособность организма. Таким способом идентифицируют жизненно важные гены. Число летальных локусов близко к числу дисков. В определенных областях хромосом отношение индивидуальных дисков к числу летальных локусов составляет 1:1. Из этого можно сделать вывод, что у *D. melanogaster* около 5000 летальных локусов.

В среднем диске содержится 25 т.п.н. гаплоидной ДНК. А priori нельзя предсказать, скольким генам это соответствует: нескольким тесно расположенным генам или

одному большому гену с большими интронами и протяженными фланкирующими последовательностями. Однако анализ последовательностей ДНК индивидуальных генов показывает, что в ряде случаев один диск содержит более одной транскрипционной единицы. Если такие диски представляют типичное большинство, общее число генов может быть в несколько раз больше, чем 5000. Мы еще не знаем, находятся ли различные единицы транскрипции в таких дисках под общим контролем или нет. Есть все основания думать, что достаточно эффективная экспрессия любого гена внутри диска может привести к появлению пуфа.

В настоящее время существует некоторое расхождение между генетическими и биохимическими данными. Первые свидетельствуют в пользу того, что каждому диску соответствует только один жизненно важный ген, тогда как, согласно вторым — их несколько. Это противоречие может быть урегулировано, если многие гены окажутся не жизненно важными, по крайней мере по критерию летальности. При этом остается загадкой, почему летальные локусы распределяются равномерно, примерно по одному на диск.

Исходя из особенностей хромосом типа ламповых щеток и политеменных хромосом можно, очевидно, сделать общий вывод: чтобы транскрипция могла осуществиться, генетический материал, обычно плотно упакованный, должен деконденсироваться.

Однако еще остается вопрос, соответствует ли такое локальное расплетение, происходящее на уровне целой хромосомы, тем событиям, которые происходят на молекулярном уровне в общей массе обычного интерфазного эухроматина.

Рекомендуемая литература

Данные по морфогенезу фагов детально изложены в книге Льюина (*Lewin, Gene Expression, 3, Plasmig and Phages Wiley, New York, 496–535, 642–673, 1978*). Необычные хромосомы, обсуждавшиеся в этой главе, рассмотрены в той же книге (*Lewin, Gene Expression 2, Eucaryotic Chromosomes, 455–475*). О последних достижениях в исследовании центромер сообщается в работе Фитцджеральде-Хейса, Кларка и Карбона (*Fitzgerald-Hayes, Clarke, Carbon, Cell, 29, 235–244, 1982*). Теломерам посвящена работа Блэкбёрна (*Blackburn, Cell, 37, 7–8, 1984*); центромерам — статья Карбона (*Carbon, Cell, 37, 351–353, 1984*). Мёррей и Шостак описали конструирование «синтетической» хромосомы из ее функциональных элементов (*Murray, Szostak, Nature, 305, 189–193, 1983*).

Глава 29

НУКЛЕОСОМНЫЕ ЧАСТИЦЫ И СТРУКТУРА ХРОМАТИНА

Одним из наиболее замечательных достижений за последние десять лет явилось установление общей структуры хроматина. До этого мы знали только, что хроматин содержит ДНК, РНК и белок, организованные

в компактную форму, в которой большинство последовательностей ДНК оказываются структурно недоступными и функционально неактивными. Считали, что причина этой недоступности заключается в связи с белком (часто

писали, что ДНК «покрыта» белками), однако не было никакой информации о структуре этого нуклеопротеина на уровне взаимодействия ДНК—белок.

Огромный прогресс был достигнут с тех пор, как в 1974 г. удалось установить тот существенный факт, что хроматин сконструирован из субъединиц, имеющих одинаковый тип организации у всех эукариот. Эта субъединица состоит примерно из последовательности ДНК в 200 п. н., которая структурно организована с помощью небольших основных белков (гистонов) в частицу, похожую на бусинку. Оказалось, что ДНК лежит на поверхности этой частицы, тогда как белковые компоненты находятся внутри.

С точки зрения фундаментальных представлений о ДНК-белковых взаимодействиях мы можем описать структуру отдельной частицы, проследив прохождение нити ДНК и ее взаимодействие с белками, а также взаимодействие отдельных белков между собой. Можно задать вопрос: существуют ли вариации в конформации частиц; как такие вариации сказываются на функциях хроматина? Что происходит с частицами при транскрипции и репликации хроматина? Для этих процессов могут быть необходимы другие (негистоновые) белки, чтобы модулировать поведение частиц.

При репликации происходит воспроизведение определенной серии частиц. Кроме удвоения количества ДНК хроматина должно также удвоиться количество белка. Наряду с вопросом о том, как устроены частицы, необходимо выяснить, что происходит с другими белками, присутствующими в хроматине. Поскольку репликация, вероятно, нарушает структуру хроматина, при этом, с одной стороны, возникает проблема сохранения участков со специфической структурой, с другой же, создается возможность внести в нее изменения.

Общая высокая плотность упаковки генетического материала говорит о том, что ДНК не может непосредственно укладываться в хроматин с окончательной структурой. Должны существовать *иерархические* ступени организации.

Первый уровень—это закручивание ДНК в бусовидные частицы с плотностью упаковки около 6. Эти частицы являются неизменной составной частью эухроматина, гетерохроматина и хромосом.

Второй уровень организации заключается в закручивании серий бус в спиральную структуру с образованием волокон размером около 30 нм, которые найдены в интерфазном хроматине и в митотических хромосомах (см. рис. 28.7). При этом плотность упаковки хроматина достигает 40. Пока точно не известно, увеличивается ли

плотность упаковки фибрилл в митотической хромосоме или она остается прежней.

Окончательная плотность упаковки зависит от третьего уровня организации, т.е. от укладки самой фибриллы. При этом плотность упаковки в эухроматине возрастает до ≥ 1000 . Это значение периодически изменяется при образовании митотических хромосом и достигает ≤ 10000 . Данный процесс, вероятно, также регулируется негистоновыми белками (и, возможно, также включает химическую модификацию гистонов), как и в случае образования различий между эухроматином и гетерохроматином.

Белковые компоненты хроматина

Относительная доля компонентов хроматина варьирует в зависимости от источника и метода выделения. Масса белка почти в 2 раза превышает массу ДНК. Белки подразделяются на два типа: гистоны и негистоновые белки. Масса РНК составляет менее 10% массы ДНК. Большая часть РНК состоит из новосинтезированных цепей, которые еще связаны с ДНК-матрицей.

Гистоны составляют большинство основных белков хроматина и находятся примерно в том же количестве, что и ДНК. По относительной доле основных аминокислот каждого типа, которую выражают отношением лизин/аргинин, сначала охарактеризовали пять типов гистонов. След этой классификации до сих пор остается в названиях гистонов. Практически у всех эукариот обнаруживают одни и те же классы гистонов. Их свойства суммированы в табл. 29.1.

Гистоны четырех классов прямо взаимодействуют с ДНК и образуют в хроматине серию частиц первого уровня организации. Консервативность типов гистонов на протяжении эволюции можно объяснить необходимостью сохранения этой важнейшей реакции. Пятый класс гистонов принимает участие во взаимодействиях между частицами. Постоянство классов гистонов позволяет предполагать, что взаимодействия типа ДНК—гистоны, гистон—гистоны и гистон—негистоновые белки могут быть в основном похожими у разных видов. Отсюда мы можем сделать заключение об общих механизмах образования как первичных частиц, так и последующих структур более сложного порядка, состоящих из серий частиц.

Гистоны первых четырех классов имеют значительное количество как кислых, так и основных аминокислот. Поэтому эти белки несут высокий заряд. Отношение основных аминокислот к кислым находится в диапазоне 1,4–2,5. Эти гистоны подразделяются на две группы.

К аргинин-богатым относятся два вида гистонов: Н3 и Н4. Они принадлежат к наиболее консервативным из всех известных белков. Аминокислотные последовательности этих белков идентичны даже у таких удаленных видов, как корова и горох. В гистонах Н3 и Н4 других видов обнаружены только редкие аминокислотные замены. Консервативность целой последовательности говорит о том, что все ее аминокислоты имеют существенное значение для выполнения функции белка. По логике вещей эта функция должна быть одинаковой у огромного большинства различных видов, что свидетельствует в пользу концепции об общей основе для структуры хроматина.

К гистонам, умеренно обогащенным лизином, относятся два белка. Их называют **Н2А** и **Н2В** (в противоположность их номенклатурному обозначению это не род-

Таблица 29.1
Гистоны—это сильноосновные белки

Гистон	Основные аминокислоты, %		Кислые аминокислоты, %	Отношение основных к кислым	Молекулярная масса
	лизин	аргинин			
Н1	29	1	5	5,4	23 000
Н2А	11	9	15	1,4	13 960
Н2В	16	6	13	1,7	13 774
Н3	10	13	13	1,8	15 342
Н4	11	14	10	2,5	11 282

Эти данные относятся к гистонам млекопитающих (крупный рогатый скот); остальные гистоны имеют аналогичный состав.

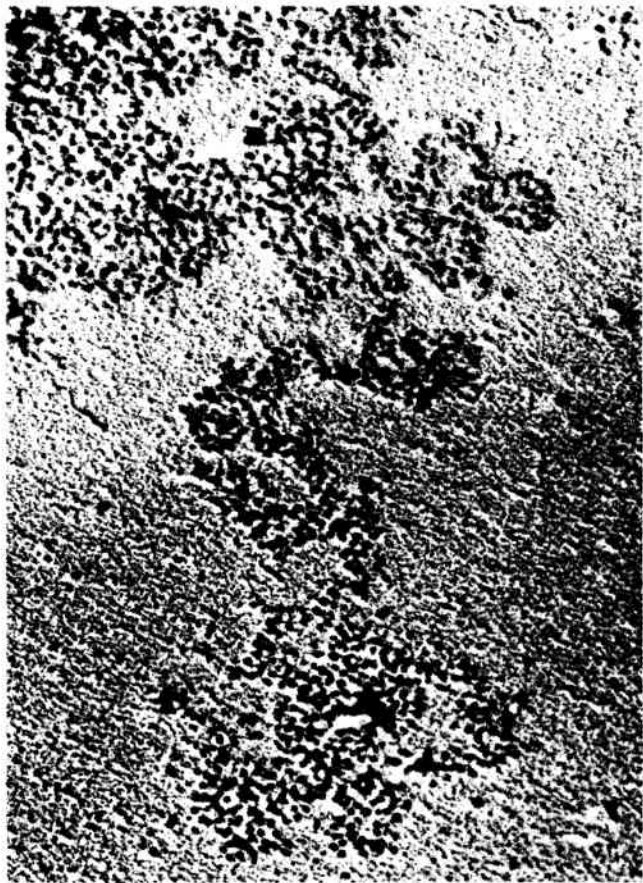


Рис. 29.1. Хроматин, вытекающий из лизированного ядра, состоит из серии плотноорганизованных частиц. (Фотография любезно предоставлена Pierre Chambon.)

ственные, а независимые белки). У различных эукариот находят те же самые два типа гистонов, но у них обнаружены заметные межвидовые вариации в аминокислотной последовательности.

Пятый класс представлен **гистонами**, очень богатыми **лизином**; он состоит из нескольких достаточно близкородственных белков с перекрывающимися последовательностями аминокислот. Это гистоны **H1** (в эритроцитах птиц существует вариант, названный **H5**). У этих гистонов обнаружены значительные межвидовые и межтканевые вариации (у дрожжей, по-видимому, гистонов данного класса нет). Хотя эти гистоны являются самыми основными гистонами, их легко можно выделить из хроматина, полностью растворив в солевом растворе (0,5M).

Как и следует из названия, **негистоны** — это все другие белки хроматина. Предполагается поэтому, что они обладают большими видовыми и тканевыми различиями, хотя строгих данных о степени их разнообразия пока нет. Эти белки составляют меньшую долю от всей массы белков хроматина, чем гистоны. Кроме того, сюда относится намного большее число белков, так что любой индивидуальный белок присутствует в значительно меньшем количестве, чем любой гистон.

В класс негистоновых белков могут попасть белки, связанные с экспрессией генов, и белки, участвующие в организации структур высшего порядка. Так, в числе наиболее выдающихся негистонов можно назвать РНК-полимеразу. НМГ-белки (высокомобильная группа) составляют отдельный, хорошо различимый подкласс неги-

стонов. Основная проблема, возникающая при работе с другими негистоновыми белками — их загрязнение другими ядерными белками.

Хроматин содержит дискретные частицы

Если интерфазные ядра суспендировать в растворе с низкой ионной силой, они разбухнут и в местах разрывов из них высвободятся нити хроматина. На **рис. 29.1** показано лизированное ядро, из которого вытекают нити. В некоторых местах нити хроматина состоят из плотноупакованного материала, но в тех местах, где они вытянуты, можно видеть, что они состоят из отдельных частиц. Эти частицы называют **нуклеосомами**. В особенно вытянутых участках видно, что индивидуальные нуклеосомы соединены тонкой нитью — это свободная двухцепочечная ДНК. Таким образом, непрерывная двухцепочечная ДНК проходит через серию частиц.

Индивидуальные нуклеосомы можно получить, обработав хроматин ферментом **нуклеазой микрококков**. Это эндонуклеаза, которая разрезает нить ДНК в местах соединения между нуклеосомами. Сначала освобождаются группы частиц, а потом отдельные нуклеосомы. Мономерные нуклеосомы отчетливо видны на **рис. 29.2** в виде компактных частиц (настоящая форма которых похожа на диск; см. ниже). Они седиментируют примерно со скоростью 11S, что соответствует общей массе в диапазоне 250 000–300 000 дальтон. Отношение белок/ДНК составляет около 1,25. Димеры, тримеры и т. д. имеют соответствующие свойства при биохимическом анализе или при наблюдении под электронным микроскопом.

Мономерные нуклеосомы содержат ДНК (~200 п. н.), связанную с гистоновым октамером. Этот октамер содержит гистоны H2A, H2B, H3 и H4 — по две копии каждого. Иногда их называют гистоновой сердцевинкой (гистоновым кором). Такие комплексы схематически изображены на **рис. 29.3**.

Эта модель объясняет стехиометрию гистоновой сердцевинки в хроматине, в котором H2A, H2B, H3 и H4 находятся в эквимолярных количествах, по одной молекуле каждого на ~100 п. н. ДНК. Прямое экспериментальное

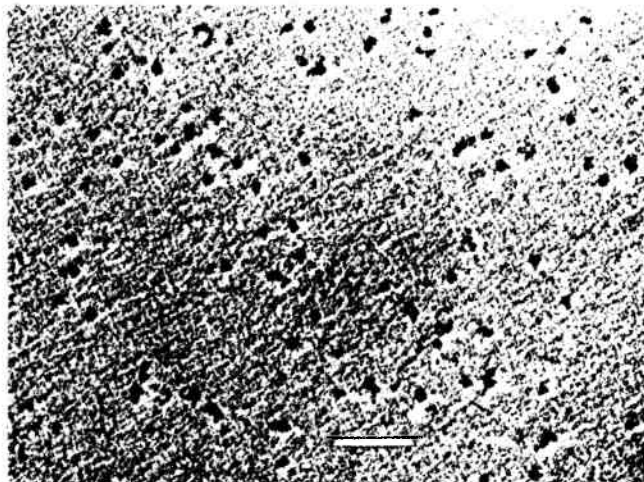


Рис. 29.2. Индивидуальные нуклеосомы высвобождаются в результате расщепления хроматина нуклеазой микрококков. Размер отрезка 100 нм.

(Фотография любезно предоставлена Pierre Chambon.)

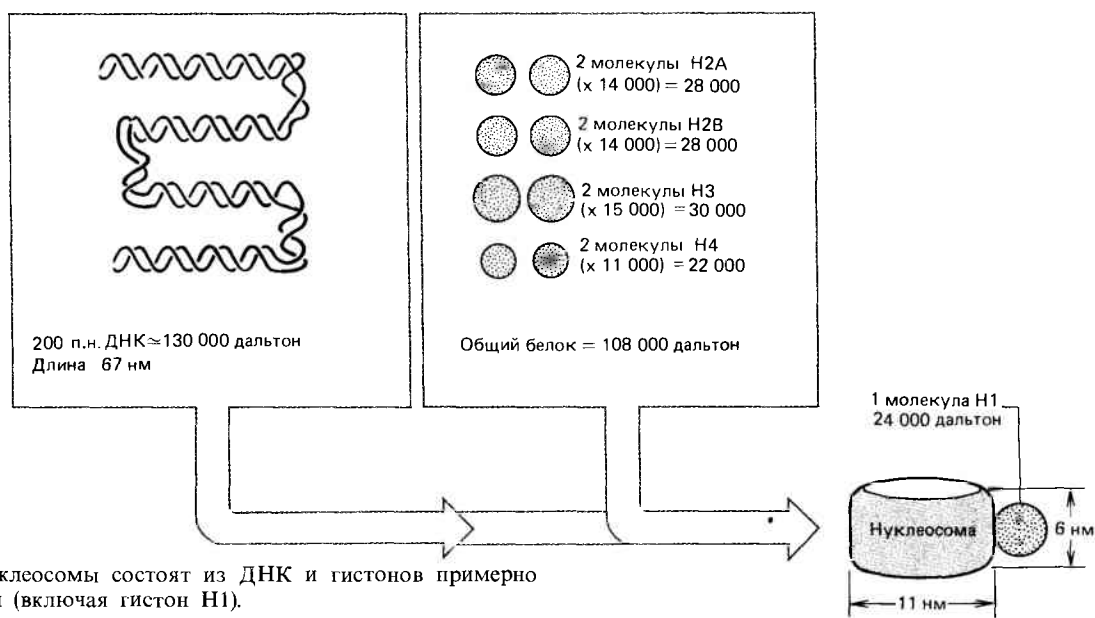


Рис. 29.3. Нуклеосомы состоят из ДНК и гистонов примерно равной массы (включая гистон Н1).

доказательство ассоциации в гистоновые октамеры получено методом образования перекрестных сшивок. В результате четыре из гистонов сердцевинки сшиваются перекрестно и образуют серию полос, по размеру достигающих октамер.

Одна молекула гистона Н1 связывается с каждой нуклеосомой. Его общее стехиометрическое количество в хроматине в два раза меньше, чем количество других гистонов. При выделении мультимерных нуклеосом он обычно присутствует в эквимольном количестве, но в препаратах мономеров его содержание меньше. Весь Н1 можно удалить, не нарушая структуры нуклеосом; из этого следует, что белок локализован снаружи.

Общая масса ДНК и белка, включая Н1, в расчете на нуклеосому составляет около 262 000 дальтон. При этом количество ДНК примерно равно количеству белка. Это приближенная оценка, полученная из экспериментальных значений. Любое дополнительное увеличение содержания белка зависит от присутствия небольшого количества негистоновых белков, которые связываются с нуклеосомами.

Гистон Н1 может лежать в месте соединения двух соседних нуклеосом. Это предположение согласуется со старыми данными о том, что Н1 достаточно легко удаляется из хроматина, причем такой оставшийся хроматин легче «растворяется». Более современные данные также подтверждают, что после удаления Н1 легче получить структуры, похожие на вытянутые нитки бус (см. ниже).

Нуклеосома – основная субъединица всего хроматина

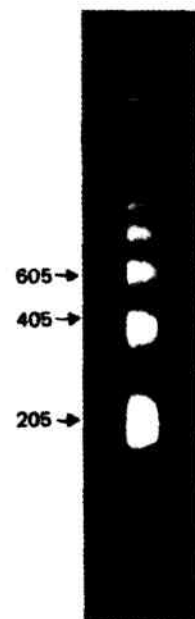
Простой диагностический тест на присутствие нуклеосом был разработан на основе данных о том, что ДНК в хроматине организована в регулярно повторяющуюся структуру. Если выделенные ядра или хроматин обработать нуклеазой микрококков (или в некоторых случаях просто выделить и обработать экзогенным ферментом), ДНК расщепляется на интегральное множество единиц длины. При этом в результате фракционирования методом электрофореза в геле образуется «лестница», пока-

званная на рис. 29.4. Такие лестницы состоят из 10 или более ступенек, и единица длины, измеренная по возрастанию последовательных ступенек, равна примерно 200 п.н.

Как показано на рис. 29.5, это наблюдение коррелирует со свойствами отдельных нуклеосом. Анализ выделенных частиц представлен в верхней части рисунка. Препарат нуклеосом был расфракционирован на мономеры, димеры, тримеры и т.д. методом седиментации в градиенте концентрации сахарозы. Затем очищали ДНК из каждой фракции и анализировали с помощью электрофореза в акриламидном геле.

Электрофорез образцов ДНК показан в нижней части рисунка. Каждая фракция образует полосу ДНК определенного размера. Отдельные полосы соответствуют определенным ступеням лестницы, полученной в результате расщепления хроматина. (Справа на фотографии представлена для сравнения «лестница».) Нуклеосомы-

Рис. 29.4. Нуклеаза микрококков расщепляет ядерный хроматин на мультимерные серии отрезков ДНК, которые можно разделить с помощью электрофореза. (Фотография любезно предоставлена Markus Noll.)



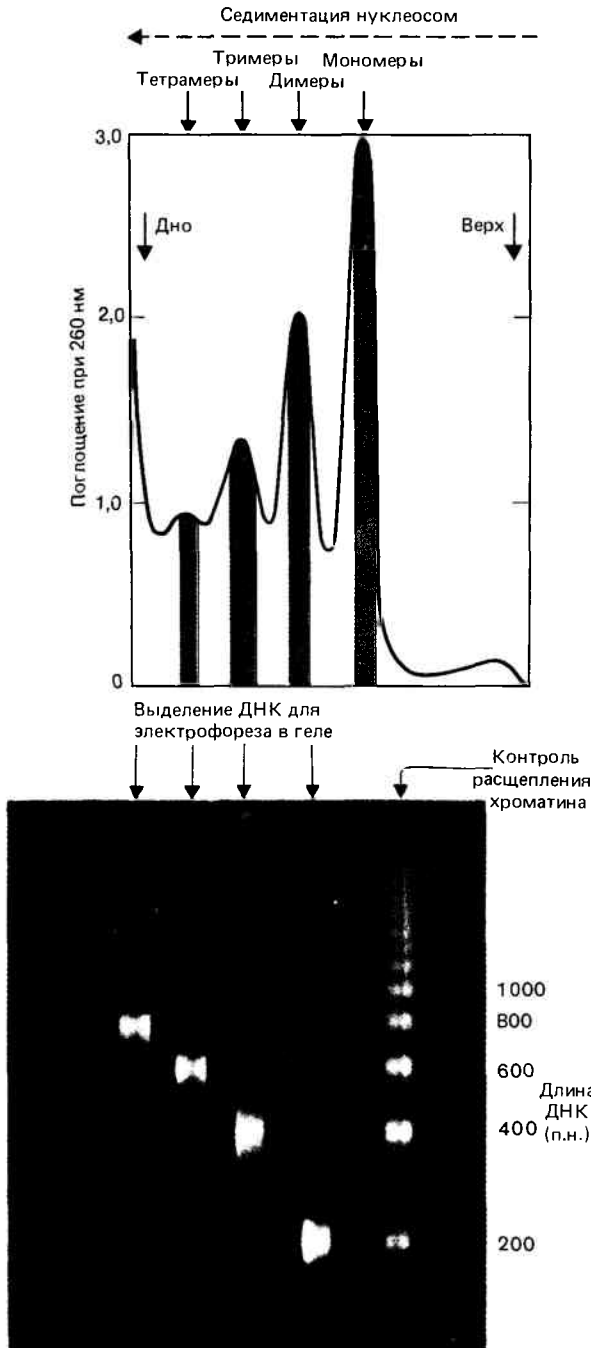


Рис. 29.5. Каждый мультимер нуклеосом содержит соответствующее число единиц длины ДНК.
(Фотография любезно предоставлена John Finch.)

мономеры содержат одну единицу длины ДНК, димеры нуклеосом содержат две единицы длины ДНК и т.д.

Каждая ступенька лестницы представляет собой ДНК, соответствующую определенному числу нуклеосом. Таким образом, можно принять, что существование «лестницы», кратной 200 п. н., указывает на нуклеосомную организацию ДНК. При обработке нуклеазой микрококков образуется «лестница» только примерно из 2% ядерной ДНК, которая переходит в кислоторастворимую фракцию (деградирует до мелких фрагментов). Таким образом, реагирует только часть ДНК, по-видимому соответствующая особенно чувствительным областям.

При высвобождении хроматина из ядра или при конструировании нуклеосом *in vitro* (см. ниже) мы часто видим серии нуклеосом, соединенных нитью свободной ДНК (бусинки на нитке). Разделены ли обычно нуклеосомы свободной ДНК, или это артефакт, возникающий *in vitro*? Поскольку *in vivo* ДНК должна быть плотно упакована, можно а priori предположить, что существует (если вообще есть) только небольшое количество свободной ДНК. Это предположение подтверждается данными о том, что более 90% ДНК хроматина оказывается в форме «лестницы» из фрагментов, кратных 200 п. н. Из этого следует, что большая часть ДНК находится в нуклеосомах, а не между ними. В естественном состоянии нуклеосомы, вероятно, так плотно упакованы, что ДНК непосредственно переходит из одной нуклеосомы в другую. Свободная же ДНК, по-видимому, образуется в результате утраты некоторых гистоновых октамеров в процессе выделения.

Длина ДНК, находящейся в одной нуклеосоме, может несколько отличаться от величины 200 п. н., которая была установлена в первых экспериментах. После того как были исследованы полные геномы из отдельных клеток, оказалось, что в каждом случае полученное значение достаточно точно соответствовало средней величине (± 5 п. н.). Средняя величина в большинстве случаев находилась между 180 и 200 п. н., но были обнаружены и крайние значения, такие, как 260 п. н. (в сперматозоидах морского ежа). Среднее значение может меняться в течение эмбрионального развития. В случае морского ежа оно уменьшается до более типичного уровня во время ранних делений эмбриональных клеток. Среднее значение может быть различным в разных тканях взрослого организма. Кроме того, различия могут наблюдаться в разных частях генома в клетках одного типа. Например, известны случаи отличия от среднегеномной величины для сателлитной ДНК и генов 5S РНК.

Частицы минимальной нуклеосомы высококонсервативны

Белковый компонент нуклеосомы представлен гистоновым октамером, вероятно, у всех эукариот. Такое постоянство структуры и функции объясняет строгую консервативность аминокислотных последовательностей гистонов. Чрезвычайная консервативность гистонов H3 и H4 объясняется тем, что их роль может быть центральной и неменяющейся, тогда как гистоны H2A и H2B обеспечивают видоспецифические вариации.

Все нуклеосомы состоят из гистонического октамера, связанного с ДНК определенной длины. Какие же факторы отвечают за вариацию длины ДНК в нуклеосомах, полученных из различных источников? На этот вопрос легче ответить, определив то, что *не* вызывает вариаций. Вариации не зависят от изменения в связывании ДНК с гистоновым октамером. Всегда образуются частицы **минимальной нуклеосомы** (нуклеосомного кора, кор-частицы, core particle), содержащие 146 п. н. ДНК, независимо от общей длины ДНК в нуклеосоме. Таким образом, варьирует размер ДНК, которая присутствует в нуклеосоме сверх основной структуры кора.

Минимальная нуклеосома была идентифицирована по действию нуклеазы микрококков на нуклеосомный мономер. Реакция этого фермента начинается с введения разреза между нуклеосомами. Но если продолжить реакцию после того, как образованы мономеры, то расщепление

захватывает часть ДНК самой нуклеосомы. Эта реакция протекает путем «подрезания» ДНК с концов нуклеосомы.

Как показано на рис. 29.6, длина ДНК уменьшается ступенчато. Например, в ядрах печени крысы нуклеосомные мономеры сначала содержали отрезок ДНК в 205 п.н. Потом были обнаружены мономеры, у которых длина ДНК была уменьшена до 160–170 п.н. Затем длина этих мономеров сократилась до размера ДНК в минимальной нуклеосоме—146 п.н. (эта величина достаточно стабильна, однако, если продолжать гидролиз, разрезы образуются и в пределах такой ДНК. При этом образуется набор фрагментов, из которых самые длинные соответствуют фрагменту ДНК минимальной нуклеосомы, а самые короткие имеют размер 20 п.н.).

На основе такого анализа можно разделить нуклеосомную ДНК на две части. ДНК минимальной нуклеосомы имеет постоянную длину 146 п.н. и относительно устойчива к расщеплению нуклеазами. (Другие нуклеазы, в том числе и нуклеаза микрококков, останавливаются перед этим отрезком ДНК.) Линкерная ДНК включает в себе остаток повторяющейся единицы. Ее длина может варьировать *in vivo* от такого малого размера, как 8 п.н., до такого большого, как 114 п.н., на 1 нуклеосому.

Четко ограниченный размер полосы ДНК, образованной в результате первого расщепления нуклеазой микрококков, говорит о том, что область, непосредственно доступная для действия фермента, ограничена. К ней относится только часть каждого линкера. (Если бы вся линкерная ДНК была чувствительна, размер полосы колебался бы в интервале от 146 п.н. до отрезка, превышающего размер повторяющейся единицы). Но, после того как линкерная ДНК разрезана, оставшаяся часть становится чувствительной и довольно быстро расщепляется ферментом.

Хотя частицы минимальной нуклеосомы меньше, чем сами нуклеосомы, их свойства похожи. Они напоминают нуклеосомы по форме и размеру, из чего можно заключить, что основная геометрия такой частицы устанавливается в результате взаимодействия между ДНК и гистоновым октамером. Из-за того что минимальные частицы легче получить в виде гомогенного препарата, их использовали для многих структурных исследований вместо препаратов нуклеосом. Мономеры нуклеосом больше варьируют в размерах из-за того, что трудно получить препараты, в которых не происходило бы «подрезания» концов ДНК.

Какова физическая природа минимальной нуклеосомы (нуклеосомного кора) и линкера? Эти термины являются *операциональными* определениями для обозначения областей, относительно более и менее чувствительных к обработке нуклеазой. Из этого нельзя делать каких-либо выводов об их действительной структуре. Это не означает, в частности, что линкерная ДНК имеет более вытянутую форму. С другой стороны, путь ДНК в нуклеосоме может быть непрерывным без каких-либо четких различий между этими областями мономера. В действительности это условное рабочее предположение, которое часто делают, пытаясь перейти от структуры минимальной нуклеосомы к структуре нуклеосом. С другой стороны, возможно, что путь линкерной ДНК отличается от пути ДНК в минимальной нуклеосоме, о чем свидетельствуют заметные вариации в длине линкеров.

Существование линкерной ДНК обеспечивается ка-

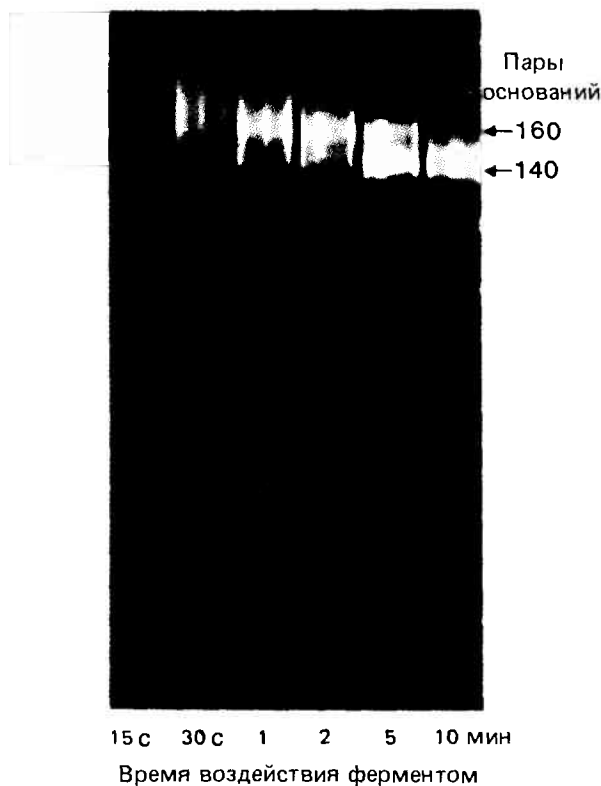


Рис. 29.6. Нуклеаза микрококков ступенчато уменьшает длину ДНК нуклеосомных мономеров. (Фотография любезно предоставлена Roger Kornberg.)

ким-то другим фактором, а не четырьмя гистонами сердцевин (гистонового кора). Эксперименты по реконструкции *in vitro* показывают, что гистонам присуща способность организовывать ДНК минимальной нуклеосомы, но не полную нуклеосому, длина ДНК в которой характерна для состояния *in vivo*. В этот процесс должны быть включены гистон Н1 и (или) негистоновые белки, которые связаны с гистоновым октамером в природных сериях нуклеосом. Первичная реакция может зависеть от «белков сборки», которые участвуют *in vivo* в образовании нуклеосом из гистонов и ДНК, но которые не становятся частью нуклеосомной структуры (см. ниже).

Мы уже говорили, что гистон Н1 теряется при деградации мономерных нуклеосом. Он еще остается в мономерах, содержащих 160–170 п.н. ДНК, но всегда теряется при последующем уменьшении до минимальной хромосомы размером 146 п.н. На основании этого факта можно предположить, что гистон Н1 располагается в области линкерной ДНК, непосредственно прилегающей к ДНК минимальной нуклеосомы.

ДНК закручена вокруг гистонового октамера

Два типа данных независимо говорят о том, что ДНК, очевидно, лежит на поверхности нуклеосомы, обвиваясь снаружи вокруг гистонового октамера. Согласно биофизическим данным, диаметр белкового компонента нуклеосомы меньше, чем диаметр витка ДНК. Биохимические данные показывают, что ДНК чувствительна к нуклеазам в участках, расположенных через определенные интервалы (см. ниже).

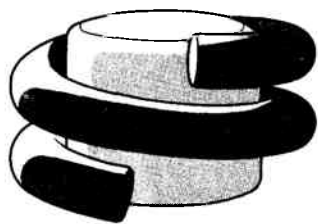


Рис. 29.7. Нуклеосому можно представить в виде цилиндра с двумя витками ДНК, закрученными снаружи вокруг него.

Нуклеосомы имеют в основном форму эллипса. Отношение осей равно 0,5 и соответствует размерам $11 \times 11 \times 6$ нм. По своей форме нуклеосома обычно представляет собой плоский цилиндр диаметром 11 нм и высотой 6 нм. Даже на основе измерений кажется вероятным, что ДНК лежит снаружи. Окружность частицы, составляющая примерно 34 нм, сравнима с длиной ДНК – 67 нм (200 п. н.). Трудно представить, каким образом ДНК могла бы втиснуться в такую частицу.

Свойства отдельных компонентов можно измерить с помощью рассеяния нейтронов. Этот метод позволяет различить рассеяние, обусловленное ДНК, от рассеяния, обусловленного белком. Измерения показали, что белковый компонент (гистоновый октамер) имеет радиус вращения около 3,2 нм, но радиус вращения ДНК-компонента составляет примерно 5,2 нм. Разница в 2 нм соответствует диаметру двойной спирали ДНК. На основе этих данных было высказано предположение, что белок организован в компактное тело, вокруг которого намотана ДНК. Существуют различные модели, описывающие способ укладки ДНК в нуклеосоме. Наиболее общими чертами всех моделей является то, что структура должна быть симметричной и что ДНК проходит вокруг октамера дважды. На рис. 29.7 схематически изображена ДНК, лежащая в виде двух витков спирали. Из этой схемы следует, что ДНК «входит» и «выходит» из нуклеосомы в точках, близко расположенных друг к другу. Однако до сих пор не уделяется достаточно внимания вопросу о том, с помощью какой модификации в укладке можно объяснить вариации длины ДНК в нуклеосоме.

Рассматривая модель, изображенную на рис. 29.8 в виде поперечного сечения, можно видеть, что два витка ДНК лежат близко один к другому. Это, возможно, имеет функциональное объяснение. Один виток вокруг нуклеосомы захватывает 80 п. н., так что точки, отстоящие на это расстояние в свободной двойной спирали, могут оказаться расположенными достаточно близко друг

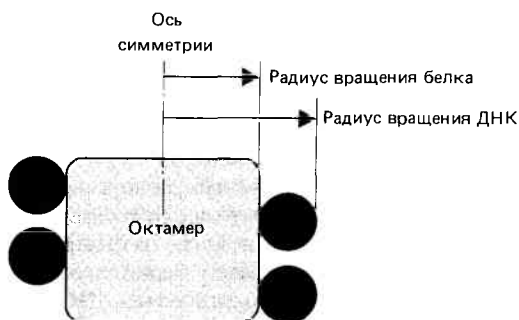


Рис. 29.8. Два витка ДНК в нуклеосоме должны лежать совсем близко друг к другу.

к другу на нуклеосоме. Таким образом, если ДНК-связывающий белок одновременно контактирует с двумя витками ДНК, как показано на рис. 29.9, узнаваемые им последовательности могут отстоять на двойной спирали ДНК гораздо дальше, чем величина соединяющего их участка в белке. Часто обсуждается вопрос о том, может ли стартовая точка транскрипции находиться поблизости от положения – 80 так чтобы обе цепи одновременно контактировали с РНК-полимеразой. (Это внесло бы конкретную деталь в обобщенную модель, показанную на рис. 11.8).

Плотность упаковки отдельной нуклеосомы равна примерно 6 (так как 67 нм ДНК упакованы в частицу длиной около 11 нм). Можно упаковать серию нуклеосом достаточно плотно, чтобы сохранить это отношение. Множество работ по изучению наборов нуклеосом было выполнено на вирусе SV40.

ДНК вируса SV40 представляет собой кольцевую молекулу в 5200 п. н., контурная длина которой равна примерно 1500 нм. И в состоянии вириона, и будучи инъецированной в ядро, она упакована в серию нуклеосом. В этой форме ее называют **мини-хромосомой**. При обычном выделении контурная длина мини-хромосомы равна примерно 210 нм, а плотность ее упаковки составляет примерно 7. Изменение в концентрации соли может превратить ее в гибкую нитку бус со значительно более низкой плотностью упаковки. Из этого следует, что нуклеосомные нити *in vitro* в зависимости от условий могут находиться в более чем одной форме.

Важным параметром в описании структуры хроматина является степень суперспирализации, которая может возникнуть на нескольких уровнях. Во-первых, суперспирализация может быть результатом упаковки ДНК на нуклеосоме. Во-вторых, суперспирализация может быть следствием укладки нуклеосом в структуру более высокого уровня. Состояние суперспирализации во многом может удерживаться белками (по принципу, описанному в гл. 28). Прямые измерения плотности суперспирализации позволяют узнать среднее напряжение скручивания ДНК, возникающее в результате образования свободных супервитков, но таким образом нельзя обнаружить суперспирализации, которая удерживается в ходе упаковки ДНК.

Для мини-хромосомы вируса SV40 можно прямо измерить степень суперспирализации в самой нуклеосоме. Мини-хромосома может иметь свободные супервитки в гирлянде нуклеосом, а также супервитки, удерживаемые на нуклеосоме. Процедура измерения суперспирализации, обусловленная только структурой нуклеосом, показана на рис. 29.10. Сначала освобождаются свободные супервитки самой мини-хромосомы, так что гирлянда нуклеосом образует кольцо с нулевой суперспирализацией. Затем экстрагируют гистоновые октамеры. В результате этой процедуры освобожденная ДНК свободно расправляется. Таким образом каждый супервиток, который содержится в мини-хромосоме, проявится в депротенинизированной ДНК как –1 оборот. Так можно измерить общее число супервитков в ДНК вируса SV40.

Действительно, наблюдаемое значение близко к числу нуклеосом. Противоположный результат наблюдают при сборке нуклеосом *in vitro* на суперспирализованной ДНК SV40: образование каждой нуклеосомы удаляет ~1 отрицательный супервиток. Таким образом, если удалить удерживающий белок, то ДНК, уложенная на поверхности нуклеосомы, образует ~1 отрицательный суперви-

Рис. 29.9. Через белок может осуществляться контакт последовательностей ДНК, находящихся в различных витках на нуклеосоме.



ток. Следует отметить, что существует потенциальное расхождение между измеренным значением, равным -1 оборот на нуклеосому, и моделью, согласно которой укладка ДНК в нуклеосоме эквивалентна -2 сверхспиральным оборотам (см. ниже).

ДНК, симметрично обработанная нуклеазами

Ферменты ДНКазы I и ДНКазы II вносят одноцепочечные разрывы в ДНК. Это означает, что они разрывают связь в одной цепи, тогда как вторая цепь остается в этом месте интактной. В результате цепь в целом кажется непрерывной, однако при денатурации вместо цепей обычной длины высвобождаются более короткие фрагменты (рис. 29.11).

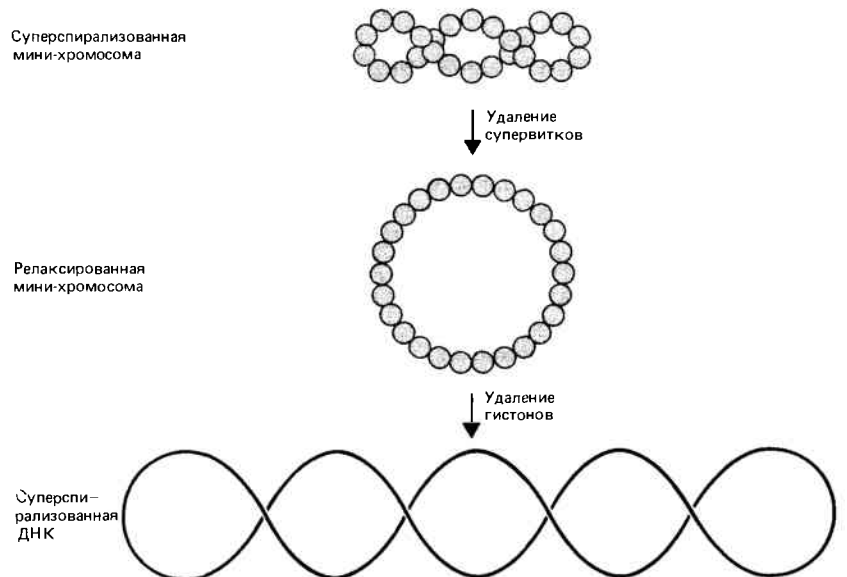
Если ДНК находится в растворе, одноцепочечные разрывы (ники) возникают случайно. В ДНК, находящейся на нуклеосомах, такие разрывы могут быть образованы с помощью ферментов, но только в определенных точках. Разрывы образуются через определенные промежутки так, что при электрофорезе денатурированной ДНК получается лестница. В экспериментах этого типа фрагменты денатурированной ДНК могут соответствовать расстоянию от конца нуклеосомы до места разрыва или могут быть образованы в результате двух внутренних разрывов. Настоящие точки разрыва можно установить, используя концевую радиоактивную метку с последующей радиоавтографической идентификацией фрагментов. Как показано на рис. 29.12, будут обнаружены только фрагменты с концевой меткой. (Этот метод аналогичен методу секвенирования, показанному на рис. 3.4, или методу отпечатков – рис. 11.4.)

Лестница, полученная в результате обработки ДНКазой I, показана на рис. 29.13. Похожие результаты получают при аналогичной обработке хроматина. Действительно, для таких экспериментов *in vitro* обычно используют препарат минимальных нуклеосом, так как трудно получить препараты полных нуклеосом с гомогенным распределением по размеру фрагментов ДНК.) Интервал между последовательными ступенями лестницы составляет примерно 10 оснований. Лестница вытягивается на всю длину ДНК минимальной нуклеосомы; сайты расщепления пронумерованы от S1 до S13 (где S1 расположен через ~ 10 оснований от меченого 5'-конца ДНК, S2 – через ~ 20 оснований и т.д.). Из этого следуют два важных вывода. Во-первых, ДНК не защищена на нуклеосоме и остается чувствительной к ДНКазе I; она не «покрыта» белками. Во-вторых, чувствительные участки расположены периодически.

Не все участки разрезаются с одинаковой частотой: некоторые разрезаются достаточно эффективно, а другие совсем редко. Ферменты ДНКазы I и ДНКазы II образуют одинаковую лестницу, хотя с некоторыми различиями в интенсивности полос. Это говорит о том, что картина разрезания обусловлена уникальной серией мишеней в ДНК, определяемых ее организацией. Отдельные ферменты имеют только слабое предпочтение к определенному сайту в этой мишени.

Поскольку в минимальной нуклеосоме находятся две цепи ДНК, метка вводится в оба 5'-конца (или 3'-конца), так что метится один конец каждой цепи. Поэтому при разрезании образуются фрагменты, относящиеся к обеим цепям. Это изображено на рис. 29.12, где каждый фрагмент происходит из разных цепей. В таком эксперименте каждая меченая полоса в действительности представляет

Рис. 29.10. Супервитки мини-хромосомы могут релаксировать и превратиться в кольцевую структуру, которая после удаления гистонов образует суперспирализованную свободную ДНК.



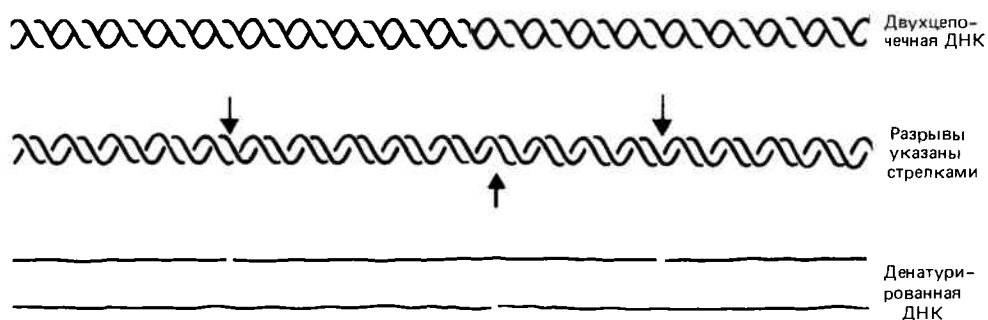


Рис. 29.11. Одноцепочечные разрывы (ники) в двухцепочечной ДНК обнаруживают по образованию фрагментов после денатурации ДНК и образования отдельных цепей.

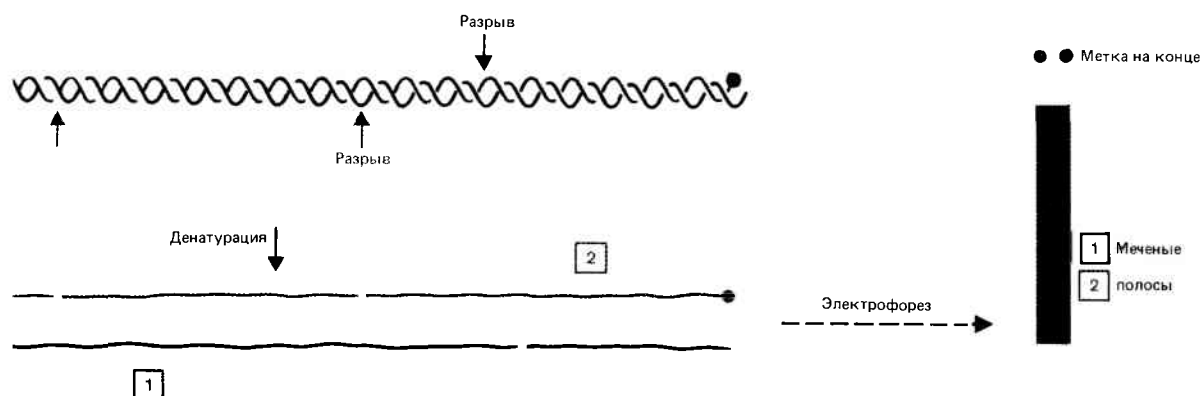


Рис. 29.12. С помощью метода концевого мечения определяют расстояние от одноцепочечных разрывов до меченых концов.

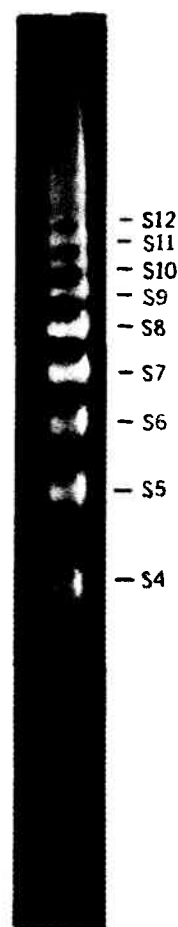


Рис. 29.13. На электрофореграмме ядерного гидролизата, образованного под действием ДНКазы I, видно, что сайты одноцепочечных разрывов располагаются на ДНК минимальной нуклеосомы через определенные интервалы. (Фотография любезно предоставлена Leonard Lutter.)

собой два фрагмента, отрезанные на *одном и том же* расстоянии от любого из меченых концов. Как же тогда можно объяснить дискретное предпочтение определенных сайтов? Обычная точка зрения заключается в том, что фермент действует симметрично на обе цепи. Таким образом, если, например после обработки ДНКазой I не образуется фрагмента в 80 п. н., это должно означать, что положение, соответствующее 80-му основанию с 5'-конца *любой* из цепей, нечувствительно к ферменту.

Какова природа участков-мишеней? Первый вопрос, который нужно решить, — это взаимоотношение между сайтами разрезания на двух цепях ДНК. Это удалось выяснить, используя условия, при которых фермент образует двухцепочечные разрезы. Исследуя концы образовавшихся двухцепочечных фрагментов, можно понять взаимоотношение между разрезами, сделанными на двух разных цепях. Результаты такого исследования суммированы на рис. 29.14.

Фермент ДНКазы I разрезает две цепи на расстоянии 2 п. н. Фермент ДНКазы II разрезает их на расстоянии 4 п. н. По отношению к точке симметрии разрез, сделанный ДНКазой I, отклоняется на одно основание в сторону 3'-конца, а ДНКазы II делает разрез, отступая на два основания в направлении 3'-конца. Нуклеаза микро-

кокков делает один разрез в обеих цепях, отступая на одно основание в направлении 5'-конца. Вероятнее всего, что все ферменты узнают одну и ту же точку, но разрезают ДНК на определенном расстоянии от этой точки, определяемом индивидуальной геометрией ферментов.

Наиболее существенный вывод, который можно сделать из этих результатов, состоит в следующем: сайт разрезания представляет собой короткий отрезок (порядка 3–4 пар оснований), в котором фосфодиэфирные связи в обеих цепях открыты для действия нуклеазы. Аналогичный вывод был сделан при детальном исследовании в гелях высокого разрешения фрагментов одной цепи с концевой меткой, полученных при разрезании минимальной нуклеосомы ДНКазой I или II. Пример такого эксперимента показан на рис. 29.15. В каждом сайте действительно существуют 3–4 положения, в которых может произойти разрез, т. е. сайт разрезания определяется с точностью ± 1 п. н. Относительная интенсивность разрезания указывает на то, что некоторые положения оказываются предпочтительнее других. На основе полученной картины можно подсчитать «среднюю» точку разрезания. При этом видно, что нуклеотидные пары сайтов от S1 до S4 лежат на расстоянии 10,0 п. н. друг от друга, сайты от S4 до S10 разделены на 10,7 п. н., а для сайтов

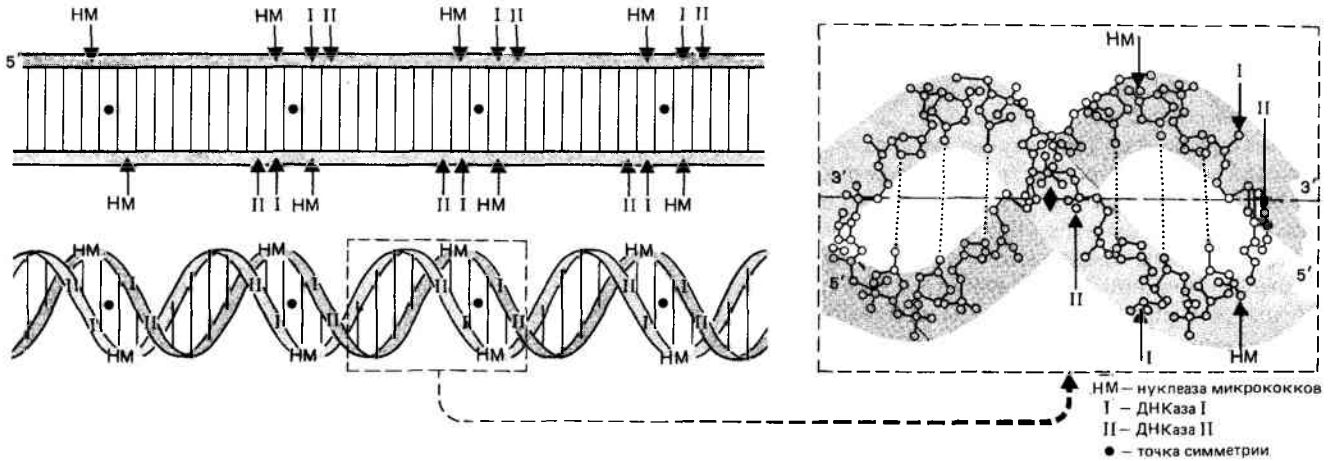


Рис. 29.14. Каждая нуклеаза разрезает ДНК на определенном расстоянии между сайтами на обеих цепях.

от S10 до S13 расстояние снова становится равным 10,0 п.н. (Поскольку мы анализируем *среднее* положение, сайты оказываются разделены не целым числом оснований.)

Нерешенный вопрос о периодичности ДНК

Эксперименты, в которых анализируют ДНК на нуклеосомах по ее чувствительности к нуклеазам, проводят по методу, близкому к методу отпечатков (footprint). Таким образом, мы можем связывать утрату реакции в определенном сайте-мишени с такой структурой нуклеосомы, в которой данные положения на ДНК стали нечувствительными. Но какова причина периодического разрезания через 10,7 п.н.?

Эта периодичность близка к периодичности В-формы ДНК. Означает ли это, что повторяемость чувствительного сайта просто отражает иммобилизацию ДНК на гистоновом октамере? Это может происходить так, как изображено на **рис. 29.16**, где пик чувствительных сайтов встречается в каждой цепи с периодичностью, определяемой числом пар оснований на виток двойной спирали.

Некоторый свет на этот вопрос может пролить сравнительный анализ ДНК, иммобилизованной на плоской поверхности (ситуация, рассмотренная на рис. 29.16). В этом случае сайты нарезаются на расстоянии ровно 10,6 п.н. Это, вероятно, означает, что свободная ДНК (т.е. выделенная в растворе), имеет двуспиральную структуру В-типа, у которой повторяющаяся длина равна 10,6 п.н. С помощью измерения числа пар на виток, выполненных другим методом, получено очень близкое значение – 10,4 п.н.

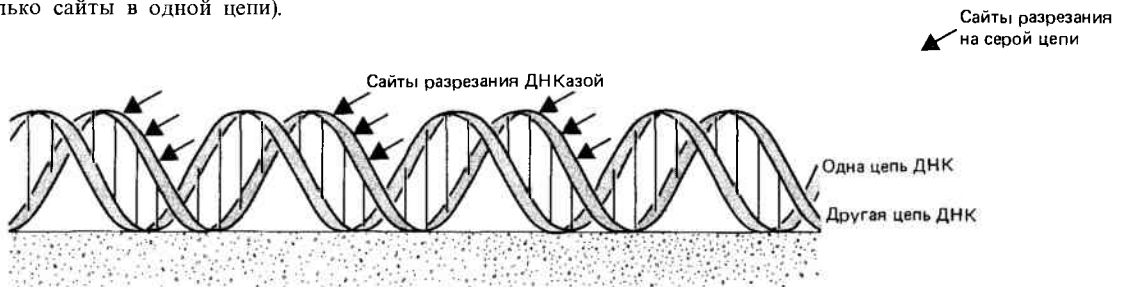
Мы считаем, что при иммобилизации ДНК на плоской поверхности **периодичность разрезания** (расстояние между точками расщепления) действительно соответствует **структурной периодичности** (число пар на виток двойной спирали). Вариации в периодичности разрезания ДНК минимальной нуклеосомы (10,0 на концах и 10,7 посередине) затрудняют интерпретацию этого явления.

Одно из предположений заключается в том, что структурная периодичность ДНК минимальной нуклеосомы действительно варьирует. В средней части она более близка к В-форме, а на концах может быть закручена слабее. Другая возможность состоит в том, что сама ДНК имеет постоянную периодичность, но геометрия нуклеосомы вызывает смещение мест разрезания в средней части. Если в некоторых положениях доступ ферменту



Рис. 29.15. Высокорастворяющий анализ показывает, что каждый сайт для ДНКазы I состоит из нескольких соседних чувствительных фосфодиэфирных связей, как это видно на представленном примере сайтов S4 и S5 в ДНК частиц минимальной нуклеосомы, исследованных с введением концевой метки. (Фотография любезно предоставлена Leonard Lutter.)

Рис. 29.16. Большинство открытых положений в ДНК повторяется с периодичностью, отражающей структуру двойной спирали (для четкости показаны только сайты в одной цепи).



ограничен, он может изменить свой выбор среди 3–4 доступных связей. В этом случае периодичность разрезания не будет соответствовать структурной периодичности.

Если это так, то такое структурная периодичность? При отсутствии необходимых данных нужно построить модель, для того чтобы увидеть, какое из предположений может обеспечить доступ фермента к ДНК. Идеи о меняющейся структурной периодичности или о фиксации расстояния в 10,6 п.н. оказываются менее подходящими, чем предположение о существовании в нуклеосоме постоянной периодичности с 10,0 п.н. на виток.

Некоторые интересные выводы вытекают из данных, дающих основание думать, что структурная периодичность ДНК в нуклеосоме (10,0) и ДНК в растворе (10,6) может быть различной. При освобождении ДНК из нуклеосомы она должна становиться более скрученной, так как у нее больше пар оснований на виток. Это изменение уменьшит степень ее суперспирализации. Предположим, что ДНК проходит в нуклеосоме путь, равный двум оборотам суперспирали. Затем удалим гистоновый октамер. Некоторое напряжение скручивания приведет к большему закручиванию ДНК, и только остаточное напряжение должно измеряться как суперспиральное. Это один из возможных способов согласовать модели для -2 суперспиральных витков на нуклеосоме с данными, в которых определен только -1 супервиток. Поскольку различие в периодичности (0,6 на оборот спирали), умноженное на число супервитков, приходящихся на нуклеосому (> 15), примерно равно 10 п.н. и соответствует одному обороту двойной спирали, то таким образом есть потенциальная возможность поглотить 1 отрицательный супервиток.

И еще одно важное замечание нужно сделать относительно укладки ДНК в нуклеосоме и ее возможной степени суперспирализации. Все обсуждавшиеся данные были получены с использованием препаратов минимальной нуклеосомы. В моделях, объясняющих противоречивость данных о числе супервитков, предполагается, что линкерная ДНК уложена таким же способом, что и ДНК минимальной нуклеосомы. Однако способ укладки линкерной ДНК может быть иным. Анализ «лестниц», полученных при расщеплении ДНКазой I фрагментов длиной более 1000 п.н., свидетельствует в пользу того, что выходящая из кора ДНК уложена так же, как и ДНК кора. Это означает, что одна и та же периодичность разрезания сохраняется не только в отдельной нуклеосоме, но и между нуклеосомами. Другими словами, ДНК следует от одной нуклеосомы к другой, не нарушая способа организации, по крайней мере насколько это видно из периодичности разрезания.

Организация гистонов и ДНК

До сих пор мы рассматривали конструкцию нуклеосомы, исходя из того, как организованы повторяющиеся единицы длины ДНК на поверхности нуклеосом. Если же исходить из организации белка в нуклеосоме, то нужно знать, каким образом гистоны взаимодействуют друг с другом и с ДНК. Взаимодействуют ли гистоны только в присутствии ДНК или способны независимо собираться в октамеры?

Мы еще многого не знаем о структуре индивидуальных гистонов в нуклеосоме, но уже стала проясняться их относительная локализация. Большинство данных о взаимодействиях гистонов получено при изуче-

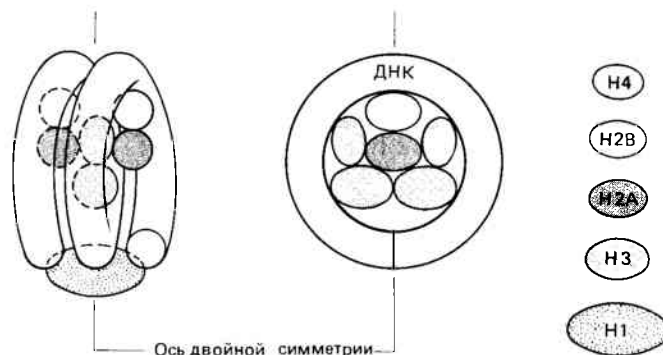


Рис. 29.17. Модель нуклеосомы спереди и сверху. Видна симметричность конструкции с сердцевинной из тетрамера $H_3H_4_2$.

нии способности к агрегированию и к образованию перекрестных сшивок в нуклеосоме.

Способность гистонов агрегировать друг с другом изучена хорошо. Гистоны, богатые аргинином, можно получать в виде хорошо различимого тетрамера ($H_3H_4_2$). Гистоны, умеренно богатые лизином, образуют продукт, который менее четко охарактеризован, но который, очевидно, образует димер ($H_2A \cdot H_2B$), имеющий тенденцию к дальнейшей агрегации. Одной из форм образующегося агрегата может быть тетрамер ($H_2A_2 \cdot H_2B_2$). Два указанных тетрамера называют **гомотиписческими** (название отражает то, что каждый из них содержит только гистоны, относящиеся к одному из исходных классов).

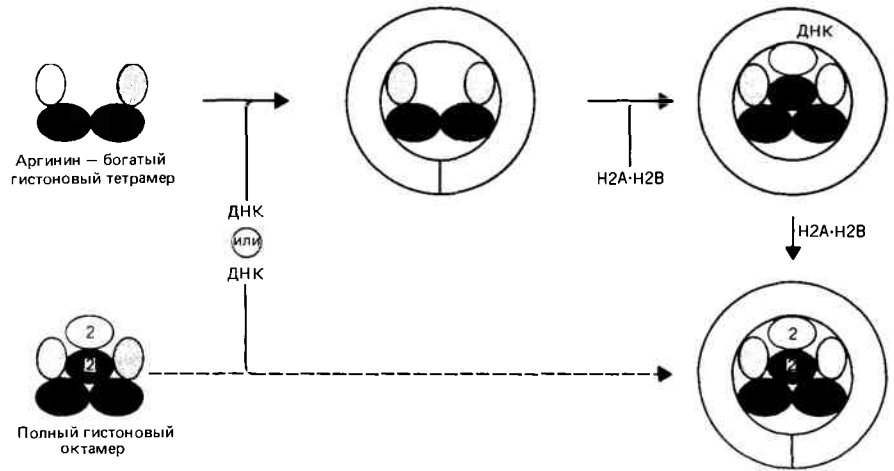
Интактные октамеры гистонов можно получить либо экстракцией хроматина, либо (с большими трудностями) путем соединения гистонов *in vitro* в условиях высокой концентрации соли и белка. Октамер может диссоциировать с образованием гексамера гистонов и свободного димера $H_2A \cdot H_2B$. В результате отделения второго димера $H_2A \cdot H_2B$ образуется тетрамер $H_3H_4_2$. Все эти формы можно экстрагировать из хроматина. Исходя из этих данных, можно сделать вывод, что в нуклеосоме имеется центральное «ядро» (kernel), состоящее из тетрамера $H_3H_4_2$, связанного с двумя независимыми димерами $H_2A \cdot H_2B$.

Перекрестные шивки позволяют установить, какие пары гистонов лежат ближе друг к другу в нуклеосоме. (Сложность в интерпретации этих данных заключается в том, что сшивается только небольшая часть гистонов. Поэтому следует оценивать результаты с осторожностью, учитывая типичность большинства взаимодействий.) На основе полученных данных была сконструирована модель организации нуклеосомы. В достаточно схематичной форме она изображена на рис. 29.17.

Структурные исследования показывают, что по своей общей форме изолированный гистоновый октамер похож на минимальную нуклеосому. Из этого следует, что общая структура зависит от гистон-гистоновых взаимодействий. Положение индивидуальных гистонов в структуре октамера было установлено на основе их способности к агрегированию и к образованию поперечных сшивок.

Диаметр всего октамера определяется размером тетрамера $H_3H_4_2$. Этой особенностью объясняется способность тетрамера *in vitro* организовывать ДНК в частицы, проявляющие некоторые свойства минимальной нуклеосомы (см. ниже). Белки располагаются в тетрамере

Рис. 29.18. В определенных условиях *in vitro* ДНК может связываться с тетрамером $H3_2H4_2$, к которому присоединяются два димера $H2A \cdot H2B$. В других условиях ДНК может прямо взаимодействовать с интактным гистоновым октамером (содержащим перекрестные сшивки).



так, что их порядок отвечает наблюдаемым поперечным сшивкам: $H3-H4$ и $H3-H4$. Пары $H2A \cdot H2B$ присоединяются в виде двух димеров. ДНК дважды обматывается вокруг октамера. Гистон $H1$ может «запечатывать» ДНК в нуклеосоме, связывая ее в точке, где сближаются входящая и выходящая нити. Этим объясняется способность гистона $H1$ образовывать поперечные связи с гистонами $H3$. Описанная модель имеет симметрию второго порядка.

Несмотря на то что общая форма октамера теперь определена достаточно точно (хотя на рисунке этого не видно), индивидуальные гистоны изображены на рисунке в виде аморфных шариков, поскольку мы не располагаем данными об их структуре. Распределение аминокислот на N-конце, несущем большой заряд, одинаково у всех гистонов. Остальная часть молекулы содержит гидрофобные аминокислоты, которые, вероятно, образуют глобулярную структуру и участвуют в белок-белковых взаимодействиях. По этой причине гистоны иногда воспринимаются как глобулярные белки с заряженными N-концевыми «хвостами». Можно было бы думать, что у «хвостов» преобладает ДНК-связывающая активность, тогда как глобулярные области входят внутрь сердцевинки. Однако против этой модели свидетельствуют данные о том, что N-концевые области можно отщепить (обработав трипсином) от гистонов сердцевинки, не вызывая при этом сколько-нибудь существенных нарушений структуры нуклеосомы. Кроме того, гистоны без N-концевых «хвостов» могут участвовать в сборке нуклеосомы *in vitro*. В настоящий момент мы не можем приписать индивидуальных функций определенным участкам гистоновых молекул.

Сборка нуклеосом и репродукция хроматина

Гистоны сердцевинки обладают способностью накручивать ДНК вокруг белкового октамера. Однако существует несколько способов, посредством которых ДНК может связаться с гистонами.

Первоначально реконструкцию нуклеосом *in vitro* можно было осуществить только в условиях, далеких от физиологических; отдельные компоненты смешивали при высокой концентрации солей и мочевины, которые потом удаляли с помощью диализа. Но теперь реакцию частично контролируют, регулируя последовательность добавления гистонов и ДНК и их относительную концентрацию. В этом случае самосборка происходит в условиях,

близких к физиологическим. Процесс этот, медленный сам по себе, еще и ограничивается тенденцией собирающихся частиц к преципитации. Два возможных способа сборки нуклеосом показаны на рис. 29.18.

Один из них обусловлен способностью тетрамера $H3_2H4_2$ организовывать ДНК в частицы, которые несколько напоминают минимальную нуклеосому (по их чувствительности к нуклеазе микрококков). При добавлении димера $H2A \cdot H2B$ эти тельца могут превращаться в минимальную нуклеосому. Именно отсюда возникла идея о том, что в структуре нуклеосомы существует «ядро» из аргинин-богатых гистонов. Возможно, такой путь используется *in vivo*, поскольку это согласуется с наблюдением, что гистоны $H3$ и $H4$ включаются в реплика-

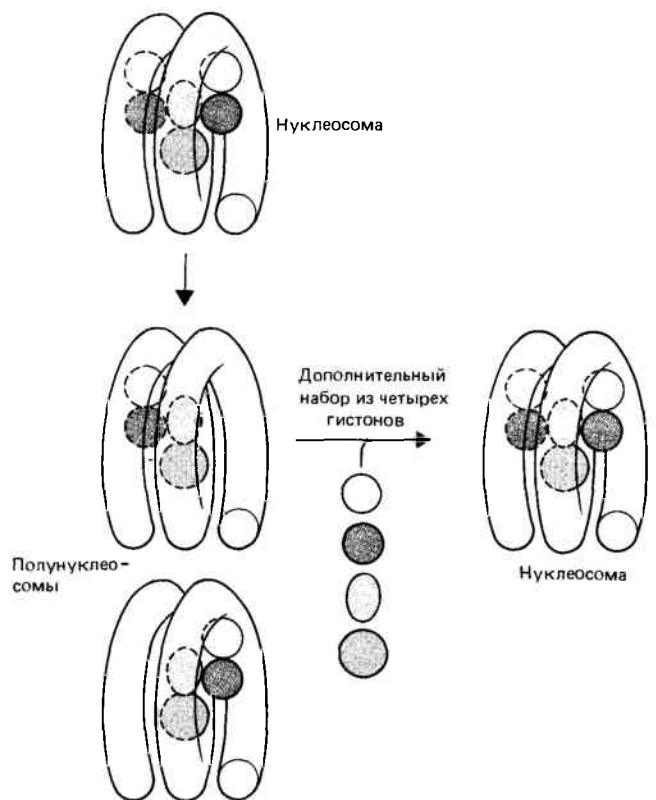
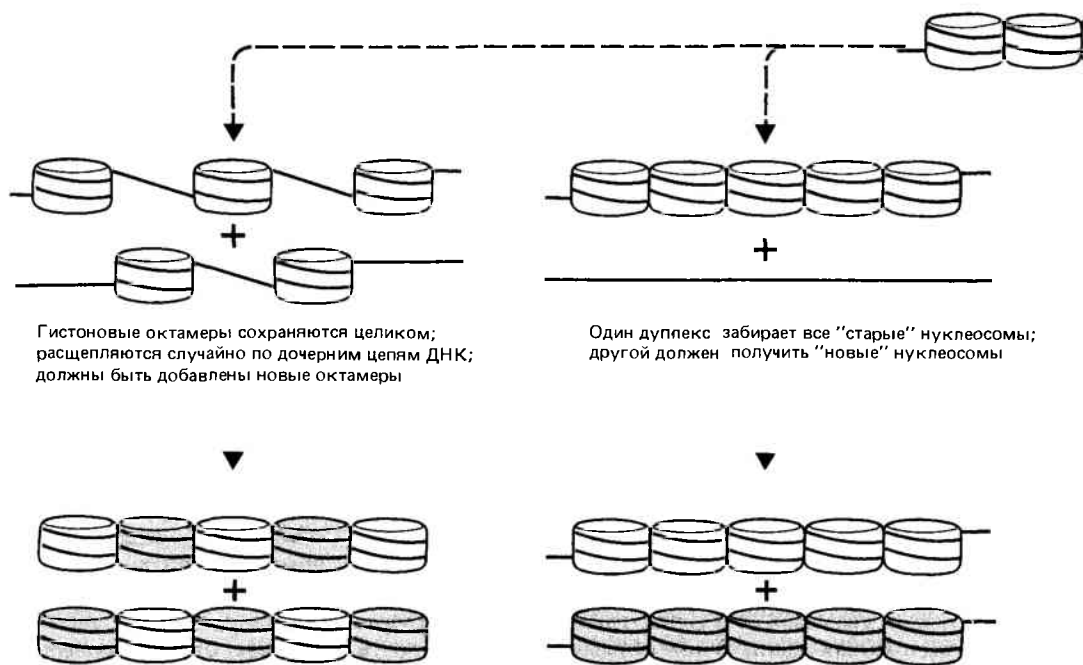


Рис. 29.19. Может ли нуклеосома разделяться на две полунуклеосомы (гетеротипические тетрамеры), которые могут стать полными нуклеосомами путем добавления второго гетеротипического тетрамера гистонов?

Рис. 29.20. Существует много возможных вариантов поведения нуклеосом при репликации ДНК.



рующийся хроматин, до того как в его состав входят гистоны H2A и H2B.

Другие данные были получены при использовании поперечно-сшитых гистоновых октамеров, которые не могут диссоциировать на отдельные белки, но тем не менее сохраняют способность связывать ДНК с образованием минимальной нуклеосомы. Это говорит о том, что в принципе ДНК может обвиваться вокруг предварительно сформированного октамера.

Далее было сделано допущение, что благодаря своей симметричной структуре нуклеосома, возможно, разделяется на две «полунуклеосомы». С этой точкой зрения согласовывалось наблюдение о том, что мини-хромосома вируса SV40 способна изменять свою структуру при низкой концентрации соли. При этом образуется вдвое больше бусинок, каждая из которых меньше по размеру, чем нуклеосома. Могут ли они быть полунуклеосомами, состоящими из **гетеротипичного тетрамера** (H2A·H2B·H3·H4)? Могут ли такие промежуточные формы использоваться при репликации хроматина, соединяясь с другим набором из четырех гистонов для восстановления полной нуклеосомы, как это гипотетически представлено на **рис. 29.19**?

При попытке собрать нуклеосомы *in vitro* процесс сборки в основном рассматривают как соединение свободной ДНК с гистонами. Но в действительности *in vivo* хроматин репродуцируется. Отрезок ДНК, уже связанный с нуклеосомами, реплицируется, давая начало двум дочерним дуплексам. Что происходит в этот момент с предсуществующими нуклеосомами? Диссоциируют ли гистоновые октамеры на свободные гистоны, которые затем вновь собираются, или же они остаются в собранном виде? Некоторые такие возможности показаны на **рис. 29.20**.

Эксперименты с использованием поперечных сшивок говорят о том, что гистоновый октамер может быть **консервативным**, сохраняясь целиком на протяжении всего цикла репликации. Остаются ли «старые» октамеры связанными в каком-либо определенном порядке с дублированными ДНК? Например, все ли старые октамеры

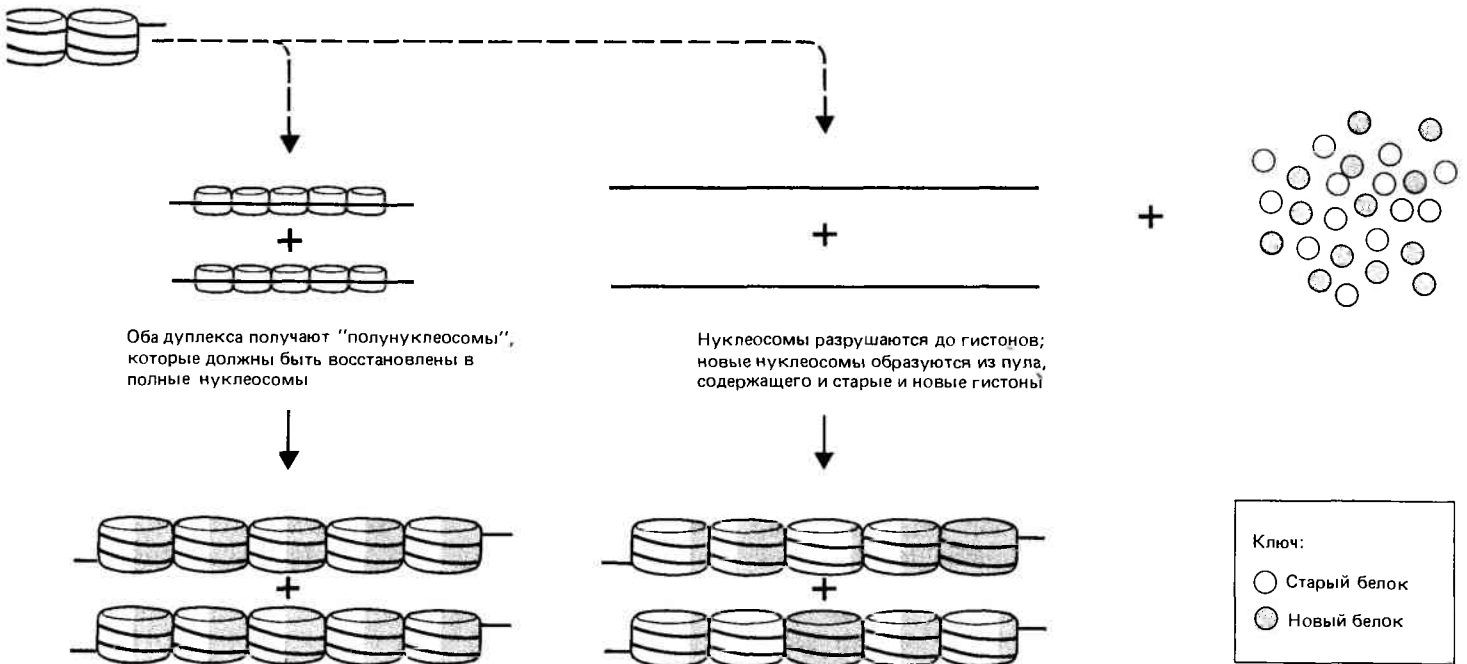
остаются на одном дочернем дуплексе, тогда как «новые» октамеры собираются на другом. Однако, если предотвращена сборка новых гистонов (например, путем подавления синтеза белка циклогексимидом), «старые» октамеры разделяются между обоими дочерними дуплексами.

Таким образом нуклеосомы могут образовываться двумя способами. При репликации хроматина уже находящиеся на нем гистоновые октамеры удаляются с ДНК, что делает возможным репликацию. Эти октамеры сохраняются и могут реассоциировать с любым из дочерних дуплексов. Однако такое же число октамеров должно быть образовано из новосинтезированных гистонов. Происходит ли сборка этих октамеров также раньше, чем они связываются с ДНК или же наоборот, в данном случае они собираются на ДНК? Или, может быть, октамеры распределяются случайно между дочерними дуплексами? Эксперименты, поставленные для выяснения этого вопроса, не дали однозначного ответа, поскольку трудно отличить заново реплицированный материал от массы предсуществовавшего хроматина.

Очевидно, что в процессе редупликации хроматина ДНК не остается свободной от нуклеосом в течение сколько-нибудь продолжительного периода. Как только ДНК реплицируется, нуклеосомы сразу образуются на обеих репликах независимо от того, каков точный механизм процесса. Это проиллюстрировано на электронной микрофотографии (**рис. 29.21**), на которой показано, что у только что реплицированного отрезка ДНК оба дочерних дуплекса уже усыпаны нуклеосомами.

Структура участка, в котором непосредственно происходит репликация ДНК, отличается от других областей. Этот участок более устойчив к нуклеазе микрококков и при ферментативном расщеплении образует полосы, отличающиеся по размеру от нуклеосомной ДНК. Следовательно, можно предположить, что в репликации ДНК участвует достаточно большой белковый комплекс, причем почти сразу после того, как он передвигается дальше, восстанавливается нуклеосомная структура.

В состав хроматина входят как нуклеосомные ги-



стоны, так и негистоновые белки, и все эти дополнительные компоненты должны репродуцироваться. Поскольку комплект негистоновых белков, по-видимому, варьирует в зависимости от конкретного фенотипа клетки, при его репликации сохраняются особенности клеточной специфичности. Таким образом, возможность существования механизма сегрегации белков в процессе репликации ДНК имеет значение, выходящее за рамки вопроса о сборке нуклеосом. Одним из наиболее принципиальных вопросов, на который хотелось бы ответить, является вопрос о том, каким образом различные состояния структуры хроматина наследуются дочерними клетками.

Рассмотрим ген, который активирован (или репрессирован) путем связывания с ДНК какого-то специфического регуляторного белка и (или) каким-то изменением структуры хроматина. Каким путем это конкретное состояние будет унаследовано дублированными хромосомами дочерних клеток, образовавшихся в результате деления? Если во время репликации все белки отделяются от ДНК, специфическое состояние должно заново устанавливаться в каждом цикле клетки. Однако возможно, что определенный механизм сегрегации используется для того, чтобы передать информацию о состоянии экспрессии генов. Одна возможность заключается в том, что специфическая структура может быть *увечовечена* путем сегрегации и дубликации в процессе репликации ДНК. Например, образец, формально эквивалентный полунуклеосомной сегрегации, показан на рис. 29.20 (безотносительно к тому, используется ли такой тип сегрегации самими гистонами). Таким образом, комплекс негистоновых белков может сформироваться на ДНК, затем расщепиться на полуккомплексы при репликации и вновь достроиться до полных комплексов на каждом дочернем дуплексе

Общая структура хроматиновой нити может играть важную роль в сборке хроматина. Если в ооцит инъектировать кольцевую ДНК, в нем происходит сборка первых нуклеосом и затем нить становится суперспирализованной. Эти свойства можно воспроизвести в системе *in vitro*, содержащей АТФ в качестве источника энергии, не-

обходимой для суперспирализации. При такой «активной сборке» образуются нуклеосомы с периодичностью 200 п.н. Выделив эту систему в очищенном виде, мы можем определить взаимодействия, участвующие в определении длины ДНК, соответствующей каждой нуклеосоме.

Другая сторона этой проблемы заключается в том, каким образом происходит *изменение* типа геной экспрессии. Чаще всего возможность изменений связывают с процессом репликации ДНК, во время которого временно нарушается структура хроматина. Однажды случившееся изменение может наследоваться с помощью механизма сегрегации до тех пор, пока не произойдет другого изменения (см. также гл. 30).

В хроматине, образованном в ооците путем активной сборки, содержится материал двух типов, различающийся по чувствительности к нарушению суперспирализации. Большая часть хроматина содержит закрепленные супервитки. Они остаются после введения одноцепочечных разрывов и снимаются только после удаления гистонов. Однако часть хроматина (до ~30%) находится в напря-

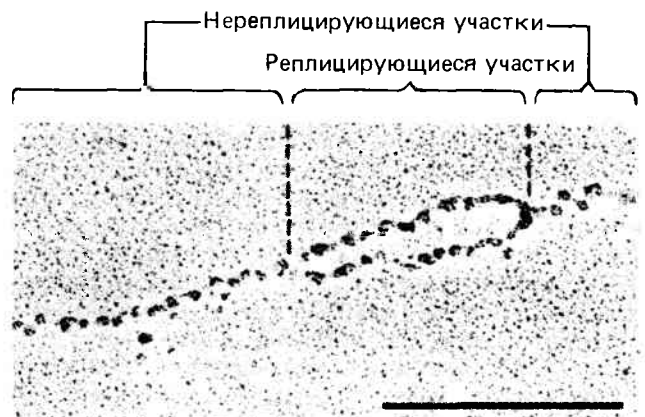


Рис. 29.21. Реплицированные участки хроматина содержат нуклеосомы в обоих дочерних дуплексах ДНК. (Фотография любезно предоставлена Steven L. McKnight.)

женном состоянии. В нем есть супервитки, которые можно снять, вводя одноцепочечные разрывы. В этой фракции содержатся транскрибирующиеся последовательно-сти. Суперспирализованная область может совпадать с доменами, ранее обнаруженными в целых геномах (см. гл. 28). Пока мы еще не знаем, чем обусловлены различия между двумя этими типами хроматина.

Для сборки нуклеосом нужны негистоновые белки

В результате экспериментов по реконструкции ДНК с гистонами образуется структура, более напоминающая частицу минимальной нуклеосомы, чем полную нуклеосому. При низких концентрациях, когда отдельные сформированные частицы соединены нитью ДНК, более 146 п.н. связываются с белковым октамером. Но это не постоянная характеристика, поскольку при увеличении концентрации, приводящей к более плотной упаковке частиц, каждый октамер оказывается связанным только со 146 п.н. ДНК, т.е. отрезком ДНК, равным по размеру ДНК минимальной нуклеосомы. Это выглядит так, как будто в отсутствие соседней минимальной нуклеосомы гистоновый октамер старается захватить больше ДНК, но при появлении ближайшего соседа отдает всю ДНК, за исключением той ее части, которая входит в состав собственной минимальной нуклеосомы.

Что же отвечает за модуляцию этой реакции *in vivo*, влияющей на расположение нуклеосом с интервалом, специфическим для соответствующего генома или даже для отдельных его частей?

Единственной системой, в которой удалось получить аутентичную плотность расположения нуклеосом в опытах по реконструкции, оказалась система ооцитов *Xenopus*. При инъекции ДНК вируса SV40 в ооциты кольцевые молекулы могут формировать мини-хромосомы. При достаточном избытке ДНК пул эндогенных гистонов истощается и сборка начинает зависеть от введения дополнительных гистонов. Характерные признаки системы сохраняются в бесклеточном экстракте, где сборка нуклеосом на свободной ДНК происходит с интервалом в 195 п.н. Это важная информация, так как она показывает, что правильная сборка нуклеосом может происходить *de novo* со свободной ДНК. Сборка не связана непременно с актом репликации; она не зависит от последовательности добавленной ДНК.

Эти признаки еще не были воспроизведены полностью во фракционированной системе реконструкции, но некоторый успех в идентификации отдельных компонентов был достигнут. Из ооцитов *Xenopus* выделен белок сборки. Это пентамер, содержащий идентичные субъединицы по 29000 дальтон. Это белок, обладающий в ооцитах, и локализован он в нуклеоплазме. Анти-тела, полученные против этого белка, реагируют с белками нуклеоплазмы многих эукариот. Следовательно, можно предположить, что этот белок соответствует эволюционно какой-то универсальной функции, закрепленной в процессе эволюции. Он был назван **нуклеоплазмином**.

В присутствии нуклеоплазмина гистоны могут связываться с ДНК, образуя в физиологических условиях (низкая концентрация соли) частицы. При расщеплении этих частиц нуклеазой микрококков образуются полосы ДНК размером в 146 и 165 п.н. Отсюда следует, что некоторые из образованных минимальных нуклеосом содержат дополнительную ДНК, протяженность которой недостаточ-

на для образования полных нуклеосом. Следовательно, нуклеоплазмин участвует в контроле реакции ДНК с гистонами таким образом, что этот процесс становится продуктивным и ведет к образованию нуклеосом, а не случайных агрегатов. Однако самого по себе нуклеоплазмينا недостаточно для образования нуклеосом со специфическим интервалом.

Какова функция нуклеоплазмина? Это кислый белок, который не связывается ни со свободной ДНК, ни с интактными нуклеосомами, но при этом он связывается со всеми индивидуальными гистонами. Реакция насыщается на уровне, равном одному пентамеру нуклеоплазмина на октамер гистона. Нуклеоплазмин, возможно, играет роль «молекулярного сопровождающего», связываясь с гистонами и передавая их ДНК более регулируемым образом, чем было бы возможно без такого конкурента. В пользу этого предположения говорит тот факт, что кислая полиглутаминовая кислота, а также РНК могут действовать сходным образом в качестве факторов сборки.

Общая особенность всех этих факторов состоит в том, что все они способны связываться с гистонами, уменьшая суммарный положительный заряд. Использование высокой концентрации соли для сборки гистонового октамера *in vitro* имитирует эту ситуацию. В этой связи следует упомянуть также прежнюю идею о том, что модификация заряженных групп гистонов может быть использована для регуляции сродства данного белка к ДНК (гл. 30). В результате таких взаимодействий гистоны могут образовывать термодинамически более стабильные агрегаты, минуя этап кинетических промежуточных продуктов (т.е. других комплексов, возникающих в результате высокого сродства гистонов к ДНК).

Из этого следует, что в процессе сборки нуклеосомы должны выполняться по крайней мере две функции, не свойственные самим компонентам нуклеосом. Первая из них касается контроля сродства гистонов к ДНК; это функция нуклеоплазмина. Второй связан с установлением длины ДНК, которая содержится в нуклеосоме; эту функцию еще предстоит выяснить.

Нуклеосомы в нитях хроматина

При исследовании хроматина под электронным микроскопом можно видеть два типа нитей: нити размером 10 нм и нити размером 30 нм. Это примерный диаметр нити (диаметр нити в 30 нм в действительности варьирует от 25 нм до 30 нм).

Нити размером 10 нм — это в основном непрерывный ряд нуклеосом. В действительности иногда они переходят в более вытянутые отрезки, на которых нуклеосомы имеют вид бусинок, нанизанных на нитку, как это показано на **рис. 29.22**. Нити размером 10 нм образуются в условиях низкой ионной силы в отсутствие гистона H1. Это означает, что образование данной структуры — функция самих нуклеосом.

Как расположены нуклеосомы в нитях размером в 10 нм? Рассматривая отдельную частицу как некий плоский цилиндр, можно расположить соседние цилиндры либо бок о бок, либо торцами друг к другу. Эти два способа расположения можно различить с помощью биофизических методов, таких, как рассеяние нейтронов или электрический дихроизм, который определяет ориентацию индивидуальных субъединиц относительно оси. На основе полученных результатов была предложена модель, изображенная на **рис. 29.23**. Согласно этой модели,

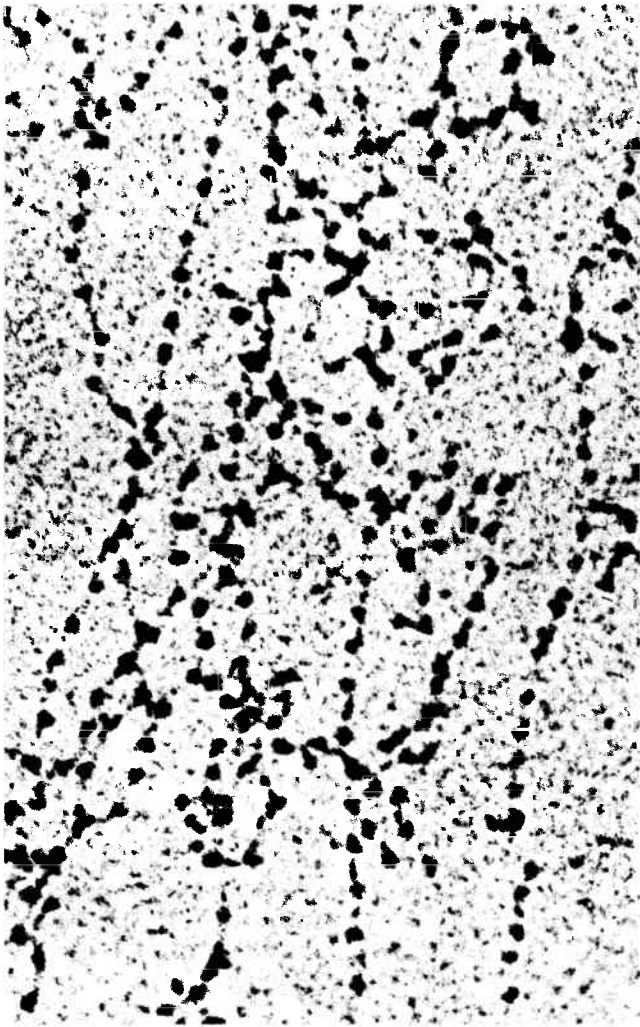


Рис. 29.22. Нить размером 10 нм в частично раскрученном состоянии представляет собой цепочку нуклеосом.
(Фотография любезно предоставлена Barbara Hamkalo.)

цилиндры ориентированы бок о бок друг к другу, а их торцы расположены параллельно (по крайней мере не сильно отклоняются) к оси фибриллы. Полученные данные подразумевают, что максимальный угол между торцами и осью может быть около 20° .

Когда хроматин исследуют в условиях высокой ионной силы и в присутствии гистона H1, получают нити размером 30 нм. Такой пример показан на рис. 29.24. Видно, что фибриллы имеют спиральную структуру. Примерно 6 нуклеосом содержится в каждом витке, плотность упаковки которого равна 40 (т.е. на каждый 1 мкм оси нити приходится 40 мкм ДНК). Такие нити служат основным компонентом интерфазного хроматина и митотических хромосом (см., например, рис. 28.7).

Нити размером 30 и 10 нм могут обратимо превращаться друг в друга при изменении ионной силы. Отсюда, очевидно, следует, что при высокой ионной силе и в присутствии гистона H1 линейный ряд нуклеосом нити размером 10 нм закручивается в структуру толщиной 30 нм.

С другой стороны, существует несколько возможных способов упаковки нуклеосомы в нити. Наиболее вероятный способ описывается радиальной моделью, пока-

занной на рис. 29.25. В этой модели нуклеосомы повернуты в спирально закрученном ряду, так что угол между торцами соседних нуклеосом составляет примерно 60° .

Похоже на то, что параметры нити размером 30 нм не жестко фиксированы и могут изменяться. Это сделало бы возможным изменение длины ДНК на нуклеосому, а также другие изменения в плотности упаковки. Пока не известно, имеют ли нити одинаковой толщины одинаковую структуру в интерфазном хроматине и в митотических хромосомах.

Хотя известно, что для образования фибрилл размером 30 нм необходимо присутствие гистона H1, данные о его локализации противоречивы. Поскольку он относительно легко экстрагируется из хроматина, можно думать, что он расположен снаружи суперспиральной оси нити. Однако другие данные свидетельствуют о его внутренней локализации, поскольку он труднее обнаруживается в нитях размером 30 нм, чем в тех отдельных нитях размером 10 нм, в которых он не утрачен.

Как перейти от нити размером 30 нм к специфическим структурам митотических хромосом? Существует ли еще какая-нибудь специфичность в расположении интерфазного хроматина? Фиксировано ли расположение бусинок в отдельных участках нитей размером 30 нм, или они располагаются случайно? Пока на эти вопросы у нас нет ответа.

Петли, домены и остов

Интерфазный хроматин кажется запутанной массой, занимающей большую часть объема ядра, в противоположность высокоорганизованной и воспроизводимой ультраструктуре митотических хромосом. Что контролирует распределение интерфазного хроматина внутри ядра?

Некоторые косвенные данные об организации интерфазного хроматина были получены при выделении гено-

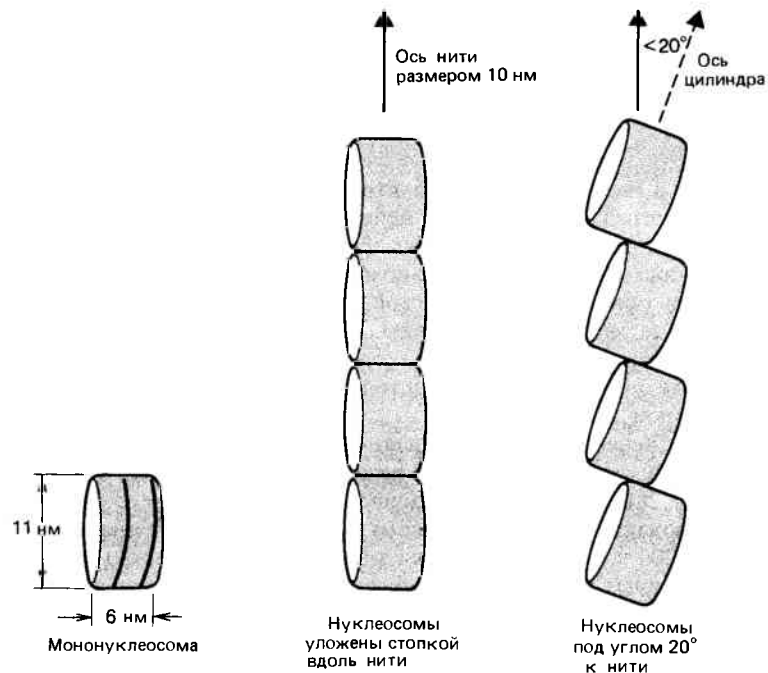


Рис. 29.23. Нить размером 10 нм состоит из серии нуклеосом, расположенных бок о бок стопкой вдоль нити или с наклоном менее 20° к оси.

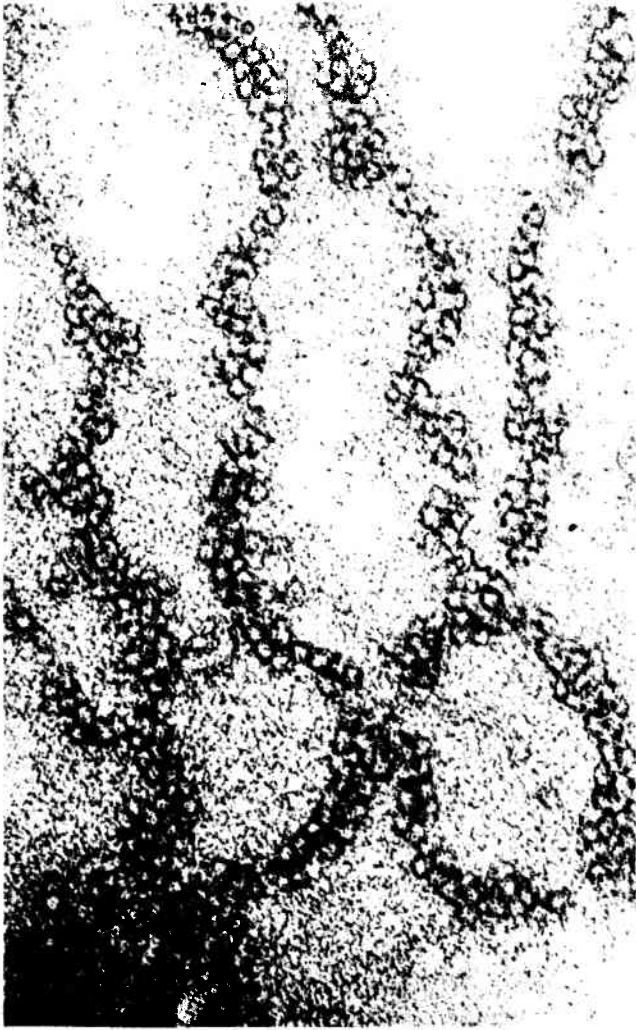


Рис. 29.24. Нить размером 30 нм имеет спиральную структуру. Эта нить показана при том же увеличении, что и нить размером 10 нм на рис. 29.22.

(Фотография любезно предоставлена Barbara Hamkalo.)

ма в виде единого компактного тела. С помощью того же метода, который был использован для выделения бактериального нуклеоида (гл. 28), можно лизировать ядра, нанося их сверху на сахарозный градиент. При этом геном освобождается в такой форме, что его можно собрать центрифугированием. Геном, выделенный из *D. melanogaster*, представляет собой компактно скрученные нити размером 10 нм, состоящие из ДНК и четырех гистонов сердцевины.

Степень суперспирализации этого компактного тела можно измерить по его реакции на бромистый этидий (гл. 28). Полученные данные свидетельствуют о существовании примерно одного отрицательного супервитка на 200 п.н. Эти супервитки можно удалить, вводя одноцепочечные разрезы с помощью ДНКазы; нуклеосомы при этом остаются. Следовательно, суперспирализация, очевидно, обусловлена расположением нуклеосом и должна заключаться в скручивании линкеров между отдельными нуклеосомами. Для полной релаксации супервитков нужен один одноцепочечный разрез на каждые 85 т.п.н. Таким образом, средняя длина «замкнутой» ДНК равна

примерно 85 т.п.н. Это может быть петля или домен, похожий по своему устройству на те, которые обнаружены в бактериальном геноме. Хотелось бы знать, соответствуют ли эти петли специфическим последовательностям и имеют ли они функциональное значение.

Если удалить большинство белков из митотических хромосом, то можно прямо наблюдать петли. Гистоны удаляют путем конкурентного замещения полианионами декстрансульфатом и гепарином. При этом также удаляется значительная часть негистоновых белков. Оставшийся комплекс состоит из ДНК, связанной с белком, который составляет примерно 8% от исходного содержания белков. Как видно из рис. 29.26, хромосомы, лишенные белка, приобретают характерную структуру в виде центрального **остова**, окруженного ореолом из ДНК.

Остов состоит из плотной сети хроматиновых нитей. Нити ДНК, исходящие из остова, выглядят как петли, длина которых в среднем составляет 10–30 мкм (30–90 т.п.н.). Если для того, чтобы уместиться в нити размером 30 нм, петли должны быть уплотнены в 40 раз, то их средняя длина будет около 0,25–1,0 мкм, т.е. немного больше, чем диаметр хромосомы.

Петли можно увидеть и другим способом. Если из хромосом удалить двухвалентные катионы, на поперечных срезах будут видны петли в форме радиальных рядов нитей размером 10 нм со средней длиной 3–4 мкм. Эти данные согласуются с моделью, в которой петли ДНК длиной около 60 т.п.н. закреплены на центральном белкосоудержающем остове.

Негистоновые белки, входящие в состав остова, могут быть металлопротеинами. Хелатирование ионов металлов разрушает остов, но он может быть специфически восстановлен при добавлении ионов Cu^{2+} . Ионы кальция также способствуют образованию остова, хотя, очевидно, менее специфически.

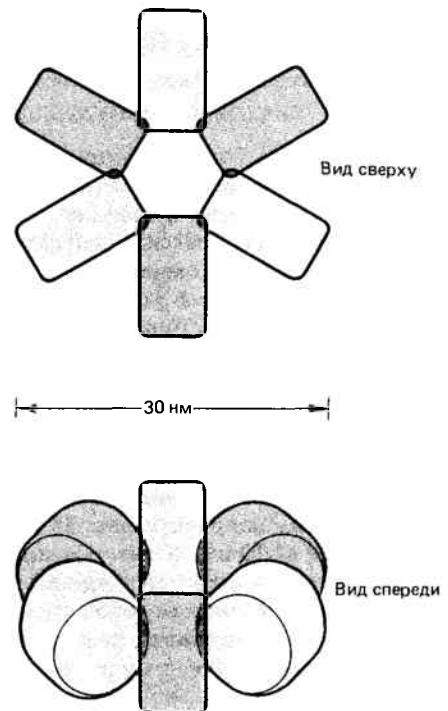
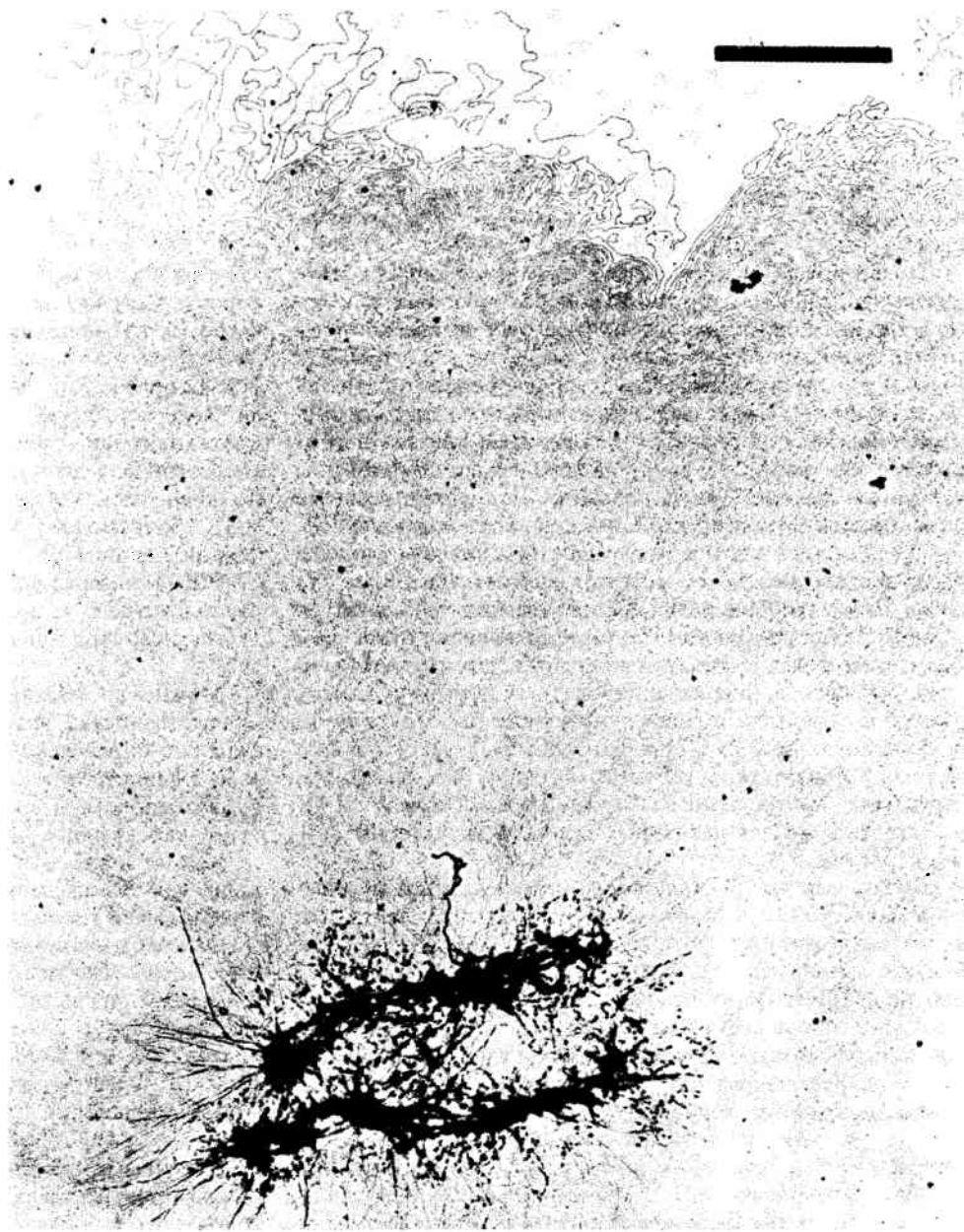


Рис. 29.25. Нить размером 30 нм может образовывать спиральный виток из 6 нуклеосом на виток, уложенных радиально.

Рис. 29.26. Лишенная гистонов хромосома состоит из белкового остова, к которому прикреплены петли ДНК. (Фотография любезно предоставлена Ulrich K. Laemmli.)



Можно расщепить ДНК, не нарушая целостности остова. Остов выглядит наподобие митотической пары сестринских хроматид. Эти сестринские половины остова обычно тесно соединены, но иногда они разделяются, оставаясь соединенными только немногими нитями. Может ли остов быть той структурой, которая отвечает за форму митотических хромосом? Образуется ли он путем соединения белковых компонентов, которые обычно закрепляют основания петель интерфазного хроматина?

Рекомендуемая литература

Развитие представлений о нуклеосомах можно проследить начиная с нескольких слабых аргументов, на основе которых была выдвинута первоначальная модель Корнберга (Kornberg, *Science*, **184**, 868–871, 1974) до вполне весомых доказательств, собранных в таких обзорах, как

(Kornberg, *Ann. Rev. Biochem.*, **46**, 931–954, 1977; McGhee, Felsenfeld, *Ann. Rev. Biochem.*, **49**, 1115–1156, 1980; Lewin в *Gene Expression*, **2**, Eucariotic Chromosomes, Wiley, New York, 332–393, 1980). Взаимоотношения между расщеплением ДНКазой I и периодичностью нуклеосомной ДНК исследовано во всех возможных деталях Клугом и Луттером (Klug, Lutter, *Nuc. Acids Res.*, **9**, 4267–4283, 1981). Вопрос о расположении ДНК на нуклеосоме кратко рассмотрен Вонгом (Wang, *Cell*, **29**, 724–726, 1982). Модель гистонового октамера разработана Клугом и др. (Klug *et al.*, *Nature*, **287**, 509–516, 1980); в этой работе цитируются также ранние работы по структурному анализу. О свойствах нуклеоплазмина сообщили Эрншоу и др. (Earnshaw *et al.*, *Cell*, **21**, 373–383, 1980), которые также рассмотрели вопросы нуклеосомной сборки (Laskey, Earnshaw, *Nature*, **286**, 763–767, 1980).

Глава 30

НУКЛЕОСОМЫ В АКТИВНОМ ХРОМАТИНЕ

Появление модели, описывающей хроматин как нить двуспиральной ДНК, закрученной вокруг серии нуклеосом, явилось первой критической вехой в выяснении способа организации генетического материала в ядре. Это несколько статический взгляд на строение отдельной субъединицы и (в некоторой степени) на его взаимоотношение со следующей субъединицей. Нам еще предстоит рассмотреть вопрос о том, происходит ли образование нуклеосом на этой нити случайным образом или это определяется специфической последовательностью ДНК.

Организация хромосом в действительности должна быть достаточно *гибкой*, для того чтобы удовлетворять различным требованиям, предъявляемым к структуре и функциям хроматина. Например, наличие нуклеосом характерно как для эухроматина, так и для гетерохроматина. Можно ли идентифицировать наборы нуклеосом, различные свойства которых объясняли бы структурные и функциональные свойства определенных участков хромосомы? Являются ли нуклеосомы единственным типом структуры, построенной из нити двухцепочечной ДНК и белка, или же для некоторых участков характерны другие структуры?

До сих пор мы рассматривали структурные изменения в хроматине, главным образом исходя из того факта, что в митотических хромосомах эухроматин обязательно должен принимать более плотноупакованное состояние. Это периодическое изменение охватывает весь эухроматин более или менее одновременно. Вероятнее всего, что этот процесс контролируется изменениями в белках, широко распространенных по всему хроматину. Изменения противоположного типа происходят при событиях двух типов, имеющих место только в таких топологических условиях, когда структура находится в растянутом состоянии. Это репликация и транскрипция.

В растущей клетке во время интерфазы весь хроматин должен удвоиться. Репликация происходит как серия индивидуальных событий в небольших участках (репликонах). При этом удваиваются соответствующие участки двухцепочечной ДНК, каждый из которых связан с набором гистоновых октамеров. Какие события происходят при удвоении нуклеосомной частицы, пока еще не установлено (гл. 29), однако разделение цепей родительской ДНК, по-видимому, должно неизбежно нарушать структуру, по крайней мере хроматиновых нитей размером 30 нм, а, возможно, также и нитей размером 10 нм.

Было бы интересно установить протяженность такого нарушения. Происходит ли оно только поблизости от точки синтеза ДНК или затрагивает и более удаленные участки? Существуют ли заметные структурные различия между участками, которые уже реплицированы, и теми, которым это еще предстоит? Быстротечность репликационного события значительно осложняет анализ структуры отдельного участка в момент его репликации.

При транскрипции также происходит раскручивание ДНК, и, вероятно, поэтому требуется расплетение нити в ограниченном участке хроматина. Нетрудно догадаться,

что для осуществления этого процесса необходимо некоторое пространство. Особенности полигенных хромосом и хромосом типа «ламповых щеток», описанные в гл. 28, дают представление о том, что более пространственная структурная организация может быть связана с экспрессией гена. Хотелось бы выяснить, какие структурные изменения происходят при транскрипции гена. Сохраняет ли он нуклеосомную структуру, и если да, то что происходит с нуклеосомой, когда РНК-полимераза транскрибирует ее ДНК? Затрагивают ли общие структурные изменения активной единицы транскрипции только участок, связанный с РНК-полимеразой, или и другие последовательности, расположенные за пределами этого участка? Чем определяется исходная чувствительность промотора к ферменту?

Можно ли идентифицировать серии нуклеосом, различные свойства которых объясняли бы структуру или функцию определенных участков? И являются ли нуклеосомы единственным типом ДНК-белковой структуры в двухцепочечной нити или существуют другие структуры, характерные для определенных участков?

Как ни важны эти вопросы, на самом деле они служат лишь как бы введением в основной вопрос, связанный с экспрессией генов. *Что изменяет состояние гена, делая его способным к транскрипции в нужное время и в нужном месте?* Из этого вопроса вытекает второй вопрос: возможно ли выключение гена, и если да, то каким образом?

Существует ли упорядоченность в расположении нуклеосом?

Мы знаем, что *in vitro* можно реконструировать нуклеосомы безотносительно к последовательности ДНК. Однако это не исключает того, что *in vivo* их образование может контролироваться последовательностью нуклеотидов. Находится ли специфическая последовательность ДНК всегда в определенном положении по отношению к топографии нуклеосомы *in vivo*? Или же нуклеосомы расположены на ДНК случайным образом так, что специфическая последовательность в одной копии генома располагается в участке минимальной нуклеосомы, а в другой приходится на межнуклеосомный промежуток?

Чтобы исследовать этот вопрос, нужно взять ДНК с известной последовательностью или, что более точно, нужно установить положение определенной точки в ДНК относительно нуклеосомы. На **рис. 30.1** изображена принципиальная схема такого эксперимента.

Предположим, что определенная последовательность ДНК организована в нуклеосомы только одной конфигурации так, что каждый участок в ДНК локализован в нуклеосоме всегда в одном и том же положении. Это называют **фазированием нуклеосом**. В серии фазированных нуклеосом линкерные области ДНК заключают в себе уникальные сайты.

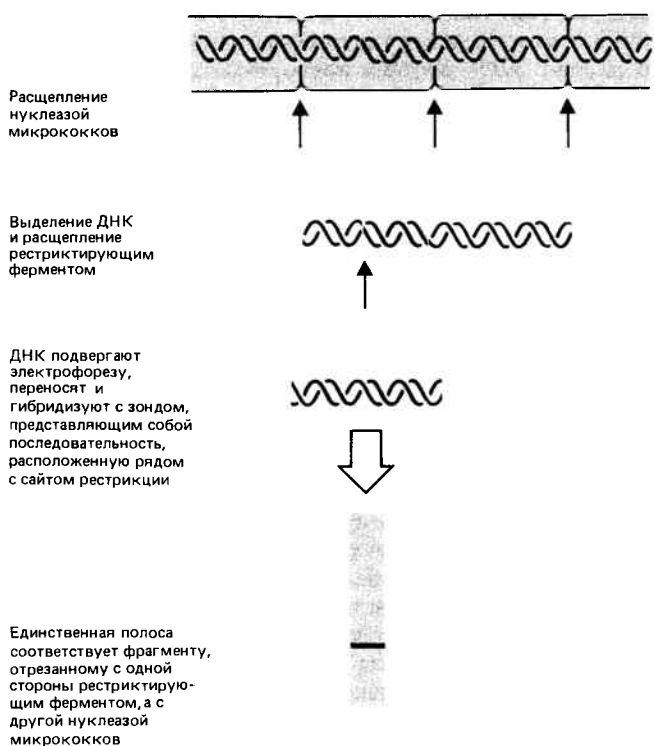


Рис. 30.1. Благодаря нуклеосомному фазированию сайты рестрикции попадают в уникальные положения линкерных участков, которые разрезаются нуклеазой микрококков.

Рассмотрим последовательность только для одной нуклеосомы. При расщеплении нуклеазой микрококков образуется мономерный фрагмент, состоящий из *специфической последовательности*. Если выделить ДНК и расщепить ее рестриктирующим ферментом, для которого имеется только один сайт рестрикции на этом фрагменте, то ДНК будет разрезана только в одной точке. При этом образуются два фрагмента, каждый с уникальным размером. Продукты двойного расщепления нуклеазой микрококков и рестриктирующим ферментом разделяют с помощью электрофореза в геле. Для идентификации фрагментов, образованных в результате двойного расщепления, используют специальный зонд, представляющий собой последовательность ДНК, непосредственно примыкающую к сайту рестрикции (с одной стороны). Этот метод называют **непрямым концевым мечением** (не совсем соответствующее название).

Иначе говоря, идентификация одной четкой полосы показывает, что положение сайта рестрикции однозначно определено по отношению к концу нуклеосомной ДНК (который соответствует разрезу нуклеазой микрококков). Таким образом, нуклеосома обладает уникальной последовательностью ДНК.

Что происходит в том случае, если нуклеосома не занимает одно-единственное положение? Эта возможность рассматривается на **рис. 30.2**, где показано пять возможных способов локализации нуклеосомы на одной последовательности ДНК. Но в действительности их может быть столько, сколько пар оснований находится в повторяющейся единице. В этом случае в каждой копии генома линкерная ДНК состоит из *разных* последовательностей ДНК. Сайты рестрикции каждый раз лежат в разном положении. В действительности они могут находиться во

всех возможных положениях по отношению к концам мономерной нуклеосомной ДНК.

Отдельный мономерный фрагмент, полученный после расщепления нуклеазой микрококков, может быть представлен разнообразием последовательностей; под действием рестриктирующего фермента этот фрагмент распадается, образуя непрерывную серию более мелких фрагментов. Если мы примем, что наименьший фрагмент, который можно идентифицировать методом гибридизации, содержит 20 пар нуклеотидов, то в описанном выше опыте мы получим размазанное пятно, содержащее фрагменты от 20 п.н. до полной длины мономерной ДНК.

При анализе реальных последовательностей интерпретация результатов оказывается несколько сложнее. В случае больших геномов невозможно достаточно точно идентифицировать один уникальный фрагмент. А нам нужно узнать, фазирована ли нуклеосома не только в одной короткой последовательности, а на протяжении некоторого важного участка. Так, при попытках исследовать фазирование нуклеосом часто используют tandemно повторяющиеся последовательности. При этом подходе нужный фрагмент эффективно амплифицируется. В таком случае изучение отдельного фрагмента позволяет устано-

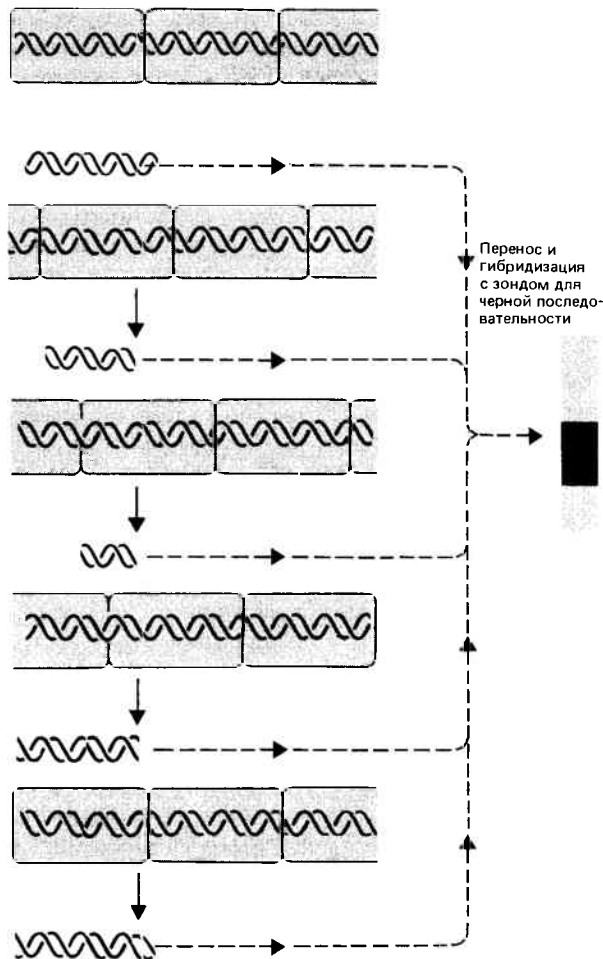


Рис. 30.2. При отсутствии нуклеосомного фазирования сайт рестрикции попадает во все возможные точки локализации в разных копиях генома. На рисунке показаны результаты образования разрывов справа от черной последовательности и в местах соединения нуклеосом.

вить, фазирована ли данная последовательность в каждой из повторяющихся единиц генома.

Такие зонды были использованы для изучения фазирования сателлитных ДНК, тРНК, 5S-ДНК, а также гистоновых генов. В этих случаях фазирование нуклеосом может иметь место только при условии простых взаимоотношений между длиной нуклеосомного повтора и длиной тандемно повторяющейся единицы.

В большинстве случаев экспериментальные данные не получаются такими четкими, как в опыте, показанном на рис. 30.1 для однозначного расположения нуклеосом. Но иногда они достаточно сильно отличаются от широких полос или размазанных пятен, предполагаемых для случайной локализации. Из этих экспериментов действительно следует, что число сайтов разрезания нуклеазой микрококков ограничено немногими положениями по отношению к конкретным сайтам рестрикции. Следовательно, можно предположить, что образование нуклеосом ограничено таким способом, который позволяет им находиться только в немногих (2–4) альтернативных фазах.

Интересная ситуация была обнаружена в некоторых генах тРНК у *Xenopus laevis*. В хроматине эритроцитов, где гены не экспрессируются, ДНК организована в нуклеосомы с длиной повтора 198 п.н. Большинство нуклеосом находится в одной фазе на протяжении по крайней мере 3 т.п.н. (внутри этого участка может быть отрезок в 250 п.н. без нуклеосом). Иная картина наблюдается в хроматине клеток печени и почек, в которых происходит экспрессия генов тРНК. Длина нуклеосомного повтора в этом случае составляет всего 185 п.н., и нуклеосомы не фазированы. Из этого следует, что размер нуклеосомного повтора и локализация нуклеосом могут, очевидно, изменяться в зависимости от обстоятельств.

Фазирование нуклеосом может осуществляться одним из двух способов. Один из них заключается в том, что каждая нуклеосома располагается специфически на определенной последовательности ДНК. Это несколько противоречит представлению о том, что нуклеосомная субъединица может образоваться в результате соединения любой последовательности ДНК с гистоновым октамером. Второй возможный способ связан с существованием некой специфической последовательности, которая предпочтительно участвует в образовании первой нуклеосомы на данном участке, а затем начинается последующее образование нуклеосом с определенным размером нуклеосомного повтора. (Если в конструкции нуклеосом существует некоторое разнообразие — например, если длина линкерного участка может варьировать, скажем, на 10 п.н., — место специфической локализации будет постоянно смещаться по мере удаления от первой фиксированной нуклеосомы.) Стартовую точку образования нуклеосом можно определить путем связывания негистоновых белков со специфическим сайтом ДНК.

Несмотря на то что фазирование возле фиксированных сайтов может быть обусловлено пограничным эффектом, могут также существовать фазированные ряды нуклеосом. В качестве примера назовем α -сателлитный хроматин африканской зеленой марышки. Эта ДНК состоит из тандемно повторяющейся последовательности в 172 п.н. Длина повтора в нуклеосомах такая же, так что одна нуклеосома приходится на один сателлитный повтор.

Широко распространенный негистоновый белок, называемый α -белком, связывается преимущественно с α -ДНК, узнавая два типа коротких А—Т-богатых по-

следовательностей. Расположение этих последовательностей в повторяющейся единице сателлитной ДНК показано на рис. 29.19. Каждая повторяющаяся единица ДНК содержит одну копию первой последовательности (сайт I) и две копии второй последовательности (сайты II и III), находящиеся в противоположной ориентации. Инвертированные последовательности отстоят на 145 п.н.: расстояние, равное длине ДНК, обмотанной вокруг гистоновой сердцевинки. Другой сайт расположен посередине между ними.

Нуклеосомы α -сателлитной ДНК фазированы таким образом, что положение сайтов II и III приходится на концы ДНК минимальной нуклеосомы. Поскольку ДНК закручена вокруг гистонного октамера, сайт I находится в непосредственной близости от них. Таким образом, кластер сайтов, связывающих α -белок, находится на одном и том же участке поверхности нуклеосомы.

Исходя из локализации этих сайтов связывания можно предположить, что α -белок отвечает за фазирование нуклеосом. Связываясь с регулярно расположенными сериями сайтов в ДНК, α -белок может обеспечивать сборку гистонов в октамеры нуклеосом, ДНК которых находится в определенной фазе. Этот механизм может использоваться для фазирования нуклеосом на протяженном участке, или, действительно, для запуска образования серии нуклеосом, начиная с фиксированного положения.

Значение фазирования нуклеосом, даже если считать его доказанным в некоторых случаях, неясно. Одна из возможных причин кажущегося фазирования заключается в том, что существуют области, в которых нет нуклеосом. Такие области могут быть связаны с контролем экспрессии генов или же с образованием структур более высокого уровня организации хроматина. вполне возможно, что область, лишенная нуклеосом, служит барьером, ограничивающим возможность образования следующей нуклеосомы в определенных положениях. В этом случае статистическое распределение нуклеосом будет таким же, как при фазировании, хотя отдельные места их расположения не обязательно должны быть фиксированы в каждой индивидуальной копии генома.

Специфичность нуклеазы микрококков

В процессе обсуждения экспериментальных данных мы говорили о нуклеазе микрококков как о ферменте, который расщепляет ДНК в незащищенных линкерных участках независимо от их нуклеотидной последовательности. Вначале так действительно и считали. Однако с тех пор было выяснено, что фермент обладает некоторой специфичностью (склонен выбирать определенные А—Т-богатые последовательности). Таким образом, мы не можем считать, что наличие специфической полосы при использовании непрямого концевое мечения указывает расстояние от места рестрикционного расщепления до линкерного участка. На деле данная полоса может отражать расстояние от места рестрикционного расщепления до сайта предпочтительного расщепления нуклеазой микрококков!

Возможность такой ситуации выяснена с помощью контрольного эксперимента, в котором очищенную ДНК обрабатывали точно так же, как и хроматин. Если в определенном участке имеются места предпочтительного действия для нуклеазы микрококков, то должны образовываться специфические полосы. Затем полученную картину

расположения полос сравнивали с той, которая получалась при обработке хроматина.

Если расположение нуклеосом на ДНК происходит случайным образом, то все сайты, чувствительные к нуклеазе микрококков, рано или поздно окажутся в линкерной области, став доступными для действия фермента. Картина расположения полос при расщеплении хроматина и ДНК будет одинаковой. Но, если нуклеосомы лежат в упорядоченных местах (фазируются), некоторые сайты, чувствительные к нуклеазе микрококков, будут вне пределов досягаемости, так как окажутся в пределах минимальной нуклеосомы. Тогда отдельные полосы, обнаруживаемые при электрофорезе расщепленной ДНК, будут отсутствовать в электрофореграммах расщепленного хроматина. Если же при образовании нуклеосом возникают новые чувствительные сайты, то в электрофореграммах расщепленного хроматина можно будет обнаружить новые полосы. Таким образом, различие в расположении полос при расщеплении контрольной ДНК и хроматина доказывает существование фазирования нуклеосом. Такие эксперименты действительно были выполнены.

Эти результаты проливают некоторый свет на природу межнуклеосомного сайта, разрезаемого нуклеазой микрококков. Вероятно, фермент разрезает линкерную ДНК не в наиболее открытой точке, а в предпочтительном сайте, ближайшем к этому положению (если такой есть). Поскольку положение таких сайтов может быть различным, ширина полос увеличивается. Способность сайтов предпочтительного расщепления привлекать к себе фермент, возможно, увеличивает количество положений в линкерной ДНК, наиболее доступных для нуклеазы микрококков.

Сохраняют ли транскрибируемые гены нуклеосомную структуру?

В процессе транскрипции локально расплетенный участок, образованный РНК-полимеразой, перемещается вдоль ДНК (как это описано в гл. 10). При этом обе комплементарные цепи ДНК непосредственно связаны с ферментом или даже окружены им, как это изображено на рис. 10.3. Примерно 50 п.н. ДНК действительно связаны с ферментом настолько прочно, что даже защищены от действия нуклеаз. Поскольку ДНК должна быть расплетена, то кажется маловероятным, чтобы участок, связанный с РНК-полимеразой, мог оставаться на поверхности нуклеосом. Однако о стереохимии работы РНК-полимеразы (особенно у эукариот) известно мало. Поэтому можно предполагать, что фермент способен использовать одну цепь ДНК в качестве затравки, тогда как другая цепь остается связанной с гистонами.

Механизм, с помощью которого ДНК расплетается, образуя матричную цепь для РНК-полимеразы, неясен. Обычно считают, что в этом процессе участвует что-то вроде шарнира, благодаря которому одна цепь ДНК может свободно вращаться относительно другой. Проблема раскручивания еще более усложняется, если ДНК во время транскрипции неподвижно закреплена на нуклеосоме.

При рассмотрении этих вопросов следует помнить об относительных размерах РНК-полимеразы и нуклеосомы. Эукариотические ферменты — это большие белки, обычно более 500 000 дальтон. Сравним это с молекулярной массой нуклеосомы, составляющей 264 000. На рис. 30.3 показан момент приближения РНК-полимеразы к нуклеосомной ДНК. Даже не зная в деталях механизма

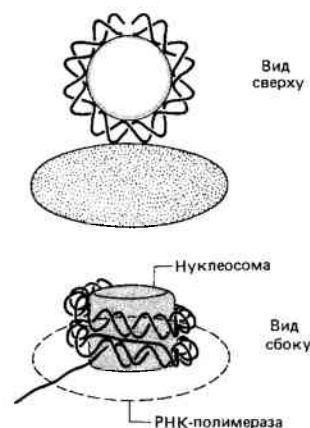


Рис. 30.3. РНК-полимераза сравнима по размеру с нуклеосомой и может испытывать затруднения, продвигаясь по ДНК вокруг нуклеосомы.

их взаимодействия, нетрудно представить, что при этом происходит сближение двух сравнимых по размерам тел. Кажется несомненным, что в процессе их взаимодействия должны происходить структурные изменения.

Нуклеосомная ДНК относительно нечувствительна к нуклеазе микрококков. Даже в изолированных частицах минимальной нуклеосомы только часть связей в ДНК оказывается чувствительной к ДНКазе I, делающей одноцепочечные разрывы (гл. 29). Напомним, что каждая нуклеосома существует не изолированно, а по соседству с другими, и рассмотрим два оборота ДНК, закрученные вокруг нуклеосомы. Все это выдвигает на первый план вопрос о том, имеет ли РНК-полимераза достаточный доступ к ДНК, если нуклеиновая кислота как обычно закручена вокруг нуклеосомы. Трудно себе представить, что в процессе транскрипции полимеразы может следовать по ДНК вокруг нуклеосомы.

Таким образом, первый вопрос, касающийся структуры активных генов, заключается в том, организована ли транскрибируемая ДНК в виде нуклеосом.

Гены, кодирующие рРНК, можно исследовать под электронным микроскопом. Структура ДНК при этом плохо видна из-за плотного расположения молекул РНК-полимеразы. Это отчетливо видно на рис. 23.4 и на другом примере, показанном на рис. 30.4. Плотность упаковки ДНК можно подсчитать, разделив известную длину единицы транскрипции, которую измеряют по оси транскрипционной матрицы. Она равна примерно 1,2. Таким образом, ДНК почти полностью вытянута и не может быть организована в нуклеосомы.

Нетранскрибируемые спейсеры между транскрипционными матрицами также почти полностью вытянуты; их плотность упаковки составляет около 1,4. При соблюдении в процессе приготовления препаратов определенных условий в спейсерах обнаруживают «бусины», которые, возможно, являются нуклеосомами. Если это так, то должна быть и свободная спейсерная ДНК, которая их соединяет. При других условиях получения препаратов спейсеры состоят из свободной ДНК. Все это свидетельствует о том, что состояние ДНК в интенсивно транскрибируемых tandemных генах рРНК сильно отличается от компактной организации, которая видна даже в отдельной цепочке соседних нуклеосом в хроматиновой нити размером 10 нм (плотность упаковки которых равна ~ 6).

С другой стороны, из инфицированных клеток могут



Рис. 30.4. Вытянутая вдоль оси единица транскрипции рДНК переходит в несколько менее вытянутый нетранскрибируемый спейсер.

(Фотография любезно предоставлена Charles Laird.)

быть экстрагированы транскрипционные комплексы мини-хромосом SV40. Они содержат обычный комплект гистонов и имеют структуру бусинок. Можно видеть цепи РНК, тянущиеся из мини-хромосом, как это показано на рис. 30.5. Из этого следует, что на ДНК вируса SV40 может происходить транскрипция, не нарушая нуклеосомной организации. Мини-хромосома SV40 транскрибируется менее интенсивно, чем гены рРНК. Более редко транскрибируемые клеточные гены также можно наблюдать в виде частично вытянутых структур, в которых содержится меньше нуклеосом, чем в нетранскрибируемых областях.

Другой подход к изучению этого вопроса заключается в расщеплении хроматина нуклеазой микрококков с последующим использованием зонда для определенного гена (или генов), чтобы установить наличие соответствующих фрагментов в обычной лестнице с шагом в 200 п.н. В таких экспериментах гены рРНК могут обнаруживаться во фракции нуклеосом. Однако всегда остается возможность, что эти определенные фрагменты происходят из некоторых нетранскрибируемых участков, tandemно входящих в исследуемый набор фрагментов.

Аналогичная проблема возникает при исследовании уникальных генов. В случае генов, которые транскрибируются не непрерывно, проблема заключается в том, что в клеточной популяции в каждый данный момент транскрибируются только некоторые гены. Для того чтобы преодолеть эту трудность, сравнивают сумму гибридирующихся с зондом последовательностей в мономерных фрагментах и в целой ДНК. При использовании зондов, соответствующих индивидуальным генам или популяциям мРНК (или в этом случае сателлитной ДНК), одни

и те же результаты получают при гибридизации с нуклеосомными фрагментами и с целой ДНК.

Из этих экспериментов можно сделать ограниченные, но важные выводы. В генах, которые в данный момент транскрибируются, нуклеосомы расположены с такой частотой, что и в нетранскрибируемых последовательностях. Таким образом, для транскрипции гены не должны обязательно принимать другую форму организации. Но, поскольку часть гена, связанная с молекулами РНК-полимеразы, может быть довольно мала, мы не можем судить о том, что происходит в местах, непосредственно связанных с ферментом. Возможно, что нуклеосомы там сохраняются. Но, может быть, в местах следования РНК-полимеразы нуклеосомная структура временно исчезает, а затем сразу восстанавливается.

Эти эксперименты показывают также, что активные гены быстрее расщепляются до мономерных фрагментов, чем неактивные последовательности. Это может свидетельствовать о том, что, несмотря на нуклеосомную организацию, структура активных генов менее защищена.

На примере некоторых генов теплового шока у *D. melanogaster* можно проиллюстрировать, что во время транскрипции в структуре нитей хроматина происходят определенные изменения. В период, предшествующий тепловому шоку, эти гены транскрибируются не слишком часто, а обработка нуклеазой микрококков приводит к появлению типичной лестницы (указывающей на нуклеосомную организацию соответствующих районов). После теплового шока они активируются и затем начинают интенсивно транскрибироваться. Как видно из рис. 30.6, лестница в результате оказывается довольно смазанной, что указывает на изменение нуклеосомной организации.

Природа такого изменения непонятна. Такие воздействия, как разрезание хроматина под действием сил сдвига, могут уничтожить характерную реакцию на нуклеазу микрококков даже при сохранении нуклеосомной структуры. Таким образом, мы не можем сказать, исчезает ли лестница из-за того, что присутствие РНК-полимеразы (или других белков) затемняет картину, поскольку при этом нарушается высокая упорядоченность структуры, или оттого, что нуклеосомы видоизменяются или даже совсем исчезают.

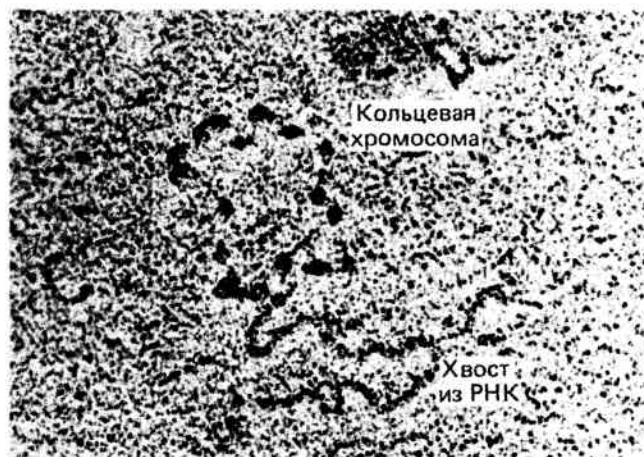


Рис. 30.5. Мини-хромосома вируса SV40 может транскрибироваться.

(Фотография любезно предоставлена Pierre Chambon.)

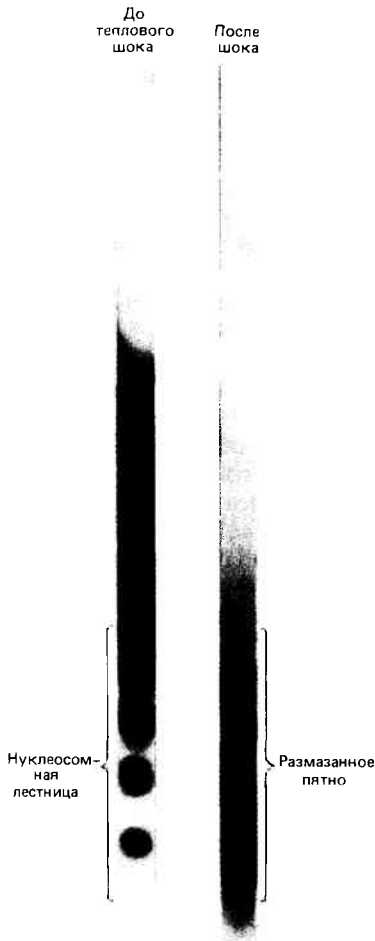


Рис. 30.6. До активации генов теплового шока в результате обработки нуклеазой микрококков возникает отчетливая лестница, которая становится смазанной после начала интенсивной транскрипции. (Фотография любезно предоставлена Sarah Elgin.)

Очевидно, что при интенсивной транскрипции гена происходят существенные структурные изменения. В случае генов рРНК это проявляется в исчезновении нуклеосом. Но, возможно, это исключительный случай. С присутствием нуклеосом в умеренно транскрибируемых генах вполне согласуется предположение о том, что РНК-полимераза нарушает нуклеосомную структуру в точке транскрипции, но сразу после этого гистоновый октамер снова занимает свое место, если только другая молекула РНК-полимеразы не мешает ему это сделать.

Домены, чувствительные к ДНКазе в транскрибируемом хроматине

Важно дифференцировать две возможные причины изменения структуры хроматина в транскрибируемых участках.

Поскольку в любой клетке транскрибируется только характерное меньшинство последовательностей ДНК, должны существовать отличительные черты, по крайней мере у промоторов (а возможно, и по всей длине), этих единиц транскрипции. Поэтому первым изменением должно быть некоторое событие, которое указывает аппарату транскрипции, что локус готов к экспрессии.

В транскрибируемом гене должно произойти некоторое изменение структуры, хотя бы только в результате передвижения РНК-полимеразы по ДНК. Таким образом, структурные изменения могут быть следствием акта транскрипции, а не его причиной. Оценивая свойства

транскрибируемых локусов, нам нужно определить, какие из его признаков существовали до транскрипции в качестве ее предпосылки, а какие были индуцированы потом как результат событий, вовлеченных в синтез РНК.

При расщеплении хроматина ДНКазой I он деградирует в конце концов до кислоторастворимого материала (очень мелких фрагментов ДНК). За общим ходом реакции можно проследить, определяя долю ДНК, перешедшей в кислоторастворимую фракцию. За судьбой же отдельного гена можно проследить, исследуя количество нерасщепившейся ДНК по ее способности вступать в реакцию со специфическим зондом (для этого обычно используются метод рестрикционного расщепления, блоттинг и гибридизацию). Схема такого опыта приведена на рис. 30.7. Основной подход заключается в том, что потеря определенной полосы свидетельствует о том, что именно этот участок переварен ферментом.

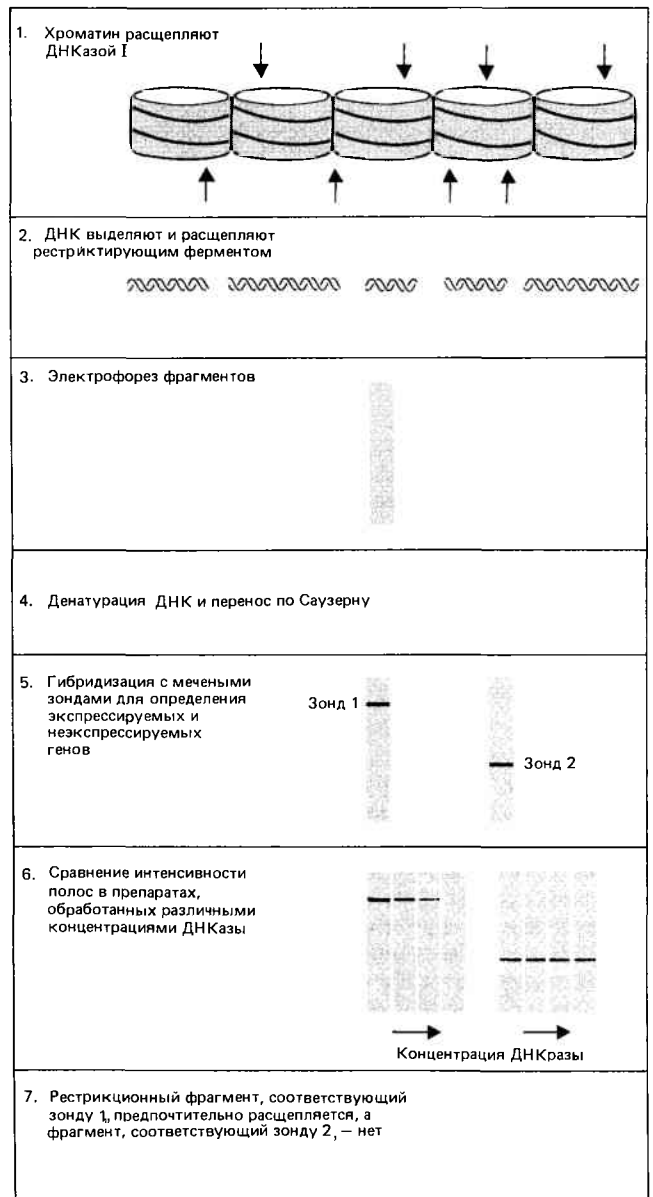


Рис. 30.7. Чувствительность к ДНКазе I можно измерить, определяя скорость исчезновения материала, гибридизующегося с данным зондом.

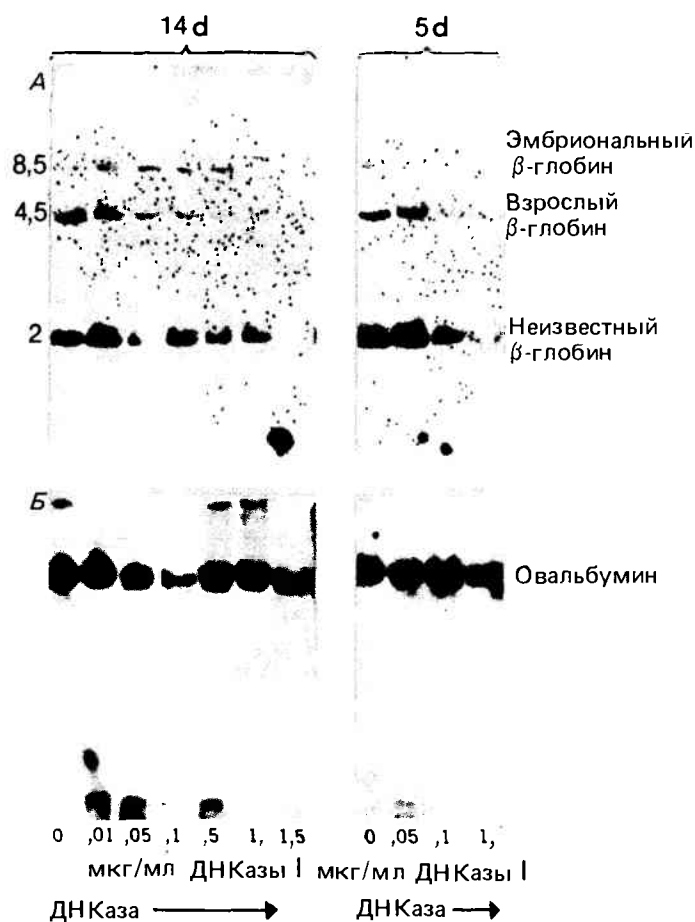


Рис. 30.8. В хроматине 14-дневных клеток эритроидного ряда β -глобиновый ген взрослого типа высокочувствителен к ДНКазе I, ген эмбрионального β -глобина чувствителен, а ген овальбумина нечувствителен. В 5-дневных клетках ген эмбрионального глобина высокочувствителен, ген глобина взрослого типа чувствителен, а ген овальбумина нечувствителен.

Изменения в чувствительности генов глобинов эмбрионального и взрослого типов обсуждаются ниже в тексте. (Данные любезно предоставлены Harold Weintraub.)

Оказывается, что расщепление до кислоторастворимого состояния *только 10% общей ДНК* приводит к исчезновению *более 50% ДНК активных генов*. Следовательно, можно предположить, что при таком расщеплении происходит преимущественная деградация активных генов. Исчезновение активных генов можно подтвердить прямым сравнением судьбы двух генов, активного и неактивного. На рис. 30.8 показано, что происходит с генами β -глобина и с геном овальбумина в хроматине, выделенном из куриных эритроцитов (в этом хроматине гены глобинов экспрессируются, а ген овальбумина неактивен). Рестрикционные фрагменты, соответствующие генам β -глобина, быстро теряются, тогда как фрагменты, представляющие ген овальбумина, расщепляются незначительно. (Ген овальбумина фактически деградирует с той же скоростью, что и вся ДНК, хотя это и не следует из показанного на рисунке анализа ранних стадий реакции методом блотт-гибридизации. Противоположная картина наблюдается в ткани яйцевода, в которой преимущественно деградирует ген овальбумина, тогда как глобиновые гены устойчивы, как весь остальной хроматин. Это коррели-

рует с активностью овальбумина и неактивностью глобина.

Таким образом, большая часть хроматина относительно устойчива к ДНКазе I и содержит неэкспрессируемые гены (так же, как и другие последовательности). Гены овальбумина и глобинов становятся относительно чувствительными к ферменту специфически в тех тканях, в которых они экспрессируются.

Характерна ли преимущественная чувствительность только для достаточно активных генов, таких, как гены глобинов и овальбумина, или для всех активных генов? Эксперименты, в которых использовали зонды для кДНК, соответствующие популяции всей клеточной мРНК, свидетельствуют о том, что все активные гены предпочтительно чувствительны к ДНКазе I независимо от того, кодируют ли они часто или редко встречающиеся мРНК. (Однако они могут различаться по степени чувствительности.) Поскольку редко экспрессируемые гены в любой момент времени, вероятно, содержат небольшое число активно транскрибирующей РНК-полимеразы, то очевидно, что чувствительность к ДНКазе I не является результатом акта транскрипции, а присуща готовым к транскрипции генам.

Как велика область, преимущественно чувствительная к ДНКазе I? Это можно определить, используя серию зондов, соответствующих фланкирующим участкам, а также самой единице транскрипции. Оказывается, что чувствительная область всегда захватывает достаточно протяженные участки по обе стороны от единицы транскрипции. Поэтому было высказано предположение, что чувствительность к ДНКазе I характерна для хромосомного домена, т.е. области, включающей в себя по крайней мере одну активную единицу транскрипции, но при этом присущее ей структурное изменение простирается за пределы этой единицы. (Заметим, что при использовании здесь термина «домен» не подразумевается какой-либо связи со структурными доменами, соответствующими петлям хроматина или хромосом.)

При исследовании вопроса о том, в равной ли мере сохраняется чувствительность к ДНКазе I на протяжении всего домена, были получены некоторые расхождения в результатах отдельных опытов. В кластере β -глобиновых генов цыпленка чувствительный домен тянется на 6–7 т.п.н. с 5'-конца кластера и по крайней мере на 8 т.п.н. с 3'-конца. Внутри самого кластера генов транскрибируемая область более чувствительна к ДНКазе I, чем ее нетранскрибируемые фланкирующие участки (которые называются участками «промежуточной» чувствительности.) Сходным образом ген овальбумина и родственные гены X и Y могут находиться в домене, чувствительном к ДНКазе, который занимает около 100 т.п.н. Но в этом случае, по-видимому, нет каких-либо различий в чувствительности между кодирующей и некодирующей ДНК внутри домена.

Наиболее существенной частью в определении домена является концепция о том, что область, чувствительная к ДНКазе I, захватывает по крайней мере длину единицы транскрипции. Часто мы думаем, что регуляция заключается в ряде событий, возникающих в дискретной последовательности ДНК, — например, в способности инициировать транскрипцию на промоторе. Даже если это и так, такая регуляция должна определяться или сопровождаться структурными изменениями более широкого масштаба. В этом может быть различие между эукариотами и прокариотами.

Чувствительность к ДНКазе обусловлена негистоновыми белками

Чем обусловлено состояние чувствительности к ДНКазе? Чтобы выявить различия в чувствительности, которые могут быть связаны с изменениями в структуре хроматина, следует обратиться к негистоновым белкам. Чувствительный хроматин отличается по поведению некоторых HMG-белков. Два негистоновых белка, HMG14 и HMG17, преимущественно высвобождаются из хроматина при расщеплении ДНКазой I. (HMG – от англ. high mobility group, т.е. высокоподвижная группа маленьких негистоновых белков. Эти белки экстрагируются из хроматина с помощью 0,35 М NaCl, что указывает на их относительно непрочное связывание.)

После экстракции этих белков глобиновый ген в хроматине эритроцитов не проявляет предпочтительной чувствительности к ДНКазе I. Однако чувствительность восстанавливается при обратном добавлении этих белков к хроматину, из которого они были высолены. Эти белки не придают предпочтительной чувствительности глобиновым генам, если их добавить к хроматину мозга, обработанному таким же способом. Отсюда следует, что данный эффект зависит не только от HMG-белков. Глобиновые гены эритроцитов и мозга могут, по-видимому, различаться по некоторому фактору, сохраняющемуся в хроматине после солевой экстракции. Именно это различие позволяет хроматину эритроцитов связывать белки HMG14 и HMG17, образуя структуру, обуславливающую чувствительность глобиновых генов к ДНКазе.

Один из этих двух белков (либо HMG14, либо HMG17) эффективен сам по себе. Добавления примерно одного моля HMG достаточно для того, чтобы придать чувствительность каждой из 10–20 нуклеосом, выделенных из общей массы хроматина. Из этого следует, что HMG может выбирать соответствующую мишень в препарате хроматина. Такое же соотношение остается и для активно транскрибируемых генов, таких, как ген глобина, и для редко транскрибируемых.

Возникает ли чувствительность к ДНКазе на уровне индивидуальной нуклеосомы? Мономеры, выделенные в результате обработки нуклеазой микрококков, проявляют точно ту же чувствительность, что и целые ядра или хроматин. При удалении белков HMG14 и HMG17 этот эффект исчезает и снова появляется при их добавлении. Такие же результаты наблюдают при исследовании частиц *минимальной нуклеосомы*. Из этого следует, что минимальные нуклеосомы из активных генов должны иметь некие характерные признаки, позволяющие им связывать белки HMG14 и HMG17. Однако попытки идентифицировать эту характерную особенность пока остались без успеха. Отметим, что в этой работе было отчетливо показано, что ген глобина сохраняет во время транскрипции нуклеосомное строение.

Промежуточный уровень чувствительности к ДНКазе, характерный для участков, фланкирующих транскрибируемые гены глобина у кур, теряется в результате воздействия, нарушающих высокоупорядоченную структуру хроматина, и не сохраняется в отдельных нуклеосомах. Его нельзя восстановить добавлением белков HMG14 и HMG17. При иммобилизации HMG14 и HMG17 на колонках их можно использовать для выделения подвергнутых солевой обработке нуклеосом, обладающих реакционной способностью. В этих экспериментах удалось

выделить только высокочувствительные транскрибируемые области.

Из этих опытов следует, что чувствительность к ДНКазе может быть обусловлена образованием высокочувствительного участка, охватывающего с обеих сторон последовательность, способную транскрибироваться. Протяженность этого участка с каждой стороны составляет, очевидно, менее 1 т.п.н. Соседние участки, на которые распространяются структурные изменения, могут проявлять промежуточную чувствительность (хотя область чувствительности достаточно велика относительно длины самого транскрибируемого участка).

Каковы временные отношения между приобретением состояния чувствительности к ДНКазе и началом процесса транскрипции? Хотя мы и располагаем некоторыми данными, полученными на системе глобина курицы, их нельзя четко интерпретировать. Клетки эритроидного ряда 5-дневного эмбриона продуцируют только эмбриональный глобин, тогда как эритроидные клетки 14-дневного эмбриона производят глобины взрослого типа. Однако в клетках 5-дневных эмбрионов гены глобина *обоого* типа (эмбрионального и взрослого) чувствительны к ДНКазе I (хотя, как видно из рис. 30,8, гены эмбрионального глобина более чувствительны). Аналогично в клетках 14-дневных эмбрионов ген взрослого глобина высокочувствителен (более чувствителен, чем в 5-дневных клетках), а ген эмбрионального глобина проявляет промежуточную чувствительность.

Сначала думали, что состояние чувствительности к ДНКазе может отражать активацию генной структуры, происходящую перед началом транскрипции. Тогда чувствительность может служить отличительным признаком генов, потенциально способных к транскрипции, а также уже транскрибируемых генов. С другой стороны, если чувствительность распространяется из максимально чувствительного участка активного гена так, что данный домен простирается до фланкирующей области, чувствительность глобиновых генов взрослого типа в эмбриональных клетках (или наоборот) может свидетельствовать о группировании этих генов в кластеры, а не о функциональном значении. Для того чтобы установить, предшествует ли появление чувствительности началу транскрипции, нужно исследовать изолированный ген, так чтобы не было оснований предполагать наличие эффекта распространения от соседних локусов.

Из двух дальнейших примеров станет очевидным, что состояние чувствительности может сохраняться и после остановки транскрипции. Эритроциты курицы – это зрелые эритроциты, в которых продолжается трансляция глобиновой мРНК, но не происходит транскрипции генов. Однако глобиновые гены остаются в чувствительном состоянии. Удаление эстрогена из курицы приводит к тому, что транскрипция овальбуминового гена прекращается, но кодирующая область остается при этом в активном состоянии. У нас пока нет данных о том, чтобы ранее активный ген превращался в неактивную форму (устойчивую к ДНКазе), например в ходе эмбрионального развития.

Очевидно, что изменение чувствительности к ДНКазе необходимо для осуществления транскрипции, однако оно не является ни ее причиной, ни следствием. Сохранение чувствительного состояния после того, как прекращается транскрипция, показывает, что структура гена может находиться в активном состоянии и в отсутствие активной молекулы РНК-полимеразы. Эти данные наря-

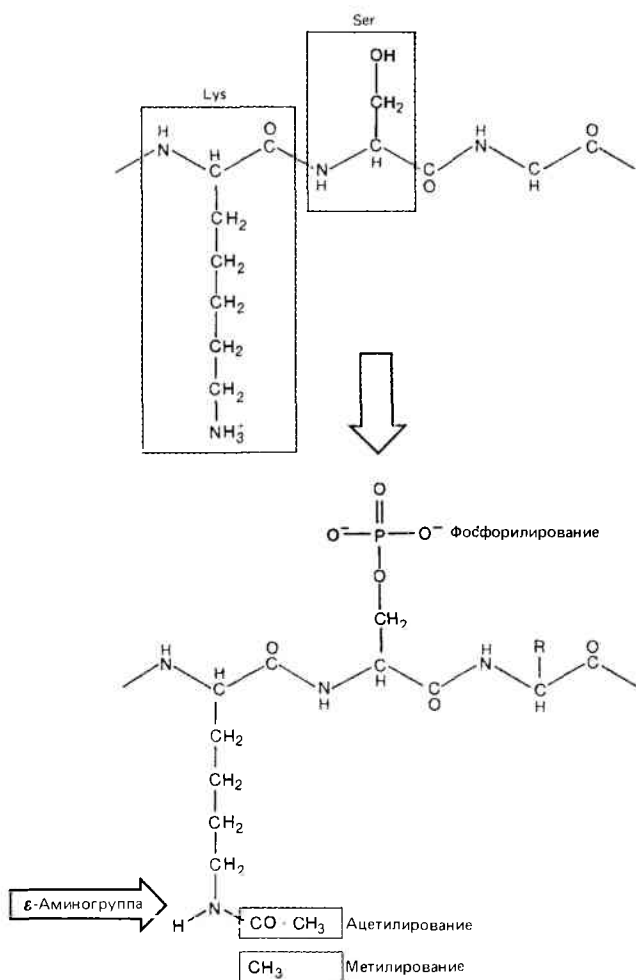


Рис. 30.9. Ацетилирование лизина или фосфорилирование серина уменьшает общий положительный заряд белка.

ду с данными о чувствительности редко транскрибируемых генов позволяют предположить, что изменение в структуре не зависит от самого акта транскрипции. Следовательно, чувствительность к ДНКазе — условие, необходимое для осуществления транскрипции, но отнюдь не достаточное. Хотелось бы получить больше сведений о природе чувствительного домена, особенно о том, как устанавливаются его границы. Следует задать вопрос: чем объясняется необходимость изменения структуры каждой нуклеосомы в единице транскрипции? Ответ, возможно, состоит в том, что такое изменение позволяет РНК-полимеразе передвигаться вдоль матрицы. Можно представить, например, что гистоны могут быть удалены с ДНК только при условии, что нуклеосомы модифицированы.

Гистоны подвергаются кратковременной модификации

В соматических клетках взрослого организма каждый гистон имеет инвариантную аминокислотную последовательность. Но все гистоны подвергаются модификации, при которой дополнительные остатки ковалентно присоединяются к свободным группам определенных аминокислот. Эти модификации приводят к уменьшению общего положительного заряда (то есть основности) белковой молекулы. Поэтому немодифицированные формы

и формы с различными модификациями можно разделить, например, с помощью гель-электрофореза.

Реакции ацетилирования и метилирования происходят в свободной (ϵ) аминогруппе лизина. Как видно на рис. 30.9, при этом удаляется положительный заряд, находящийся на NH_3^+ -группе. Реакция метилирования происходит также по аргинину и гистидину.

Фосфорилирование гистонов происходит по гидроксильной группе серина, а также гистидина. При этом вводится отрицательный заряд в виде фосфатной группы (см. рис. 30.9).

Все эти модификации затрагивают внутренние остатки и являются кратковременными. Противоположную картину представляет собой стабильное ацетилирование N-концов некоторых гистонов, происходящее во время их синтеза.

Кратковременные модификации вносятся на одной стадии клеточного цикла и (обычно) снимаются на другой стадии. Поскольку модификации приводят к уменьшению положительного заряда белковой молекулы, их рассматривают как потенциальную возможность изменения функциональных свойств гистонов. В настоящий момент еще не имеется данных, свидетельствующих о связи этих изменений с функциями хроматина, хотя и обнаружены некоторые интересные корреляции.

Цикл ацетилирования и деацетилирования гистонов хорошо изучен на сперме лосося (где вместо гистонов находятся основные белки, называемые протаминами). В этой системе гистоны H3 и H4 служат основной мишенью; к каждому из них добавляется до четырех ацетильных групп. Гистоны H2A и H2B также ацетилируются, но в меньшей степени. В этом и некоторых других случаях установлены положения, подвергающиеся ацетилированию (или метилированию). У различных органов эти положения неидентичны, хотя они и образуют небольшое число перекрывающихся сайтов.

В синхронизированной культуре клеток и ранее существовавший и новосинтезированный гистоны ацетилируются и метилируются в фазе S. (Это часть клеточного цикла, во время которой происходит репликация ДНК; тогда же синтезируются гистоны.) В более поздней части клеточного цикла модифицированные группы удаляются. Эти события повторяются в каждом клеточном цикле, и их зависимость во времени суммирована на рис. 30.10.

Совпадение времени модификации и репродукции хроматина дает основание думать, что ацетилирование (и метилирование) могут быть связаны со сборкой нуклеосом. Одна из возможностей состоит в том, что уменьшение положительного заряда гистонов может способствовать уменьшению их сродства к ДНК. Это позволило бы лучше контролировать реакцию. Данная идея несколько утратила свою обоснованность ввиду того, что была продемонстрирована способность нуклеосом к реконструкции, по крайней мере *in vitro*, с немодифицированными гистонами (но см. гл. 29).

Недолговечность ацетилированного состояния оказалась препятствием для его изучения. Однако это препятствие можно преодолеть, добавляя масляную кислоту к клеткам, растущим в культуре. При этом ингибируется фермент гистон-деацетилаза и ацетилированные нуклеосомы накапливаются. Все гистоны сердцевинны ацетилированы. При ацетилировании в хроматине происходят изменения, подобные тем, которые наблюдаются при активации гена. Такой хроматин более чувствителен к ДНКазе I (возможно) к нуклеазе микрококков. Одна-

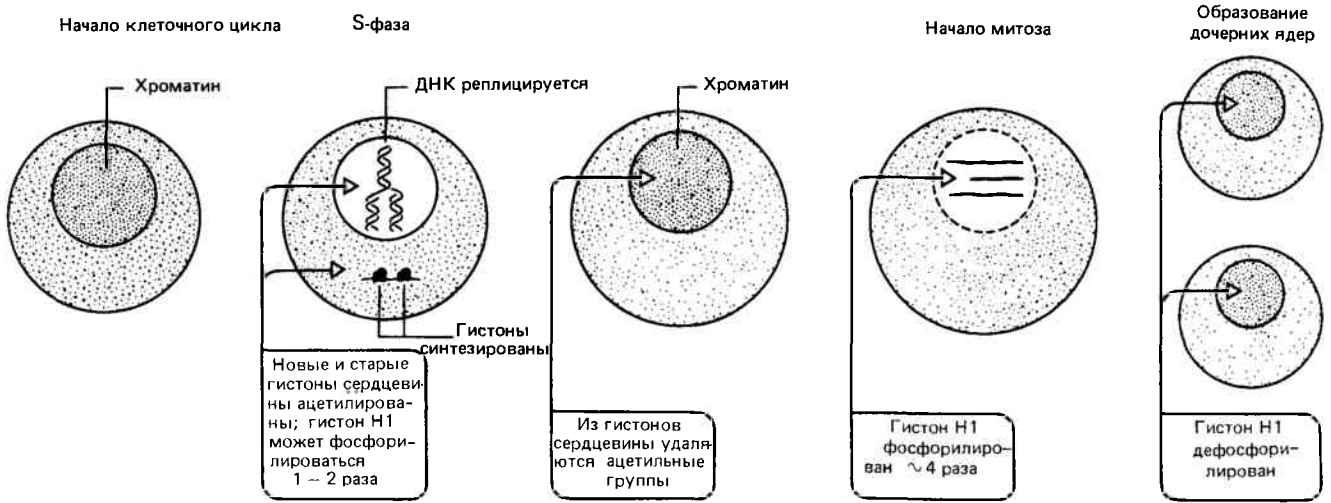


Рис. 30.10. В период синтеза ДНК происходит ацетилирование гистонов сердцевинны с последующим деацелированием. Основ-

ное фосфорилирование гистона Н1 происходит в начале митоза; в конце митоза гистон Н1 дефосфорилируется.

ко каких-либо убедительных корреляций обнаружить не удалось, преимущественного ацетилирования активных генов не обнаружено. Из этого можно сделать вывод, что ацетилирование действительно затрагивает структуру хроматина. Однако значение этого изменения еще неясно.

Цикл фосфорилирования и дефосфорилирования обнаружен у гистона Н1, однако по времени он отличается от цикла модификации других гистонов. В культуре клеток млекопитающих одна или две фосфатные группы могут быть введены в фазе S. Но основное фосфорилирование происходит позднее, когда добавляется еще несколько фосфатных групп, так чтобы общее их число достигло шести. Как показано на рис. 30.10, это происходит в митозе. Все фосфатные группы удаляются в конце процесса деления. Введение некоторых фосфатных групп катализируется ферментом фосфокиназой, активность которой резко возрастает в самом начале митоза. О фосфатазе, под действием которой позднее происходит удаление фосфатных групп, известно немного.

Время основного фосфорилирования гистона Н1 навело на мысль о том, что он может участвовать в митотической конденсации хромосом. Это согласуется с данными о необходимости гистона Н1 для образования нитей хроматина размером 30 нм (гл. 29). Некоторые температурочувствительные мутанты, дефектные по Н1-фосфорилированию, не способны закончить репликацию ДНК и завершить деление клетки. Но такая корреляция не говорит ничего о том, является ли фосфорилирование причиной или сопутствующим событием митотической конденсации.

В некоторых нуклеосомах гистон Н2А связывается с убиквитином

Необычная модификация гистона была обнаружена при изучении белка, первоначально обнаруженного в негистоновой фракции хроматина печени крысы. Этот белок, известный сначала как А24, имеет ту же самую С-концевую аминокислотную последовательность, что и Н2А. Однако, как изображено на рис. 30.11, у него две N-концевые последовательности. Одна из них принадле-

жит гистону Н2А, а другая – белку убиквитиину. Название этого белка отражает его повсеместное (ubiquitous) присутствие в клетках – от бактерий до млекопитающих. Конъюгат Н2А-убиквитин теперь обозначают как UН2А.

Белок убиквитин состоит из 76 остатков (сравним с гистоном Н2А, содержащим примерно 130 остатков). Изопептидная связь образуется между С-концевым глицином убиквитина и свободной ε-NH₂-группой лизина в положении 119 гистона Н2А. (Названием *изопептидная связь* подчеркивается, что данная NH₂-группа – это не обычная аминогруппа, участвующая в образовании пептидной связи.) Убиквитин – это кислый белок, в котором содержание глутаминовой и аспарагиновой кислот таково, что общее отношение основные/кислые аминокислоты во вновь образованном конъюгированном белке оказывается пониженным.

Около 5–15% гистона Н2А может находиться в форме UН2А. Обычно только одна из двух молекул Н2А в гистоновом октамере связана с убиквитином. Следовательно, убиквитин может присутствовать в 10–30% нуклеосом. Он находится, вероятно, на поверхности нуклеосомы. Довольно маленькая доля гистона Н2В также может соединяться с убиквитином.

Нуклеосомы, содержащие гистон Н2А, связанный с убиквитином, можно отделить от немодифицированных

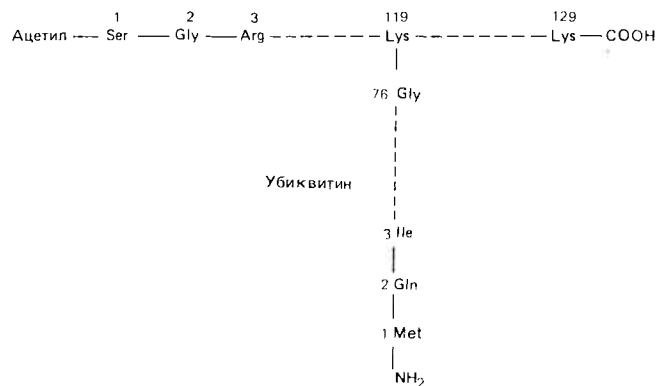


Рис. 30.11. UН2А состоит из убиквитина, связанного с Н2А.

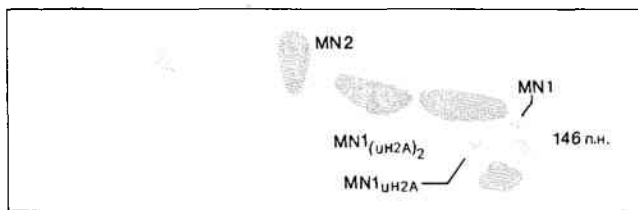


Рис. 30.12. Нуклеосомы, частицы минимальной нуклеосомы и частицы, связанные с убиквитином, можно разделить с помощью двумерного гель-электрофореза.

нуклеосом с помощью двумерного электрофореза. Сначала нуклеосомы разделяют электрофоретически на одной вертикальной дорожке. Затем электрофорез проводят под углом вправо к первоначальному направлению через всю пластинку геля. Соответствующие фракции отмечены на рис. 30.12. Здесь представлены нуклеосомы, содержащие около 170 п. н. ДНК и Н1 (фракция MN2), частицы минимальной нуклеосомы (фракция MN1) и частицы минимальной нуклеосомы, содержащие один или два конъюгата UH2A.

Образуют ли частицы с UH2A особый класс нуклеосом? Около 20% нуклеосом в хроматине *D. melanogaster* содержат UH2A. Их ДНК-последовательность можно исследовать методом блотт-гибридизации со специфическими зондами. При использовании зондов, представляющих транскрибируемые гены, около 50% материала оказывается во фракции, содержащей убиквитин. При использовании в качестве зондов сателлитной ДНК во фракцию с убиквитином попадает менее 4% исследованного препарата. Из этого следует, что UH2A отсутствует в гетерохроматине. (Негистоновый белок обнаружен только в нуклеосомах гетерохроматина *D. melanogaster*.) Конъюгат UH2A имеет тенденцию накапливаться в нуклеосомах транскрибируемых последовательностей.

Функция убиквитина изучена очень плохо. По-видимому, он высвобождается из хроматина в митозе. Этот факт приводит в замешательство, поскольку возникает вопрос: каким образом убиквитинированное состояние определенных участков хроматина наследуется дочерними клетками?

Недавно появились сообщения о возможной функции убиквитина в цитоплазме, где он принимает участие в системе деградации белков. Одна (или более) молекул убиквитина связывается с белком-мишенью в реакции, использующей АТФ. Затем белок-мишень деградирует. Следовательно, убиквитин служит маркером, используемым для идентификации субстрата системой деградации. Относится ли это и к событиям в ядре, не известно.

Экспрессия гена связана с деметилированием

В ДНК животных клеток от 2 до 7% С-остатков метилированы (конкретное значение варьирует в зависимости от вида). Сателлитная ДНК часто бывает метилирована очень сильно. Остальные метильные группы разбросаны по всему геному. Большинство метильных групп найдено в дуплетах CG (CG-GC), и действительно большинство CG-последовательностей может быть метилировано. Эта короткая последовательность представляет собой палиндром, и, как правило, нуклеотид С метилируется в обеих

цепях, образуя структуру:



Дуплет, который в отличие от этого метилирован только по одной из двух цепей, называют **полуметилированным**.

Распределение метильных групп можно исследовать, используя рестриктирующие ферменты, которые расщепляют сайты-мишени, содержащие CG-дуплеты. На рис. 30.13 сопоставляются два фермента с рестриктирующей активностью. Это **изосхизомеры**, т.е. ферменты, которые расщепляют в качестве мишени одну и ту же последовательность ДНК, но по-разному реагируют на ее метилирование.

В этой паре фермент HpaII расщепляет последовательность CCGG (записана последовательность только одной цепи ДНК). Но, если второй нуклеотид С метилирован, фермент перестает узнавать этот сайт. Однако фермент MspI расщепляет тот же самый сайт-мишень независимо от метилирования нуклеотида С. Следовательно, фермент MspI можно использовать для идентификации всех CCGG-последовательностей, а HpaII для того, чтобы определить, метилированы ли эти последовательности. Таким образом, если субстратом служит неметилированная ДНК, оба фермента образуют одни и те же рестрикционные полосы. Но если в ДНК имеются метилированные сайты, то расщепления по этим положениям под действием HpaII происходить не будет. В каждом таком положении один большой фрагмент, образующийся под действием рестриктазы HpaII, будет соответствовать двум фрагментам, образующимся под действием рестриктазы MspI. Такой пример показан на рис. 30.14.

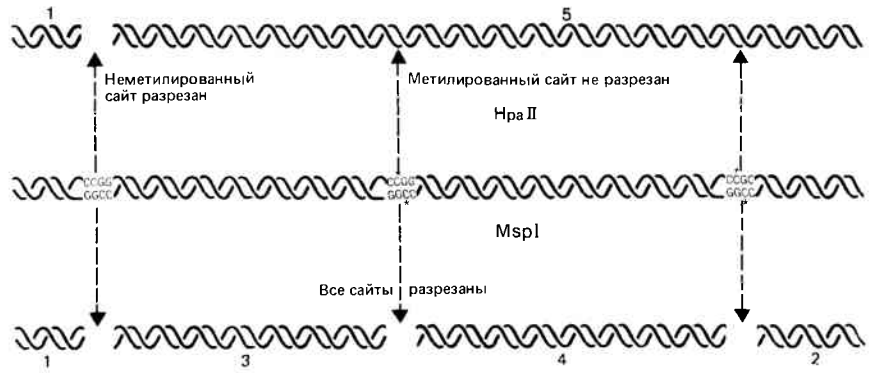
Инвариантна ли картина метилирования, или она меняется в зависимости от конкретных условий? Отдельные сайты были исследованы в нескольких случаях, в том числе в генах, кодирующих клеточный белок в тандемном кластере рДНК и в последовательности нескольких интегрированных или свободных вирусных геномов. Сайты, идентифицированные с использованием рестриктирующих ферментов, — это лишь некоторые из метилированных последовательностей, но мы полагаем, что их поведение типично для всех таких сайтов.

Полученные результаты во всех случаях были схожи. Некоторые сайты были метилированы во всех исследованных тканях, тогда как некоторые другие во всех тканях оставались неметилированными. *Меньшая часть сайтов* оказывалась метилированной в таких тканях, в которых данный ген не экспрессировался; в тех же тканях, в которых данный ген экспрессировался, эти сайты были метилированы. Таким образом, можно говорить, что активный ген «недометилирован».

В случае кластера β-глобиновых генов человека наилучшая корреляция наблюдается в генах эмбриона. Большинство (~15) идентифицированных сайтов оказывается неметилированными в эритроидной ткани с активной экспрессией, частично метилированными во взрослой эритроидной ткани и метилированными в других тканях. У взрослого организма в тканях с активной экспрессией β-глобиновых генов метилирование некоторых сайтов уменьшено лишь частично. (Под «частичным» метилированием подразумевается, что данный сайт метилирован в одних клетках, но не метилирован в других клетках того же фенотипа.)

В генах рРНК *X. laevis* большинство CG-сайтов остаются метилированными во всех случаях. Два сайта

Рис. 30.13. Рестриктурирующий фермент MspI разрезает все последовательности CCGG независимо от того, метилированы они или нет, а HpaII разрезает только неметилированные CCGG-тетрамеры.



(расположенные около промотора) деметилируются в процессе эмбрионального развития, когда начинается транскрипция рРНК.

Наряду с изучением состояния метилирования резидентных генов можно сравнивать результаты введения метилированной и неметилированной ДНК в новые клетки-хозяева. Такие эксперименты впервые были выполнены на вирусных генах, а затем на нескольких очищенных клеточных генах. В них обнаружена четкая корреляция: метилированные гены неактивны, а неметилированные гены активны.

Протяженность неметилированных областей *in vivo* была определена для кластера α -глобинового гена курицы. В клетках эритроидного ряда взрослого организма недометилирование этого гена ограничено сайтами, расположенными на участке протяженностью 500 п.н. в направлении против хода транскрипции (начиная от первого из двух β -глобиновых генов взрослого). Сайты недометилирования присутствуют во всей области, включая спейсер между генами. Область недометилирования достаточно хорошо совпадает с областью максимальной чувствительности к ДНКазе I. Это свидетельствует о том, что недометилирование может быть признаком домена, содержащего транскрибируемый ген или гены.

Интерпретация связи между недометилированием и активацией гена осложняется тем, что в этом процессе участвует меньшинство (иногда совсем незначительное) метилированных сайтов. Однако можно предположить, что критическое значение имеет метилирование специфических сайтов. Например, деметилирование промотора может оказаться достаточным для инициации транскрипции. Возможно также, что уменьшение уровня метилирования является частью некоторого структурного изменения, необходимого для начала процесса транскрипции.

В некоторых генах метилирование, связанное с экспрессией, происходит на 5'-конце. Некоторые гены не метилированы на 5'-конце, но остаются метилированными на 3'-конце во время экспрессии. Как и в случае других изменений в хроматине, кажется вероятным, что отсутствие метильных групп связано со способностью к транскрипции, а не с самим актом транскрипции. Насколько важно деметилирование промотора, было бы относительно легко проверить непосредственно, поскольку можно сконструировать *in vitro* гены, различающиеся только метилированием 5'-конца.

Существуют исключения из описанной нами общей картины. Некоторые гены могут экспрессироваться даже в сильно метилированном состоянии. Ярким исключением является локус гистосовместимости H2 в линии ранних эмбриональных клеток, который, по-видимому,

активируется в результате метилирования. Таким образом, связь между метилированием и экспрессией не универсальна.

Некоторые модели контроля метилирования

Большое внимание уделялось тому, каким образом состояние метилирования может передаваться в ряду клеточных поколений или быть изменено. В ДНК половых клеток, например сперматозоидов, каждый ген находится в неактивном состоянии, т.е. метилированы и постоянные сайты (модифицированные во всех тканях), и переменные сайты, т.е. те, которые специфически не метилированы в тканях с экспрессируемыми генами. Таким образом, отсутствие определенных метильных групп в активном состоянии представляет собой *потерю* ранее существовавших модификаций. Мы не знаем, сохраняют ли клеточные гены метильные группы после того, как они перестают экспрессироваться. Критический вопрос, на который хотелось бы получить ответ, заключается в следующем: какие последовательности выбираются в качестве мишени для тканеспецифических изменений в состоянии метилирования?

В основе простой модели, объясняющей воспроизведение метилированных сайтов, лежит предположение, что ДНК-метилаза действует на полуметилированную ДНК. Как видно из рис. 30.15, при репликации полностью метилированного дуплета CG образуются два полуметилированных дочерних дуплекса. Узнавая каждый полуметилированный сайт, фермент может переводить его в обычное, полностью метилированное состояние. Такая модель согласуется с тем фактом, что при введении мети-

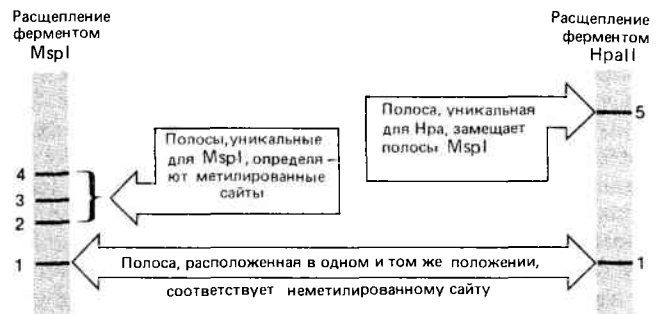


Рис. 30.14. Результаты расщепления ферментами MspI и HpaII сравниваются по положению фрагментов в гель-электрофорезе. Число полос соответствует фрагментам, идентифицированным на рис. 30.13.

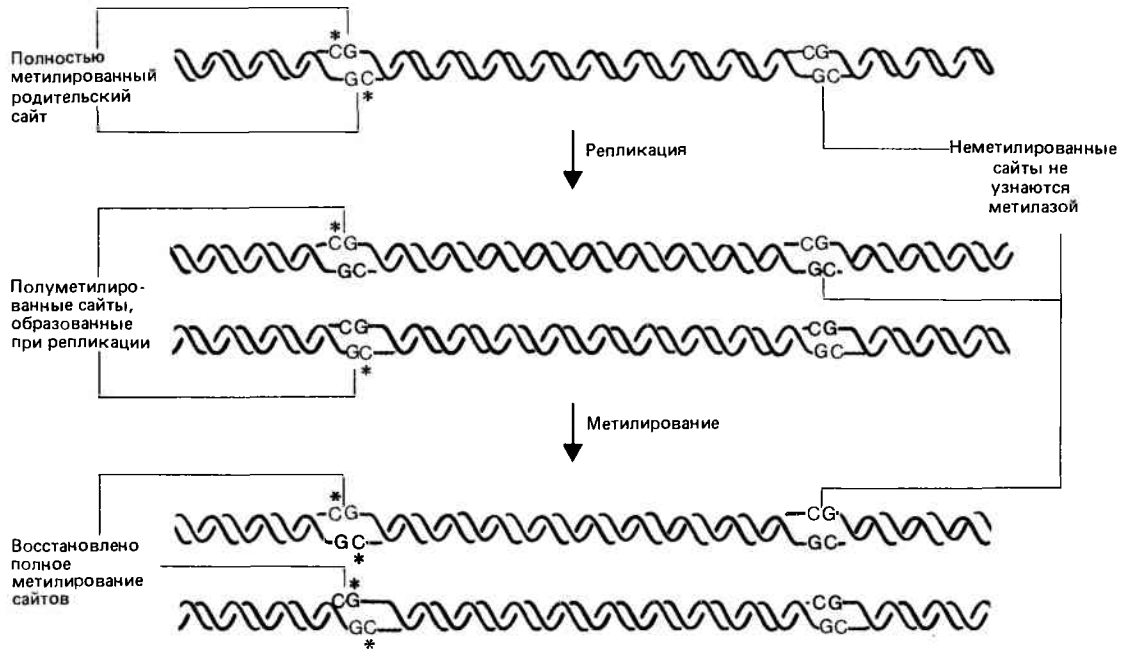


Рис. 30.15. Положение метилированных сайтов может воспроизводиться в потомстве с помощью фермента, который узнает в качестве субстрата только полуметилированные сайты.

лированной ДНК в клетку она продолжает метилироваться на протяжении неопределенного числа циклов репликации с точностью, достигающей 95% на сайт. Если вводится неметилированная ДНК, она не метилируется *de novo*. Из этого следует, что фермент узнает только полуметилированные сайты. При таком действии фермента состояние дуплета $\frac{CG}{GC}$ — метилированное или неметилированное — может наследоваться дочерними клетками.

Если это правильная модель, то совсем другая ферментативная активность должна участвовать в любом акте формирования новых сайтов метилирования (например, когда восстанавливается метилированное состояние гена, необходимое для прекращения транскрипции).

Существуют два пути, ведущие к тому, что сайт может утратить свое метилированное состояние: 1) метильная группа может быть активно удалена деметилирующим ферментом, или 2) может просто не происходить метилирования в полуметилированном сайте после репликации. Тогда в одном из дуплексов ДНК, образованных при репликации, не будет сайта метилирования.

В опытах с лекарственным препаратом 5-азациитидином получены данные, косвенно свидетельствующие о том, что деметилирование может вызвать экспрессию гена. Это соединение включается в ДНК вместо цитидина и не может быть метилировано, так как у него блокировано 5'-положение. В результате в ДНК появляются неметилированные сайты. Фенотипический эффект 5-азациитидина включает в себя индукцию изменений, нарушающих клеточную дифференцировку. Например, индуцируется развитие мышечных клеток из немuscularных клеток-предшественников. Это соединение также активирует гены в молчащей X-хромосоме, из чего можно сделать вывод о возможной роли метилирования в инактивации хромосомы.

Хотя пока мы еще не до конца разобрались в эффекте метилирования, уже очевидно, что отсутствие метильных групп ассоциируется с экспрессией гена. Однако предпо-

ложение об общей регуляторной роли метилирования в генной экспрессии имеет свои слабые стороны. Так, например, в случае *D. melanogaster* (и других двукрылых насекомых) не обнаружено метилирования ДНК. Между тем другие различия между активным и неактивным хроматином у этих организмов такие же, как и у видов, которым свойственно метилирование. Таким образом, у *Drosophila* метилирование либо не нужно, либо оно заменено каким-то другим механизмом.

Сайты, сверхчувствительные к ДНКазе, расположены перед активными промоторами

Первый эффект, наблюдаемый при расщеплении хроматина очень низкой концентрацией ДНКазы, — это индукция разрывов в двухцепочечной ДНК в специфических сайтах. В случае целого генома, например у *D. melanogaster*, в результате воздействия образуется серия специфических фрагментов, варьирующих в размере от 2 до 20 т.п.н. Эти фрагменты соответствуют участкам ДНК, расположенным между сверхчувствительными сайтами. Такие фрагменты не образуются при расщеплении свободной ДНК; они возникают (тканеспецифично) при расщеплении хроматиновой структуры.

Где локализованы эти сверхчувствительные сайты? Их положение можно установить для индивидуальных локусов, используя метод непрямого концевое мечения, о котором мы говорили ранее в связи с фазированием нуклеосом. Схема опыта воспроизведена на рис. 30.16. Хроматин расщепляют ДНКазой I, ДНК выделяют и затем расщепляют рестриктирующим ферментом. Смесь, полученную после двойного гидролиза, подвергают электрофорезу и гибридизуют с зондом, который соответствует участку, расположенному рядом с сайтом рестрикции. ДНКазы I делает разрез в определенном сайте,

образуя дискретный фрагмент, размер которого указывает расстояние между сверхчувствительным сайтом и сайтом рестрикции.

Многие сверхчувствительные сайты связаны с экспрессией генов. Каждый активный ген имеет один или (иногда) более сайтов разрезания, расположенных в участке, непосредственно прилегающем к промотору со стороны, противоположной направлению транскрипции (с 5'-конца). Некоторые примеры локализации сверхчувствительных сайтов суммированы на рис. 30.17.

5'-Сайт разрезания обнаруживается *только* в хроматине клеток с экспрессирующими генами. В неактивных генах он не обнаруживается. Например, у глобиновых генов сверхчувствительные сайты расположены в 5'-направлении от генов эмбрионального глобина в эмбриональных клетках, но не во взрослых клетках, и наоборот. Такие сайты иногда могут различаться по чувствительности (судя по интенсивности соответствующих полос). При близком расположении множественных сайтов некоторые из них оказываются значительно чувствительнее других. Помимо этих сайтов могут быть обнаружены сайты в других местах, в пределах транскрибируемого гена или на некотором расстоянии от него по ходу транскрипции (в сторону 3'-конца). Значение сайтов, расположенных в других местах, неясно.

Корреляция между транскрипцией гена и сверхчувствительными сайтами, располагающимися в направлении 5'-конца, убедительно подтверждается данными, полученными при изучении гена *Sgs4 D. melanogaster*, кодирующего клейкий белок слюнных желез. В этой ткани в данном гене существуют два особенно сверхчувствительных сайта, лежащие на расстоянии 330 и 405 п.н. от 5'-конца гена. В мутантной линии мух есть делеция протяженностью 100 п.н., захватывающая оба сверхчувствительных сайта. Несмотря на то что еще остался отрезок более 250 п.н. «выше» стартовой точки транскрипции, у этого мутанта не образуется клейкого белка. Исходя из этих результатов, можно предположить, что удаление сверхчувствительного сайта нарушает активацию гена.

Присутствие сверхчувствительного сайта хотя и необходимо, но недостаточно для того, чтобы обеспечить транскрипцию. В естественных условиях трудно установить, через какое время после появления сверхчувствительных сайтов начинается транскрипция, поскольку нелегко получить клетки, находящиеся на одной и той же стадии развития. Однако это можно выполнить на определенных клетках эритроидного ряда курицы, иммортализованных в культуре с помощью РНК-содержащего онкологического вируса. Различные клеточные линии представляют разные стадии эритропоэза. В одной из таких линий уже имеются сверхчувствительные к ДНКазе сайты, но еще не начата транскрипция глобинового гена. Эта линия, возможно, захвачена на стадии между активацией сверхчувствительного сайта и началом генной экспрессии. Все указания, однако, свидетельствуют о том, что сверхчувствительный 5'-сайт (или сайты) появляется перед тем, как начинается транскрипция, и, очень возможно, служит предпусковой инициации.

Каким образом возникают сверхчувствительные сайты? Для анализа их структуры нужна система *in vitro*. Сверхчувствительная область гена взрослого β -глобина цыпленка была воссоздана на плазмиде. Когда плаزمид, несущая этот ген, соединялась с гистонами в присутствии экстракта из ядер эритроцитов, соответствующие участки

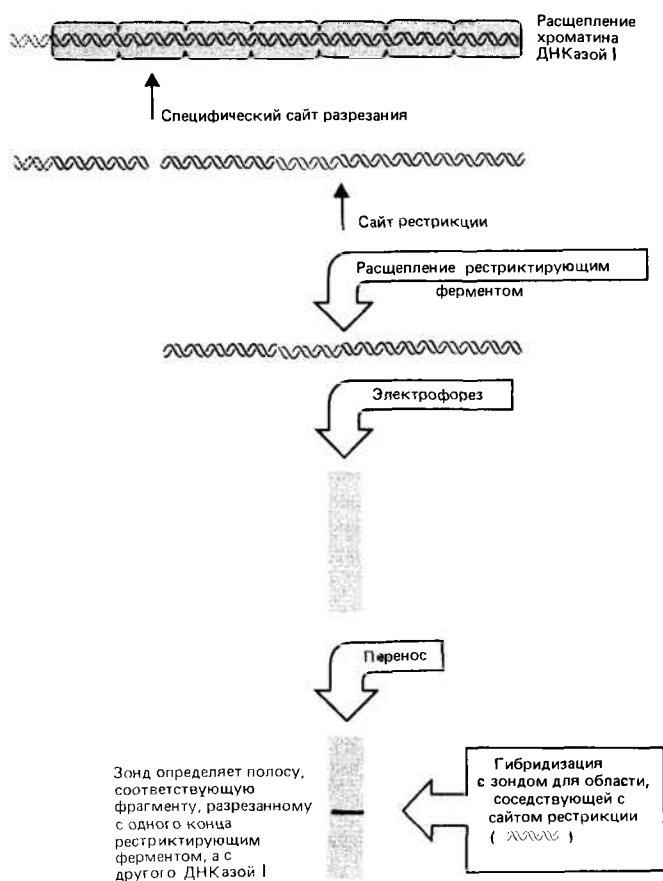


Рис. 30.16. С помощью метода непрямого концевого мечения определяют расстояние от сайта сверхчувствительности к ДНКазе до сайта рестрикционного расщепления.

становились сверхчувствительными. Очистка активного фактора должна быть достаточно простой: для определения можно использовать саму способность образования сверхчувствительных участков.

Экстракт не может придавать сверхчувствительность, если его добавлять *после* гистонов. Из этого следует, что соответствующий фактор в экстракте узнает ДНК непосредственно и каким-то образом меняет организацию данного участка, лишенного нуклеосом. По крайней мере, в этих условиях соответствующий компонент не может удалить нуклеосомы после того как они сформировались. Таким образом, мы возвращаемся обратно к исходному вопросу: каким образом изменяется структура хроматина, когда ген должен перейти в активное состояние?

Стабильность сверхчувствительного сайта была выявлена при изучении свойств куриных фибробластов, трансформированных температурочувствительными онкологическими вирусами. В этих опытах используется необычное явление: хотя фибробласты не относятся к клеткам эритроидной линии, трансформация этих клеток при нормальной температуре приводит к активации глобиновых генов. В активированных генах имеются сверхчувствительные сайты. Если трансформацию проводят при более высокой (непермиссивной) температуре, не происходит ни активации глобиновых генов, ни появления сверхчувствительных сайтов. В тех случаях, когда глобиновые гены были активированы трансформацией при низкой температуре, их можно инактивировать повыше-

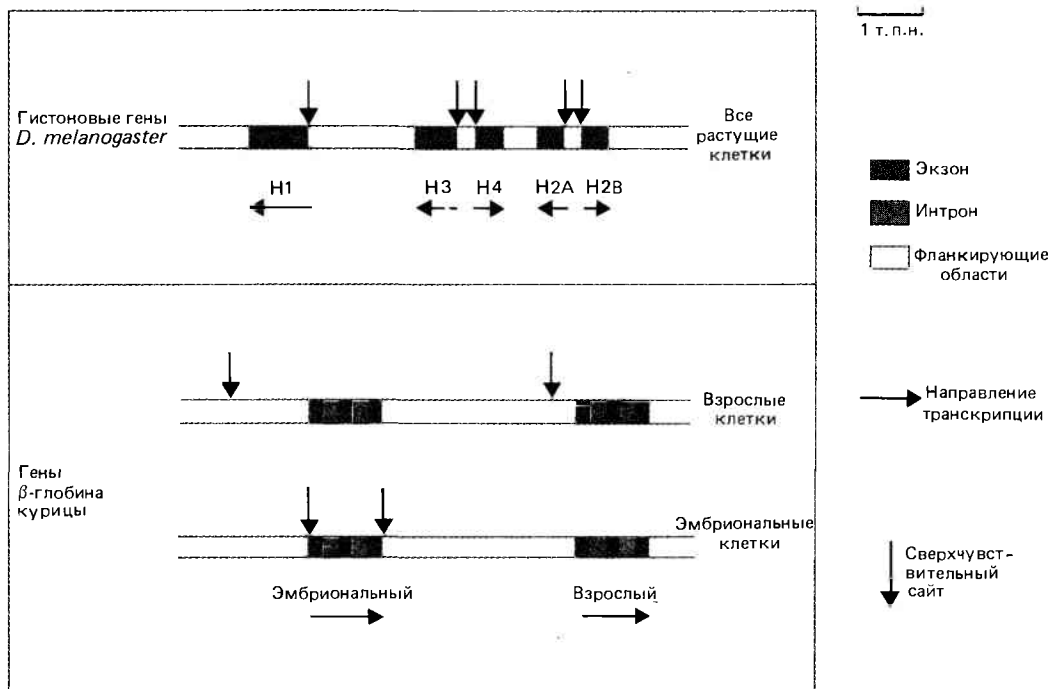


Рис. 30.17. Сверхчувствительные сайты располагаются около 5'-конца активных генов, а (иногда) также в других местах.

нием температуры. Но сверхчувствительные сайты сохраняются на протяжении по крайней мере 20 следующих клеточных делений.

Этот результат еще раз показывает, что образование сверхчувствительного сайта — это лишь одно из условий, необходимых для инициирования транскрипции. Из этого следует, что события, связанные с образованием сверхчувствительного сайта, отличаются от событий, связанных с его воспроизведением в потомстве. Однажды образованный сайт воспроизводится при репликации в отсутствие условий, необходимых для его индукции. (Вероятно, поэтому требуется специфическое вмешательство, для того чтобы удалить из молекулы сверхчувствительный сайт.)

Каким образом сверхчувствительные (или другие) сайты передаются следующему поколению при репликации ДНК? Можно представить разнообразные механизмы. Наследственное изменение может произойти в самой ДНК — например, деметилирование. В регулярной двойной спирали может происходить структурное изменение (при участии белков, восстанавливающих структуру после каждой репликации). Высококооперативные белки могут образовывать комплекс (с двухцепочечной ДНК), который способен распадаться на отдельные компоненты при репликации, а потом восстанавливаться (см. рис. 29.20).

Мы рассмотрели три изменения, обнаруженные в активных генах: 1) появление сверхчувствительного сайта (сайтов) около промотора; 2) более высокая чувствительность нуклеосомного домена, включающего транскриби-

руемую область к ДНКазе I; 3) недометилирование ДНК в транскрибируемой области. Все эти изменения могут быть необходимы для транскрипции.

В сверхчувствительных участках нет нуклеосом

Какова природа сверхчувствительных сайтов? Каждый из них представляет собой область размером 100–200 п. н., доступ к которой не ограничивается типичным для нуклеосом способом. Некоторые из этих сайтов изучены детально.

Сайт на 5'-конце β-глобинового гена взрослого типа у курицы предпочтительно расщепляется несколькими ферментами, включая ДНКазу I, ДНКазу II и нуклеазу микрококков. На карте, изображенной на рис. 30.18, показано, что сайты, предпочитаемые этими ферментами, лежат довольно близко друг от друга в пределах одной области.

Рестриктирующий фермент MspI расщепляет два сайта этой области. Часть фрагментов, образованных этим ферментом, оказывается в виде свободной ДНК. Таким образом, в период транскрипции область, расположенная от –60 до –260, становится предпочтительно чувствительной к нуклеазам.

В сходных экспериментах показано, что 5'-сверхчувствительный сайт гена тимидин-киназы вируса герпеса занимает область от –4 до –182. Сайты-мишени этой

Рис. 30.18. Сверхчувствительный сайт β-глобинового гена курицы занимает область, чувствительную к нескольким нуклеазам.

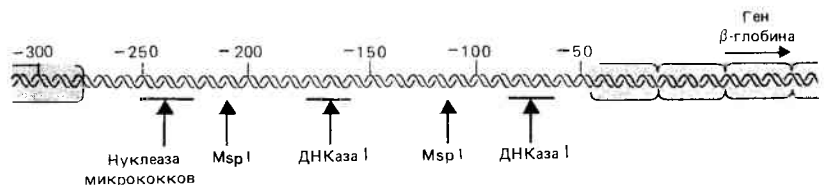




Рис. 30.19. В мини-хромосоме вируса SV40 может быть безнуклеосомный пробел.

(Фотография любезно предоставлена Moshe Yaniv.)

области предпочтительно расщепляются рестриктирующими ферментами.

ДНК сверхчувствительных сайтов активных генов глобина у кур расщепляются нуклеазой S1, субстратом которой служит одноцепочечная ДНК. Исходя из этого, было высказано предположение, что само состояние двухцепочечной ДНК в сверхчувствительном сайте может быть модифицировано.

В сверхчувствительной области гена β -глобина цыпленка нуклеаза S1 расщепляет сайты, лежащие в участках с необычной последовательностью. Это отрезки ДНК, одна из цепей которых целиком состоит из пуринов, а другая из пиримидинов. Некоторые особенности сайтов расщепления свидетельствуют о том, что они представляют собой не одноцепочечные ДНК: их узнают рестриктирующие ферменты, действующие только на двухцепочечную ДНК, и не узнает фермент, специфичный к одноцепочечной ДНК. Из этого можно, очевидно, сделать вывод, что такие сайты имеют несколько необычную структуру, что позволяет нуклеазе S1 узнавать их, хотя при этом не образуется настоящих одноцепочечных участков.

Сверхчувствительный сайт представляет собой изменение структуры хроматина. Есть основания думать, что такой сайт состоит из последовательности, не участвующей в образовании обычной нуклеосомной структуры, возможно, потому, что с этой ДНК связываются другие белки. Если нуклеосомы все-таки образуются в сверхчувствительной области, их структура должна отличаться от обычной.

Важная особенность сверхчувствительных сайтов заключается в том, что они могут проявлять свою активность на различных расстояниях от точки начала транскрипции. В гене *Sgs D. melanogaster* сайт удален более чем на 250 п.н. в 5'-направлении; в гене β -глобина курицы он расположен непосредственно рядом с 5'-сторонами, а в гене тимидин-киназы эта область включает в себя все последовательности, необходимые для функции промотора. Сверхчувствительные сайты могут участвовать в промоторной функции непосредственно, вероятнее всего создавая необходимые условия для узнавания РНК-полимеразного участка, расположенного перед промото-

ром (гл. 11). Но можно предположить также, что эти сайты играют более общую роль в активации гена — например индуцируя локальное структурное изменение, необходимое для связывания РНК-полимеразы или даже для ее синтетической активности.

Особенно хорошо изучена чувствительная к нуклеазе область мини-хромосомы вируса SV40. Короткий сегмент около точки начала репликации, непосредственно предшествующий промотору для поздней единицы репликации, предпочтительно расщепляется ДНКазой I, нуклеазой микрококков и другими нуклеазами (в том числе рестриктирующими ферментами). По минимальной оценке область предпочтительного расщепления имеет в длину 400 п.н., причем этот фрагмент может быть вырезан в виде свободной ДНК.

Состояние мини-хромосомы вируса SV40 можно увидеть под электронным микроскопом. В участке, составляющем до 20% всего образца, заметен пробел в нуклеосомной организации. Это хорошо видно на электронной микрофотографии (рис. 30.19). Пробел длиной около 120 нм (примерно 350 п.н.) с обеих сторон окружен нуклеосомами, занимающими весь остальной геном. Положение пробела можно определить, расщепив кольцевую мини-хромосому с помощью рестриктирующего фермента, для которого известен лишь один сайт-мишень. Видимый пробел соответствует области, чувствительной к нуклеазам. Это прямо показывает, что повышенная чувствительность к нуклеазам коррелирует с отсутствием нуклеосом.

Отсюда, однако, не следует, что данная ДНК свободна от нуклеосом *in vivo*, поскольку вполне возможно, что при выделении мини-хромосомы была потеряна какая-то другая белковая структура.

Подобный нуклеазочувствительный пробел в нуклеосомной организации найден у родственного вируса полиомы. В этом случае было показано, что различные участки этой области не в одинаковой степени чувствительны к нуклеазам. Внутри чувствительной области размером 260 п.н. находятся два сверхчувствительных сайта и «защищенный» участок. Карта этой области показана на рис. 30.20.

Защищенный участок, вероятно, представляет собой ДНК, связанную с белком. И действительно, вполне возможно, что наличие таких комплексов обуславливает отсутствие нуклеосом в данной области.

Точно такой же результат наблюдают у мутанта вируса полиомы с существенными перестройками в этой области. Из сказанного следует, что наличие без нуклеосомного пробела определяется не всей последовательностью ДНК в самой чувствительной области, а должно быть «отмерено» некоторым образом от фиксированной

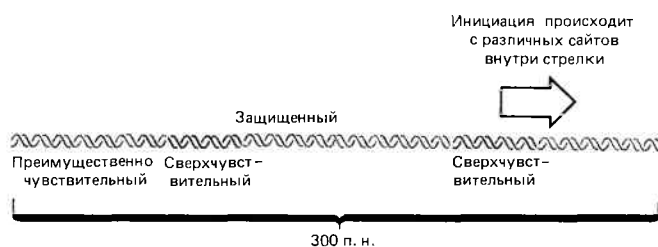


Рис. 30.20. В безнуклеосомном пробеле вируса SV40 имеются сверхчувствительные сайты, чувствительные участки и «защищенная» область ДНК.

точки, являющейся либо частью этой области, либо находящейся рядом с ней.

Область вируса SV40 или полиомы, окаймляющая чувствительный к нуклеазе пробел, имеет несколько потенциальных функций, и, следовательно, ее нельзя просто считать структурой, активирующей промотор. Наличие пробела может быть связано с последовательностями элементов усилителя транскрипции (энхансер, *enhancer*), которые находятся в этой области и которые необходимы для функции промотора (гл. 11).

Интересная структура была обнаружена во вне хромосомной рДНК макронуклеуса *Tetrahymena pyriformis*. Напомним, что молекулы рДНК представляют собой линейные палиндромы, в которых две транскрипционные единицы рДНК расположены на некотором расстоянии друг от друга и транскрибируются в противоположных направлениях. Изучение организации рДНК показало, что близко от центра палиндрома расположены два участка начала репликации. Это короткие последовательности, с которых начинается репликация ДНК. Структура хроматина в этой области показана в утолщенной части карты.

В этой области находятся несколько сверхчувствительных сайтов, по одному сайту на каждый промотор и каждый участок начала репликации. Между двумя участками начала репликации располагается серия из 5 нуклеосом. Участки начала репликации разделены промежутками в 1000 п.н., и участок, приходящийся на каждую нуклеосому, равен точно 200 п.н., так что каждая нуклеосома находится в точно определенном положении. У организма, родственного *T. pyriformis*, центральная область между участками начала репликации составляет 1400 п.н. и заполнена семью нуклеосомами.

рДНК *Tetrahymena* служит совершенным примером нуклеосомного фазирования. Тот факт, что две копии рДНК различаются целым числом нуклеосомных повторов, дает основание думать, что, по-видимому, давление отбора способствовало отбору вариантов, у которых центральная область упакована в строго определенное число нуклеосом. Локализация гистоновых октамеров, вероятно, определяется границами соседних сверхчувствительных сайтов, а не свойствами самой фазируемой последовательности.

Подобные структурные изменения хроматина могут происходить у дрожжей в области центромеры. В данном случае кажется вероятным, что вместо видимого кинетохора сама нить хроматина выполняет функцию центромеры. Строение центромерной области было исследовано методом непрямого концевое мечения с использованием коротких фрагментов ДНК в качестве зондов. С помощью зондов были идентифицированы два сверхчувствительных к ДНКазе I сайта, локализованные с обеих сторон от областей I и III. Это короткие консервативные последовательности центромерной ДНК, приведенные на рис. 28.11. Область размером 220–250 п.н. между двумя сверхчувствительными сайтами защищена от нуклеазного расщепления.

Роль консервативной области, возможно, заключается в связывании специфических белков и образовании безнуклеосомной структуры.

При расщеплении хроматина нуклеазой микрококков в том же самом эксперименте наблюдается другой эффект. В этом случае также обнаружено наличие защищенной области, но, кроме того, с обеих сторон от нее располагаются фазированные серии нуклеосом. Фазирование

захватывает более 12 нуклеосом, каждая из которых содержит ДНК длиной 160 п.н. Этот результат находится в противоречии с очевидной утратой фазирования нуклеосом в других участках дрожжевого генома.

Распространяется ли фазирование нуклеосом от центромерной последовательности, или же это свойство самих фазированных последовательностей? В плазмиде, содержащей эти последовательности, но не имеющей центромерной области, фазирование сохраняется. Из этого следует, что естественная центромерная область характеризуется двумя особенностями. Консервативная последовательность образует некую безнуклеосомную структуру. А фланкирующим последовательностям присуща способность фазировать свои нуклеосомы.

Что происходит с нуклеосомами, расположенными на конце хромосомы? При расщеплении теломер инфузории *Oxytricha* нуклеазой микрококков получается серия повторяющихся полос на расстояниях 100, 300, 500, 700, 900 п.н. от конца.

На основе такого разделения делают вывод, что самые последние 100 п.н. теломеры защищены нуклеосомным белковым комплексом, рядом с которым располагается серия фазированных нуклеосом.

Наличие сверхчувствительных сайтов может отражать более общее явление. Возможно, что особые (пока еще неизвестные) белки модифицируют короткие участки ДНК так, что нуклеосомы не образуются. Структуры, сформированные в каждом отдельном случае, не обязательно должны быть похожи между собой (за исключением того, что каждая, по определению, образует сайт сверхчувствительности к ДНКазе I). Мы видели, что сайты, характеризующиеся сверхчувствительностью, предшествуют транскрибируемым генам в нуклеазочувствительном пробеле вирусов SV40 и полиомы и в дрожжевой центромере. Такие сайты могут быть связаны с транспозицией генов типа спаривания у дрожжей (гл. 37).

Эти различные структуры могут служить многим целям, и будет неудивительно, если их обнаружат в других подобных ситуациях.

Предположения о природе активации гена

Можно лишь рассуждать о том, будет ли справедливым суждение, обратное этому принципу: *гены, регуляторные участки которых организованы в нуклеосомы, не могут экспрессироваться*. Предположим, что образование нуклеосом происходит независимо от последовательности в любом участке ДНК, из которого специально удалены гистоны. Тогда в отсутствие специфических регуляторных белков промоторы, усилители транскрипции и другие регуляторные участки будут организованы с помощью гистоновых октамеров в такое состояние, в котором они, возможно, *не могут быть активированы*. (Доказательств существования какого-либо белка, способного удалять гистоны с ДНК, нет.)

Имеются некоторые основания думать, что для активации гена необходимо нарушение структуры хроматина. Фактор транскрипции гена 5S-РНК не активирует *in vitro* гены, если они находятся в комплексе с гистонами. Однако этот фактор может образовывать необходимый комплекс со свободной ДНК, после чего добавление гисто-

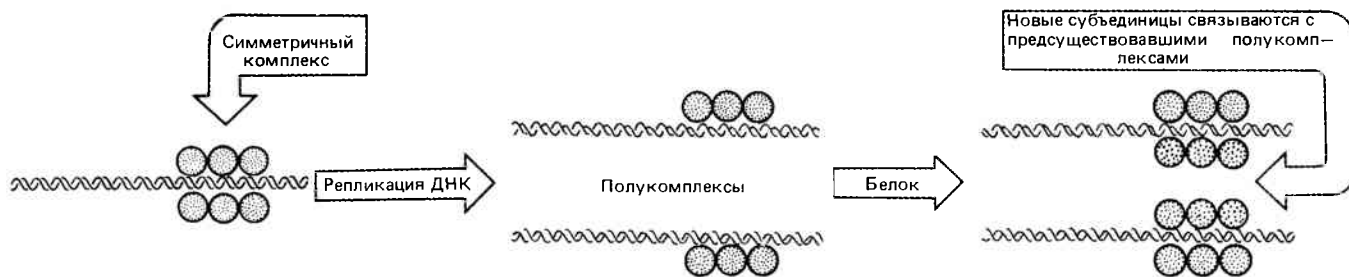


Рис. 30.21. Белковый комплекс может воспроизводиться на ДНК, расщепляясь на полукомплексы при репликации и ко-

оперативно подбирая затем недостающие белковые молекулы для восстановления полных комплексов.

нов не снимает остаточной активности генов. Уже связанный с ДНК фактор остается на своем месте, позволяя молекулам РНК-полимеразы последовательно инициировать транскрипцию. Критичным в этой ситуации может оказаться то, кто первым (фактор транскрипции или гистоны) свяжется с регуляторным сайтом.

При репликации происходит очевидное нарушение структуры хроматина, так как нуклеосомы должны временно отсутствовать по крайней мере в одной из дублированных хромосом. При этом другие белки могут связываться с регуляторным сайтом и препятствовать образованию нуклеосом. Безнуклеосомный участок может быть узан по сверхчувствительному сайту и может служить границей, с которой начинается фазирование нуклеосом.

Каким образом можно проверить предположение, что изменение состояния хроматина часто связано с репликацией? Системой нужного нам типа могла бы быть (к примеру) индукция в генах сверхчувствительных сайтов в ответ на действие гормонов. Однако четкий ответ на вопрос о том, как образуются активные районы, будет зависеть от разработки таких систем, которые можно будет использовать для активации структуры хроматина *in vitro* (что может оказаться трудным, если репликация ДНК действительно необходима).

Сверхчувствительный сайт (или какая-то другая активизирующая структура) сформированный однажды, может сохраняться при репликации. Это может происходить в том случае, если фактор способен сегрегировать при репликации.

На рис. 30.21 показан гипотетический случай: белок, кооперативно связывающийся с ДНК, образует комплекс, который расщепляется при репликации. Каждый полукомплекс затем достраивается в полный на дочерних хромосомах.

Поскольку реконструкция комплекса обязательна при каждой репликации, можно предположить, что ген, возможно, выключается в делящейся клетке просто из-за ограниченного количества фактора. Если при делении фактор разводится, комплекс не сможет образовываться снова и гистоны образуют нуклеосомы в регуляторных участках. Этим можно объяснить, как происходит выключение генов в процессе эмбрионального развития. Отсюда возникает вопрос, каким образом могут выключаться гены в неделящихся клетках. Действительно, могут ли выключаться гены в терминально дифференцированных клетках (т.е. в клетках, достигших конечной стадии фенотипического выражения)?

Точный механизм активации (или инактивации) хроматина пока еще слишком гипотетичен, однако общая нить рассуждений сводится к тому, что активное и неактивное состояния хроматина не могут находиться в динамическом равновесии: чтобы перевести одно состояние в другое нужны неожиданные, внезапные события. Исходя из рассмотренных нами типов моделей можно сделать вывод, что для инициирования активации, возможно, нужна гораздо большая концентрация регуляторного белка, чем для сохранения активного состояния хроматина.

События, участвующие в активации гена, не должны совпадать с инициированием транскрипции. Образование соответствующего комплекса может обусловить способность гена к транскрипции, но чтобы транскрипция действительно произошла, может понадобиться дополнительное условие. На языке биологии развития образование комплекса может означать процесс детерминации, происходящий несколько раньше процесса дифференцировки, который действительно включает ген. Возможным примером могут служить, гены, экспрессия которых зависит от гормонов.

В соответствующей клетке они сначала могут прийти в состояние детерминированности, а затем при появлении гормона может начаться их транскрипция. (В таком случае гены прокариот контролируются по механизму типа дифференцировки, так как их активация совпадает с началом транскрипции, а не приводит сначала к образованию компетентного состояния).

Рекомендуемая литература

Темы, затронутые в этой главе, рассматриваются Льюином (Lewin, Gene Expression, 2, Eucariotic Chromosomes Wiley, New York, 394–427, 1980). Данные по функциональным аспектам хроматина представлены Иго-Кеменесем, Хорцем и Цахау (Igo-Kamenes, Horz and Zachau, Ann. Rev. Biochem., 51, 89–121, 1982). Существование сверхчувствительных сайтов кратко обсуждается в работе Элгин (Elgin, Cell, 27, 413–415, 1981). Их наследованию дочерними клетками посвящена статья Граудина и Вайнтрауба (Groudine, Weintraub, Cell, 30, 131–139, 1982). Серию наиболее результативных экспериментов, посвященных исследованию изменений в транскрибируемых генах, описали Вайнтрауб, Ларсен и Граудинг (Weintraub, Larsen, Groudine, Cell, 24, 333–344, 1981), а также Вайсброт, Грау-

дин и Вайнтрауб (*Weisbrod, Groudin, Weintraub, Cell, 19, 289–301, 1980*). Моделям метилирования посвящен обзор Разина и Риггса (*Rasin, Riggs, Science, 210, 609–610, 1980*). Результаты исследования вирусных безнуклеосомных пробелов собраны и подытожены в сообщении Хербоме-

ла и др. (*Herbomel et al., Cell, 25, 651–658, 1981*). Структуру центромеры выяснили Блум и Карбон (*Bloom, Carbon, Cell, 29, 305–317, 1982*). Великолепный обзор, посвященный вопросам формированного и поддержанию структуры активного гена приведен Брауном (*Brown, Cell, 37, 359–365, 1984*.)

Часть IX

СОХРАНЕНИЕ ДНК В РЯДУ ПОКОЛЕНИЙ

Если истинно утверждение, что сущность жизни состоит в накоплении и передаче опыта от поколения к поколению, то ключевой проблемой биологии, по-видимому, можно считать вопрос о том, как увековечивает свой опыт живая материя. Взгляните на бактерию, находящуюся в большом объеме питательной жидкости. Она ассимилирует растворенные в среде вещества, растет и делится. Две дочерние особи делают то же самое. Иногда при этом возникает ошибка и появляется клетка с несколько отличающимися свойствами. Она дает начало таким же измененным потомкам. Открытие репродукции и мутаций позволяет верить, что игра эволюции продолжается и разнообразие форм будет безгранично умножаться.

Макс Дельбрюк, 1949

РЕПЛИКОН: ЕДИНИЦА РЕПЛИКАЦИИ

Независимо от того, содержит клетка только одну хромосому (как у прокариот) или много хромосом (как у эукариот) за период времени, соответствующий одному клеточному делению, весь геном должен быть реплицирован только один раз. Каким образом акт репликации связан с клеточным циклом?

Инициация репликации неизбежно влечет за собой дальнейшее деление прокариотической и эукариотической клетки. С этой точки зрения число потомков клетки определяется серией принимаемых ею решений, инициировать или не инициировать репликацию ДНК. Если репликация началась, то до ее завершения последующее деление клетки не сможет произойти. Именно завершение репликации, по-видимому, служит сигналом для клеточного деления. Дублированные геномы сегрегируют по одному в каждую дочернюю клетку. У эукариот это происходит в процессе митоза, у прокариот используется какой-то другой механизм. В обоих случаях единицей сегрегации является хромосома.

Единица, с помощью которой клетка контролирует отдельные акты репликации, получила название **репликона**. Каждый репликон в каждом клеточном цикле «возбуждается» только один раз. В нем обязательно должны присутствовать необходимые для репликации контролирующие элементы: **точка начала** (*origin*), в которой инициируется репликация, **точка окончания** (*terminus*), в которой репликация останавливается. Любая последовательность, присоединенная к точке начала (или, более точно, не отделенная от точки начала точкой окончания), реплицируется как часть этого репликона. Говоря языком генетики, точка начала представляет собой *цис-*действующий сайт, способный влиять только на ту молекулу ДНК, физической частью которой он является.

Событие репликации контролируется на стадии инициации. Однажды начавшись, репликация продолжается до тех пор, пока весь репликон не будет дублирован. Частота инициации контролируется взаимодействием регуляторного белка (белков) с точкой репликации. Поэтому первоначальное определение репликона (у прокариот) рассматривало его как обладателя точки начала и гена, кодирующего регуляторный белок. Теперь, однако, термин «репликон» часто применяется в отношении эукариотических хромосом для обозначения единицы репликации, содержащей только одну точку начала; любой необходимый регуляторный белок (или белки) может быть обеспечен другой контролирующей единицей.

Бактериальная хромосома содержит один репликон: следовательно, единицы репликации и сегрегации совпадают. Инициация в одной точке начала репликации ведет к репликации всего генома. Наряду с хромосомой бактерии могут содержать плазмиды; каждая плазида представляет собой автономный кольцевой геном и является отдельным репликоном. Некоторые плазмиды разделяют строгость репликации бактериальной хромосомы: они представлены в бактериальной клетке в виде одной копии на каждую копию хромосомы (**однокопийные плазмиды**). Другие плазмиды являются **многокопийными**, количество их копий в клетке превышает 1. ДНК каждой фаговой или вирусной частицы также составляет репликон, спо-

собный многократно инициироваться во время инфекционного цикла. По-видимому, лучше всего характеризует репликон прокариот следующее определение: любая молекула ДНК, которая содержит точку начала репликации, способна автономно реплицироваться в клетке. Число репликационных событий будет зависеть от взаимодействия точки начала с соответствующими регуляторными белками. (Исключение составляют случаи, когда некоторые независимые репликоны физически встраиваются в бактериальную хромосому; при этом действие их собственных точек начала репликации подавляется и такая единица реплицируется как часть бактериальной хромосомы.)

Эукариотическая хромосома содержит большое число репликонов, т. е. единица сегрегации включает много единиц репликации. Этот факт заставляет по-иному взглянуть на проблему контроля. За один клеточный цикл все репликоны хромосомы должны быть активированы, однако они не становятся активными одновременно. Это происходит на протяжении определенного периода. В то же время каждый из этих репликонов в течение клеточного цикла должен быть активирован *только один раз*. По-видимому, должен существовать определенный сигнал для того, чтобы отличить реплицированный репликон от нереплицированного и показать, когда именно процесс репликации должен быть завершён.

Нам хотелось бы знать последовательности нуклеотидов, которые функционируют в качестве точек начала репликации, и способы их узнавания соответствующими белками аппарата репликации. Кроме того, необходимо ответить на вопрос, что именно узнается: двухцепочечная нуклеотидная последовательность или какая-то альтернативная вторичная структура. Очень важно иметь представление и о том, с помощью какого механизма регуляторным белкам удастся инициировать только один цикл репликации на каждый цикл клеточного деления.

Синтез ДНК является последовательным и полуконсервативным

Синтез ДНК начинается с расхождений цепей, каждая из которых служит матрицей для синтеза дочерней цепи (как мы уже видели при рассмотрении структуры ДНК в гл. 2). Продуктами раунда репликации являются две дочерние двухцепочечные молекулы, каждая из которых состоит из одной родительской (исходной) и одной (вновь синтезированной) дочерней цепи. Таким образом, от одного поколения к другому передается одна из двух цепей, составляющих родительскую молекулу ДНК. Такой способ репликации носит название **полуконсервативного**. Схематически он изображен на **рис. 31.1**.

Если родительские цепи отличаются от вновь синтезируемых по плотности — старые цепи «тяжелые», а новые «легкие», — то первое поколение реплицировавшихся ДНК должно быть представлено «гибридными» двухцепочечными молекулами. Второе поколение, согласно этой модели, будет состоять из одинакового количества гиб-

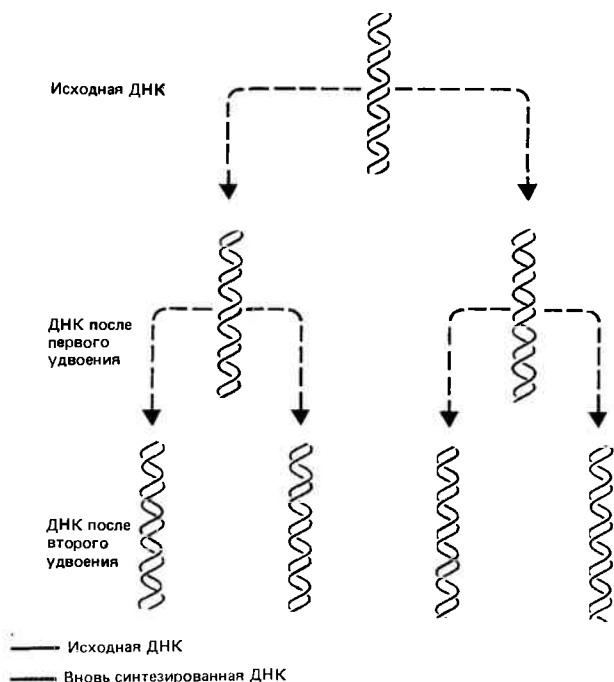


Рис. 31.1. Полуконсервативный механизм репликации ДНК.

ридных и легких двухцепочечных молекул, причем их индивидуальные отдельные цепи останутся тяжелыми или легкими. Полуконсервативный механизм был подтвержден экспериментально Меселсоном и Сталем в 1958 г. при изучении репликации ДНК *E. coli*.

Рассмотрим молекулу ДНК, вступающую в репликацию. Ее нереплицированная область представляет собой родительскую двухцепочечную ДНК, которая открывается в уже реплицировавшейся области, состоящей из двух дочерних двухцепочечных ДНК. Точка, в которой происходит репликация, получила название **репликационной вилки** (иногда называемой **точкой роста**). Репликационная вилка движется *последовательно* вдоль ДНК от ее стартовой точки (рис. 31.2).

Репликация может осуществляться либо в одном, либо в двух направлениях. На рис. 31.3 показано, что это зависит от количества репликационных вилок, которые отходят от точки начала репликации. При **однонаправленной** репликации вдоль ДНК движется одна репликационная вилка. При **двунаправленной** репликации от точки начала в противоположных направлениях расходятся две репликационные вилки.

С помощью электронной микроскопии реплицирующейся ДНК удалось установить, что область, которая уже реплицирована, имеет вид «**глазка**» внутри нереплицировавшейся ДНК. Этот глазок выглядит одинаково

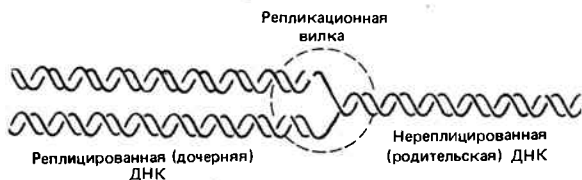


Рис. 31.2. Репликационная вилка—это область между нерасплетенной родительской двухцепочечной ДНК и вновь синтезированными дочерними двойными цепями ДНК.

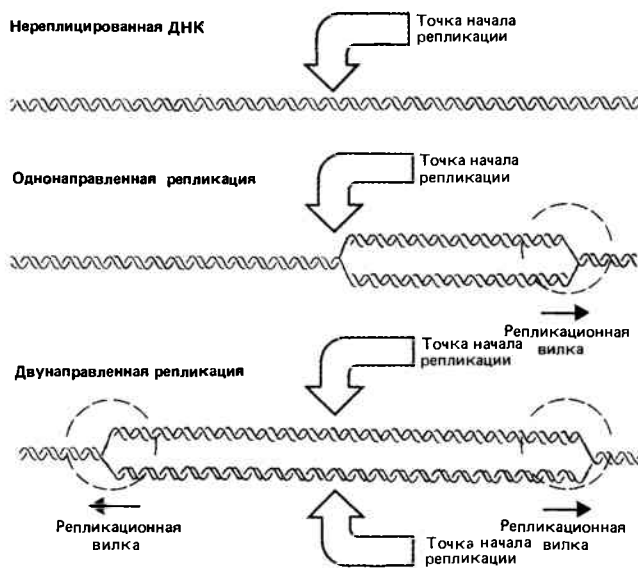


Рис. 31.3. Репликация может быть однонаправленной или двунаправленной в зависимости от того, одна или две репликационные вилки образуются в точке начала репликации.

при однонаправленной и двунаправленной репликации. Как изображено на рис. 31.4, он может представлять любую из двух структур. При **однонаправленной репликации** глазок имеет фиксированную точку начала и движущуюся репликационную вилку; при **двунаправленной** репликации он представлен двумя репликационными вилок. В любом случае продолжающаяся репликация расширяет глазок до тех пор, пока он не включит в себя весь репликон.

В случае кольцевого репликона присутствие глазка способствует образованию θ -структуры, изображенной на рис. 31.5. Последовательные стадии репликации кольцевой ДНК вируса полиомы, зафиксированные под электронным микроскопом, показаны на рис. 31.6. На микрофотографиях хорошо видно, что **глазок** увеличивается,

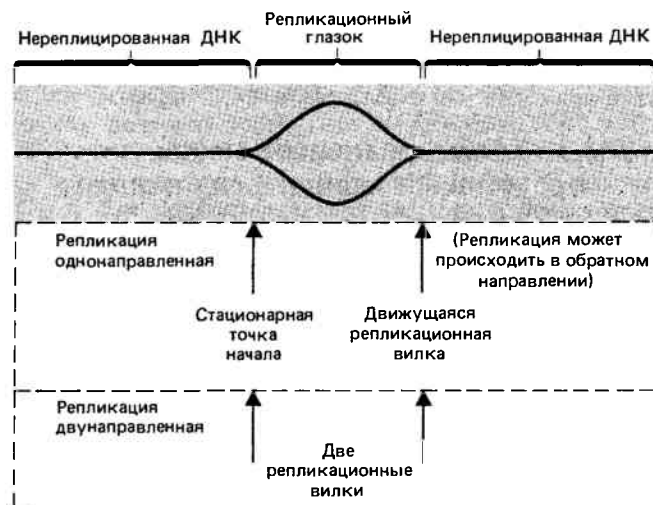


Рис. 31.4. Репликационный глазок может присутствовать при однонаправленной и при двунаправленной репликации.

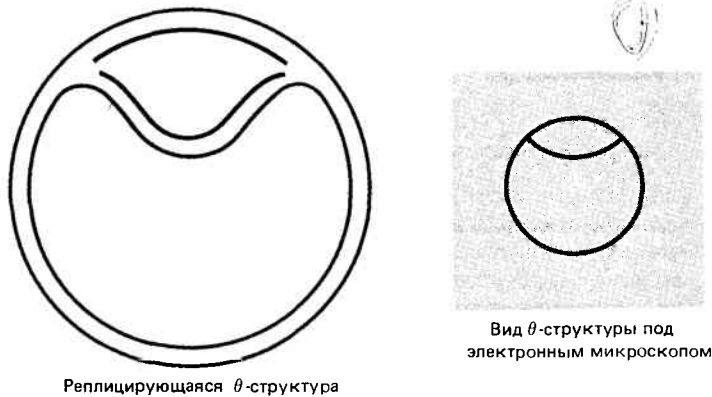
Реплицирующаяся θ -структура

Рис. 31.5. В кольцевой ДНК репликационный глазок образует θ -структуру.

а нереплицированный сегмент уменьшается. (В скрученных структурах можно определить обе стороны глазка, так как они имеют одинаковую длину.)

Содержит ли реплицирующийся глазок одну или две вилки, можно узнать двумя способами. Выбор метода зависит от того, является ли ДНК определенной молекулой или неопределенной областью клеточного генома.

В случае определенной линейной молекулы мы можем использовать электронную микроскопию для измерения расстояния от каждого конца глазка до конца ДНК. Затем положения концов глазков в молекулах, имеющих эти структуры различных размеров, можно сравнить. Если репликация однонаправленная, только один из концов будет подвижным, другой окажется фиксированным в точке начала репликации. Если репликация двунаправленная, перемещаться будут оба конца и точка начала репликации окажется равноудаленной от них. Определенную кольцевую молекулу ДНК необходимо перевести в линейную форму до электронной микроскопии (путем обработки рестрицирующим ферментом, разрезающим ее в уникальном сайте).

Для того чтобы пометить движение репликационных вилок неопределенных областей больших геномов, используют следующие друг за другом радиоактивные импульсы разной интенсивности. Радиоавтография таких препаратов показывает, что однонаправленная репликация ведет к появлению одного типа метки, сменяемого другим, в одном конце «глазка». Двунаправленная репликация приводит к появлению симметричных меченых участков с обеих сторон этой структуры (рис. 31.7).

Бактериальный геном представлен одним репликоном

Геном *E. coli* реплицируется двунаправленно от одной точки начала, определенной как генетический локус *oriC*. Первоначально точка начала была идентифицирована непрямими методами, но в настоящее время выделена в виде фрагмента ДНК.

В популяции бактерий число копий генов, находящихся вблизи точки начала, увеличено, так как эта область реплицируется рано в клеточном цикле. Следовательно, локализацию точки начала репликации и способ репликации можно распознать по увеличенному числу генов, расположенных поблизости. Если репликация двунаправленная, число их уменьшается по мере увеличения удаленности от точки начала в любом направлении.

Если клетки синхронизированы и репликация в них начинается в одно и то же время, точку начала можно идентифицировать более точно, как сайт, в который включается короткая радиоактивная метка. Этот метод используют для выявления истинных стартовых точек репликации.

ДНК области начала репликации может быть выделена благодаря ее способности поддерживать репликацию любой последовательности ДНК, к которой она присоединена. Принцип такого подхода состоит в клонировании ДНК из области, несущей точку начала репликации, в такой молекуле ДНК, которая имеет подходящие генетические маркеры, но утратила точку начала. Подобная реконструкция приведет к образованию плазмиды, способной автономно реплицироваться лишь в том случае, если ДНК из области точки начала репликации будет содержать все необходимые для функционирования последовательности. (Такой подход был использован для идентификации центромерной или теломерной ДНК у дрожжей по их влиянию на выживаемость плазмиды; гл. 28.)

Анализ последовательностей, соответствующих точкам начала репликации, связан с определенными трудностями, поскольку плазмиды, несущие *oriC*, отличаются значительной нестабильностью. Однако в некоторых бактериальных штаммах такие плазмиды способны сохраняться.

Рекомбинантные плазмиды, в которые включен участок ДНК длиной около 1 т.п.н. из области *oriC*, поддерживаются в клетках *E. coli* в количестве 1–2 копий на бакте-

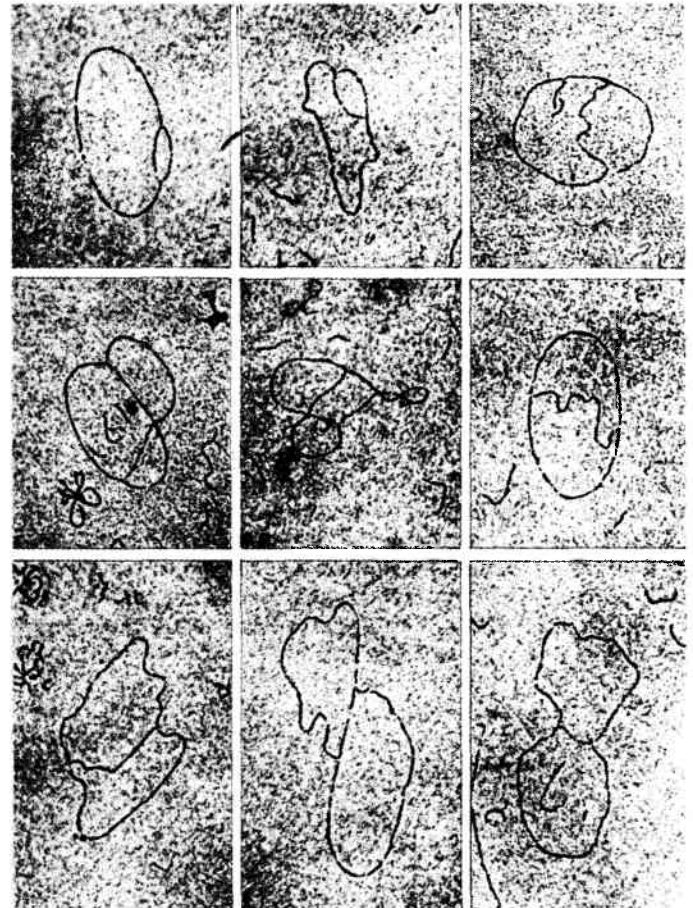


Рис. 31.6. Репликационный глазок увеличивается по мере продвижения репликационных вилок вдоль репликона.

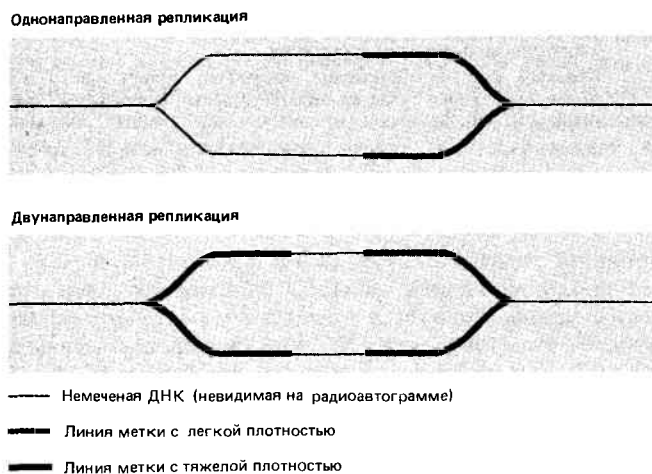


Рис. 31.7. Для определения однонаправленного или двунаправленного характера репликации можно использовать радиоактивную метку разной плотности.

риальную хромосому, т.е. ведут себя как сама хромосома. (Однако присутствие дополнительных точек начала может иметь неблагоприятные последствия для клетки; если нет возможности как-то различать точки начала, плазмидный и хромосомный *oriC*-локусы будут вынуждены разделять между собой число инициационных событий, в результате чего число хромосом может уменьшаться. Из этого следует, что истинные плазмидные точки начала, по-видимому, отличаются от точки начала репликации хромосомы.)

Выделенный фрагмент ДНК, несущий точку начала репликации, можно использовать для исследования отдельных функций репликации, таких, как акт инициации репликации, контроль частоты событий инициации и сегрегация реплицированных хромосом в дочерние клетки. Выделение мутантов, дефектных по любой из этих функций, дает возможность идентифицировать последовательности, контролирующие каждое репликационное событие.

Плазмиды, несущие в своем составе последовательность *oriC*, сегрегируют неправильно; для их стабилизации необходимо введение дополнительных последовательностей. Такой результат позволяет сделать два вывода: сама точка начала репликации не несет информации, требуемой для распределения дублированных хромосом между дочерними клетками; функции, необходимые для правильной сегрегации можно идентифицировать при выявлении последовательностей, которые придают плазмиде сегрегационную стабильность.

Путем уменьшения размера клонированного фрагмента удалось установить, что функциональная область *oriC* содержит только 422 пары оснований. Наличие такого фрагмента в плазмиде гарантирует ее выживаемость. Следует отметить, однако, что число копий таких плазмид в клетке может достигать 20. Следовательно, они утратили некоторые свойства, которые ограничивают частоту инициационных событий. По-видимому, число копий и акт инициации репликации могут зависеть от различных последовательностей.

Точка начала репликации бактерии *Salmonella typhimurium* была локализована во фрагменте из 296 пар оснований. Сравнение с *E. coli* показало, что 86% оснований в этих областях совпадают. Рассматривая последовательность одной цепи, можно предположить, что образование

вторичной структуры, напоминающей клеверный лист, происходит так, как это показано на рис. 31.8. О потенциальном значении указанной структуры свидетельствует тот факт, что почти все 42 основания, которыми отличаются эти районы у *E. coli* и *S. typhimurium*, попадают в неспаренные области. Весьма вероятно поэтому, что ДНК участка начала репликации имеет не правильную двухцепочечную структуру, а такую вторичную структуру, когда каждая цепь напоминает клеверный лист.

Для некоторых лямбдоидных фагов (один из которых лямбда) была также определена нуклеотидная последовательность участков начала репликации. Оказалось, что они родственны между собой, однако отличаются от соответствующей области бактериальной ДНК. Теоретически фаговые области начала репликации способны образовывать «клеверные листья» (с более тесно расположенными ответвлениями). Как показано на рис. 31.8, консервативная последовательность находится справа от потенциального места этой структуры. Такая последовательность присутствует в ДНК *E. coli*, лямбда и другого фага, G4. Если допустить, что именно консервативные последовательности выполняют функции участков начала репликации, то следует отметить их вторичную структуру, которая скорее всего узнается именно благодаря своей структуре, а не нуклеотидной последовательности.

Маловероятно, что нарушение структуры ДНК, необходимое для образования «клеверного листа», происходит спонтанно, поскольку такая форма существенно менее стабильна, чем обычная двухцепочечная ДНК. Ее появлению могут способствовать посторонние факторы, такие, как напряжение от суперспирализации и стабилизация белками. Если вся эта последовательность ДНК действительно необходима для выполнения функции начала репликации, то ее с уверенностью можно назвать самым большим *cis*-действующим сайтом среди когда-либо найденных. (Двухцепочечные последовательности промоторов и операторов намного короче.)

Согласованность процессов репликации ДНК и клеточного деления

Репликация и клеточный рост у бактерий тесно связаны. Частота инициации циклов репликации определяется скоростью роста клетки. И завершение цикла репликации согласовано с делением клетки на две, в процессе которого дочерние хромосомы сегрегируют.

Клетки *E. coli* способны расти с различными скоростями; время удвоения варьирует от 18 до более чем 180 мин. Так как бактериальная хромосома представляет собой один репликон, частота репликационных циклов контролируется числом событий инициации в единственной точке начала репликации. Скорость синтеза ДНК при постоянной температуре более или менее постоянна. Репликация происходит с одинаковой скоростью до тех пор, пока не наблюдается ограничений в снабжении предшественниками.

Цикл репликации *E. coli* может быть определен в значениях двух констант. D соответствует промежутку времени (примерно 20 мин) между завершением раунда репликации и клеточным делением, с которым она связана. Вероятно, этот период необходим для получения компонентов, принимающих участие в делении. С представляет

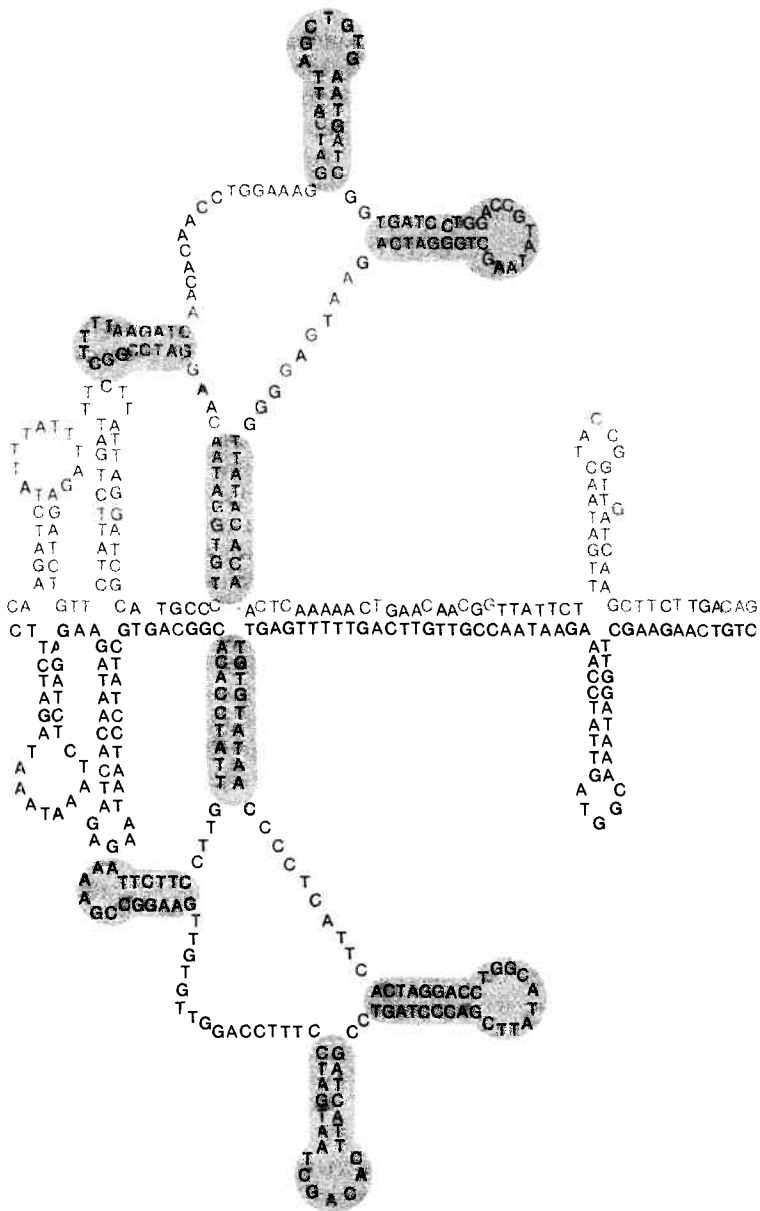


Рис. 31.8. Точка начала репликации у *E. coli* имеет палиндромные последовательности, способные формировать сложную вторичную структуру.

Красным цветом обозначены основания, отличающиеся у *S. typhimurium*, серые участки соответствуют основаниям, идентичным у фага лямбда.

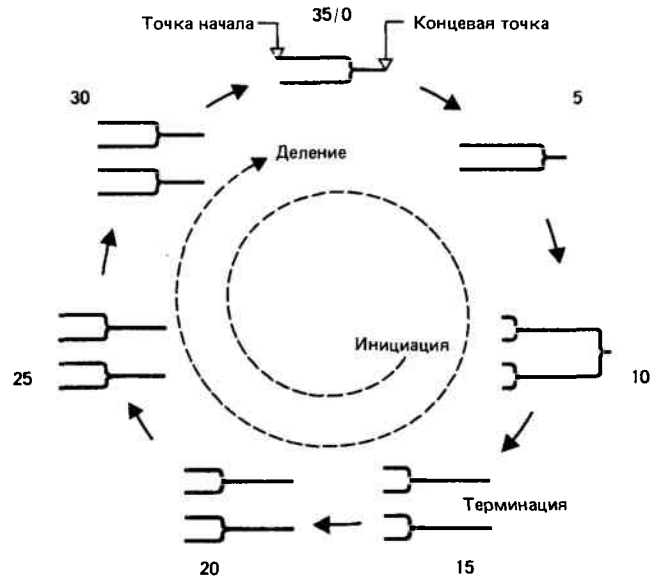


Рис. 31.9. Постоянный интервал 60 мин между инициацией репликации и клеточным делением ведет в быстро растущих клетках к образованию хромосомы со многими вилокми.

собой фиксированное время, равное примерно 40 мин, которое необходимо для репликации всей бактериальной хромосомы. Продолжительность этого периода соответствует скорости движения отдельной репликационной вилки с присоединением около 50 тыс. пар оснований в одну мин. (Константы С и D отражают максимальную скорость, с которой бактерия способна расти; они применимы ко всем значениям скорости роста, соответствующим времени удвоения от 18 до 60 мин, однако обе константные фазы становятся длиннее, если клеточный цикл длится более 60 мин.)

Цикл репликации хромосомы должен иницироваться в фиксированное время, $S + D = 60$ мин, перед клеточным делением. Если бактерии делятся более часто, чем в 60-минутные интервалы, репликация у них должна иницироваться до окончания предыдущего цикла деления. Рассмотрим пример с клетками, делящимися каждые 35 мин. Цикл репликации, связанный с делением, должен иницироваться на 25 мин раньше предшествующего деления. Такая ситуация изображена на рис. 31.9, показывающем состояние хромосом в бактериальной клетке через каждые пять минут клеточного цикла.

При делении (35/0 мин) клетка получает частично реплицированную хромосому. Репликационная вилка продолжает двигаться. Через 10 мин, когда «старая» репликационная вилка еще не достигла терминатора, в обеих точках начала репликации частично реплицированной хромосомы происходит инициация репликации. Начало движения этих «новых» репликационных вилок создает хромосому со многими вилокми. Через 15 мин, т.е. за 20 мин до следующего деления, старая репликационная вилка достигает терминатора. Две дочерние хромосомы разделяются. Каждая из них уже частично реплицирована с помощью новых репликационных вилок (теперь это единственные репликационные вилокми). Эти вилокми продолжают двигаться. В точке деления две частично реплицированные хромосомы сегрегируют. В результате восстанавливается то состояние, которое было вначале. Одна репликационная вилка становится «старой», она достигает точки терминатора через 14 мин, а спустя 20 мин происходит деление. Мы видим, что событие инициации происходит перед событием деления, с которым оно связано, за период, соответствующий $1^{25}/35$ клеточных циклов. Основной принцип связи между инициацией и ци-

клом клеточного деления сводится к следующему: чем быстрее растут клетки (чем короче цикл), тем большее количество циклов перед соответствующим делением отделяет событие инициации. Соответственно существует большее число хромосом в отдельной бактериальной клетке. Такая взаимосвязь может рассматриваться как ответ клетки на ее неспособность завершать периоды С и D более быстро в соответствии с более коротким циклом.

Как узнает клетка, когда именно следует инициировать цикл репликации? Событие инициации происходит при постоянном соотношении клеточной массы и числа точек начала репликации. Клетки, растущие быстрее, имеют большую массу и поэтому обладают большим числом точек начала репликации. Как показано на рис. 31.9, в точке, соответствующей 10 мин после деления, клеточная масса увеличена достаточно, для того чтобы поддержать событие инициации в обеих доступных точках начала репликации.

Для объяснения зависимости инициации репликации от клеточной массы были предложены две модели. Одна из них предполагает непрерывный синтез белка-инициатора в течение клеточного цикла. Накопление критического количества такого белка может служить сигналом для инициации. В соответствии со второй моделью в определенном периоде цикла в клетке синтезируется белок-ингибитор, а при увеличении объема клетки его уровень падает ниже эффективного. Существуют данные, свидетельствующие о том, что инициация действительно регулируется клеточной массой. Однако пока не установлено, происходит ли это в результате накопления инициатора или разведения ингибитора. В настоящее время предпочтение отдается модели, предполагающей синтез инициатора, поскольку она согласуется с данными, свидетельствующими о том, что для события инициации требуется синтез белка.

Инициация циклов репликации может быть нарушена мутациями в нескольких генах *E. coli*; существует также множество генов, продукты которых необходимы для клеточного деления. Однако белки, ответственные за контроль частоты инициации или за связь репликации с циклом деления, при анализе этих мутаций еще не идентифицированы.

Независимо от того, какой тип регуляции используется, рост бактерий может быть описан в значениях клеточных единиц, имеющих длину 1,7 мкм. Бактерия содержит одну точку начала репликации на клеточную единицу; быстро растущая клетка с двумя точками начала репликации должна иметь длину более 3,4 мкм. Топологическая связь между событием инициации и структурой клетки может определяться наличием сайта роста в клетке, единственного участка, в котором может происходить инициация. На одну клеточную единицу должен приходиться один сайт роста.

Было высказано предположение, что между бактериальной ДНК и мембраной существует физическая связь. Бактериальную ДНК можно обнаружить в мембранных фракциях. Она содержит увеличенное число копий генетических маркеров, локализующихся вблизи точки начала репликации, репликационную вилку, терминатор. Вполне вероятно, что сайт роста — это структура на мембране, к которой прикрепляется своим участком с точкой начала репликации хромосома, чтобы инициировать репликацию. Репликационные вилки хромосом млекопитающих могут быть связаны с ядерным матриксом.

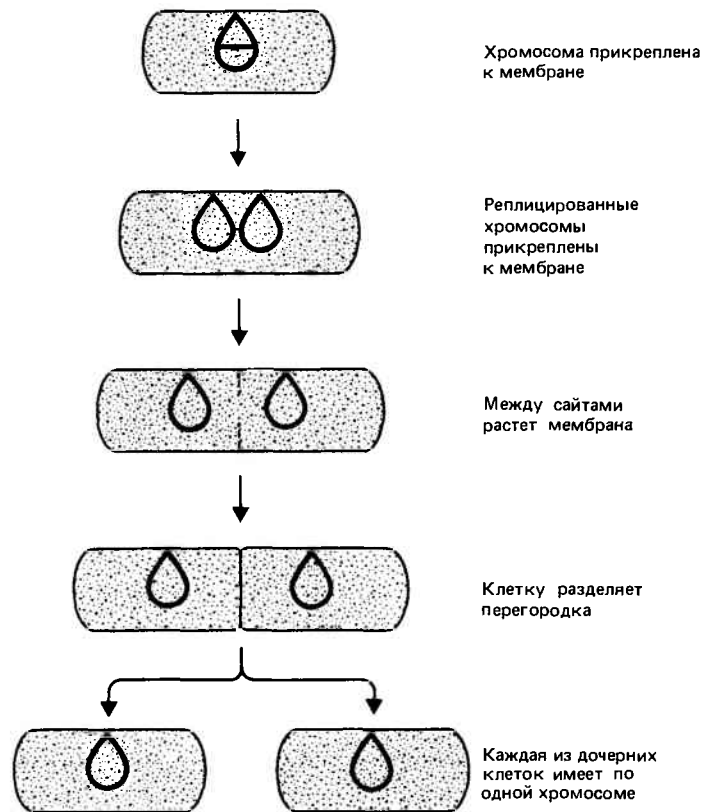


Рис. 31.10. Прикрепление бактериальной ДНК к мембране может обеспечивать механизм сегрегации.

Связь между ДНК и мембраной может объяснять и сегрегацию. Если дочерние хромосомы прикреплены к мембране, они могут быть физически отделены друг от друга при росте мембраны между ними. Благодаря образующейся перегородке хромосомы оказываются в различных дочерних клетках (рис. 31.10). Если мы вспомним событие сегрегации, изображенное на рис. 31.9, то следует ожидать, что и терминатор репликации хромосомы должен быть вовлечен в этот процесс.

До сих пор мы имели дело с бактериальной хромосомой, предполагая, что она линейна. Поскольку на самом деле хромосома бактерии замкнута в кольцо, каждая из двух репликационных вилок движется вокруг генома к точке встречи. Мы не знаем, что происходит при их столкновении. Разрушаются ли вилки, наталкиваясь одна на другую, или имеется специфический терминатор, в котором они останавливаются? Каким образом ферменты, вовлекаемые в процесс репликации, освобождаются от хромосомы, если мы знаем, что ДНК должна быть реплицирована и воссоединена в области, где вилки встречаются? Свойства мутации *dnaT* позволяют предположить, что она относится к гену, который кодирует белок *E. coli*, участвующий в терминации раунда репликации. Однако продукт этого гена еще не охарактеризован.

Область хромосомы *E. coli*, которая может обеспечить остановку движения репликационных вилок, т. е. служить терминатором процесса репликации, была идентифицирована. Оказалось, что она не локализована точно на полпути по кругу хромосомы. Такая несоразмерность указывает на то, что две вилки должны проходить различные расстояния. В ДНК плазмиды R6K терминатор был идентифицирован как область, которая останавливает

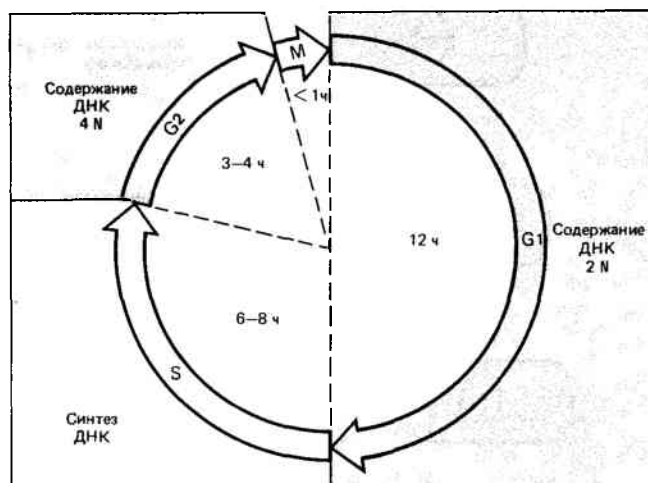


Рис. 31.11. Цикл эукариотической клетки делится на четыре фазы, G1, S, G2 и M. Указана продолжительность фаз в культивируемой линии клеток млекопитающих.

процесс репликации. Эта область остается эффективной, если она внедрена в другую молекулу ДНК. Потенциальная последовательность терминатора, по-видимому, не способна образовывать вторичную структуру, так что способ ее действия неизвестен.

Каждая хромосома эукариот содержит много репликонов

В эукариотических клетках репликация ДНК занимает часть клеточного цикла. Период между митозами (интерфазу) можно разделить на определенные фазы (рис. 31.11). Типичная клетка начинает свой цикл в диплоидном состоянии. Она остается в этом состоянии в течение периода G1, который, как правило, занимает основную часть цикла (и наиболее вариабелен по продолжительности у клеток с различным фенотипом). Фаза S характеризуется синтезом ДНК. У высших эукариотических клеток этот период длится в течение нескольких часов. Репликация создает тетраплоидное состояние, в котором клетка остается в течение периода G2. Затем митоз (M) уменьшает набор хромосом до диплоидного числа в каждой дочерней клетке. (Периоды, предшествующие и сменяющие S-фазу, получили названия G1 и G2; они представляют «интервалы» в синтезе ДНК.)

События, ответственные за инициацию S-фазы, происходят во время периода G1. Известно, что в этом периоде происходит синтез белка, однако точная природа событий до сих пор остается неясной; в частности, не решен вопрос относительно того, являются ли эти события внезапными или постепенно накапливаемыми. Мы мало знаем о том, каким образом на молекулярном уровне клетка принимает решение перейти к репликации генома. Возможно, это происходит в результате регуляторного события, отличающегося по природе от тех событий, которые включены в последующий синтез ДНК. Репликация большого количества ДНК, содержащейся в эукариотической хромосоме, осуществляется посредством разделения хромосомы на множество отдельных репликонов. Такие репликоны активируются не все одновременно; в любой точке S-фазы только некоторые из них вступают в репликацию. По-видимому, каждый ре-

пликон активируется в специфичное время S-фазы; однако убедительных доказательств этого еще не получено.

Очень важно, что сигналом для вступления в S-фазу служит активация первых репликонов. В течение следующих нескольких часов происходят инициационные события в других репликонах. Поэтому контроль S-фазы включает два процесса: выведение клетки из периода G1 и инициацию репликации в отдельных репликонах упорядоченным способом. Большинство наших сведений о свойствах отдельных репликонов получено в результате радиоавтографических исследований, проведенных по схеме, представленной на рис. 31.7. Хромосомные репликоны обычно проявляют двунаправленную репликацию, о чем свидетельствуют симметричные пары треков.

Какова величина среднего репликона и сколько их в геноме? Трудность в определении отдельных единиц обусловлена тем, что смежные репликоны могут сливаться, образуя большие реплицируемые «глазки» (рис. 31.12). Используемый обычно подход для определения отдельных репликонов в такой структуре основан на измерениях последовательностей ДНК, в которых несколько репликонов можно различить как активные, по-видимому захваченные на стадии, когда все они уже инициированы, но репликационные вилки смежных репликонов еще не встретились. Существуют некоторые данные, свидетельствующие, что такой тип активации обусловлен «региональным» контролем, при котором группы репликонов иницируются более или менее координированно. Эти данные противоречат модели, при которой отдельные репликоны в различных областях генома активируются независимо друг от друга.

В группах активных репликонов средний размер реплицирующейся единицы определяется по расстоянию между точками начала репликации (т.е. между средними точками смежных репликонов). У многих высших эукариот они удалены друг от друга на 100–200 тысяч пар нуклеотидов. Следовательно, в гаплоидном геноме млекопитающих должно быть 20 000–30 000 репликонов. У *D. melanogaster* или у *S. cerevisiae* репликоны меньше, их размер достигает в среднем 40 тысяч пар нуклеотидов. Это соответствует примерно 3500 репликонам на гаплоидный набор хромосом плодовой мушки и примерно 500 репликонам у дрожжей. Колебания в размерах отдельных репликонов большие, более чем десятикратные, так что указанные средние величины дают лишь приблизительные значения числа репликонов.

Скорость перемещения репликационных вилок можно оценить по максимальному расстоянию, которое проходят радиоавтографические треки в течение определенного времени. Она колеблется от 1000 до 3000 пар оснований в 1 мин для млекопитающих; у растений эта величина не превышает 1000 пар оснований в 1 мин, возможно, из-за их более низкой температуры. Эти значения гораздо ниже, чем соответствующие показатели движения для бактериальной репликационной вилки (примерно 50 000 пар оснований в 1 мин).

Что представляет собой точка начала репликации каждого репликона? Является ли она специфической последовательностью ДНК, возможно относящейся к семейству повторяющихся последовательностей? Или это хроматиновая нить, имеющая структуру более высокого порядка, не обязательно связанную с одним типом последовательности? Далее мы должны ответить на вопрос: каким образом выбираются точки начала для инициации репликации в разные периоды S-фазы? Заключена ли со-

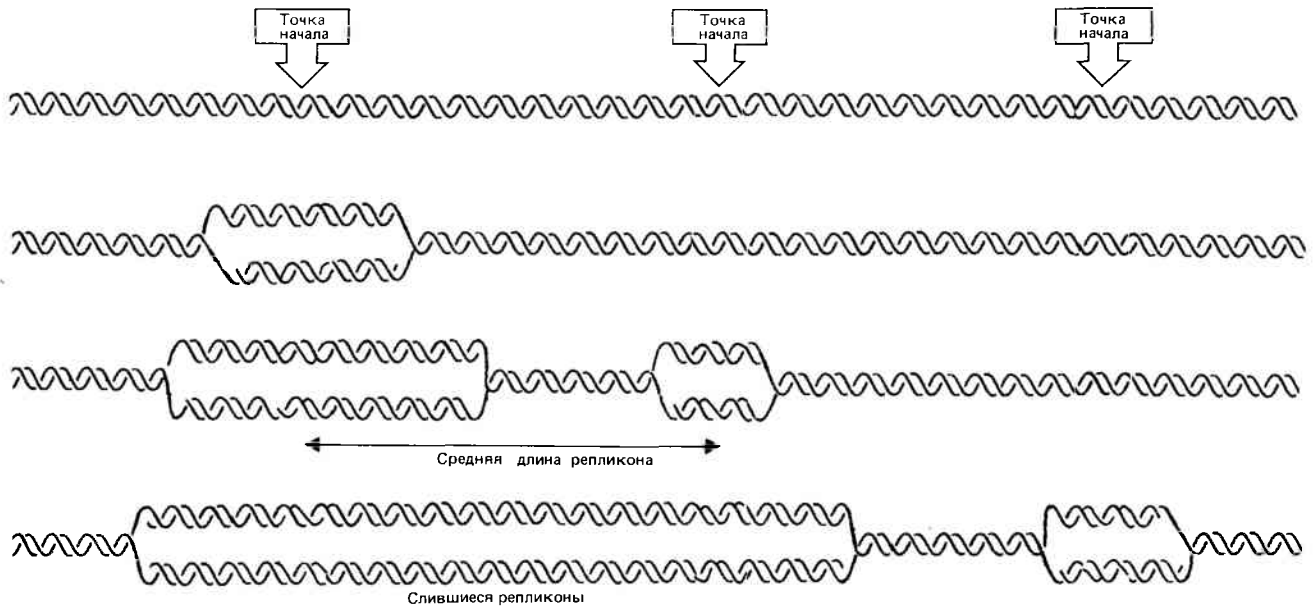


Рис. 31.12. Измерение размера репликона производится на последовательности ДНК, в которой активны смежные репликоны.

ответствующая информация в структуре области начала репликации, или это является следствием какой-то упорядоченной перестройки? До сих пор остается загадкой, как репликационный аппарат отличает уже использованные точки начала от тех, которые еще должны быть реплицированы. Чем обусловлено однократное использование каждой точки начала во время S-фазы – свойством ДНК (например, состоянием ее метилирования) или белков, связанных с ней?

Известен один случай, когда точки начала не фиксированы, а зависят от типа клеток. Репликоны ранних эмбрионов у *D. melanogaster* оказываются значительно меньшими, чем репликоны соматических клеток. Их одновременная активация уменьшает время, затрачиваемое на репликацию хромосом, – необходимость, обусловленная скоростью клеточных делений в этот период. Возможное объяснение сводится к тому, что геном содержит ряд тканеспецифичных точек начала репликации, активируемых только при эмбриональных делениях и не узнаваемых при делениях соматических клеток. Однако мы не располагаем достаточной информацией, чтобы судить, являются ли случаи тканеспецифичной репликации общими.

Не ясно, имеют ли хромосомные репликоны дискретные концы, в которых репликационные вилки прекращают движение и (по-видимому) элиминируются. Альтернативным является предположение о движении вилки, продолжающемся от точки начала до встречи с вилкой, движущейся к ней из смежного репликона. Мы уже упоминали потенциальную топологическую проблему воссоединения вновь синтезированной ДНК в месте стыковки репликационных вилок.

Известно, что эухроматин и гетерохроматин реплицируются в разное время. Как конститутивный, так и факультативный гетерохроматин относится к поздно реплицирующимся структурам. Именно гетерохроматин обычный (хотя и не всегда) последним в клетке синтезирует ДНК во время S-фазы. Это может быть следствием его более конденсированной структуры.

Выделение точек начала дрожжевых репликонов

Любой сегмент ДНК, имеющий точку начала репликации, способен реплицироваться. Отсюда следует, что, хотя плазмиды редко встречаются в клетках эукариот, их можно сконструировать путем соответствующих манипуляций *in vitro* и ввести в соответствующие реципиентные клетки. (В гл. 38 мы рассмотрим вопрос о судьбе ДНК, введенной в клетки эукариот.)

Одной из систем, в которой внехромосомная ДНК способна реплицироваться, является ооцит *Xenopus*. Инъекцированная ДНК реплицируется в клеточном цикле только один раз, что свидетельствует о наличии соответствующего контроля. Однако в ооците *Xenopus* способна реплицироваться любая кольцевая ДНК, и можно сделать вывод, что специфическая точка начала в этой системе обязательна.

Другая система предоставляет возможность для выделения дискретных точек начала репликации. Клетки дрожжей *S. cerevisiae*, мутантные по какой-то функции, могут быть «трансформированы» путем добавления ДНК, которая несет копию гена дикого типа. Схема эксперимента представлена на рис. 31.13. Мутация клетки-хозяина должна затрагивать ген, продукт которого можно селективировать. ДНК клеток дикого типа выделяют, фрагментируют и клонируют в составе плазмид *E. coli*. Гибридные плазмиды инкубируют с мутантными клетками дрожжей в таких условиях, в которых эти клетки способны выжить только в случае экспрессии гена дикого типа. В зависимости от частоты возникновения трансформированных клеток различают два типа трансформации.

Любой ген может быть трансформирован с низкой частотой. Выживающие клетки имеют встроенную в хромосому копию донорного гена дикого типа. Вероятно, этот ген замещает исходный мутантный в результате гомологичной рекомбинации. Отсутствие какого-либо другого потомства доказывает, что донорный ген не спосо-

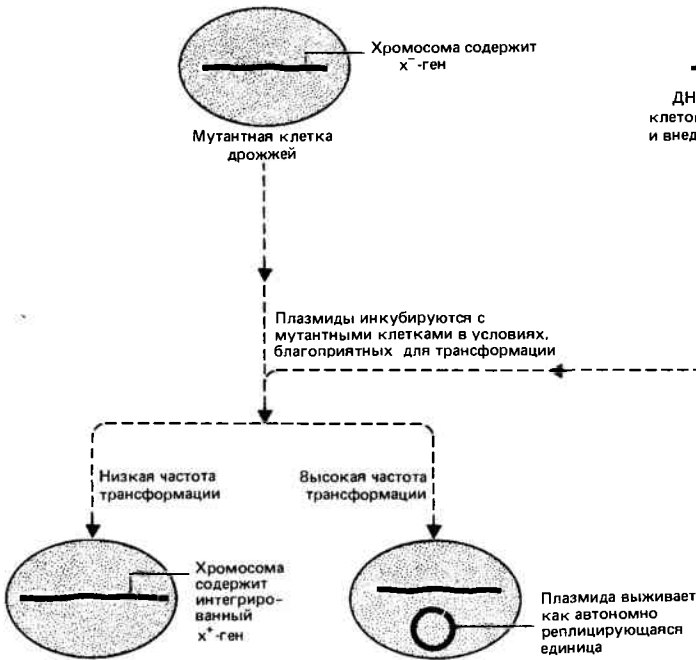


Рис. 31.13. Частота трансформации у дрожжей зависит от того, происходит ли внедрение фрагмента в хромосомную ДНК или он сохраняется в виде плазмиды.

бен сохраниться в реципиентной клетке, если он не интегрирует с хромосомой. Следовательно, низкая частота трансформации объясняется, по-видимому, необходимостью интеграции. Некоторые фрагменты дрожжевой ДНК способны трансформировать дефектные клетки с *более высокой* (примерно 1000-кратной) *частотой*. Такие фрагменты выживают в клетке в неинтегрированном (автономном) состоянии, т.е. в форме исходно введенной кольцевой молекулы ДНК. Именно это, отсутствие необходимости в интеграции, и обуславливает высокую частоту трансформации. Фрагмент, трансформирующийся с высокой частотой, обладает последовательностью, без которой эффективная репликация гибридной плазмиды в клетках дрожжей невозможна. Этот сегмент получил название *ars*-последовательности (от англ. *autonomously replicating sequence* – автономно реплицирующаяся последовательность). Возможно, именно она и есть подлинная точка начала репликации. Известно, что *ars*-последовательности встречаются примерно с той же частотой, которая характерна для точек начала репликации (одна на каждые 30–40 т.п.н.).

У дрожжей точки начала репликации не относятся к семейству повторяющихся последовательностей. Вполне возможно, что много разных последовательностей способны обеспечивать функцию точки начала репликации. Уменьшая размер клонированных фрагментов, можно определить минимальную последовательность, необходимую для обеспечения *ars*-функции. Единственная область гомологии *ars*-элементов представлена канонической последовательностью из 11 п.н. (в каждом элементе может быть не больше трех замен). Обычно эти последовательности находятся в окружении А–Т-богатой ДНК.

Точечные мутации *ars*-элемента приводят к тому, что область коры длиной 14 п.н. становится точкой начала репликации. Кор содержит каноническую последовательность. Не исключено, что именно она является сайтом узнавания для белка, участвующего в проявлении функции точки начала репликации. Если в осуществление этой функции вовлекаются и другие последовательности, то у каждого из известных *ars*-элементов они разные. Извест-

на и другая точка начала репликации в составе 2 мкм-плазмиды дрожжей. Она также не имеет протяженной гомологии с *ars*-последовательностями. Вполне вероятно, поэтому, что функцию точки начала репликации у дрожжей способны обеспечить много разных последовательностей.

Репликация может происходить по типу «глазков», «катящихся колец» или D-петель

Вид структуры, формируемой в процессе репликации, зависит от взаимосвязи между матрицей и репликационной вилкой. Критическими факторами при этом оказываются циркулярность или линейность матрицы и осуществление синтеза в репликационной вилке обеих или только одной цепи ДНК.

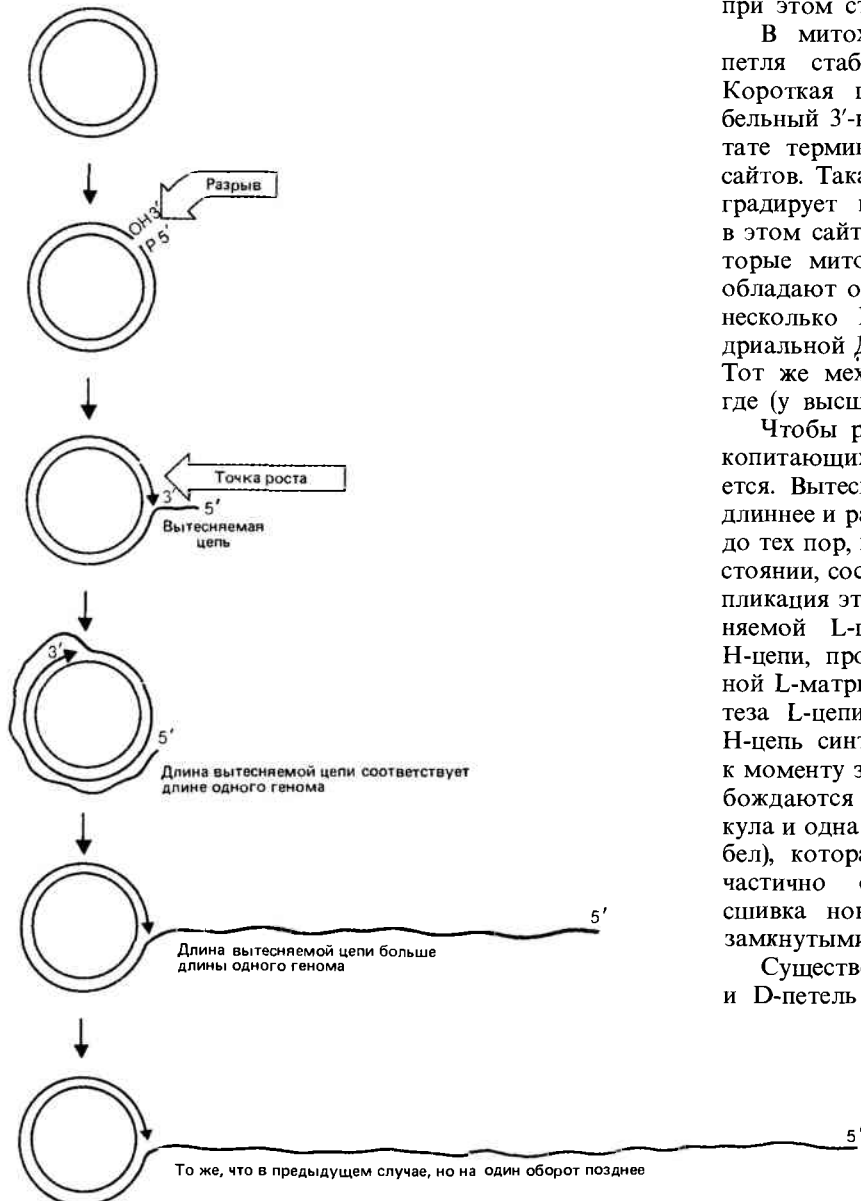
До сих пор мы условно допускали, что репликационная вилка представляет собой сайт, в котором синтезируются обе новые цепи ДНК. В линейной молекуле независимо от того, идет репликация в одном или в двух направлениях, движение вилки (вилок) формирует «глазок» (рис. 31.4). Если матрицей служит кольцевая ДНК, реплицирующаяся молекула принимает форму θ -структуры (рис. 31.5), однако возможен и другой исход. Предположим, что репликация одной цепи двухцепочечной молекулы начинается в точке начала. Разрез открывает одну цепь, и свободный 3'-ОН-конец, образовавшийся при этом, наращивается с помощью ДНК-полимеразы. Вновь синтезируемая цепь вытесняет исходную родительскую. Последующие события изображены на рис. 31.14.

Такой тип структуры получил название *катящегося кольца*, поскольку точка роста «скользит» вокруг кольцевой матричной цепи. В принципе это может продолжаться неопределенно долго. При перемещении точки роста репликационная вилка наращивает внешний конец цепи и вытесняет предшествующего партнера. Так как вновь синтезируемый материал ковалентно сцеплен с исходным материалом, вытесняемая нить достигает размера исход-

ного генома в ее 5'-конце. За исходной единицей следует любое число единиц геномов, синтезируемых путем непрерывного вращения матрицы. Каждый цикл ведет к вытеснению материала, синтезированного в предыдущем цикле. Катящееся кольцо используется *in vivo* несколько раз.

У некоторых фагов геном представлен одноцепочечной кольцевой ДНК. Такое кольцо сначала превращается в двухцепочечное, а затем реплицируется посредством механизма катящегося кольца. Высвобождаемая цепь приводит к образованию серии геномов; они могут быть фрагментированы и внедрены в фаговые частицы; кроме того, возможно их использование в дальнейших репликационных циклах. (Такие репликационные системы обсуждаются в гл. 33.)

Освобождаемая отдельная цепь катящегося кольца может быть превращена в двухцепочечную ДНК путем синтеза комплементарной цепи. Благодаря этому образуются конкатемерные двухцепочечные молекулы, необходимые для созревания определенных фаговых ДНК (гл. 28), например ДНК фага лямбда. Отметим, что данный способ образования конкатемерных молекул не



единственный; для этого может использоваться и рекомбинация, как в случае фага Т4.

С помощью механизма катящихся колец образуется амплифицированная рДНК в ооцитах *Xenopus*. Именно этим объясняется большое число идентичных повторяющихся единиц в амплифицированном материале (см. рис. 23.6). Одна геномная повторяющаяся единица превращается в катящееся кольцо, при этом вытесняемый конец превращается в двухцепочечную ДНК. В какой-то точке она отрывается от кольца, два конца соединяются, образуя амплифицированное кольцо рДНК.

Другой тип поведения репликационной вилки идентифицирован в некоторых митохондриях. Репликация начинается в специфической точке начала в кольцевой двухцепочечной ДНК. Но первоначально только одна из двух родительских цепей (Н-цепь в митохондриальной ДНК млекопитающих) используется в качестве матрицы для синтеза новой цепи. Синтез происходит только на коротком участке и вызывает вытеснение исходной комплементарной (L) цепи, которая остается одноцепочечной (рис. 31.15). Происходящие в этой области события были названы «вытеснением» (displacement), а образующаяся при этом структура D-петлей.

В митохондриях млекопитающих единственная D-петля стабильна и состоит из 500–600 оснований. Короткая цепь имеет уникальный 5'-конец и переменный 3'-конец, образующийся, по-видимому, в результате терминации синтеза в любом из 3–4 дискретных сайтов. Такая короткая цепь нестабильна; она часто деградирует и повторно синтезируется для сохранения в этом сайте двойной спирали в открытой форме. Некоторые митохондриальные ДНК, например у *X. laevis*, обладают одной длинной D-петлей. Другие могут иметь несколько D-петель; например, в линейной митохондриальной ДНК *Tetrahymena* обнаружено шесть D-петель. Тот же механизм используется в хлоропластной ДНК, где (у высших растений) существуют две D-петли.

Чтобы реплицировать митохондриальную ДНК млекопитающих, короткая цепь в D-петле вначале удлиняется. Вытесняемая область исходной L-цепи становится длиннее и расширяет D-петлю. Расширение продолжается до тех пор, пока не достигает точки, находящейся на расстоянии, составляющем примерно 67% длины кольца. Репликация этой области раскрывает точку начала в вытесняемой L-цепи. В этом сайте инициируется синтез Н-цепи, протекающий вдоль вытесненной одноцепочечной L-матрицы в противоположном направлении от синтеза L-цепи. Из-за задержки в начале такого синтеза Н-цепь синтезируется только на 30–40% длины кольца к моменту завершения синтеза L-цепи. В результате освобождаются одна полная двухцепочечная кольцевая молекула и одна кольцевая молекула, содержащая брешь (пробел), которая до завершения синтеза Н-цепи остается частично одноцепочечной. Наконец, осуществляется сшивка новых цепей, которые становятся ковалентно замкнутыми.

Существование механизмов катящегося кольца и D-петель способствует пониманию общего принципа

Рис. 31.14. Катящееся кольцо порождает мультимерный одноцепочечный конец (который может быть превращен в двухцепочечную ДНК путем синтеза комплементарной цепи; это на рисунке не показано).

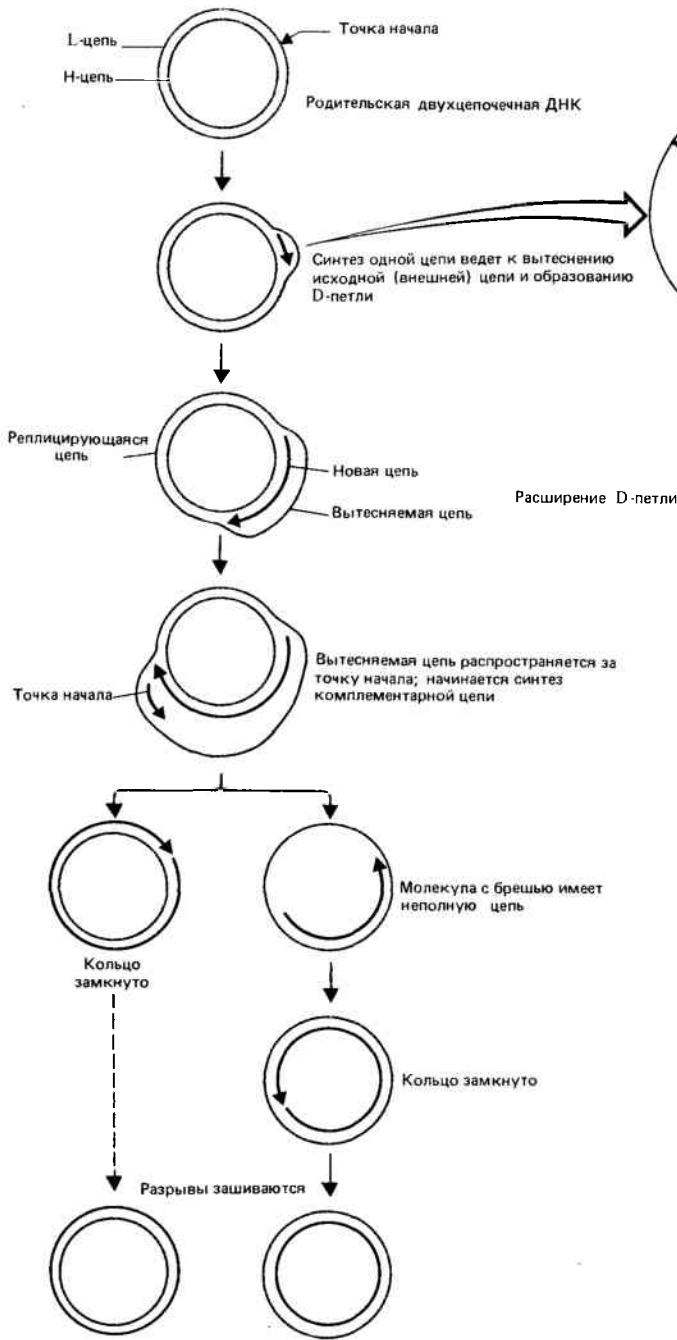


Рис. 31.15. В митохондриальной ДНК млекопитающих, имеющей отдельные точки начала для репликации каждой из цепей, D-петля представляет собой стабильную структуру.

Несовместимость плазмид связана с числом их копий

Плазмиды эгоистичны. Завладев бактериальной клеткой, плаزمида старается предотвратить проникновение и размножение в ней любой другой плазмиды того же типа. Для этого используются два независимых механизма: **поверхностное исключение** и **плазмидная несовместимость**. Поверхностное исключение определяет неспособность плазмид внедряться в клетки, уже несущие другую плазмиду того же типа. Эффект проявляется на поверхности бактерии и характеризует систему, используемую при переносе плазмид. В случае полового фактора *E. coli* поверхностное исключение достигается путем подавления выхода ДНК из клеток бактерии-хозяина. Таким образом, это связано скорее с контролем эмиграции, а не иммиграции.

Несовместимость имеет отношение к регуляции числа копий плазмид. **Группы совместимости** определяются как группы плазмид, члены которых не способны сосуществовать в одной и той же клетке бактерии. Присутствие плазмиды, относящейся к одной группе совместимости, не влияет на выживаемость плазмиды, принадлежащей к другой группе. Таким образом, только один репликон данной группы совместимости может быть сохранен в бактериальной клетке, но он не взаимодействует с репликационными другими группами (хотя в неблагоприятных условиях они могут конкурировать за «жизненное пространство»).

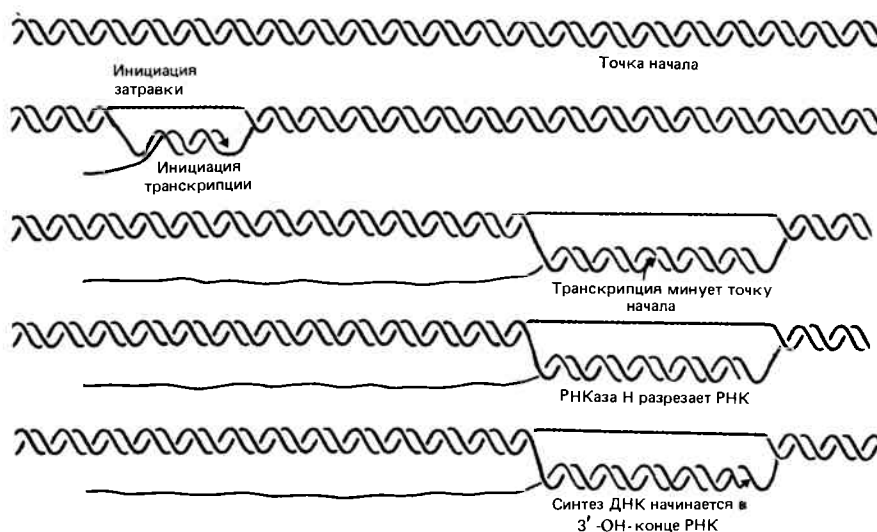
Модель негативного контроля плазмидной несовместимости основывается на предположении, что контроль числа копий достигается путем синтеза репрессора, определяющего число точек начала репликации. (Формально она не отличается от модели «титрования», предложенной для объяснения регуляции репликации бактериальной хромосомы.) Введение новой точки начала репликации в составе плазмиды той же группы совместимости имитирует результат репликации резидентной плазмиды; теперь присутствуют две точки начала репликации. В результате любая дальнейшая репликация предотвращается до тех пор, пока две плазмиды не разойдутся в разные клетки и таким образом не будет восстановлено точное дорепликационное число их копий.

Число копий может быть увеличено или уменьшено в результате мутаций. Такие мутации важны, так как с их помощью удается идентифицировать генетический локус, контролирующий число копий плазмиды и изучать используемый для этого механизм. Мутации, ответственные за число плазмидных копий, влияют также на несовместимость, указывая на ее связь с числом копий.

Система контроля числа копий и несовместимость наиболее изучены у плазмиды ColE1 *E. coli*. Прежде чем рассмотреть эту систему, мы должны кратко остановиться

репликации. Точкой начала репликации может быть последовательность ДНК, которая служит для инициации синтеза ДНК при использовании одной цепи в качестве матрицы. Раскрытие спирали не обязательно ведет к инициации репликации на другой цепи. Она может иметь свою собственную точку начала репликации. D-петля образуется в тех случаях, когда точки начала репликации двух цепей отделены некоторым расстоянием. На основании этого естественно заключить, что точка начала, которая служит для инициации репликации *обеих* цепей, так специфически ведет себя благодаря своей собственной организации, т.е. благодаря тому что точки начала репликации *обеих* цепей локализованы в одной и той же области. Такая организация также определяет, будут ли репликационные вилки двигаться однонаправленно или двунаправленно.

Рис. 31.16. Репликация ДНК плазмиды ColE1 инициируется разрезом в РНК-затравке в результате которого образуется 3'-ОН-конец.



ся на механизме инициации репликации в точке начала ColE1. События, относящиеся к этому процессу, показаны на рис. 31.16. Репликация начинается с транскрипции РНК, которая инициируется на расстоянии 555 пар оснований, расположенных против течения репликации, от точки начала. Транскрипция продолжается через точку начала репликации. Фермент РНКазы Н (чьё название отражает специфичность в отношении субстрата: РНК, сгидризованной с ДНК) отрезает транскрипт в точке начала репликации. Образуется 3'-конец, использующийся в качестве затравки для синтеза ДНК.

РНК-затравка находится под влиянием двух регуляторных систем. Одна из них использует белок, кодируемый расположенным рядом локусом, другая — включает синтез РНК, комплементарной затравке.

Небольшой белок (63 аминокислотных остатка), кодируемый геном *rop*, локализован на некотором расстоянии от точки начала репликации по направлению хода репликации. Белок Rop подавляет образование затравки. В настоящее время механизм этого подавления неизвестен, однако вероятно, что в данный процесс вовлекаются последовательности, находящиеся по ходу репликации сразу же за точкой начала инициации затравки. Независимо от характера молекулярного взаимодействия с этими последовательностями, его результатом является уменьшение числа копий плазмиды.

В той же области ДНК ColE1 синтезируется другой вид РНК. Это РНК I, молекула размером около 108 оснований, кодируемых цепью, противоположной той, которая кодирует РНК-затравку. Рис. 31.17 иллюстрирует взаимосвязь между РНК-затравкой и РНК I. Молекула РНК I имеет конец вблизи сайта, в котором инициируется РНК-затравка. Следовательно, РНК I комплементарна 5'-концевой области РНК-затравки.

Молекула РНК, которая выполняет некую функцию за счет комплементарности с другой РНК, кодируемой в той же области, может быть названа *контртранскриптом*.

Мутации, которые уменьшают или элиминируют несовместимость, могут быть получены при отборе плазмид той же группы по их способности к сосуществованию. Мутации группы ColE1 картируются в области перекрытия между РНК I и РНК-затравкой, причем

четыре из них обусловлены заменой G—C на A—T, обратным изменением. Так как эта область представлена в двух различных РНК, любая (или обе) из них может быть вовлечена в выражение этого эффекта.

Добавление РНК I к системе репликации ДНК плазмиды ColE1 *in vitro* вызывает подавление синтеза РНК-затравки. В то же время присутствие РНК I инициации или элонгации синтеза РНК-затравки не подавляет. Следовательно, РНК I предотвращает образование 3'-конца РНК-затравки, порождаемого при участии фермента РНКазы Н. Наиболее вероятно, что в основе этого явления лежит спаривание оснований РНК I и РНК-затравки.

Обе молекулы РНК имеют одну и ту же вторичную структуру в этой области, отличающуюся двумя двухцепочечными шпильками с одноцепочечными петлями. Мутации несовместимости локализованы именно в этих петлях. Можно предположить, что первой ступенью в спаривании оснований между РНК I и РНК-затравкой служит контакт между неспаренными петлями. Такое спаривание может быть нарушено из-за мутации в одной из родительских плазмид, приводящей к появлению различий в последовательностях РНК I и РНК-затравки. Каждая РНК I будет продолжать спариваться с РНК-затравкой, кодируемой той же плазмидой, но может оказаться неспособной спариваться с РНК-затравкой, кодируемой другой плазмидой.

Каким образом спаривание с РНК I препятствует разрезанию, способствующему образованию РНК-затравки? Возможная схема этого процесса приведена на рис. 31.18. В отсутствие РНК I РНК-затравка имеет вторичную структуру, при которой ее 5'-концевая область спарена с какой-то другой частью молекулы. Вмешательство РНК I изолирует 5'-конец потенциальной РНК-затравки,



Рис. 31.17. Последовательность РНК I комплементарна 5'-области РНК-затравки.

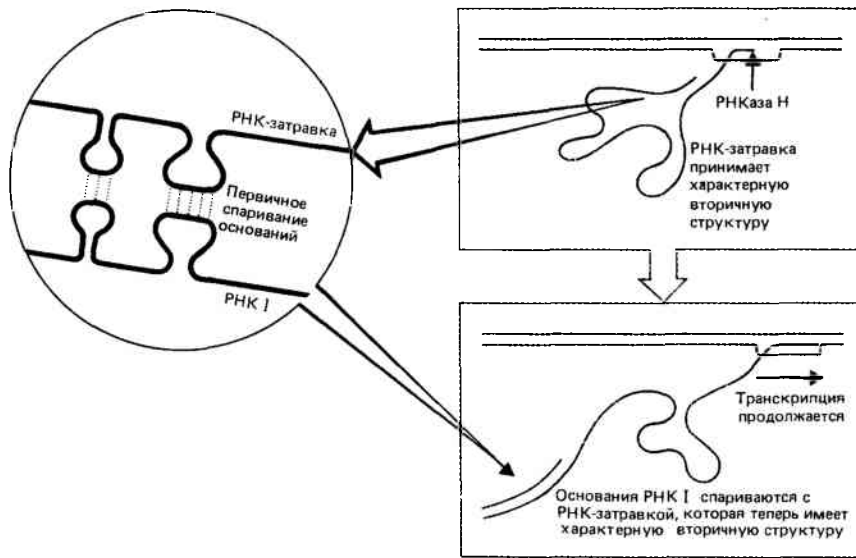


Рис. 31.18. Спаривание оснований с РНК I может изменить вторичную структуру последовательности РНК-затравки и таким образом предотвратить разрез и образование 3'-ОН-конца.

изменяя ее вторичную структуру. Такая новая вторичная структура каким-то образом мешает осуществлению реакции разрезания, хотя природа такого эффекта остается непонятной. (Предполагалось, что действие РНКазы H зависит от узнавания гибридной области РНК—ДНК, а не структуры РНК.)

Модель в какой-то степени напоминает механизм, участвующий в аттенуации транскрипции, при котором альтернативные способы спаривания последовательности РНК позволяют или предотвращают образование вторичной структуры, необходимой для терминации транскрипции, осуществляемой РНК-полимеразой (гл. 15). Формально эта модель равнозначна постулированию присутствия в клетке репрессора, который подавляет функционирование вновь введенной ДНК, аналогично репрессору фага лямбда (гл. 16). Вместо белка-репрессора, который связывает новую ДНК, РНК связывает вновь синтезированный предшественник РНК-затравки. Способность РНК I подавлять инициацию репликации может быть частью цикла негативного контроля, с помощью которого несовместимость связана с контролем числа копий. Однако мы еще не знаем роли этих событий в поддержании характерного числа копий ColE1 ДНК (примерно 20 на 1 клетку). Возможно, она определяется соотношением между частотой инициации РНК-затравки и способностью затравки «запускать» синтез ДНК. Этот тип несовместимости может быть следствием событий, используемых для регуляции репликации. Вполне вероятно также, что несовместимость является результатом механизмов, с помощью которых при делении плазмиды распределяются между дочерними клетками.

В составе плазмиды pSC101 имеется область размером 370 п. н., получившая название *par*, которая необходима для распределения реплицировавшихся плазмид между дочерними клетками. Известно, что эта область не кодирует белок, активна в *цис*-положении и состоит из

трех дискретных сегментов. Центральная последовательность способна образовывать шпильчатую структуру при спаривании оснований с любой из фланкирующих последовательностей этой же цепи. Делетирование любого из этих сегментов ведет к нарушению распределения копий плазмид.

В противоположность этому у профага P1 функция *par* связана с кодированием белка, узнающего *цис*-активный сайт, названный *incB*. Мутации несовместимости картируются именно в этом локусе.

Роль *цис*-активных сайтов распределения эквивалентна роли эукариотических центромер. Эти сайты обеспечивают прикрепление ДНК к клеточным структурами. Повидимому, несовместимость обусловлена конкуренцией за сайты прикрепления. Связь между распределением и клеточным делением может осуществляться специфическими белками.

Рекомендуемая литература

Обзоры, посвященные бактериальным и эукариотическим репликам, можно найти в книгах Льюина (*Lewin, Gene Expression, 1, Bacterial Genomes, Eucaryotic Chromosomes, Wiley, New York, 1974, 1980*). О выделении точек начала репликации ДНК бактерий и фагов сообщают Ясуда и Хирота (*Yasuda, Hirota, Proc. Nat. Acad. Sci. USA, 74, 5458–5462, 1977*), Зискинд и Смит (*Zyskind, Smith, Proc. Nat. Acad. Sci. USA, 77, 2460–2464, 1980*) и Гроссчедл и Хобом (*Grosschedl, Hobom Nature, 282, 39–43, 1979*). Ласки и Харланд (*Laskey, Harland, Cell, 24, 283–284, 1981*) обсуждают структуру точки начала репликации ДНК эукариот. Митохондриальной ДНК посвятил свой обзор Клейтон (*Clayton, Cell, 28, 693–705, 1982*). Томизава и Ито описали систему ColE1 (*Tomizawa, Itoh, Proc. Nat. Acad. Sci. USA, 78, 6096–6100, 1981*).

Глава 32

ТОПОЛОГИЯ РЕПЛИКАЦИИ ДНК

Топологические перестройки ДНК играют важную роль в проявлении ее функциональных активностей (репликации и транскрипции), а также в организации структуры более высокого порядка. Исходное описание ДНК как двухцепочечной спиральной молекулы предусматривало, что ее синтез включает разделение родительских цепей. Это предположение получило экспериментальное подтверждение, после того как был доказан полуконсервативный способ репликации ДНК (гл. 31). Цепи способны разделяться при поглощении энергии, достаточной для разрыва нековалентных связей, стабилизирующих двойную спираль. Следует учесть, однако, что цепи не просто лежат рядом друг с другом; они взаимозакручены, поэтому для их разделения необходимо вращение. Некоторые возможные способы раскручивания показаны на рис. 32.1.

Рассмотрим структуру ДНК, исходя из предположения, что ее молекула имеет свободный конец, который позволяет цепям вращаться вокруг оси для раскручивания (рис. 32.1, А). При данной длине спирали, однако, это привело бы разделяющиеся цепи к значительному количеству вращений, которое, по-видимому, невозможно в условиях клетки. Аналогичный результат был бы получен при наличии в области свободного конца специального механизма, контролирующего вращение (рис. 32.1, Б). Во вращение должен вовлекаться непомерно большой по длине материал, поскольку эффект необходимо передать на значительное расстояние. К тому же следует помнить, что репликация превращает каждую из отдельных вращающихся цепей в дочерний дуплекс, имеющий еще более громоздкую структуру.

ДНК ведет себя как замкнутая структура, утратившая свободные концы (гл. 28). Установление этого факта опровергло предложенные модели и со всей серьезностью поставило проблему топологии ДНК. Рассмотрим эффекты разделения двух цепей в молекуле, концы которой не способны свободно вращаться (рис. 32.1, В). Если две взаимозакрученные цепи разделяются в одном конце, это приводит к увеличению их взаимозакрученности дальше по длине молекулы. Следовательно, само по себе движение репликационной вилки увеличивает положительное суперскручивание впереди нее и делает дальнейшее продвижение репликационной вилки невозможным.

Проблема может быть решена введением временного одноцепочечного разрыва. Схема, представленная на рис. 32.1, Г, показывает, что наличие внутреннего свободного конца дает возможность разрезанной цепи вращаться вокруг интактной цепи, после чего разрез должен быть ликвидирован. Реакция разрезания и сшивки может повторяться по мере продвижения репликационной вилки. Такие реакции имеют место в случае использования в качестве субстратов замкнутых молекул ДНК. Они осуществляются с помощью ферментов, получивших название топоизомераз. Их действие при репликации (а также транскрипции, см. гл. 11) является решающим.

Ферменты, способные синтезировать новые цепи ДНК на матричной цепи, называют ДНК-полимеразами. Как прокариотические, так и эукариотические клетки имеют

несколько полимеразных активностей. Но только один-единственный фермент обладает функцией репликазы в бактериальном или эукариотическом ядре. Другие выполняют вспомогательную роль в репликации и (или) участвуют в репарирующем синтезе ДНК при замещении поврежденных последовательностей (гл. 34).

Принято думать, что ДНК-полимеразы обладают ДНК-синтезирующей активностью; однако вопрос о том, существует ли в виде дискретной единицы фермент ДНК-репликазы, остается спорным. (Мы уже обсуждали в главах 10 и 13 подобный вопрос в отношении РНК-полимеразы.) У бактерий может быть идентифицирован комплекс с активностью ДНК-репликазы, но он выделяется в виде отдельных (перекрывающихся) агрегатов из различных субъединиц. Не ясно, какая из них может рассматриваться как компонент самой репликазы, а какая служит вспомогательным фактором, необходимым для нормальной работы этого фермента. В процессе нормальной репликации ДНК-синтезирующая активность представляет только одну из нескольких взаимосвя-

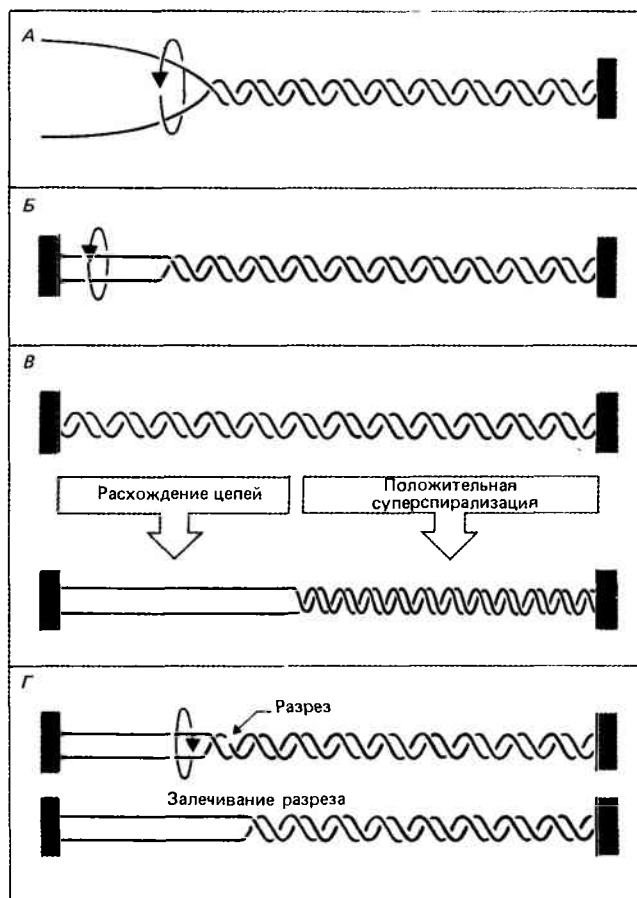


Рис. 32.1. Разделение цепей дуплексной ДНК может осуществляться посредством одного из указанных механизмов.

А. Вращение вокруг свободного конца. Б. Аппарат, удерживающий цепи, во время вращения. В. Компенсирующая положительная суперспирализация в другой области. Г. Разрезание, вращение и воссоединение.

занных функций. Вероятно, необоснованно считать ее отдельной функцией только потому, что мы можем выделить соответствующую ей ступень в общем процессе.

Еще не открыта ДНК-полимеразная активность, которая способна инициировать синтез дезоксирибонуклеотидной цепи. Все известные ферменты удлиняют ранее начатые цепи. Следовательно, необходимым звеном аппарата репликации является обеспечение затравочной активности, которая способна запустить синтез цепи ДНК. Эта стадия обычно включает синтез короткой рибонуклеотидной последовательности, которая позднее удаляется.

Описание топологии ДНК

До сих пор мы имели дело с организацией двухцепочечной ДНК в пространстве, имея в виду плотность суперспирализации, или число суперспиральных витков, приходящихся на один виток спирали. Суперспиральный виток образуется, когда двухцепочечная ДНК делает один оборот около своей оси в пространстве либо в том же направлении, в котором закручена сама двойная спираль (положительная суперспирализация), либо в противоположном направлении (отрицательная суперспирализация). Положительная суперспирализация «перекручивает» двойную спираль, отрицательная суперспирализация «недокручивает» ее (гл. 2).

После того как было введено понятие суперспирализации, для описания такого состояния приняли новую нomenclатуру. В ее основе лежит понятие о **числе зацеплений**, которое определяет, сколько раз две цепи замкнутой молекулы пересекают друг друга в пространстве. Следовательно, число зацеплений равно числу оборотов, которое одна цепь ДНК делает вокруг другой, если молекула ДНК расправлена (гипотетически) в одной плоскости. Принято определять число зацеплений в значениях числа правосторонних пересечений над левосторонними, поэтому для правозакрученной спирали указанная величина всегда положительна и выражается целым числом. Число зацеплений включает два компонента, T (*twist*) и W (*writhe*), определяемые уравнением

$$L = W + T.$$

T (**число оборотов**) характеризует молекулу ДНК, а именно степень закрученности одной ее цепи вокруг другой. T определяется числом пар оснований на виток. Для релаксированной замкнутой кольцевой ДНК, расправленной в одной плоскости, T соответствует отношению общего числа пар оснований к числу пар оснований в витке. Следовательно, T —это число витков в спирали.

W характеризует число оборотов оси спирали в пространстве. Для релаксированной молекулы $W = 0$. В этом случае число зацеплений равно числу оборотов одной цепи вокруг другой ($L = T$).

Параметр, измеряемый в определенных экспериментах, отражает изменение в числе зацеплений, ΔL . Величина этого параметра определяется уравнением

$$\Delta L = \Delta W + \Delta T.$$

Из него следует, что любое изменение в общем числе оборотов одной цепи ДНК вокруг другой складывается из изменений закрученности молекулы в пространстве (ΔW) и изменений взаимозакрученности самой спирали (ΔT). Пользуясь терминологией нашего предыдущего обсуждения, можно сказать, что ΔW аналогично изменению в су-

перспирализации, а ΔT характеризует недокручивание или перекручивание ДНК. Уменьшение в L , т.е. отрицательная величина ΔL , соответствует введению комбинации отрицательных суперспиралей и (или) недокручиванию или уменьшению положительной суперспирализации (перекручиванию). Положительная величина ΔL свидетельствует об увеличении положительной суперспирализации (перекручивания) или уменьшении отрицательной суперспирализации (недокручивания).

Важной особенностью числа зацеплений является инвариантность этого параметра. Он характеризует любую отдельную замкнутую молекулу ДНК (т.е. утратившую свободные концы). Число зацеплений не может быть изменено никакой деформацией молекулы, за исключением одной, включающей разрыв и воссоединение цепей. В результате кольцевая молекула с определенным числом зацеплений может выражать это число в значениях различных комбинаций T и W , но не может изменять их сумму, пока цепи остаются целыми. (В самом деле, распределение L между T и W предотвращает существование фиксированных значений последних параметров для молекулы ДНК в растворе.)

Теперь нам ясно значение числа зацеплений. Оно связано с действительными ферментативными событиями, с помощью которых осуществляются топологические изменения в ДНК. Фермент разрезает цепь или цепи, реорганизует их взаимосвязь в пространстве и затем соединяет между собой. Осуществляя это, он должен изменять число зацеплений на целое число. Его значение может характеризовать реакцию. Затем мы можем рассматривать эффекты этого изменения в значениях ΔW и ΔT .

Топологические перестройки ДНК

Перестройки, происходящие при репликации, могут быть связаны с реакциями, которые вызывают в отдельных молекулах ДНК структурные изменения более высокого порядка (в суперспирализации). Действительно, топологические перестройки ДНК—это основной аспект, который необходимо учитывать при рассмотрении всех ее функциональных активностей—рекомбинации, репликации и транскрипции. Следует напомнить, что для всех этих процессов дуплексная структура ДНК служит препятствием, которое преодолевается какой-либо реакцией, связанной с разделением цепей. Таким образом, все рекомбинационные события представляют собой топологические перестройки разных типов.

Разделение цепей требуют и все синтетические активности, связанные с двухцепочечной ДНК. Эта реакция может быть осуществлена при поглощении энергии, достаточной для разрушения нековалентных связей, которые стабилизируют двойную спираль.

Одинаковые молекулы, отличающиеся только числом зацеплений, называются **топологическими изомерами**. Ферменты, с помощью которых топологические изомеры могут превращаться из одной формы в другую, получили название **ДНК-топоизомераз**. Это название заменило многие более ранние, такие, как ДНК-релаксирующий фермент, шарнираза, фермент раскручивания, фермент разрезания-сшивки. Некоторые топоизомеразы могут релаксировать (удалять) только отрицательные супервитки в ДНК; другие способны релаксировать как положительно, так и отрицательно суперскрученные молекулы. Некоторые могут вводить отрицательные супервитки. Топо-

изомеразы делятся на два класса, в соответствии с природой механизмов, которые они используют. Ферменты типа I временно надрезают одну цепь ДНК. Ферменты типа II производят временный двухцепочечный разрез.

Наиболее полно охарактеризована топоизомераза типа I из *E. coli* (первоначально названная белком ω). Этот фермент эффективно релаксирует ДНК, отрицательно суперспирализованную в высокой степени. Скорость процесса замедляется по мере приближения числа зацеплений к значению, характерному для полностью релаксированной ДНК. Фермент не действует на положительно суперспирализованную ДНК. Он кодируется геном *top*; мутации в нем вызывают увеличение уровня суперспирализации в нуклеоиде (и могут оказывать влияние на транскрипцию, как описано в гл. 11).

Наряду с релаксацией отрицательных суперспиралей в двухцепочечной ДНК фермент взаимодействует и с одноцепочечной ДНК. Это показывает, что причина его специфичности в отношении отрицательных суперспиралей заключена в способности последних стабилизировать одноцепочечные участки, обеспечивающие субстрат для связывания с ферментом. Фермент не узнает определенную последовательность мишени, но проявляет неслучайное узнавание сайтов на ДНК. Природа такой предпочтительности в отношении определенных сайтов неизвестна.

При связывании топоизомеразы I *E. coli* с ДНК образуется стабильный комплекс, в котором одна цепь ДНК разрезается и ее 5'-фосфатный конец ковалентно связывается с тирозиновым остатком фермента. Это указывает на механизм действия фермента, который перемещает фосфодиэфирную связь к белку, манипулирует структурой двух цепей ДНК и затем воссоединяет исходную цепь. Перенос связи от нуклеиновой кислоты к белку объясняет, каким образом функционирует фермент, не требуя какого-либо источника энергии. В процессе реакции не происходит необратимого гидролиза связей; их энергия сохраняется после реакции переноса.

Модель действия топоизомеразы I представлена на рис. 32.2. Фермент связывается с областью, в которой произошло разделение двух цепей ДНК, разрезает одну из них, протаскивает другую через образовавшийся пробел и затем этот пробел залечивает. Такое событие изменяет число зацеплений на единицу за один раз. В каждом случае, когда одна цепь проходит через разрыв в другой, ΔL равно +1. На рисунке показано действие фермента на примере перемещения отдельных цепей. Однако в высоко суперскрученной молекуле свободная взаимозаменяемость W и T должна приводить к изменению числа зацеплений за счет изменения $\Delta W = +1$, т.е. за счет сбрасывания одного отрицательного сверхвитка. Реакция эквивалентна вращению, показанному на рис. 32.1, Г, с тем исключением, что данный фермент обеспечивает единственное протягивание цепи за один раз. (Наоборот, введение разреза в суперскрученную молекулу позволяет цепи свободно вращаться и уменьшать таким образом напряжение.)

Топоизомераза I способна также протаскивать один сегмент одноцепочечной ДНК через другой. Последствия такого одноцепочечного перемещения показаны на рис. 32.3.

Две одноцепочечные комплементарные кольцевые структуры способны ренатурировать, образуя двухцепочечное кольцо. Образование дуплексной области, соединяющей две кольцевые молекулы, должно породить экви-

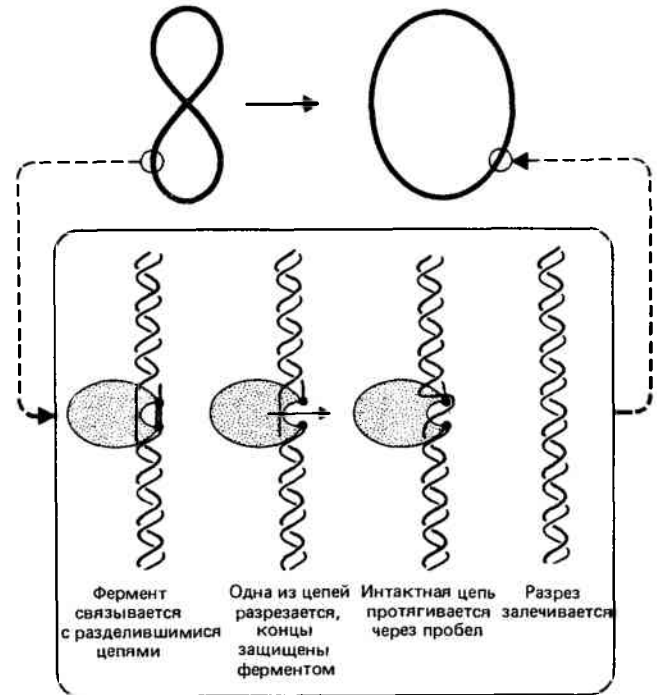


Рис. 32.2. Бактериальные топоизомеразы типа I узнают частично расплетенные сегменты ДНК и протягивают одну цепь через разрез, сделанный в другой цепи.

валентное недокручивание, т.е. отрицательные T и W в каждом кольце. В отсутствие фермента это предотвращает дальнейшее спаривание. Топоизомераза I уменьшает напряжение, протягивая обе цепи в направлении, противоположном направлению вращения, вовлекаемого в спаривание. Это дает возможность реакции продолжаться.

Если одноцепочечное перемещение осуществляется в двух частях одной и той же кольцевой молекулы, образуется узел.

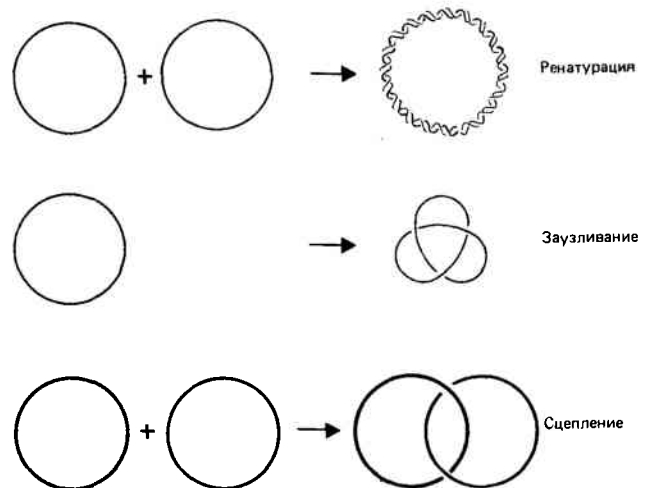


Рис. 32.3. Реакция одноцепочечного перемещения, осуществляемая топоизомеразой типа I, может способствовать ренатурации комплементарных колец одноцепочечной ДНК, заузливанию одноцепочечных ДНК и сцеплению кольцевых (разрезанных) дуплексных молекул ДНК.

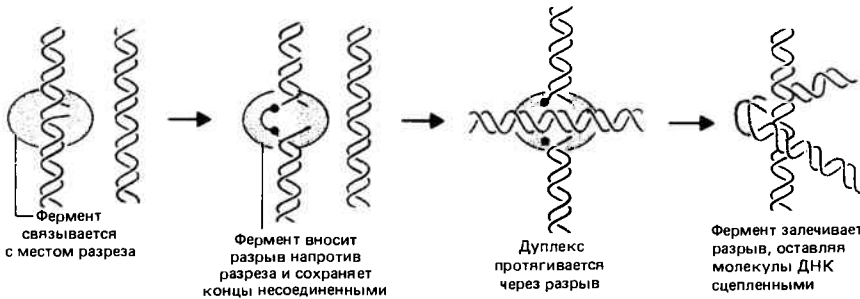


Рис. 32.4. Топоизомеразы типа I способны протягивать двухцепочечную ДНК через интактную отдельную цепь напротив разреза в другой двухцепочечной ДНК.

Возможности топоизомеразы I расширяются за счет присущей ей другой активности, открытой позднее. Фермент способен действовать на две кольцевые двухцепочечные молекулы ДНК, делая разрез в одной из цепей одной молекулы. При протягивании одного кольца через другое они становятся взаимосцепленными. Структура получила название **катенана**; активность описывается как сцепление ДНК. (Формально реакции сцепления и заузливания топологически подобны, так как первая включает сцепление различных кольцевых молекул, вторая – сцепление в пределах одного кольца.)

Предполагается, что сцепление двухцепочечных кольцевых молекул происходит согласно модели, проиллюстрированной на рис. 32.4. Фермент связывается напротив разреза, вводит разрез в другую цепь, протаскивает вторую двухцепочечную ДНК через образовавшийся провал, после чего этот провал заливается. Единственным отличием от процедуры, приводящей к заузливанию одноцепочечной ДНК, является то, что реакция перемещения в этом случае включает перенос двухцепочечной ДНК. Поскольку двухцепочечные и одноцепочечные ДНК отличаются по структуре, кажется удивительным, что фермент способен протаскивать любую из них. Возможно, однако, что фермент на самом деле протягивает через образующийся провал одну цепь дуплексной ДНК за один раз (предполагаемые интермедиаты имели бы сложную топологию).

Формально топоизомеразы типа I эукариот имеют похожие свойства, однако они способны релаксировать как положительные, так и отрицательные суперспирали.

Топоизомеразы типа II были охарактеризованы как у прокариот, так и у эукариот. Большинство из них релаксирует как отрицательные, так и положительные суперспирали. (Реакция АТР-зависима.) Вероятно, для каждого каталитического события гидролизует одна единица АТР. Реакция осуществляется посредством двухцепочечного разреза в одной области двухцепочечной ДНК и продевания другой дуплексной области через провал. Эта реакция показана на рис. 32.5. Формальным следствием двухцепочечного переноса является изменение

числа зацеплений на величину, кратную двум. Активность топоизомеразы II может быть также использована для получения сцепленных дуплексных кольцевых молекул и заузленных молекул.

Реакция, вероятно, включает неспецифическое узнавание дуплексной ДНК, в результате которого фермент присоединяется к любым двум двухцепочечным пересекающимся сегментам. Гидролиз АТР вызывает конформационные изменения фермента, которые обеспечивают энергию, необходимую для протягивания одной дуплексной ДНК через провал, сделанный во второй. Благодаря топологии суперспирализованной ДНК взаимоотношение пересекающихся сегментов должно быть таким, чтобы снималась суперспирализация как в положительно, так и отрицательно суперспирализованных кольцах. Наиболее охарактеризованным ферментом этого типа является топоизомераза II фага T4, которая кодируется тремя генами; ее активность оказывается существенной для репликации фаговой ДНК.

Гириза вводит отрицательные суперспирали в ДНК

Бактериальная ДНК-гириза (наиболее изученная у *E. coli*) представляет собой топоизомеразу типа II, способную вводить отрицательную суперспирализацию в релаксированную замкнутую кольцевую молекулу. ДНК-гириза связывается с кольцевой двухцепочечной ДНК и суперспирализует ее процессивно и каталитически: она продолжает вводить суперспирали в одну и ту же молекулу ДНК. Одна молекула ДНК-гиризы способна ввести примерно 100 суперспиралей в минуту.

Суперспирализованная форма ДНК обладает большей свободной энергией, чем релаксированная форма, и энергия, необходимая для завершения превращения, восполняется при гидролизе АТР. В отсутствие АТР гириза способна релаксировать отрицательные, но не положительные суперспирали со скоростью, в 10 раз более низкой, чем скорость введения суперспиралей.

Гириза *E. coli* состоит из двух типов субъединиц. Она

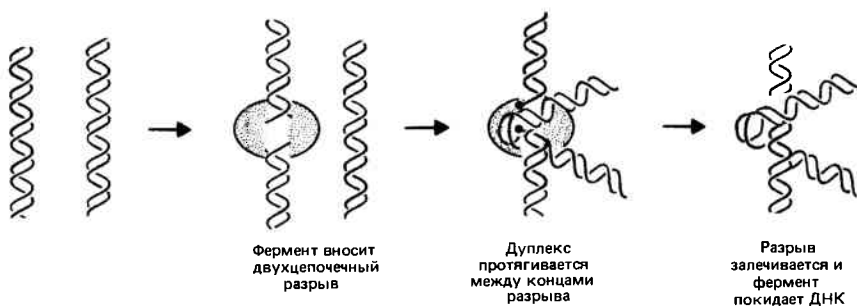


Рис. 32.5. Топоизомеразы типа II способны протягивать двухцепочечную ДНК через двухцепочечный разрыв на другой дуплексной молекуле ДНК.

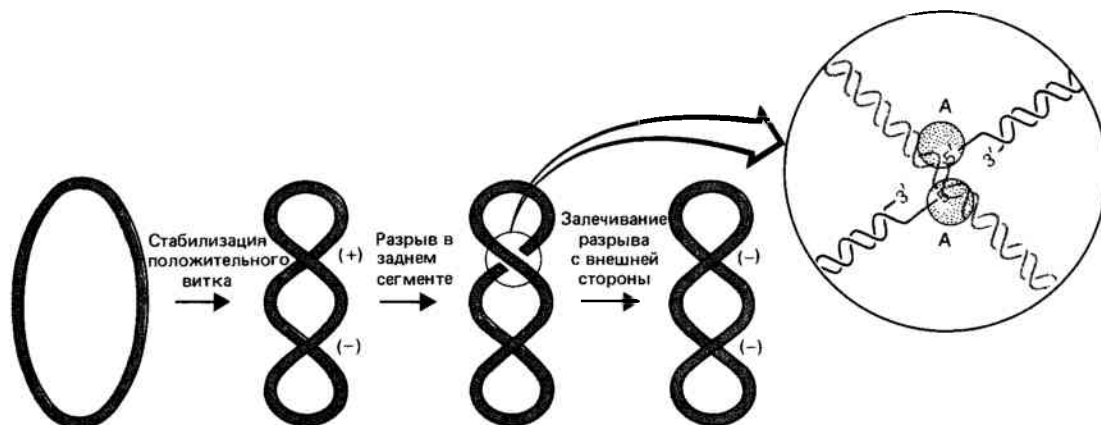


Рис. 32.6. ДНК-гираза способна вносить отрицательные супервитки в двухцепочечную ДНК посредством инверсии положительной суперспирали.

ингибируется двумя видами антибиотиков, каждый из которых действует на определенную субъединицу. Мутации, обуславливающие устойчивость к антибиотикам, способствовали идентификации локусов, кодирующих субъединицы ДНК-гиразы. Локус *gyrA* (первоначально носивший название *nalA*) контролирует устойчивость к налидиксовой и оксолиновой кислотам. Локус *gyrB* (ранее называвшийся *soi*) детерминирует устойчивость к каувермицину А1 и новобиоцину. Эти антибиотики подавляют репликацию, что, вероятно, свидетельствует об участии ДНК-гиразы в синтезе ДНК.

Активный фермент ДНК-гираза имеет структуру A_2B_2 с общей мол. массой около 400 000 дальтон. Субъединица GyrA имеет мол. массу, равную приблизительно 105 000, субъединица GyrB — около 95 000 дальтон. (Клетки *E. coli* содержат субъединицы А в количестве, в 10 раз превышающем количество субъединиц В. Причиной этого является дополнительное связывание субъединицы А с белком, названным *v*, имеющим мол. массу 50 000 дальтон который может быть фрагментом субъединицы В. Агрегат Av получил название топоизомеразы II *E. coli*; по-видимому, он является природным компонентом клетки.)

Гириза связывается с субстратной ДНК, оборачивая ее вокруг тетрамерного белка. Фермент защищает примерно 140 пар оснований ДНК от переваривания нуклеазой микрококка (подобно защите, обеспечиваемой значительно меньшим по размеру гистоновым октамером).

Модель «инверсии знака», объясняющая действие ДНК-гиразы, представлена на рис. 32.6. Фермент связывается с ДНК в месте перекреста, эквивалентного положительной суперспирали. Это событие индуцирует образование компенсирующей отрицательной спирали в несвязанной части ДНК. Затем фермент разрезает двойную цепь в области пересечения положительной суперспирали, продевает через образующийся пробел другую дуплексную цепь и залечивает разрыв. В результате изменяется знак спирали от +1 к -1, а число зацеплений изменяется на величину $\Delta L = -2$. Последнее изменение происходит, согласно правилу изменения числа зацеплений, на величину, кратную двум, для случая двухцепочечных переносов. Один из пересекающихся сегментов образовавшейся отрицательно суперспирализованной молекулы освобождается от гиразы; по длине ДНК происходит перераспределение отрицательных супервитков (с изменением значений либо T, либо W, либо и того и другого),

и цикл повторяется снова. Тот же тип топологических перестроек ответствен за образование сцепленных и заузленных молекул, хотя роль таких перестроек *in vivo* неизвестна.

При отделении гиразы от инвертированной суперспирали ее конформация изменяется. Для того чтобы предпринять новый цикл суперспирализации, фермент должен восстановить свою исходную конформацию. Такой процесс получил название **оборота фермента**. Полагают, что этот оборот осуществляется благодаря гидролизу АТФ, поскольку замещение АТФ негидролизуемым аналогом позволяет гиразе ввести только одну инверсию (-2 суперспирали) на субстрат. Следовательно, АТФ не требуется для реакции суперспирализации, но необходим для осуществления второго цикла этой реакции. Новобиоцин препятствует осуществлению АТФ-зависимых реакций гиразы, предотвращая связывание АТФ с субъединицей В.

Реакция релаксации (АТФ-независимая) подавляется налидиксовой кислотой. Это свидетельствует в пользу того, что субъединица А вовлекается в осуществление разрыва и воссоединения. Обработка гиразы налидиксовой кислотой позволяет выделить ДНК в виде фрагментов, образованных в результате ступенчатого разреза поперек дуплексной молекулы. Концы содержат свободную 3'-ОН-группу и одноцепочечные 5'-концы протяженностью в 4 основания, ковалентно связанные с субъединицей А (рис. 32.6). Ковалентная сшивка сохраняет энергию фосфатной связи; она может быть использована для проведения повторной реакции, что объясняет ее способность осуществлять релаксацию в отсутствие АТФ. Сайты разрезания довольно специфичны и встречаются примерно один раз на каждые 100 пар оснований.

ДНК-полимеразы эукариот

Эукариотические ДНК-полимеразы наиболее полно изучены у млекопитающих. Первый эукариотический фермент такого типа был выделен из тимуса теленка еще в 1960 г. Как правило, ДНК-полимеразы эукариот обладают только ДНК-синтезирующей активностью, в то время как большинство бактериальных ферментов могут выполнять также функции экзонуклеаз.

У эукариот известны три класса ДНК-полимераз. Их свойства суммированы в табл. 32.1. Может ли одна из

Таблица 32.1

В клетках млекопитающих содержатся три ДНК-полимеразы

	ДНК-полимераза α	ДНК-полимераза β	ДНК-полимераза γ
Локализация	Ядерная, однако при делении переходит в цитоплазму	Ядерная	Ядерная и митохондриальная
Степень активности, %	~ 80	10–15	2–15
Размер (в дальтонах)	110–220 000	45 000	60 000
Субъединицы	Несколько	1	?

них быть идентифицирована в качестве подлинной репликазы?

ДНК-полимераза α является единственным ферментом, уровень которого возрастает во время S-фазы. Она устойчива к 2'3'-дидезокситимидинтрифосфату; аналог не подавляет репликационные системы *in vitro*, но ингибирует полимеразы β и γ . Антибиотик афидиколин, выделенный из грибов, подавляет репликацию эукариотической ДНК *in vivo*; единственный фермент, который он ингибирует *in vitro* – ДНК-полимераза α . Следует отметить, что ДНК-полимераза α , выделенная из мутантных клеток, устойчивых к афидиколину, также нечувствительна к этому антибиотику *in vitro*. Полученные результаты позволили идентифицировать ДНК-полимеразу α как ядерную ДНК-репликазу.

Кроме репликации ДНК-полимеразы выполняют репарационные функции, с помощью которых замещаются сегменты поврежденной ДНК. По-видимому, в этом принимает участие ДНК-полимераза β .

Роль ДНК-полимеразы γ предположительно сводится к осуществлению репликации митохондриальной ДНК. Ее роль в ядре (если она в ядре функционирует) неизвестна.

Репликация ДНК наряду с ДНК-полимеразой требует участия многих других ферментов. Трудности при анализе роли эукариотических ДНК-полимераз обусловлены отсутствием систем, содержащих все необходимые для репликации компоненты. В последнее время, однако, появилась возможность осуществлять *in vitro* реплика-

цию определенных матриц. Были созданы ядерные системы, содержащие все компоненты, необходимые для продолжения движения репликационной вилки. Однако инициацию синтеза в стартовой точке репликации в системе *in vitro* еще не удалось воспроизвести.

ДНК-полимеразы прокариот проявляют несколько ферментативных активностей

У *E. coli* обнаружены три ДНК-полимеразы. ДНК-полимераза I, вовлекается в репарацию повреждений ДНК и участвует в процессе полуконсервативной репликации, выполняя вспомогательную роль. Она кодируется локусом *polA*. Второй фермент, ДНК-полимераза II, детерминруется геном *polB*. О его функции *in vivo* ничего не известно. Третий фермент, ДНК-полимераза III, представляет собой сложный белок, состоящий из многих субъединиц, включая субъединицу, кодируемую локусом *polC* (известным также как *dnaE*). Фермент выполняет функции репликазы и отвечает за синтез *de novo* новых цепей ДНК. Данные об активностях трех ферментов представлены в табл. 32.2.

Все ДНК-полимеразы про- и эукариот обладают одной и той же основной синтезирующей активностью. Каждая из них способна удлинять цепь ДНК, последовательно наращивая по одному нуклеотиду к 3'-ОН-концу. Таким образом, синтез новой цепи осуществляется в направлении 5' → 3'. Реакция схематически изображена на рис. 32.7. Предшественником синтеза ДНК служит нуклеозидтрифосфат, который в процессе реакции утрачивает две конечные фосфатные группы. Выбор нуклеотида, добавляемого к цепи, диктуется спариванием основания с матричной цепью.

Точность процесса репликации зависит от тех же причин, которые мы уже обсуждали при рассмотрении процесса трансляции. Она определяется специфичностью спаривания оснований. Ожидаемая при этом частота встречаемости ошибок находится в пределах от 10^{-4} до 10^{-5} на реплицируемую пару оснований. Действительная частота ошибок в клетках *E. coli* или при репликации ДНК фага T4 колеблется в пределах от 10^{-8} до 10^{-10} на реплицируемую пару оснований. (Это соответствует приблизительно одной ошибке на 1000 репликационных циклов

Таблица 32.2

В клетках *E. coli* содержатся три ДНК-полимеразы

	ДНК-полимераза I	ДНК-полимераза II	ДНК-полимераза III
Размер (в дальтонах)	109 000	120 000	250 000
Состав	Одна цепь	—	Гетеромультимер
Количество молекул на 1 клетку	~ 400	—	10–20
Ферментативная активность	Элонгация в направлении 5' → 3' от 3'-ОН-затравки 3' → 5'-экзонуклеаза 5' → 3'-экзонуклеаза	Элонгация в направлении 5' → 3' от 3'-ОН-затравки 3' → 5'-экзонуклеаза	Элонгация в направлении 5' → 3' от 3'-ОН-затравки 3' → 5'-экзонуклеаза 5' → 3'-экзонуклеаза
Мутанты	<i>polA</i> -мутанты дефектны в репарационных активностях, повреждение 5' → 3'-экзонуклеазной активности снижает их жизнеспособность	<i>polB</i> -мутанты не имеют фенотипического проявления	<i>polC</i> , <i>dnaN</i> , <i>dnaX</i> , <i>dnaZ</i> – условно летальные мутанты по репликации ДНК

бактериальной ДНК, или примерно 10^{-6} на один ген на поколение.)

Фермент может контролировать выбор комплементарного основания на любой (или в обеих) из двух стадий. Путем специфического узнавания определенного химического признака он может тщательно проверить поступающее основание относительно его комплементарности с матричным основанием. Таким образом осуществляется контроль за ошибками на стадии, предшествующей синтезу. Проверка правильности подбора основания возможна и после добавления основания к цепи. В тех случаях, когда была допущена ошибка, удаляется самое последнее из присоединенных оснований. Все бактериальные ферменты обладают 3'—5'-экзонуклеолитической активностью, которая, по-видимому, и осуществляет функцию исправления ошибок. Отсутствие такой активности у ферментов эукариот свидетельствует о том, что в их клетках должен существовать какой-то другой механизм контроля частоты ошибок.

3'—5'-Экзонуклеолитическая активность проявляется в направлении, обратном синтезу ДНК. Действие фермента, исправляющего ошибки, схематически показано на рис. 32.8. На стадии роста цепи нуклеотид предшественник занимает положение в конце растущей цепи. Формируется связь. Фермент передвигается дальше на одну пару оснований, готовую к спариванию со следующим нуклеотидом-предшественником. Если происходит ошибочное спаривание, фермент перемещается в обратном направлении и вырезает последнее добавленное основание, на место которого может присоединиться правильный нуклеотид-предшественник. Такого рода активности была наиболее полно охарактеризована у фермента ДНК-полимеразы I, однако вероятно, что другие ДНК-полимеразы прокариот функционируют аналогично.

ДНК-полимераза I представляет собой необычный фермент с несколькими активностями. Она содержит один полипептид с мол. массой 109 000 дальтон. При протеолитической обработке полипептидная цепь делится на два отрезка. Большой фрагмент (76 000 дальтон) проявляет 5'—3'-полимеразную и 3'—5'-экзонуклеазную активности. Малый фрагмент (36 000 дальтон) обладает другой активностью—5'—3'-экзонуклеолитической. Последняя поэтому оказывается независимой от синтезирующей (исправляющей ошибки) активности. С помощью 5'—3'-экзонуклеазной активности вырезаются небольшие участки, содержащие до 10 нуклеотидов.

ДНК-полимераза I проявляет уникальную способность инициировать репликацию *in vitro* в месте разрыва ДНК. В точке, где разрушена фосфодиэфирная связь в двухцепочечной ДНК, фермент удлиняет 3'-ОН-конец. Когда новый сегмент ДНК синтезирован, он *замещает существующую гомологичную цепь в двойной цепи*. Такой процесс «смещения разрыва» показан на рис. 32.9. Замещаемая цепь деградирует с помощью 5'—3'-экзонуклеолитической активности фермента. В результате сохраняется ДНК с неизменными свойствами, за исключением того сегмента ДНК, который был замещен вновь синтезированным материалом, где положение разрыва оказалось сдвинутым по длине двойной цепи. Такой эффект имеет большое практическое значение, так как именно он позволяет внедрить в ДНК *in vitro* радиоактивно меченные нуклеотиды. Смещение разрыва (ник-трансляция)—это основной метод, используемый для получения радиоактивно меченной ДНК.

Хотя в бактериальной клетке содержится большое ко-

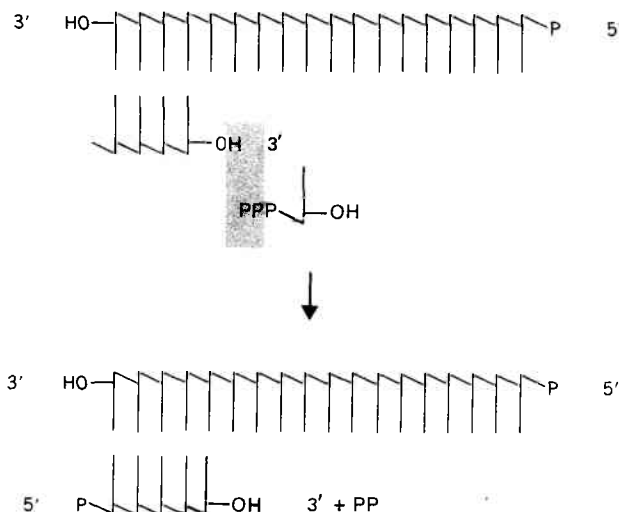


Рис. 32.7. Синтез ДНК осуществляется посредством добавления нуклеотидов к 3'-ОН-концу растущей цепи.

личество ДНК-полимеразы I, она не отвечает за репликацию, которая продолжается в мутантах *polA*. Ее 5'—3'-синтезирующая/3'—5'-экзонуклеолитическая активности, вероятно, используются *in vivo* для превращения одноцепочечных участков в двухцепочечные. Такие участки возникают во время репликации (см. ниже и гл. 33) и в случае удаления из ДНК поврежденных оснований (гл. 34).

Фермент ДНК-полимераза III был открыт при изучении условно-летальных мутаций в *polC*-локусе, который кодирует общий компонент всех препаратов фермента—его α -субъединицу с мол. массой 140 000. Очищенные препараты фермента с активностями ДНК-полимеразы III состоят из нескольких субъединиц. Мы использовали слово «активности» во множественном числе, поскольку фермент был получен в нескольких формах, и остается неизвестным, какая из них представляет собственно фермент, включающийся в репликацию *in vivo*. Данные о различных формах фермента суммированы в табл. 32.3.

Все субъединицы, перечисленные в таблице, необхо-

Таблица 32.3

Различные комбинации субъединиц приводят к образованию разных форм ДНК-полимеразы III

Субъединица	Размер (в дальтонах)	Другое наименование (или ген)	Субъединицы содержатся в			
			Кор Pol III	Pol III'	Pol III ¹⁾	Голо-фермент
α	140 000	<i>polC(dnaE)</i>	+	+	+	+
β	40 000	<i>dnaN</i> -фактор I-ко Pol III ¹⁾			+	+
δ	32 000	<i>dnaX</i> -фактор III } фактор II <i>dnaZ</i>			+	+
γ	52 000				+	+
ϵ	25 000		+	+	+	+
θ	10 000		+	+	+	+
τ	83 000			+	+	+

¹⁾ Стехиометрия различных комплексов не определена, хотя в большинстве случаев содержится только по одной копии каждой из субъединиц.

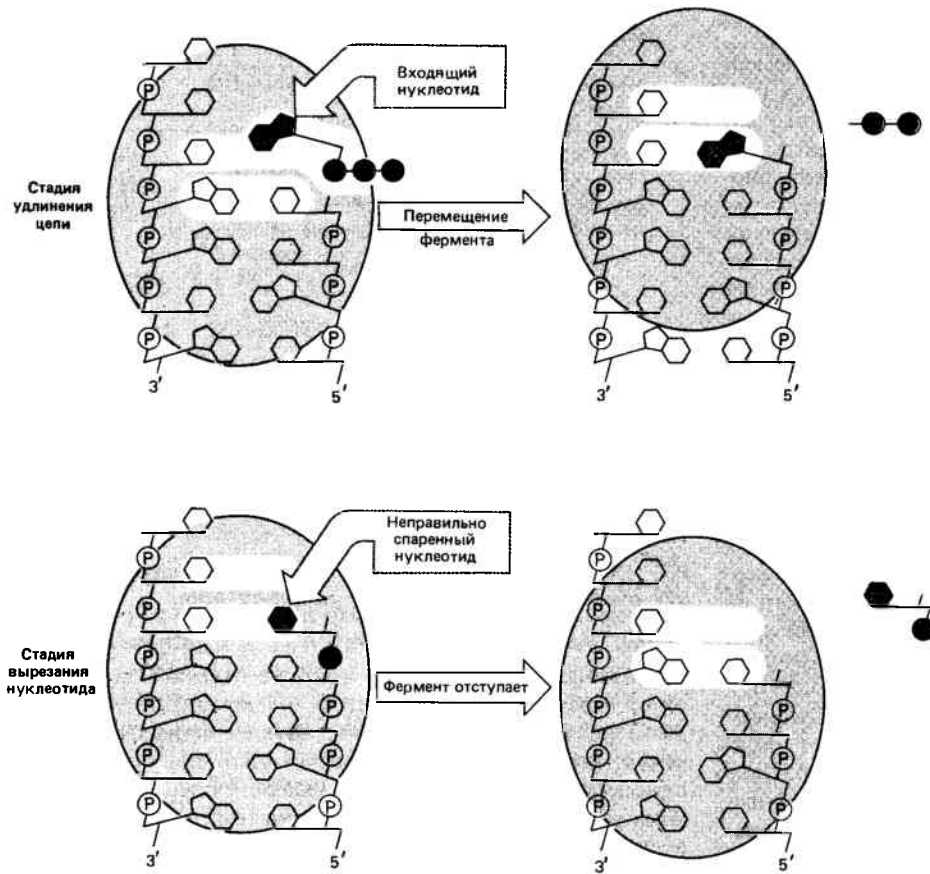


Рис. 32.8. Бактериальные ДНК-полимеразы проверяют правильность подбора пары оснований в конце растущей цепи и в случае ошибки вырезают добавленный нуклеотид.

димы для проявления полной активности в репликационной системе *in vitro* (гл. 33). Были предприняты попытки разделить субъединицы собственно ДНК-репликазы и вспомогательные белки. Однако системы *in vitro* содержат много дополнительных активностей, и не ясно, существует ли в действительности такой фермент «ДНК-репликаза». Функция синтеза ДНК может осуществляться лишь частью большого и сложного аппарата репликации, у которого выявляемая репликазная активность является активностью полипептида, непосредственно вовлекаемого в катализируемый рост цепи ДНК.

Бактерия *Bacillus subtilis* также имеет три ДНК-полимеразы, которые пронумерованы с целью сопоставления с их двойниками в *E. coli*. Существуют некоторые различия в экзонуклеолитических активностях бациллярных ДНК-полимераз по сравнению с ферментами *E. coli*. Субъединица, у которой идентифицирована ДНК-синтезирующая активность ДНК-полимеразы III *Bac. subtilis*,

стимулируется факторами I и II *E. coli* (субъединицы $\beta\delta\nu$), что свидетельствует о существовании сходства в разделении функций репликационных аппаратов двух бактерий. Уровень ДНК-полимеразы II зависит от физиологического состояния: она отсутствует в покоящихся спорах и появляется примерно через 30 мин после начала прорастания.

Некоторые фаги также кодируют ДНК-полимеразы. Среди них такие, как T4, T5, T7 и SP01. Все ферменты обладают как 5'—3'-синтезирующей, так и 3'—5'-эксонуклеолитической активностями. В каждом случае мутация в гене, кодирующем один фаговый полипептид, предотвращает развитие фага. Каждый фаговый полимеразный полипептид связан с другими белками, имеющими фаговое или бактериальное происхождение. Репликационные системы фагов T4 и T7 обсуждаются в гл. 33.

Синтез ДНК является полунепрерывным

В связи с антипараллельной структурой двух цепей дуплексной ДНК возникает проблема их репликации. При движении репликационной вилки дочерние цепи должны синтезироваться на обеих родительских цепях. Вилка перемещается в направлении от 5'-к 3'-концу на одной цепи и от 3'-к 5'-концу на другой. Однако нуклеиновые кислоты синтезируются только в направлении от 5'-к 3'-концу. Противоречие разрешается благодаря тому, что цепь в действительности синтезируется в виде коротких фрагментов с обычной 5'→3'-полярностью. Их последующее объединение приводит к росту цепи в 3'→5'-направлении.

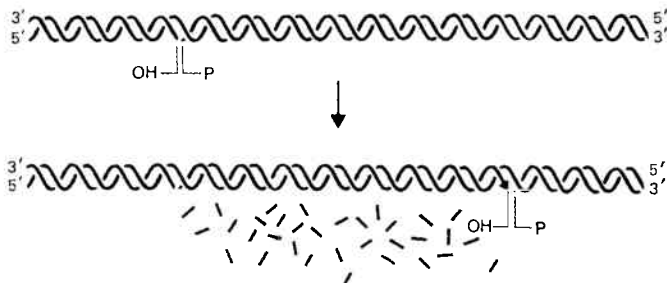
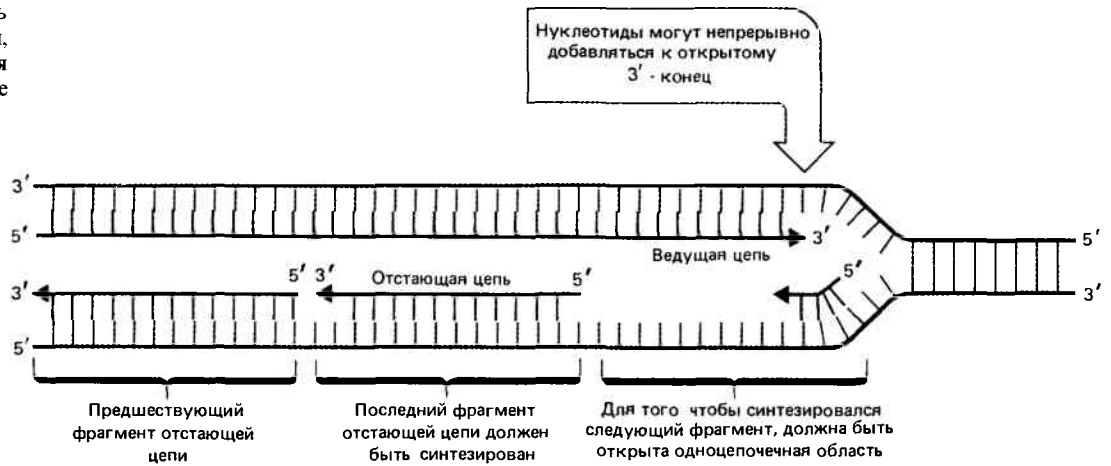


Рис. 32.9. Смещение разрыва ведет к замещению части предсуществующей цепи дуплексной ДНК на вновь синтезированный материал.

Рис. 32.10. Ведущая цепь синтезируется непрерывно, в то время как отстающая цепь синтезируется в виде фрагментов.



Рассмотрим область, находящуюся рядом с репликационной вилкой (рис. 32.10). Если молекула ДНК раскручена, то на одной цепи синтез может идти непрерывно в направлении 5' → 3'. Эта цепь получила название **ведущей**. Однако на другой цепи участок одноцепочечной родительской ДНК должен быть раскрыт, и тогда на нем синтезируется сегмент ДНК в направлении, обратном движению репликационной вилки. Эта цепь получила название **отстающей**. В направлении 5' → 3' синтезируется серия таких фрагментов, которые затем соединяются, образуя интактную отстающую цепь. Такой способ синтеза назван **прерывистой репликацией**.

Прерывистая репликация изучена путем прослеживания за судьбой вновь синтезированной ДНК, получившей очень короткую метку радиоактивности. Метка поступает во вновь синтезируемую ДНК, имеющую вид коротких фрагментов, седиментирующих в пределах 7–11S, что приблизительно соответствует последовательностям длиной 1000–2000 оснований. Такие фрагменты были названы **фрагментами Оказаки**. Они обнаружены во всех реплицирующихся ДНК, как прокариот, так и эукариот. При более длительной инкубации метка входит в большие по длине сегменты ДНК. Они образуются в результате ковалентного связывания фрагментов Оказаки. Отстающая цепь *должна* синтезироваться в форме фрагментов Оказаки. Довольно продолжительное время велась полемика относительно того, синтезируется ли ведущая цепь таким же способом или ее синтез идет непрерывно.

У *E. coli* вся вновь синтезированная ДНК существует в виде коротких фрагментов. На модели фага T4 было показано, что обе цепи представлены в виде фрагментов Оказаки. На первый взгляд это свидетельствует о прерывистом синтезе обеих цепей. Однако оказалось, что не все фрагменты являются истинными фрагментами Оказаки; многие из них псевдофрагменты, образуемые при разрывах цепи ДНК, которая на самом деле синтезирована как непрерывная цепь.

Причиной таких разрывов служит введение в ДНК урацила вместо тимина. Сначала полагали, что такое неправильное включение происходит редко. Но сейчас известно, что ДНК-полимераза III *E. coli* не способна точно различать dUTP и dTTP in vitro и, по-видимому, in vivo.

На рис. 32.11 представлена схема пути биосинтеза dUTP и dTTP. Примерно 75% dUTP образуется путем дезаминирования dCTP, остальное – при фосфорилировании dUTP. С помощью фермента dUTPазы dUTP превращается в dUMP (который образуется и путем дезами-

нирования dCMP). dUMP представляет собой промежуточное соединение, используемое для синтеза dTMP, необходимого для образования dTTP, предшественника в синтезе ДНК.

При нормальных условиях соотношение dUTP и dTTP в *E. coli* составляет примерно 1 : 300. Если dUTP и dTTP не различаются ферментом, то на каждые 1200 оснований вновь синтезированной ДНК будет происходить включение примерно одного остатка урацила. Он удаляется ферментом **урацил-ДНК-гликозидазой**, который играет важную роль, защищая ДНК от мутаций и от ошибок репликации. Активность урацил-ДНК-гликозидазы была обнаружена как у прокариот, так и у эукариот. Вероятно, существует унифицированная система защиты ДНК от присутствия урацила.

Удаление урацила происходит путем отщепления этого основания от сахара. При этом в фосфодиэфирную связь вводится разрыв. Он может быть превращен в пробел в результате удаления других нуклеотидов. Такие пробелы застраиваются синтезирующейся ДНК, замещающей утраченный материал. Серия таких событий показана на рис. 32.12. В течение периода времени между появлением и застройкой пробела разрывы во вновь синтезируемой ДНК вызывают образование фрагментов, сравнимых по размеру с фрагментами Оказаки. (Подобные события приводят также к уменьшению размера фрагментов Оказаки отстающей цепи.)

Роль двух ферментов, урацил-ДНК-гликозидазы и dUTPазы, выявлена при изучении эффектов мутаций

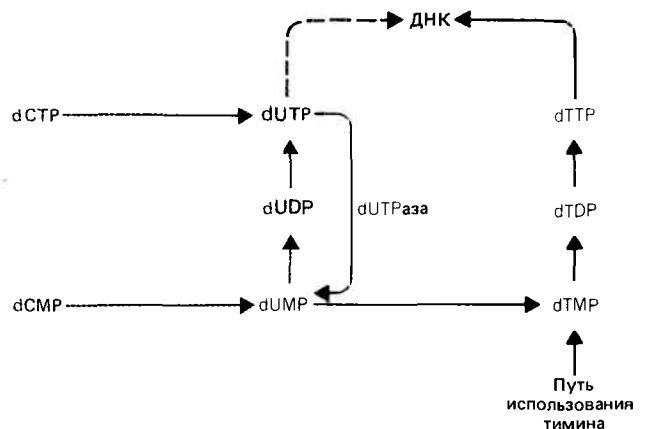


Рис. 32.11. Синтез dUTP и dTTP связан.

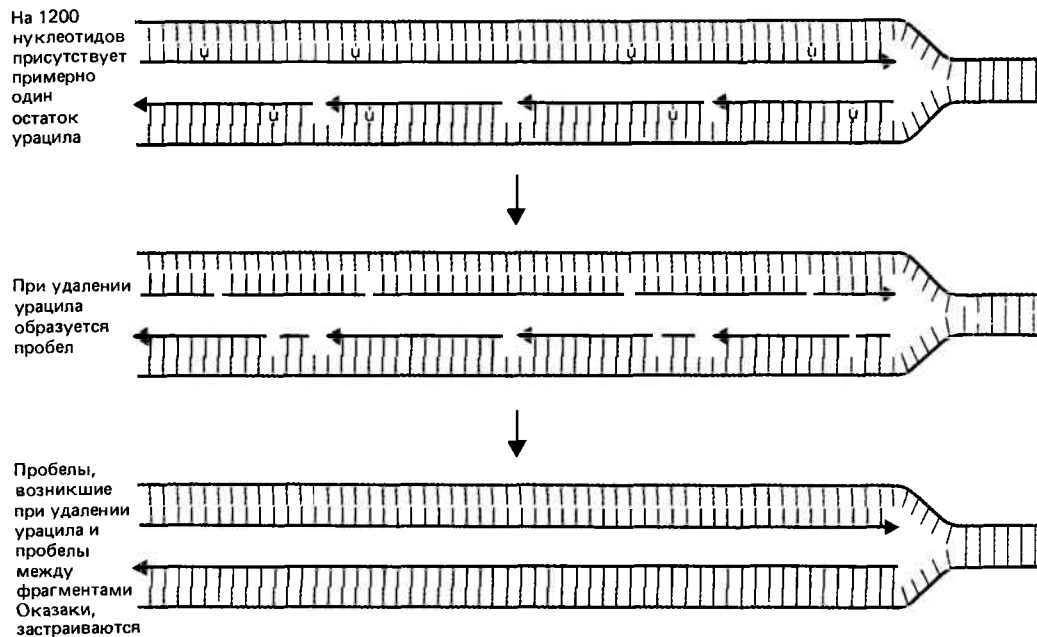


Рис. 32.12. Вырезание урацила из вновь синтезированной ДНК ведет к образованию пробелов как в непрерывно синтезируемой ведущей цепи, так и во фрагментах Оказаки отстающей цепи.

Для образования непрерывных цепей ДНК пробелы должны быть застроены и фрагменты Оказаки соединены.

в генах, кодирующих эти ферменты. Выщепление урацила предотвращается в мутантах *ung⁻*, утративших гликозидазный фермент. У таких мутантов примерно половина радиоактивной метки появляется во вновь синтезированной ДНК в форме истинных фрагментов Оказаки; другая половина включается в значительно большие по размеру фрагменты. Отсюда следует, что ведущая цепь синтезируется непрерывно.

Клеточный баланс между dUTP и dTTP в большой степени зависит от активности dUTPазы, поскольку она одновременно служит как для удаления dUTP, так и для увеличения эффективности пути, ведущего к синтезу dTTP. Фермент dUTPаза неактивен в мутантах *E. coli dut⁻* (называемых также *sof⁻*). Его отсутствие приводит к значительному увеличению количества dUTP за счет dTTP, в результате чего мутантные клетки накапливают мелкие фрагменты вновь синтезированной ДНК.

Таким образом, в клетках *E. coli* отстающая цепь синтезируется прерывисто, а ведущая — непрерывно. По-видимому, то же самое происходит в других системах. Такой способ синтеза ДНК получил название **полунепрерывной репликации**.

Синтез фрагментов Оказаки инициируется РНК

Мы уже упоминали, что все известные ДНК-полимеразы, как прокариот, так и эукариот, способны удлинять дезоксирибонуклеотидную цепь от свободного 3'-ОН-конца; однако ни одна из них не может инициировать синтез. Кроме того, события инициации должны происходить непрерывно, поскольку каждый фрагмент Оказаки отстающей цепи нуждается в своей собственной инициации de novo. Синтез каждого фрагмента Оказаки начинается с **затравки**, короткой последовательности, обеспечивающей 3'-ОН-конец для удлинения цепи с помощью ДНК-полимеразы. Затравкой служит РНК, длина которой соответствует примерно 10 основаниям.

Мы можем теперь подробнее обсудить события, которые приводят к воссоединению фрагментов Оказаки (рис. 32.13). Их порядок еще не установлен, однако очевидно, что он должен включать удаление РНК-затравки, ее замещение удлиняющейся ДНК и ковалентное связывание смежных фрагментов.

На рисунке показано, что синтез фрагментов Оказаки терминируется как раз перед стартовой точкой РНК-затравки предшествующего фрагмента. Фермент ДНК-полимераза I, вероятно, вовлекается в застройку пробела, образуемого при удалении затравки, поскольку у мутантов *polA* фрагменты Оказаки должным образом не соединяются. Действительно этот фермент может использовать свою 5'-3'-экзонуклеазную активность для удаления РНК-затравки с одновременным замещением ее последовательностью ДНК, синтезируемой с 3'-ОН-конца соседнего фрагмента Оказаки. Она эквивалентна активности при смещении разрыва, с той разницей что в этом случае новая ДНК замещает РНК, а не сегмент ДНК. После удаления РНК и ее замещения остается только соединить смежные фрагменты Оказаки. После застройки пробелов 3'-ОН-конец одного фрагмента присоединяется к 5'-фосфатному концу предыдущего фрагмента. За соединение концов отвечает фермент **ДНК-лигаза** (о котором мы уже говорили в связи с его использованием в экспериментах *in vitro* при реконструкции ДНК). Лигазы присутствуют в клетках как прокариот, так и эукариот; наиболее полно изучены лигазы *E. coli* и фага Т4. Об участии лигирующих (сшивающих) ферментов в процессе воссоединения фрагментов Оказаки *in vivo* свидетельствует неспособность к их соединению мутантов *lig⁻*.

Ферменты *E. coli* и Т4 могут залечивать разрывы, имеющие 3'-ОН- и 5'-фосфатные концы (рис. 32.14), хотя используют для этого разные кофакторы: *E. coli* — NAD (никотинамидадениндинуклеотид), а лигаза фага Т4 — АТР. Оба фермента участвуют в двухступенчатой реакции, на первом этапе которой образуется комплекс фермент-АМР. У *E. coli* это достигается путем отщепле-

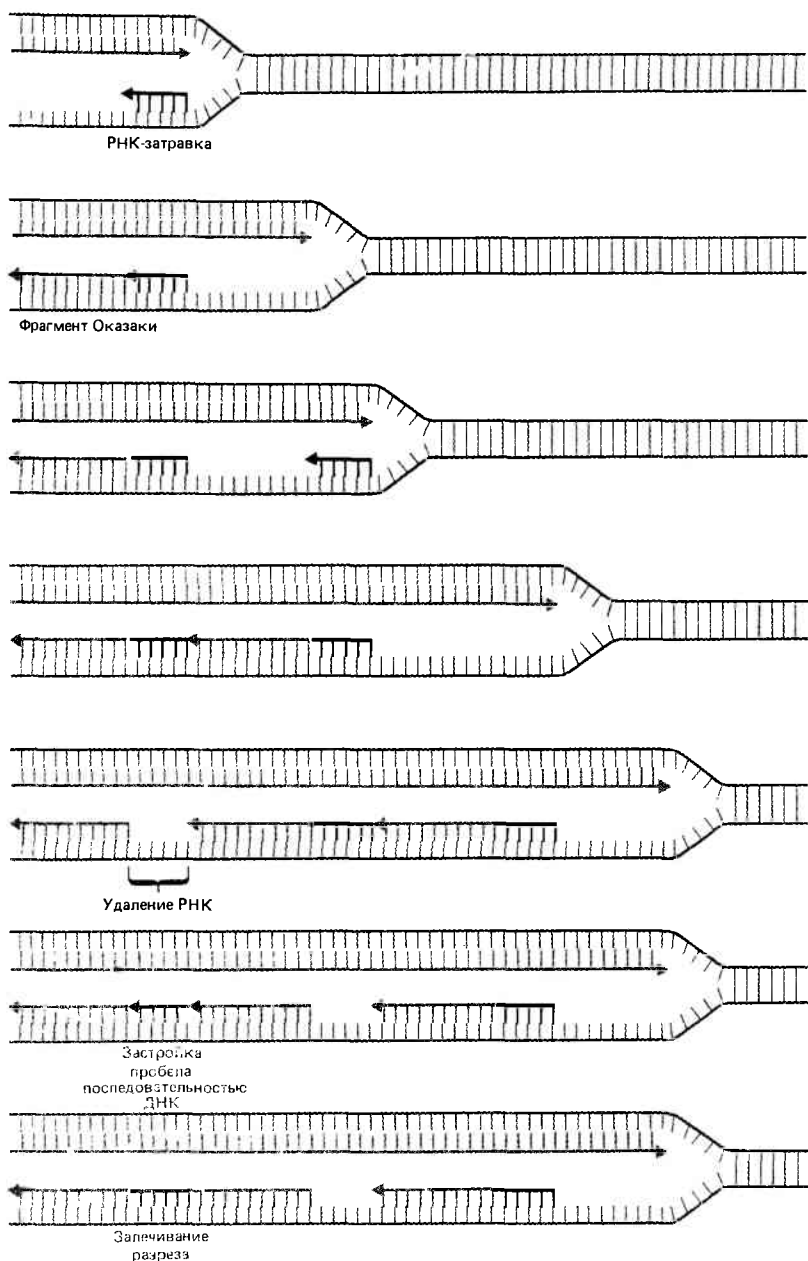


Рис. 32.13. Соединение фрагментов Оказаки между собой может быть осуществлено после удаления РНК, застройки пробелов и объединения концов.

ния NMN от NAD, а у фага Т4 – пирофосфата от АТФ. Образовавшийся комплекс присоединяется к 5'-фосфатному концу фрагмента через АМР, затем образуется фосфодиэфирная связь, соединяющая его с 3'-ОН-концом другого фрагмента, при этом фермент и АМР освобождаются.

Описанные выше факты поставили проблему, касающуюся специфичности синтеза фрагментов Оказаки. Какой фермент отвечает за синтез РНК-затравки? Осуществляется ли начало и остановка ее синтеза в специфических последовательностях, или эти события определяются организацией репликационной вилки? Продолжается ли синтез фрагмента Оказаки до встречи со смежным фрагментом, или существуют специфические точки терминации?

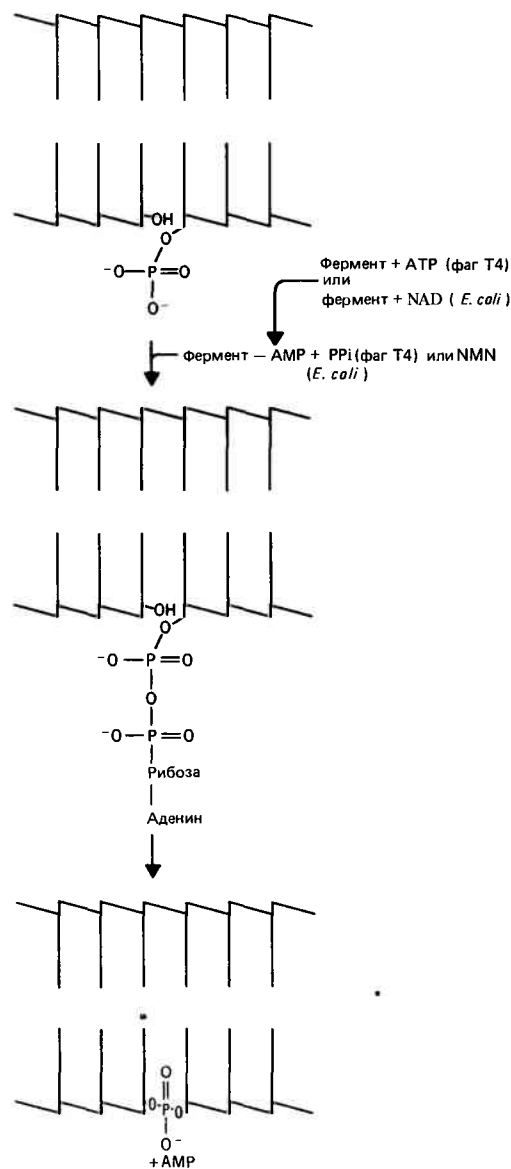


Рис. 32.14. ДНК-лигаза залечивает разрывы между смежными нуклеотидами, используя промежуточный комплекс фермент-АМР.

Рекомендуемая литература

В настоящее время топоизомеразам посвящено большое количество публикаций. Различия между типами топоизомераз рассматривают Лю с соавт. (Liu et al., Cell, 19, 697-707, 1980) и Козарелли (Cozzarelli, Cell, 22, 327-328, 1980). Функции гиразы обсуждает Козарелли (Cozzarelli, Science, 207, 953-960, 1981). Общий анализ всех топоизомераз осуществлен Геллертом (Gellert, Ann. Rev. Biochem., 50, 879-910, 1981). Эукариотические ДНК-полимеразы рассматривает Хуберман (Huberman, Cell, 23, 647-648, 1981). Сведения о репликации и ферментах, участвующих в этом процессе, можно найти в гл. 33.

Глава 33

ФЕРМЕНТАТИВНЫЙ АППАРАТ РЕПЛИКАЦИИ ДНК

Репликация двухцепочечной ДНК – это сложный процесс, требующий нескольких ферментативных активностей. Прежде всего необходимо разделение родительских цепей и их временная стабилизация в одноцепочечном состоянии. Далее должна быть иницирована репликация ведущей цепи, и затем эта цепь должна удлиниться. Кроме того, постоянно иницироваться, удлиниться и соединяться между собой должны фрагменты Оказаки отстающей цепи. В конце репликации необходимы реакции соединения и (или) терминации. Указанные события могут контролироваться отдельными белками или группами белков, но, поскольку эти белки функционируют взаимосвязано, вполне вероятно, что разные активности принадлежат дискретной структуре мультибелка, иногда называемого **реплисомой**. Реплисома не выявляется в виде свободного клеточного компонента, подобного, например, рибосоме; очевидно, она собирается из отдельных компонентов в начале репликации. Таким образом, реплисома может существовать только как белковый комплекс, связанный с определенной структурой, которую ДНК приобретает в репликационной вилке.

В настоящее время исследования процесса сборки белков, включаемых в репликацию эукариот, только начаты. При определении компонентов репликационных аппаратов бактерий и фагов использованы два подхода, дающих совпадающие результаты. Во-первых, с помощью мутаций, нарушающих репликацию, удалось идентифицировать гены, продукты которых либо являются компонентами аппарата репликации, либо выполняют вспомогательные функции. Во-вторых, для изучения репликации *in vitro* в качестве матриц могут быть использованы геномы определенных мелких фагов. В изучении репликационного аппарата достигнут заметный прогресс, поскольку каждый его компонент доступен для изучения в системе *in vitro* как биохимически чистый продукт, и участие в репликации *in vivo* различных компонентов идентифицируется по эффекту мутаций в их генах.

Сложность репликационного аппарата бактерий

Неспособность реплицировать ДНК – летальное событие для любой клетки. Поэтому можно выделить лишь условно-летальные мутанты по репликации, способные осуществлять этот процесс при перmissive условиях (при нормальной температуре инкубации) и проявляющие свой дефект только при nonpermissive условиях (при температуре 42°C). Такие термочувствительные мутанты позволили идентифицировать серию локусов, описанных как гены *dna*. Мутанты *dna* на основе их поведения при повышенной температуре разделены на два класса.

Мутанты с замедленной остановкой репликации завершают текущий раунд репликации, но оказываются неспособными начать новый раунд. Следовательно, у них нарушен цикл инициации репликации в точке ее начала (локусе *ori*). О мутантах этого класса известно немного

(см. ниже табл. 33.3). Основной вывод сводится к тому, что для инициации цикла репликации необходимо несколько различных белков.

Мутанты с быстрой остановкой репликации прекращают этот процесс при повышении температуры *немедленно*. Те из них, которые имеют отношение непосредственно к репликации, являются дефектными по компонентам реплисомы (см. ниже табл. 33.1). Следует отметить, что среди мутантов, отнесенных к классу быстро останавливающих репликацию, могут быть также и дефектные по инициации репликации, хотя это скрыто их фенотипом.

У некоторых штаммов с быстрой остановкой репликации нарушено образование предшественников, необходимых для синтеза ДНК. Например, мутации *nrdA*, первоначально названные *dnaF*, повреждают ген, кодирующий субъединицу рибонуклеотидредуктазы, которая существенна для обеспечения синтеза ДНК предшественниками. Точно так же мутации *dut⁻*, исходно названные мутациями *dnaS*, элиминируют активность dUTPазы. Наличие таких мутантов свидетельствует о важности обеспечения предшественниками процесса репликации ДНК. Следовательно, выделение мутаций с фенотипом *Dna⁻* не указывает, что продукт гена, повреждаемого мутацией, непосредственно вовлекается в процесс репликации.

Несмотря на интенсивные усилия, затраченные на выделение мутаций *dna*, до сих пор еще не идентифицированы все функции, необходимые для репликации. Известно, что для функционирования систем репликации *in vitro* (см. ниже табл. 33.2) необходимо несколько дополнительных белков, кодируемых неизвестными генами. При испытании экстрактов из клеток *E. coli* на способность синтезировать ДНК доминирующей ферментативной активностью во всех случаях была активность ДНК-полимеразы I. В ее присутствии оказывается невозможным определить активности ферментов, действительно ответственных за репликацию. Вот почему для изучения этого процесса *in vitro* используют экстракты клеток мутантов *polA*.

Определенный интерес представляют системы, в которых хотя и не удается проанализировать отдельные компоненты, однако можно более четко, чем в обычных условиях *in vivo*, охарактеризовать процесс репликации. Например, бактерии обрабатывают растворителями жиров, что делает их мембрану проницаемой для предшественников синтеза ДНК (и других малых молекул). Обработанные клетки не сохраняют жизнеспособность, но реплицируют свою ДНК со скоростью, которая лишь незначительно отличается от обычной скорости репликации *in vivo*. С той же целью бактерии можно лизировать на целлофановых дисках, находящихся на поверхности агара. Репликационная система при этом задерживается на поверхности диска, и к ней могут быть добавлены необходимые компоненты.

Для того чтобы охарактеризовать отдельные компоненты репликационных систем *in vitro*, необходима мат-

рица, которая позволила бы исследовать природу синтеза ДНК. Мы должны быть уверены, что синтез соответствует подлинному репликационному процессу *in vivo*, а не вызван альтернативной активностью, связанной, например, с репарацией поврежденной ДНК. Для получения систем репликации *in vitro* можно использовать различные матрицы, обеспечивающие синтез либо отстающих, либо ведущих цепей.

Существуют два способа изучения роли отдельных компонентов репликации *in vitro*. Первый способ заключается во фракционировании системы, очистке компонентов и повторном их объединении. В другом случае используют мутант *dna* и эксперимент осуществляют при условиях, когда продукт мутантного гена неактивен. Затем проверяют экстракты из клеток дикого типа на их способность восстанавливать активность этой системы. Такой **комплементарный анализ** *in vitro* может быть использован при очистке белка, кодируемого локусом *dna*.

Инициация синтеза одиночной цепи ДНК

Образование фрагмента Оказаки на отстающей цепи связано с синтезом ДНК комплементарной одноцепочечной области, образуемой при движении репликационной вилки. Для изучения событий инициации и продолжения синтеза такой цепи ДНК в качестве матриц были использованы три одноцепочечные ДНК фагов. Каждый из этих фаговых геномов представлен кольцевой одноцепочечной ДНК, получившей название **вирулентной** или (+)-цепи. На **рис. 33.1** показано, что синтез **комплементарной**, или (-)-цепи, превращает геном в двухцепочечную кольцевую **репликативную форму** (РФ) ДНК. Она может находиться в виде или релаксированного кольца (РФII) или суперспирализованного кольца (РФИ). Синтез комплементарной цепи рассматривается как аналог синтеза отстающей цепи двухцепочечной ДНК.

Несмотря на то что такие фаговые геномы проявляют лишь незначительное спаривание оснований в пределах одной цепи и поэтому сохраняются в основном в одноцепочечном состоянии, сама ДНК не является подходящей матрицей для реакции репликации. Нуклеиновая кислота сначала должна быть покрыта **белком, связывающимся с одноцепочечной ДНК (белок SSB)**. Этот белок представляет собой тетрамер с мол. массой, равной 74 000, который кооперативно связывается с одноцепочечной ДНК, сохраняя ее в растянутом состоянии. Суть кооперативного способа связывания состоит в том, что связывание одной молекулы белка облегчает связывание другой. В результате однажды начатая реакция связывания на определенной молекуле ДНК быстро распространяется до тех пор, пока вся одноцепочечная ДНК не будет покрыта белком SSB. Следует отметить, что этот белок не является расплетающим; его функция состоит в стабилизации одноцепочечной ДНК.

На **рис. 33.2** представлены стадии репликации ДНК фага M13, которая может служить моделью наиболее простой системы. В одноцепочечной структуре ДНК имеется участок, представленный последовательностью из 59 оснований, образующих шпильку. Это единственная часть фагового генома, не связывающаяся с белком SSB. Шпильчатая структура обеспечивает сигнал для инициации.

РНК-полимераза хозяина иницирует синтез РНК за шесть оснований перед шпильчатой структурой. Цепь



Рис. 33.1. Одноцепочечная кольцевая ДНК фага превращается в двухцепочечную в результате синтеза комплементарной цепи.

РНК распространяется на длину 20–30 остатков, разрушая эту структуру. Белок SSB связывается с нематричной половиной шпильки, освобождаемой при образовании гибрида РНК–ДНК. Именно это, вероятно, обуславливает терминацию транскрипции в верхней части шпильчатой структуры. Поскольку реакция зависит от РНК-полимераза, она подавляется рифампицином.

В настоящее время еще не известно, как РНК-полимераза распознает шпильчатую структуру в качестве сайта инициации транскрипции; возможно, в этот процесс вовлечены другие белки с вспомогательными функциями. Делетирование всей последовательности, кодирующей синтез РНК-затравки, не делает фаг нежизнеспособным, однако вызывает его более медленное развитие. Следовательно, должны существовать другие последовательности, способные использоваться в качестве точек инициации, хотя в клетках *E. coli* они менее эффективны (а в клетках других бактериальных хозяев они могут быть более эффективными).

РНК-затравка удлиняется с помощью ДНК-полимеразы III, которая синтезирует цепь ДНК, распространяющуюся по всей длине кольцевой молекулы до тех пор, пока она не достигает точки инициации РНК. РНК удаляется посредством экзонуклеолитического действия; какой фермент ответствен за эту активность, пока не ясно. Это может быть ДНК-полимераза I, которая также способна удлинить новую цепь ДНК, устраняя пробел. В результате остается только разрыв между началом и концом комплементарной цепи, который ликвидируется с помощью лигазы.

При использовании в качестве матрицы одноцепочечной кольцевой молекулы ДНК фага G4 наблюдается одно значительное отличие в реакции. Вместо РНК-полимеразы хозяина функционирует белок, кодируемый геном *dnaG*. Этот фермент представлен одним полипептидом с мол. массой 60 000 дальтон (значительно меньшей, чем у РНК-полимеразы). Фермент назван **праймазой**. Одна молекула белка DnaG связывается непосредственно

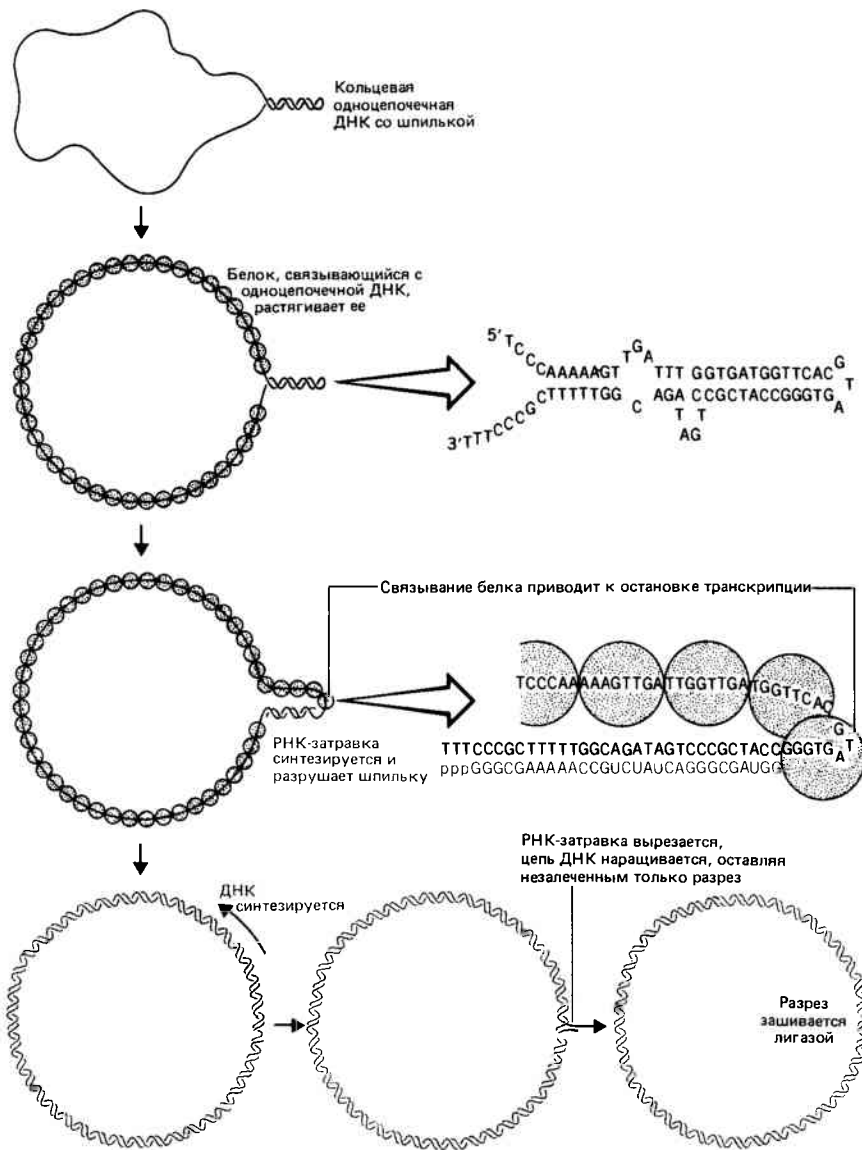


Рис. 33.2. Для образования комплементарной цепи ДНК фага М13 на матричной вирусной ДНК, покрытой белком SSB необходимы синтез РНК-затравки, наращивание ДНК, удаление затравки, заполнение пробела и залечивание разреза.

с SSB-резистентной двухцепочечной шпилькой в ДНК. Праймаза синтезирует РНК из 29 оснований, которая комплементарна одной цепи шпильчатой структуры. В том случае, когда синтез ДНК происходит одновременно с синтезом затравки (вместо того, чтобы следовать за ним), образуется более короткая затравка, иногда содержащая всего шесть оснований. Затравочная реакция не подавляется рифампицином. Праймаза отличается от РНК-полимеразы и своей способностью использовать в качестве предшественников при синтезе *in vitro* как рибонуклеотиды, так и дезоксирибонуклеотиды.

Примеры с М13 и G4 показывают, что РНК-затравка может быть синтезирована либо с помощью РНК-полимеразы, либо праймазы, причем в любом случае используется уникальный сайт, определяемый, по-видимому, вторичной структурой ДНК. Завершение реакции, вероятно, обусловлено топологическими особенностями области инициации и (или) действием ДНК-полимеразы. Удлинение цепи ДНК с 3'-ОН-конца затравки осуществляется у обоих фагов при участии одинаковых синтезирующих активностей.

Движение праймосомы

Ситуация в случае фага фХ174 может служить лучшим примером инициации фрагментов Оказаки, так как синтез ДНК иницируется в нескольких сайтах фагового генома. Сравнение системы репликации фага фХ174 с системами фагов М13 и G4 дано на рис. 33.3. В случае фага фХ174 необходимо значительно большее количество белков. Некоторые из них идентифицированы как продукты определенных генов, в частности *dna* (данные об их свойствах суммированы в табл. 33.1); о других известно только, что они необходимы для реакции *in vitro* (табл. 33.2).

Первая стадия синтеза у фага фХ174 такая же, как у фагов М13 или G4: ДНК покрывается белком SSB. Однако затем она превращается с помощью шести белков, *p*, *p'*, *p''*, *i*, *DnaV* и *DnaC*, в предшествующий затравке промежуточный комплекс. Такой «предзатравочный» комплекс собирается в специфическом сайте, имеющем вид шпильки, образованной последовательностью из 55 оснований (в участке, аналогичном уникальной стартовой точке в ДНК фага G4). После того как предзатравочный

Таблица 33.1

Некоторые компоненты репликационного аппарата, идентифицированные с помощью мутаций

Ген	Функция	Размер продукта (в дальтонах)	Число копий на 1 клетку
<i>dnaB</i>	Образование предпраймерной РНК	300 000, гексамер	~ 20
<i>dnaC</i>	Действует в комплексе с <i>dnaB</i>	Неизвестен	Неизвестно
<i>dnaG</i>	Праймаза иницирует синтез фрагментов Оказаки	60 000, мономер	~ 75
<i>polC</i> (<i>dnaE</i>)	ДНК-полимераза III	140 000, субъединица	~ 20
<i>dnaZ</i>	ДНК-полимераза III	52 000, субъединица	~ 20
<i>lig</i>	Залечивает разрывы между фрагментами Оказаки	75 000, мономер	~ 300
<i>rep</i>	АТР-зависимая геликаза	66 000, мономер	~ 50
<i>ssb</i>	Белок, связывающийся с одноцепочечной ДНК	74 000, тетрамер	~ 300

комплекс сформирован, праймаза *DnaG* способна узнать матрицу; она синтезирует различные короткие молекулы длиной от 15 до 50 оснований, причем инициация происходит во многих сайтах на ДНК фХ174. Следовательно, предзатравоочный комплекс образуется в уникальном участке, а синтез РНК иницируется праймазой в нескольких сайтах.

Белок *p'* может быть компонентом, ответственным за выбор сайта, в котором образуется предзатравоочный комплекс. Белок проявляет АТРазную активность, которая специфически зависит от 55 оснований шпильчатой структуры фага фХ174. (Образование такой структуры само по себе не обеспечивает сайт узнавания для белка *p'*, поскольку шпильчатые структуры ДНК M13 или G4 не могут заменить шпильчатую структуру фага фХ174). О количестве гидролизованной АТР и использовании выделенной энергии данных не получено.

Если ДНК фага фХ174 не покрыта белком SSB, необходимость в предзатравоочных белках исчезает. В случае чистой одноцепочечной ДНК фага фХ174 для синтеза РНК-затравки праймаза нуждается только в присутствии белка *DnaB*.

Комплекс *DnaB*—праймаза также активен при использовании непокрытых белком SSB ДНК фагов M13, G4 или матрицы *poly* (dT).

При добавлении к ДНК фага фХ174 белка *DnaB* и праймазы синтезируется множество затравок. Размер РНК колеблется в пределах от 10 до 60 оснований, и, вероятно, их появление свидетельствует о выборе ряда специфических сайтов. Следовательно, комплекс *DnaB*—праймаза способен узнавать серию сайтов инициации на свободной ДНК; такая способность утрачивается в присутствии белка SSB. Восстановление затравоочной активности требует участия других белков предзатравоочного комплекса: *p*, *p'*, *p''*, *i* и *DnaC*. Мы знаем, что последнее отражает действительные события *in vivo*, поскольку белок SSB необходим для репликации у бактерий.

Об этом свидетельствует фенотип штаммов с быстрой остановкой репликации, определяемый мутацией в гене *ssb*.

Комплекс белков, вовлекаемых в затравоочную реакцию, получил название **праймсомы**. Существующее противоречие между наличием уникального предзатравоочного сайта и множества сайтов инициации синтеза затравок можно объяснить, предположив, что праймосома первоначально собирается только в одном участке, однако затем некоторые или все белки способны перемещаться по одноцепочечной ДНК к сайтам, в которых действительно осуществляется инициация синтезов затравок. Движущийся комплекс должен включать белки *DnaB* и праймазу; белок *p'*, по-видимому, входит в состав этого комплекса, так как его способность вытеснять белок SSB с ДНК, вероятно, необходима для движения.

Белок *DnaB* также может играть важную роль в движении. На одну репликационную вилку приходится, по-видимому, один гексамер белка *DnaB*. В присутствии АТР белок с высокой эффективностью связывается с одноцепочечной ДНК. В дальнейшем праймаза стабилизирует комплекс *DnaB*—АТР на одноцепочечной ДНК. Замещение АТР на ADP вызывает аллостерическое изменение, которое уменьшает сродство белка *DnaB* к ДНК. Таким образом, для того чтобы дать возможность белку *DnaB* отделиться от ДНК и переместиться из одного сайта инициации затравки в другой (где ADP замещается на АТР), может быть необходим гидролиз АТР. В настоящее время в отдельных плаزمидах обнаружены сайты, которые формируют праймосомы. Сравнение последовательностей этих сайтов у разных штаммов, их анализ у мутантов дает возможность определить чем они отличаются от других последовательностей ДНК.

Использование линейных фрагментов ДНК фага фХ174 позволило идентифицировать связь между уникальным сайтом предзатравоочного комплекса и множественными сайтами инициации синтеза затравок. Праймосома, по-видимому, перемещается первоначально в направлении, противоположном направлению удлинения цепи—вдоль родительской цепи от 5'- к 3'-концу. На рис. 33.4 показано, что такая ситуация соответствует условиям инициации синтеза фрагментов Оказаки отстающей цепи дуплексной ДНК. При продвижении репликационной вилки со стороны 3' образуется одноцепочечная

Таблица 33.2

Другие белки, участвующие в репликации

Белок	Функция	Размер продукта (в дальтонах)	Число копий на 1 клетку
<i>i</i>	Образование предзатравоочной РНК (1–2 молекулы на репликационную вилку)	66 000, тример	~ 50
<i>p'</i> (= Y)	Узнавание сайта, АТРазы (1 молекула на репликационную вилку)	55 000, мономер	~ 80
$\left. \begin{matrix} p'' \\ p \end{matrix} \right\}$	Образование предзатравоочной РНК	Неизвестен	Неизвестно

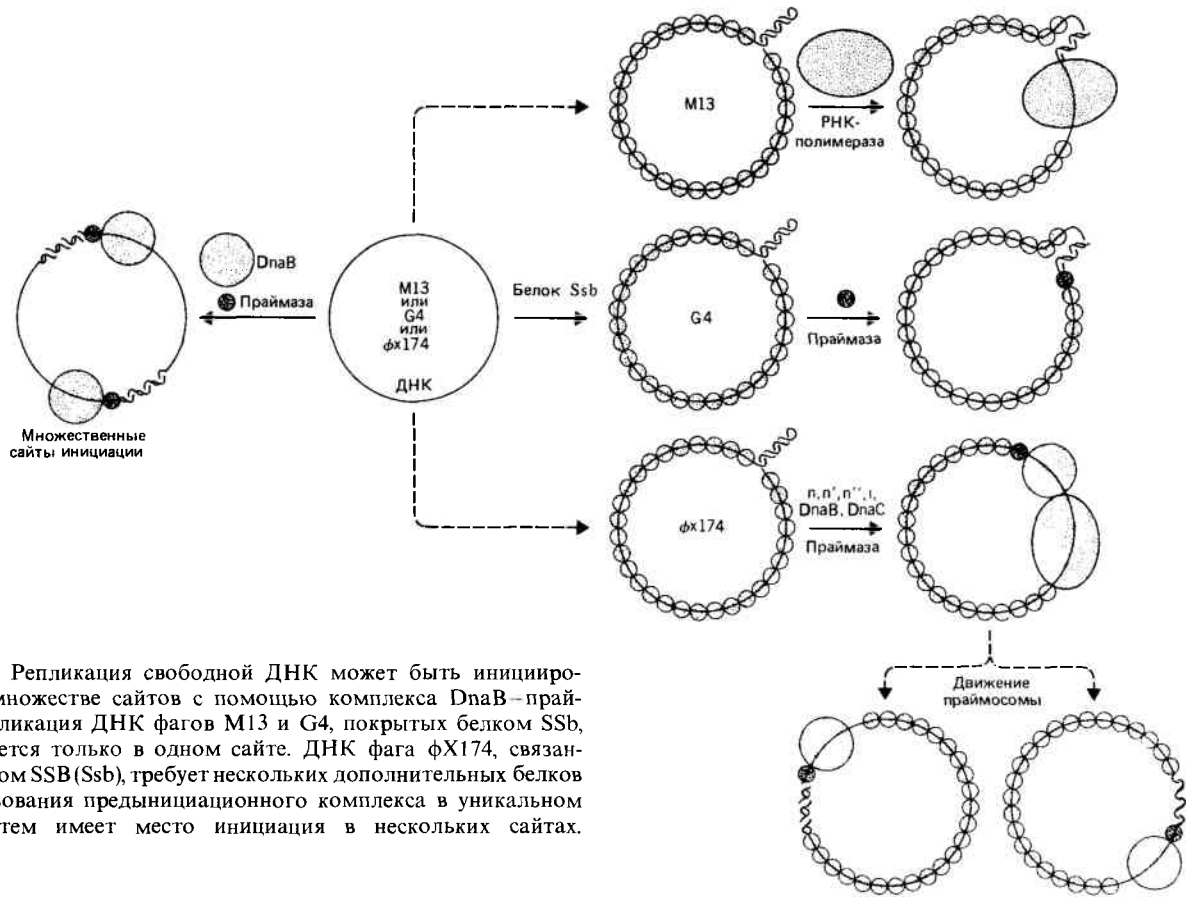


Рис. 33.3. Репликация свободной ДНК может быть инициирована во множестве сайтов с помощью комплекса DnaB–праймаза. Репликация ДНК фагов M13 и G4, покрытых белком SSB, инициируется только в одном сайте. ДНК фага фX174, связанная с белком SSB (Ssb), требует нескольких дополнительных белков для образования предынициационного комплекса в уникальном сайте; затем имеет место инициация в нескольких сайтах.

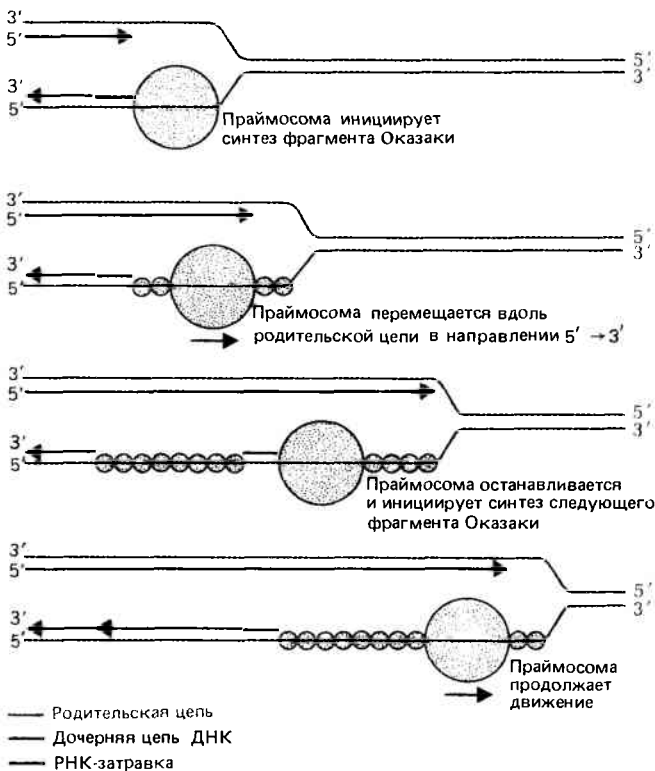


Рис. 33.4. Для образования отстающей цепи праймосома следует вдоль матрицы за репликационной вилкой в направлении 5' → 3', периодически иницируя синтез фрагментов Оказаки.

область перед праймосомой. После каждого события инициации праймосома передвигается вдоль одноцепочечного сегмента к сайту инициации синтеза следующего фрагмента Оказаки. Следовательно, праймосома движется в том же направлении, что и репликационная вилка, но в противоположном направлении синтеза ДНК отстающей цепи.

Чтобы установить соответствие между реакцией синтеза, охарактеризованной для ДНК фага фX174, и репликацией двухцепочечной ДНК, мы должны рассмотреть, каким образом одноцепочечная матрица готовится к синтезу отстающей цепи. По-видимому, по мере продвижения репликационной вилки, белок SSB связывается с ДНК, удерживая две родительские цепи разделенными. В результате обе цепи (или по крайней мере отстающая) оказываются в условиях, необходимых для их использования в качестве матриц. Белок SSB в репликационной вилке должен присутствовать в стехиометрических количествах.

Какой белок ответствен за первоначальное расплетание двух цепей родительской молекулы? Эту функцию связывают с мутациями *гер*, которые замедляют движение репликационной вилки. Белок Rep представляет собой АТФазу, зависимую от одноцепочечной ДНК. В присутствии АТФ совместное действие белков Rep и SSB приводит к разделению двухцепочечной репликативной формы ДНК фага фX174, несущей разрыв в одной из цепей, на составляющие ее отдельные цепи. Вероятно, белок Rep расплетает цепи, а белок SSB затем фиксирует их в одноцепочечном состоянии.

Возможный механизм действия белка Rep представлен на рис. 33.5. Белок связывается с одноцепочечной

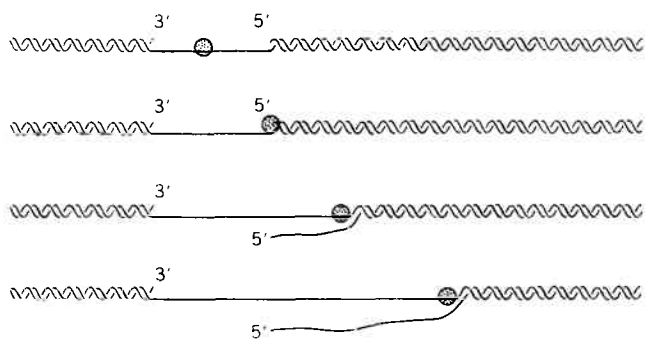


Рис. 33.5. Белок Rep может двигаться вдоль одной из цепей ДНК, вытесняя вторую цепь.

областью в двухцепочечной ДНК, перемещаясь по ней в определенном направлении. При достижении двухцепочечного сегмента он вытесняет цепь, имеющую 5'-конец. Расплетание каждой пары оснований требует гидролиза двух молекул АТФ.

Если роль расплетающего белка в репликационной вилке приписывается белку Rep правильно, то возникает вопрос, каким образом мутанты *rep* реплицируют бактериальную ДНК, не обладая способностью разделять цепи ДНК фага ϕ X174. Жизнеспособность фага M13 также зависит от Rep-белка; в мутантных клетках *rep*⁻ потомство (одноцепочечной) фаговой ДНК не может быть получено из кольцевой молекулы, имеющей репликативную форму. Однако если в ДНК фага M13 внедрен фрагмент *E. coli*, содержащий область инициации, то ДНК реплицируется. Следовательно, ДНК двух указанных фагов нуждается в белке Rep для репликации, а ДНК бактериальной клетки способна реплицироваться без него. У *E. coli*, по-видимому, существует какой-то другой (неизвестный) механизм, с помощью которого происходит инициация разделения цепей и далее цепи поддерживаются в разделенном состоянии.

Инициация репликации в точках начала репликации двухцепочечной ДНК

Начало цикла репликации двухцепочечной ДНК требует двух типов активностей наряду с теми, которые необходимы для синтеза фрагментов Оказаки. Во-первых, цепи ДНК должны разделиться. (Существовали гипотезы, согласно которым приобретение необычной вторичной структуры, изображенной на рис. 31.8, может быть связано именно с этой стадией.) Во-вторых, должен быть иницирован синтез ведущей цепи с помощью механизма, аналогичного или подобного тому, который осуществляет инициацию синтеза фрагментов Оказаки. Однако не исключена возможность существования другого механизма.

Изучение мутантов *E. coli* с замедленной остановкой репликации позволило идентифицировать несколько локусов, контролирующих события в сайте инициации репликации; однако об их молекулярных активностях известно очень мало. Из перечисленных в табл. 33.3 локусов только три имеют известные функции. Установлено, что и замедленная и быстрая остановки репликации могут обуславливаться мутациями в локусах *dnaB* и *dnaC*; мы

Таблица 33.3
Мутации, обуславливающие замедленную остановку репликации, недостаточно хорошо охарактеризованы

Ген	Функция
<i>dnaA</i>	Связана с инициацией репликации в точке начала
<i>dnaB</i>	Одна мутация обуславливает замедленную остановку репликации, функции других описаны в табл. 33.1
<i>dnaC</i>	Белок действует совместно с белком DnaB
<i>dnaI</i>	Неизвестна
<i>dnaJ</i>	»
<i>dnaK</i>	»
<i>dnaP</i>	Устойчивость к фенетиловому спирту; связана с мембраной?

имеем некоторую информацию о функциях белков, контролируемых ими в процессе репликации. Вполне допустимо, что роль этих белков в сайте инициации окажется родственной. Некоторые мутации локуса *dnaA* могут быть супрессированы мутациями в генах РНК-полимеразы, что свидетельствует об участии белка DnaA в какой-то стадии инициации, требующей синтеза РНК. Точно установить роль белка DnaA пока не удалось, поскольку он с трудом поддается очистке.

Другие локусы идентифицированы только с помощью мутаций. Исключение составляет *dnaK*, относящийся к группе локусов, выражение которых стимулируется температурным воздействием. Было замечено удивительное сходство между последовательностями белка DnaK и основного белка «теплового шока» *D. melanogaster*, который становится доминирующим в условиях воздействия высокой температуры на клетки насекомых. Такое сходство между белками прокариот и эукариот остается пока загадочным.

Об участии РНК-полимеразы в инициации репликации в сайте *ori* свидетельствовали данные, показывающие, что начало репликации бактериальной ДНК можно подавить рифампицином в период, когда ингибиторы белкового синтеза оказываются неэффективными. Это значит, что, по-видимому, важен именно синтез РНК (а не ее трансляция и образование белка). Подходящей системой для изучения этой функции может служить фаг лямбда, сайт инициации которого позволяет вести двуправленную репликацию и, кроме того, способен приобретать структуру «клеверного листа», подобную той, которая постулирована для *E. coli*. Инициация репликации в сайте начала репликации ДНК фага лямбда требует «активации» посредством транскрипции.

Репликация фага лямбда не может быть иницирована, если репрессор связан с O_R , так как связывание предотвращает транскрипцию, которая иницируется в P_R и проходит через сайт начала репликации. Карта этой области представлена на рис. 33.6.

Необходимость в синтезе РНК не обязательно связана с синтезом РНК-затравки для репликации ведущей цепи. Возможно, что синтез РНК играет роль в индукции некоторых структурных изменений в этой области.

В репликации фага лямбда участвуют белки, кодируемые фаговыми генами *O* и *P*, и, кроме того, несколько бактериальных белков, в частности тех, повреждение которых вызывает замедленную остановку репликации. Белок *O* связывается с фаговым локусом *ori*; белок *P* взаимодействует с белком *O* и бактериальными белками.

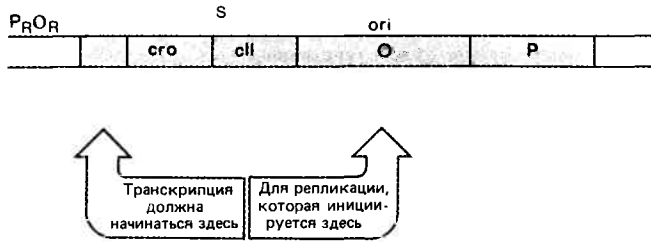


Рис. 33.6. Для активации точки начала репликации требуется инициация транскрипции в промоторе P_R .

Локус *ori* расположен в пределах гена *O*, так что белок действует вблизи места своего синтеза.

Существуют варианты фага, *λdv*, имеющие более короткие геномы, в которых заключена информация, необходимая для репликации. Тем не менее такие фаговые частицы утрачивают инфекционные функции и могут сохраняться в бактерии в виде плазмид. ДНК *λdv* может реплицироваться *in vitro* с помощью системы, содержащей кодируемые фагом белки *O* и *P*, а также репликационные белки бактерии-хозяина. Эта система дала возможность проанализировать роль транскрипционной активации.

Наличие плазмид, несущих последовательность *oriC* *E. coli*, позволило получить бесклеточную систему, в которой репликация осуществляется, иницируясь в этом локусе. Инициация репликации плазмиды подавляется рифампицином, а также соединениями, предотвращаю-

щими функционирование ДНК-гиразы. Подавляют реакцию и антитела против белков *DnaB* или *SSB*; в то же время экстракт, приготовленный из мутантных клеток *DnaA*, неактивен.

Существуют две гипотезы относительно инициации синтеза ведущей цепи. Одна из них предполагает использование РНК-затравки, которая должна действовать так же, как в случае инициации синтеза фрагментов Оказаки на отстающей цепи, за исключением того, что синтез затравки должен сопровождаться непрерывным удлинением цепи. Альтернативная возможность связана с образованием разрыва в родительской цепи и использованием ее 3'-ОН-конца в качестве матрицы, как это показано на рис. 33.7. В конечном счете ковалентная связь между дочерней и родительской цепями должна быть разрушена.

Установлено, что механизм репликации, заключающийся в разрыве и последующем удлинении родительской цепи, свойствен многим фаговым ДНК. Репликация происходит по типу катящегося кольца. Этот способ репликации (см. рис. 31.14) изучали, используя РФ-форму ДНК фага фХ174. Синтез по типу катящегося кольца требует участия меньшего количества ферментативных функций, чем в случае образования репликативной формы ДНК из исходной вирусной цепи. Суперспирализованные кольца ДНК фага фХ174 могут быть реплицированы *in vitro* с помощью системы, содержащей продукт фагового гена *A* и бактериальные белки *Rep*, *SSB* и ДНК-полимераза III.

Белок *A* разрезает вирусную (+)-цепь дуплексной молекулы в специфическом сайте, определяющем начало репликации РФ-ДНК. ДНК может быть разрезана только в том случае, если она суперспирализована. Белок *A* способен связываться с одноцепочечным фрагментом ДНК длиной десять нуклеотидов, который окаймляет сайт разрыва. Отсюда следует, что суперспирализация необходима для образования одноцепочечной области, обеспечивающей белку *A* его сайт связывания. В результате разрыва образуются 3'-ОН и 5'-фосфатный конец, и каждый из них играет определенную роль в репликации ДНК фага фХ174.

В присутствии белков *Rep* и *SSB*, а также АТФ разрезанная ДНК расплетается (рис. 33.8). Белок *Rep* обеспечивает функцию геликазы, которая разделяет цепи; белок *SSB* фиксирует их в одноцепочечной форме. Белок *A* остается ковалентно связанным с 5'-фосфатным концом; кроме того, он связан с белком *Rep* и передвигается по кольцевой молекуле вместе с ним. В результате замещаемая цепь «выплетивается» из сайта расплетания. Завершение движения по кольцу ведет к освобождению кольцевой комплементарной (-)-цепи и линейной вирусной (+)-цепи.

Если в систему включена ДНК-полимераза III, 3'-ОН-конец разрезанной цепи наращивается. Цепь растет вдоль

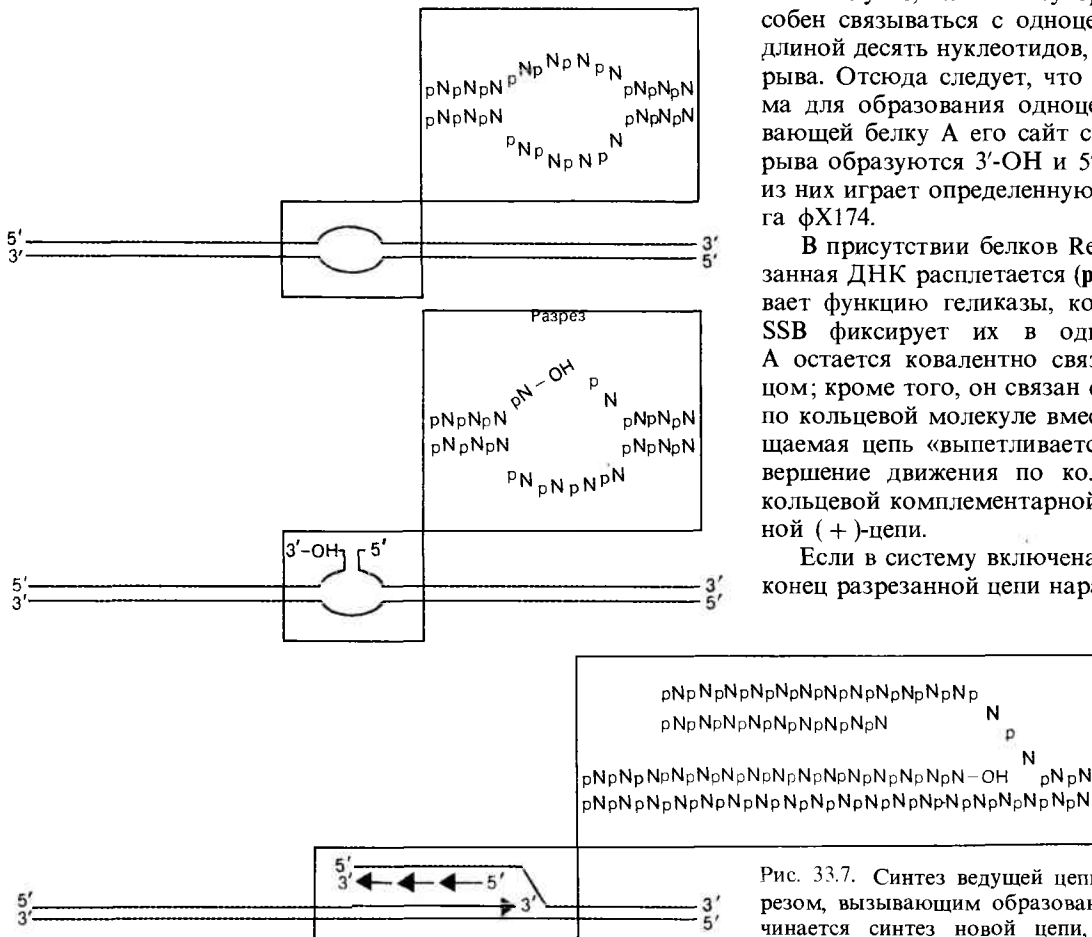


Рис. 33.7. Синтез ведущей цепи может быть иницирован разрезом, вызывающим образование 3'-ОН-конца, с которого начинается синтез новой цепи, вытесняющей исходную цепь.

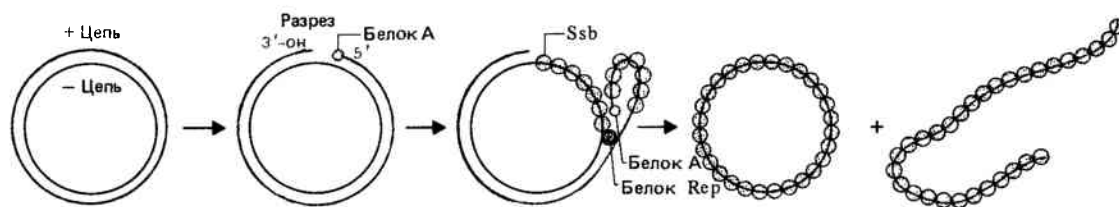


Рис. 33.8. РФ ДНК фага φX174 может быть разделена на отдельные цепи в результате разрезания, осуществляемого белком

А, раскручивания с помощью белка Rep и стабилизации отдельных цепей белком SSB(Ssb).

кольцевой (–)-цепи матрицы до тех пор, пока она не достигнет стартовой точки и не вытеснит исходную цепь, как это показано на рис. 33.9. Поскольку белок А движется по кольцевой молекуле вместе с Rep-белком, в момент достижения растущей точкой сайта инициации репликации он оказывается вблизи этого участка, в результате чего тот же белок А снова узнает точку начала репликации и разрезает цепь, прикрепившись после этого к концу образовавшегося разреза. Такой цикл может повторяться неограниченное число раз.

Вслед за событием разрезания вытесненная (+)-цепь освобождается в виде линейной молекулы, которая может повторно приобрести форму кольца. Белок А вовлекается в процесс образования кольцевой структуры. На самом деле соединение 3'- и 5'-концов (+)-цепи может осуществляться белком А в процессе реакции, с помощью которой он отделяется от цепи и прикрепляется к концу нового разреза.

Белок А обладает необычным свойством, которое может быть связано с его активностями. *In vivo* он действует только в *cis*-положении, однако *in vitro* при добавлении к бесклеточной системе активен на любой матричной ДНК. Это означает, что *in vivo* белок А, синтезированный определенным геномом, способен присоединяться только к ДНК этого генома. Мы не знаем, как это происходит. Однако активность этого белка *in vitro* показывает, каким образом он может оставаться связанным с одной и той же родительской матричной (–)-цепью.

В случае морфогенеза фага в цикле имеет место модификация. Вместо связывания вытесненной (+)-цепи с белком SSB (и повторного вступления в цикл репликации) она упаковывается в фаговый вирион.

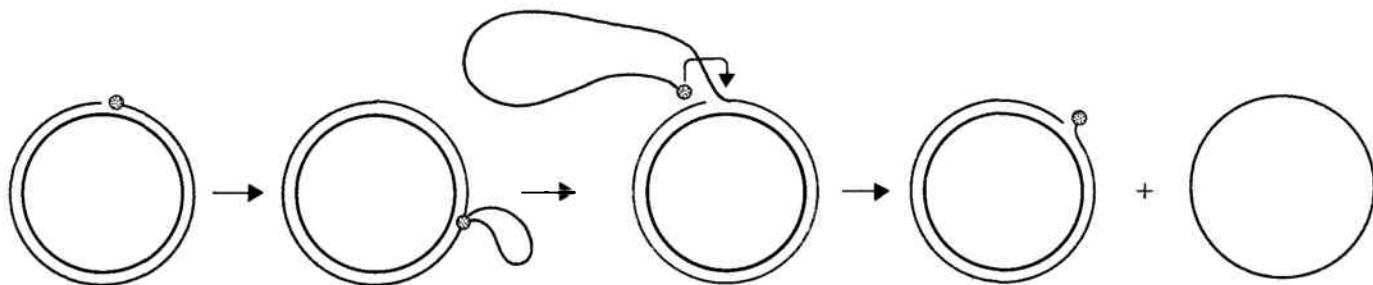


Рис. 33.9. РФ ДНК фага φX174 может использоваться в качестве матрицы для синтеза одноцепочечных кольцевых молекул ДНК вируса. Белок А остается прикрепленным к тому же геному в течение неограниченного числа циклов, каждый раз про-

Репликационный аппарат фага Т4

При проникновении фага Т4 в клетку *E. coli* он обеспечивает появление множества своих собственных функций, которые либо замещают, либо дополняют функции хозяина. Поскольку ДНК хозяина деградирует, фаг не может использовать функции, требующие экспрессии генов хозяина, и использует только те ферменты, которые обычно присутствуют в клетках бактерий в больших количествах. Деградация ДНК бактерии-хозяина необходима для освобождения нуклеотидов, которые затем используются при синтезе фаговой ДНК. (Фаговая ДНК отличается по составу оснований от клеточной ДНК: вместо цитозина она содержит гидроксиметилцитозин.)

Фаговые гены, участвующие в синтезе ДНК в инфицированной клетке, можно идентифицировать с помощью мутаций, которые предотвращают или задерживают образование зрелых фаговых частиц. Наиболее важные из них (существенные гены) идентифицированы при изучении условно-летальных мутантов, которые можно отнести к трем фенотипическим классам. Если в результате мутации синтез ДНК полностью подавлен, можно сделать вывод, что белок, кодируемый мутантным геном, либо представляет собой необходимый компонент репликационного аппарата, либо участвует в обеспечении синтеза ДНК предшественниками (особенно гидроксиметилцитозином). Мутации, вызывающие **задержку синтеза ДНК**, по-видимому, затрагивают гены, связанные с инициацией репликации ДНК. Фаги, у которых синтез иницируется, а затем останавливается, вероятно, несут мутации, затрагивающие регуляторные функции, ДНК-лигазу и некоторые ферменты, связанные с деградацией ДНК клетки-хозяина. Имеются также несущественные гены,

изводя разрез в точке начала репликации на вирусной (+)-цепи и перемещаясь к новому 5'-концу. В то же время вытесненная вирусная цепь образует кольцевую форму.

связанные с репликацией; например, те, которые участвуют в гликозилировании гидроксиметилцитозина в ДНК.

Синтез ДНК фага Т4 катализируется мультиферментным агрегатом, собираемым из продуктов семи существенных генов. Каждый из компонентов был проверен в комплементационном анализе в опытах *in vitro*. Система активна при использовании нескольких матриц ДНК *in vitro*; она действует со скоростью, близкой к природной, и частота ее ошибок колеблется в пределах от 10^{-5} до 10^{-6} (*in vivo* эта величина составляет примерно $2 \cdot 10^{-8}$). Следует отметить, что в системе *in vitro* СТР используется так же эффективно, как гидроксиметил-СТР. Белки репликационного комплекса фага Т4 охарактеризованы в табл. 33.4.

Белок гена 32 – высококооперативный, связывающийся с одноцепочечной ДНК и необходимый в стехиометрических количествах. Этот белок был первым охарактеризованным белком такого типа. По-видимому, он необходим для правильной геометрии репликационной вилки фага Т4, поскольку белок SSB из *E. coli* не замещает его. Белок гена 32 образует комплекс с ДНК-полимеразой фага Т4; это взаимодействие, по-видимому, важно для образования репликационной вилки.

Репликация ДНК фага Т4 зависит от синтеза РНК-затравки. Однако по своей природе фаговая затравка отличается от той, которая образуется в *E. coli*. Продукты генов 41 и 61, действуя совместно и используя одноцепочечную ДНК фага Т4 в качестве матрицы, синтезируют короткие затравки. Все они представляют собой пентарибонуклеотиды с общей последовательностью pppApCpNpNpNp. Следовательно, стартовая последовательность, длина которой ограничена пятью основаниями, содержит два специфических нуклеотида в начале, в то время как последние три позиции могут быть заняты любым из оснований.

В присутствии полного аппарата репликации эти затравки удлиняются в цепи ДНК. Затравки образуются в специфических сайтах, хотя эти сайты и встречаются часто – вероятно, чаще, чем один на каждые 2000 оснований,

разделяющих стартовые точки фрагментов Оказаки. Возможно, что репликационная вилка минует серию потенциальных затравочных сайтов; вероятно, это диктуется расстоянием, пройденным после последнего события инициации.

Белок гена 41 имеет ГТРазную активность, стимулируемую одноцепочечной ДНК. Связываясь с ней, белок способен перемещаться со скоростью примерно 400 нуклеотидов в секунду. Возможно, что его роль аналогична роли белка DnaB бактерии-хозяина и заключается в поиске сайтов, в которых этот белок останавливается и иницирует синтез затравки. Предполагается, что энергию для перемещения обеспечивает гидролиз ГТР.

Белок гена 61 необходим в значительно меньших количествах, чем большинство других белков, участвующих в репликации фага Т4. Это препятствовало его изучению. (Требуемое количество этого белка так мало, что первоначально он не был замечен как необходимый компонент; его было достаточно, когда он присутствовал в виде загрязняющей примеси в препарате белка гена 32.) Возможно, что белок гена 61 является праймазой, аналогичной белку DnaG *E. coli*. В клетке, по-видимому, существует не более 10 молекул белка гена 61.

ДНК-полимераза гена 43 обладает обычной 5'—3'-синтетической активностью, связанной с 3'—5'-экзонуклеазной корректирующей активностью (гл. 32.) Оставшиеся три белка отнесены к «полимеразным вспомогательным белкам». Продукт гена 45 является димером. Продукты генов 44 и 62 образуют прочный комплекс, который характеризуется АТРазной активностью и тем, что он увеличивает скорость движения ДНК-полимеразы в 3–4 раза, от 250 до примерно 800 нуклеотидов/с; последнее значение близко к скорости движения ДНК-полимеразы *in vivo* (~1000 нуклеотидов/с). Увеличение в скорости не зависит от того, используется одно- или двухцепочечная матрица. В тех случаях, когда ДНК-полимераза фага Т4 работает на одноцепочечной матрице ДНК, скорость ее движения непостоянна. Фермент быстро передвигается в пределах одноцепочечных областей, однако при прохождении через области, имеющие вторичную структуру, образуемую спарившимися внутри цепи основаниями, движение его замедляется. Прохождению ДНК-полимеразы через такие участки способствуют вспомогательные белки. Их роль заключается в увеличении скорости репликации в «трудных» областях и, следовательно, в поддержании равномерности движения ДНК-полимеразы. Присутствие белков увеличивает средство ДНК-полимеразы с ДНК, а также ее способность удерживаться на матричной молекуле без диссоциации. Возможно, что белки действуют как «зажим», удерживая ДНК-полимеразную субъединицу на матрице более прочно. Для такого эффекта необходимо совместное действие всех трех белков. Подобный тип взаимоотношений оставляет открытым вопрос о том, что является компонентом ДНК-полимеразы, а что – вспомогательным фактором.

До сих пор мы говорили о репликации ДНК, касаясь круга вопросов, связанных с движением репликационной вилки. Наличие мутантов с задержкой или прекращением синтеза ДНК свидетельствует об участии в ней генов, контролирующих другие функции. Мутации в четырех из них – 39, 52, 58 и 60 – определяют задержку синтеза ДНК. С помощью комплементационного анализа в системе *in vitro* показано, что продукты генов 39 и 52 связываются с другим белком (неизвестной природы) и образуют ком-

Таблица 33.4
Репликационный аппарат фага Т4 состоит из определенных компонентов

Ген	Функция	Размер продукта (в дальтонах)	Относительное количество молекул
43	ДНК-полимераза	110 000	1
32	Белок, связывающийся с одноцепочечной ДНК	35 000	150
44	{ АТРаза; увеличивает скорость в 3-4 раза }	{ $4 \times 34\,000$ $2 \times 22\,000$ }	{ 180 000 комплекс } 5
41	ГТРаза, стимулируемая одноцепочечной ДНК; участвует в синтезе РНК-затравки	58 000	20
61	Участвует в синтезе РНК-затравки		1
45	Неизвестна, но, по-видимому, помогает ДНК-полимеразе	$2 \times 27\,000$	10

плекс из трех субъединиц (51 000, 64 000 и 110 000 дальтон). Этот комплекс проявляет активность топоизомеразы II фага T4. Существенная роль этого фермента свидетельствует о том, что ДНК фага T4 не остается в линейной форме, а утрачивает свободные концы. Топоизомераза может быть необходимой для раскручивания ДНК перед репликационной вилкой.

Репликационный аппарат фага T7

Изучение других систем репликации показывает, что в каждой из них присутствуют одни и те же типы активности, хотя обеспечиваться они могут разными способами. Репликационный аппарат фага T7 аналогичен аппарату репликации *E. coli* или фага T4, хотя, по-видимому, отличается менее сложной организацией.

Фаг кодирует одну субъединицу ДНК-полимеразы. Вторую субъединицу представляет белок бактерии-хозяина — тиоредоксин. Этот белок состоит из полипептида с мол. массой 12 000 дальтон и используется как кофермент в окислительно-восстановительной реакции при восстановлении рибонуклеотидов. Его функция в качестве партнера ДНК-полимеразы фага T7 загадочна; по-видимому, она отличается от той, которую он выполняет в бактерии-хозяине.

ДНК-полимераза фага T7 использует в виде вспомогательного белка продукт фагового гена 4, в котором объединены функции раскручивания двухспиральной матрицы и синтеза РНК-затравки. ДНК-полимераза вместе с белком гена 4 способна завершить репликацию *in vitro*, хотя *in vivo* при удалении РНК-затравки в процесс может также вовлекаться экзонуклеаза, кодируемая фаговым геном 6. Репликационная вилка содержит, по-видимому, в расчете на одну растущую цепь 10 молекул белка гена 4 и одну молекулу ДНК-полимеразы. Продукт гена 4 гидролизует НТР свыше того количества, которое используется в качестве субстрата для синтеза РНК. На каждое основание, включаемое в ДНК, может приходиться до четырех дополнительных гидролитических событий. Вероятно, этот гидролиз обеспечивает энергией процесс разделения цепей.

Праймная активность гена 4 используется для синтеза тетрарибонуклеотидных затравок, имеющих почти постоянную последовательность rrrArCpCpCp^A . Узнаваемая последовательность в случае реакции синтеза затравки включает предшествующий остаток G. Последовательность, в которой иницируется синтез затравки, встречается в ДНК фага T7 примерно один раз на каждые 500 оснований, в то время как длина фрагментов Оказаки колеблется в пределах от 1000 до 6000 оснований. Отсюда следует, что используются только некоторые сайты затравки. Точка переклечения синтеза затравки на синтез ДНК не является строго установленной, так как размеры некоторых затравок увеличены до пентарибонуклеотидов.

Проблема линейных репликонов

В связи с тем что все известные полимеразы нуклеиновых кислот способны действовать только в $5' \rightarrow 3'$ -направлении, возникает проблема синтеза ДНК в конце линейного репликона. Рассмотрим две родительские цепи, изображенные на **рис. 33.10**. Нижняя цепь может использоваться в качестве матрицы для синтеза дочерней цепи, которая быстро наращивается вправо до

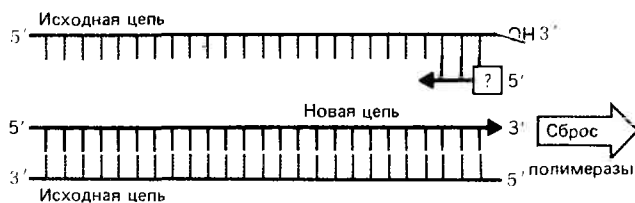


Рис. 33.10. Репликация может удлинить 3'-конец вновь синтезированной цепи, но может ли она иницировать синтез 5'-конца?

конца, где полимераза, по-видимому, отпадает. Однако, для того чтобы могла синтезироваться комплементарная цепь, в конце верхней цепи в самом последнем ее основании должно произойти затравочное событие.

Мы не знаем, возможно ли образование терминальной затравки. Принято считать, что РНК-полимераза связывается в сайте, расположенном рядом с участком, в котором осуществляется включение основания. Были предложены две гипотезы, объясняющие возможность копирования конца. Во-первых, ДНК может образовывать необычную структуру, например шпильку, что приведет к отсутствию свободного конца. Иная возможность реализуется у *Paramecium*. Показано, что при репликации ее линейной митохондриальной ДНК образуются поперечные сшивки. (В некоторых случаях проблема решается путем превращения линейного репликона в кольцевую или конкатемерную молекулу.) Во-вторых, допустимо также вмешательство белка, делающего затравочную реакцию возможной.

Некоторые линейные нуклеиновые кислоты вирусов содержат белки, ковалентно связанные с 5'-концевым основанием. Наиболее хорошо изучены ДНК аденовируса, фага $\phi 29$ и РНК полиовируса. ДНК аденовируса представляет собой большую линейную двухцепочечную молекулу; оба ее 5'-конца ковалентно связаны с белком, имеющим мол. массу 55 000 дальтон. Соединение осуществляется с помощью фосфодизфирной связи с серином (**рис. 33.11**). Тот же тип организации установлен в ДНК вируса $\phi 29$, где к каждому из 5'-концов прикреплен белок с мол. массой 27 000 дальтон. У полиовируса, содержащего одноцепочечную РНК, белок VPg из 22 аминокислот сцеплен через гидроксильную группу тирозина с 5'-концевым основанием. В каждом случае прикрепляемый белок кодируется вирусом и участвует в репликации.

ДНК аденовируса $\phi 29$ реплицируется со своих концов с помощью механизма вытеснения цепи, показанного на **рис. 33.12**. Одни и те же события происходят независимо, т.е. не одновременно в каждом из концов. Иницированная синтез новой цепи ведет к вытеснению гомологичной цепи, которая была ранее спаренной в двухцепочечной молекуле. В конечном счете либо репликационная вилка достигает другого конца молекулы, либо две репликационные вилки, двигающиеся от противоположных концов, встречаются в середине. Любое событие ведет к отделению родительских цепей друг от друга.

Белок, присутствующий на реплицирующихся цепях ДНК аденовируса, имеет мол. массу 80 000 дальтон, т.е. он крупнее белка, найденного на зрелой вирусной ДНК. Крупный белок является родственным более мелкому. Предполагается, что белок с мол. массой 80 000 дальтон участвует в инициации репликации, но на определенной

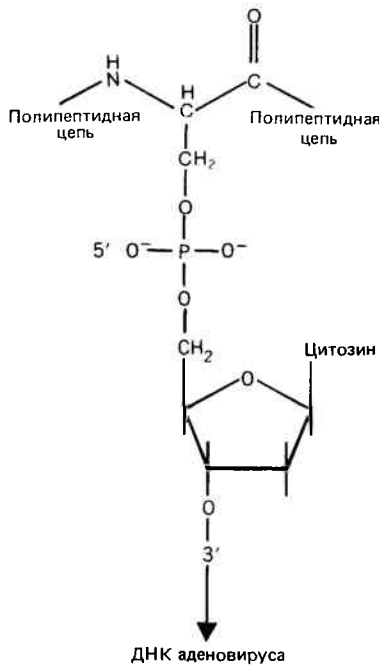


Рис. 33.11. 5'-концевой фосфат в каждом конце аденовирусной ДНК ковалентно присоединен к серину в белке (55 000 дальтон), связывающемся с аденовирусом.

стадии созревания вируса крупный белок расщепляется на белки меньшего размера.

Каким образом прикрепление белка решает проблему затравки? Способность белка ковалентно связываться с dCTP в экстрактах клеток, инфицированных аденовирусом, объясняет роль крупного белка. На рис. 33.13 изображен предполагаемый механизм его действия. Большой белок, связанный с цитидином, имеющим свободную 3'-ОН-группу, взаимодействует с концом аденовирусной ДНК. Он вытесняет имеющийся 5'-конец, который узнается благодаря его связи с белком, имеющим мол. массу 55 000. Затем свободный 3'-ОН-конец взаимодействует с поступившим нуклеотидом, который спаривается с матричной цепью при участии ДНК-полимеразы. Белок с мол. массой 80 000 дальтон может образовывать комплекс с ДНК-полимеразой. Вся серия реакции, по-види-

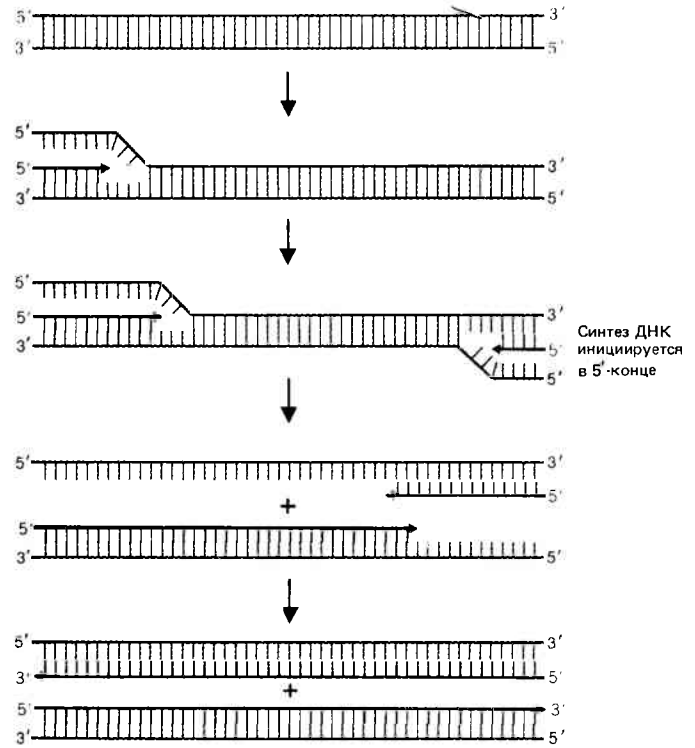


Рис. 33.12. Репликация аденовирусной ДНК инициируется независимо в двух концах молекулы и происходит путем вытеснения цепи.

тому, скоординирована, поскольку белок (80 000 дальтон) способен присоединяться к dCTP только в присутствии ДНК аденовируса. Эта модель, вероятно, применима и к репликации ДНК фага φ29.

РНК-полимераза, ответственная за репликацию полиовируса, подобно ДНК-полимеразам, строго зависит от затравки. И в этом случае найден более крупный предшественник ковалентно связанного VPg белка. Он имеет размер, равный 12 000 дальтон, и связан с мембраной. Вероятно, предшественник участвует в инициации, а затем разрезается, в результате чего образуется остаточный белок VPg.

Белок VPg отделяется от РНК полиовируса, которая предназначена для использования в качестве мРНК; он найден только на тех геномах, которые должны быть реплицированы.

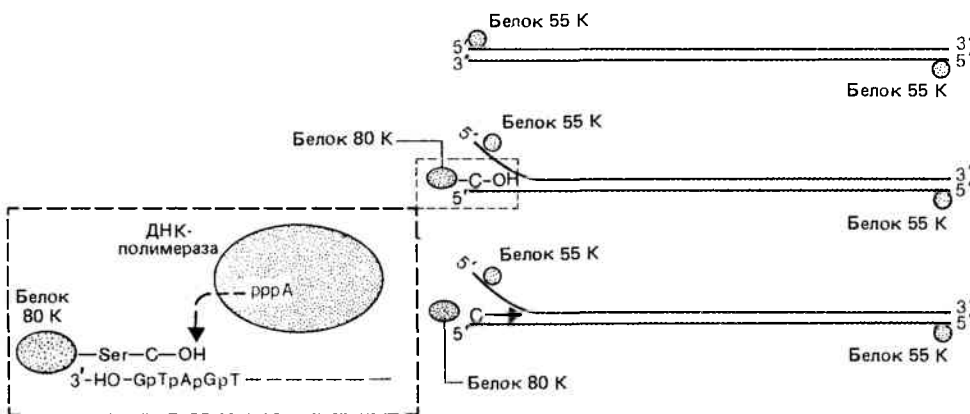


Рис. 33.13. Аденовирусный белок с мол. массой 80 000 дальтон смещает 5'-конец ДНК (который связан с белком, имеющим мол. массу 55 000 дальтон) и предоставляет dCTP для инициации синтеза новой цепи ДНК.

Рекомендуемая литература

Механизмы репликации детально анализируются в книге Корнберга (*Kornberg, DNA Replication, Freeman and Co., San Francisco, 1980*). Различные репликационные системы обсуждены Томизавой и Селзером (*Tomizawa, Selzer, Ann. Rev. Biochem., 48, 999–1034, 1979*). Система репликации *E. coli* рассматривается Араи и др. (*Arai et al., J. Biol. Chem., 256, 5239–5246, 1981*); структуру праймосомы анализируют Араи и Корнберг (*Arai, Kornberg, Proc. Nat. Acad. Sci. USA, 78, 69–73, 1981*). Функции Rep-белка посвящена работа Ярантона и Гефтера (*Yarranton, Gifter, Proc.*

Nat. Acad. Sci. USA, 76, 1658–1662, 1979). Роль белка A фага φX174 рассматривают Скотт и др. (*Scott et al., Proc. Nat. Acad. Sci. USA, 74, 193–197, 1977*). Репликации *oriC* посвящена статья Фуллера, Какунаги и Корнберга (*Fuller, Kakunagi, Kornberg, Proc. Nat. Acad. Sci. USA, 78, 7370–7374, 1981*). Обзоры исследований, проведенных на системе фага T4, опубликованы Лю (*Liu et al., Cold Spring Harbor Symp. Quant. Biol., 43, 469–487, 1979*), на системе T7 – Ричардсоном (*Richardson et al., Cold Spring Harbor Symp. Quant. Biol., 43, 427–440, 1978*). Белки, связанные с вирусами, анализирует Виммер (*Wimmer, Cell, 28, 199–201, 1982*).

Глава 34

СИСТЕМЫ ЗАЩИТЫ ДНК

Молекула ДНК осуществляет свои наследственные функции посредством репликации и транскрипции. В каждом из этих процессов участвует большое количество ферментов, среди которых обязательно присутствие соотвествующей полимеразы; оба процесса регулируются (гл. 10 и 33). Кроме того, с ДНК взаимодействуют различные ферменты, которые модифицируют ее структуру или репарируют возникающие в ней повреждения. Рассмотрение этих процессов очень важно для решения вопроса о механизмах сохранения ДНК в течение неопределенного числа поколений. Ведь сам по себе акт репликации еще не гарантирует ДНК выполнения ее роли в эволюции.

Как про-, так и эукариоты содержат ферменты, метилирующие ДНК. В каждом бактериальном штамме присутствует фермент (или ферменты), определяющий тип **модификации** его ДНК; этот признак можно использовать для решения вопросов о происхождении той или иной ДНК.

Бактерия способна отличить свою собственную ДНК от любой вторгающейся «чужеродной» именно по типу ее модификации. Различие в модификации делает чужеродную ДНК чувствительной к действию **рестрицирующих ферментов**, которые узнают отсутствие метильных групп в соответствующих сайтах. **Системы рестрикции и модификации** широко распространены у бактерий; их существование играет важную роль в защите резидентной ДНК от загрязнения последовательностями чужеродного происхождения. Следует отметить, однако, что присутствие системы модификации и рестрикции не обязательно; некоторые бактериальные штаммы утратили ее.

Метилирование у эукариот имеет другую цель: узнать гены, находящиеся в различных функциональных условиях (гл. 30); этот процесс не связан с рестрикцией немодифицированных сайтов.

Бактерия имеет несколько систем, защищающих ее ДНК от последствий повреждений внешними агентами или ошибок репликации. Любое событие, которое вызывает отклонение от правильной двухцепочечной структуры в ДНК, узнается как неприемлемое. Это изменение может быть точковой мутацией, которая превращает одно основание в другое, в результате чего образуется

пара, не соответствующая правилам Уотсона–Крика. Это может быть и структурное изменение, которое модифицирует ДНК или свяжет два основания вместе.

Сайты мутаций узнаются специальными нуклеазами, которые вырезают поврежденный участок из ДНК, а затем другие ферменты синтезируют замещающую последовательность. Вместе эти активности составляют **репарирующую систему**. Наряду с репарацией повреждений путем вырезания и замещения существуют системы, исправляющие вредные последствия репликации поврежденной ДНК. Такие **исправляющие системы** родственны системе генетической рекомбинации. Клетки *E. coli*, лишённые исправляющих систем, становятся чрезвычайно чувствительными к определенным типам повреждений. Репарирующие и исправляющие системы наиболее полно охарактеризованы у *E. coli*, однако их аналоги в клетках эукариот, вероятно, играют такую же важную роль. Можно предположить, что определенные болезни человека возникают в результате неправильного функционирования специфических репарирующих систем.

Говоря о молекулярном уровне организации таких систем, следует отметить участие в них ферментов, обладающих замечательными свойствами узнавать последовательности или структурные особенности ДНК. Ферменты рестрикции связываются со специфическими последовательностями ДНК, что дает нам дополнительную информацию о природе взаимодействия белков с нуклеиновыми кислотами. Некоторые рестриктазы связываются с ДНК в одном сайте, а разрезание производят в другом, достаточно удаленном; это свидетельствует о способности белков перемещаться вдоль двухцепочечной ДНК. По-видимому, репарирующие ферменты узнают поврежденные сайты в ДНК из-за искажения в этом участке молекулы правильной структуры. Ферменты, участвующие в рекомбинации, могут связывать две молекулы ДНК, стимулируя спаривание между ними.

Анализ различных ферментативных активностей дает два типа информации. С одной стороны, мы узнаем о взаимодействиях белка с ДНК, осуществляемых путем специфического узнавания последовательностей или структурных особенностей – спорным вопросом в проблеме биологии гена. С другой стороны, появляется общее представление о разнообразной природе систем, необхо-

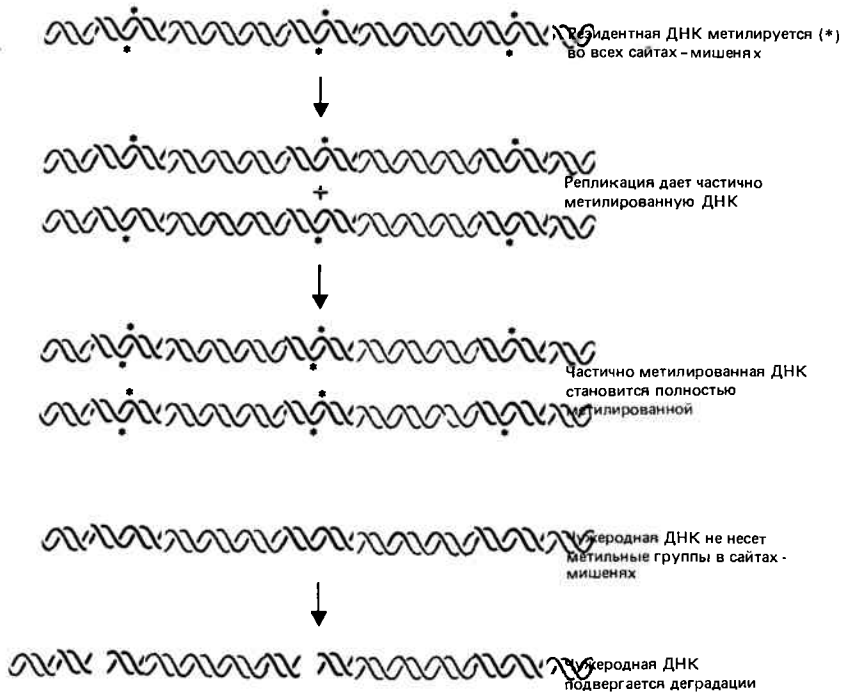


Рис. 34.1. Метилированные сайты сохраняются неопределенно долго и не повреждены рестрикцией; неметилированные сайты разрезаются.

димых для сохранения и защиты ДНК от воздействий окружающей среды или от ошибок других клеточных систем.

Процессы рестрикции и модификации

Все рестрикционные эндонуклеазы бактерий узнают специфические, довольно короткие последовательности ДНК и связываются с ними. Этот процесс сопровождается разрезанием либо в самом сайте узнавания, либо в каком-то другом, что определяется типом фермента. Подробно мы анализировали реакцию разрезания при обсуждении перспектив ее использования для картирования или реконструирования ДНК *in vitro* (гл. 3 и 19). Теперь мы рассмотрим вопрос о том, как выполняют эти ферменты свои природные функции.

Наряду с рестрикционной активностью (или активностями) бактериальный штамм обладает способностью метилировать ДНК; для этого процесса характерна такая же специфичность в отношении последовательностей ДНК, как и для рестрикции. Метилаза добавляет метильные группы к адениновым или цитозиновым остаткам в том же сайте, в котором связывается рестрикционный фермент. В результате метилирования сайт становится устойчивым к рестрикции. Следовательно, метилирование защищает ДНК от разрезания.

При метилировании бактериальной ДНК в соответствующих сайтах она становится иммунной к рестриктазам (рис. 34.1). Такая защита необходима для предотвращения деградации ДНК клетки под действием ее собственных рестриктаз. Однако этот способ защиты не распространяется на любую чужеродную ДНК, входящую в клетку. Она имеет немодифицированные сайты-мишени и поэтому атакуется ферментом рестрикции. Следовательно, комбинация модификации и рестрикции позволяет клетке отличать чужеродную ДНК от своей собственной.

В этом контексте термин «чужеродная» используется для обозначения любой ДНК, полученной из бактериального штамма, утратившего модифицирующую и рестри-

цирующую активность. Не существует различий среди типов ДНК, таких, как бактериальные, плазмидные или фаговые. Все ДНК, являющиеся резидентными в определенном штамме бактерий или временно находившиеся в нем, обладают одинаковым типом модификации. Следовательно, система рестрикции и модификации не препятствует обмену ДНК между бактериями одного и того же штамма.

Системы модификации и рестрикции были открыты благодаря их действию в отношении инфицирующей фаговой ДНК. Фаговая ДНК, выделяемая из бактерии, может успешно инфицировать другую бактерию того же штамма, поскольку обе клетки имеют один и тот же тип модификации. Однако фаговая ДНК, которая переходит из одного штамма в другой, атакуется рестрикционными эндонуклеазами. Следовательно, фаг ограничен (*restricted*) одним бактериальным штаммом, отсюда и появился термин «рестрикция». Следует отметить, что рестрикция не обязательна. Некоторые инфицирующие фаги избегают ее, что обусловлено либо мутациями в сайтах-мишенях, либо неточной работой системы клетки-хозяина. В этом случае они приобретают тип модификации нового хозяина.

Плазмиды могут вносить свой вклад в штаммовую специфичность бактерий. В геноме некоторых плазмид (и фагов) включены гены систем рестрикции и модификации, так что бактерия, несущая один из этих элементов, и такая же бактерия, но утратившая его, будут характеризоваться как принадлежащие к разным штаммам. Во всяком случае существование систем модификации и рестрикции служит противодействием передачи бактериальных последовательностей с помощью плазмид и фагов.

Рестрикционные ферменты относятся к двум основным классам, которые в свою очередь могут быть подразделены на более мелкие группы. Их свойства суммированы в табл. 34.1.

Первый класс включает ферментативные активности типа II. Сюда относят рестрикционные ферменты, уже обсуждавшиеся ранее. Они обнаруживаются примерно

Таблица 34.1

Активности рестрикции и метилирования могут быть объединены или могут существовать отдельно

Класс фермента	Тип II	Тип III	Тип I
Структура белка	Отдельно эндонуклеаза и метилаза	Бифункциональный фермент из двух субъединиц	Бифункциональный фермент из трех субъединиц
Сайт узнавания	Короткая последовательность (4–6 пар оснований), часто палиндромная	Асимметричная последовательность (5–7 пар оснований)	Двусторонняя и асимметричная (например, TGAN ₈ TGCT) последовательность
Сайт разрезания	Тот же, что и сайт узнавания или близко от него	Отделен от сайта узнавания участком в 24–26 пар оснований	Неспецифичный, примерно на расстоянии 1000 пар оснований от сайта узнавания
Рестрикция и метилирование	Отдельные реакции	Одновременные	Взаимно исключают
Необходима ли АТФ для рестрикции?	Нет	Да	Да

в одном из трех бактериальных штаммов. Каждый из этих ферментов отвечает только за акт рестрикции; метилирование той же последовательности осуществляет другой фермент. Структура белка не отличается сложностью. Наиболее подробно охарактеризованы компоненты системы EcoRI, в которой рестрикционный фермент представляет собой димер из идентичных субъединиц, а метилаза является мономером. Узнавание общей мишени могло возникнуть либо путем дубликации соответствующей части белка, либо в результате конвергентной эволюции.

В качестве сайтов-мишеней часто выступают палиндромы из 4–6 пар оснований. Симметрия подразумевает, что те из них, которые должны быть метилированы, встречаются на обеих цепях ДНК. В результате сайт-мишень может быть полностью метилирован (обе цепи модифицированы), полуметилирован (только одна цепь метилирована) или неметилирован. Полностью метилированный сайт не подвержен ни рестрикции, ни модификации. Полуметилированный сайт не узнается ферментом рестрикции, но может быть превращен с помощью метилазы в полностью метилированный. У бактерий метилирование, как правило, связано с сохранением имеющегося состояния модификации. Репликация полностью метилированной ДНК ведет к образованию полуметилированной ДНК. Вероятно узнавание полуметилированных сайтов представляет собой обычный этап функционирования метилазы *in vivo*. Неметилированный сайт-мишень представляет собой субстрат либо для рестрикции, либо для модификации *in vitro*. В клетке немодифицированная ДНК с большей вероятностью рестрицируется.

Метилаза добавляет за один раз только одну метильную группу. Так, в случае неметилированного субстрата одно метилирование сопровождается отделением метилазы от ДНК. Для введения второй метильной группы требуется отдельное связывание и событие метилирования. Поскольку репликация полностью метилированной ДНК всегда ведет к образованию частично метилированной ДНК (см. рис. 34.1 и рис. 31.15), узнавание частично метилированных сайтов является, вероятно, обычным способом действия метилазы *in vivo*. Фермент рестрикции не связывается с метилированными сайтами-мишенями.

Большинство рестрикционных ферментов типа II разрезают ДНК в неметилированном сайте-мишени. В каждой цепи ДНК разрезается одна связь; разрезание может производиться последовательно. Некоторые ферменты вносят ступенчатые разрезы, другие вызывают образова-

ние тупых концов (гл. 19). Ферменты, разрезающие ДНК с одной стороны вблизи от сайта узнавания, относят к одному подклассу. Вероятно, по мере изучения механизмов действия ферментов будут обнаружены и другие подклассы.

Ко второму классу активностей, характеризующих систему рестрикции и модификации отнесены активности ферментов типа I и III. Эти ферменты представляют собой мультимеры и выполняют в клетке как эндонуклеотическую, так и метилирующую функции. Механизмы их действия отличны друг от друга и от механизма действия ферментов типа II. Точно не известно, у какой части бактерий функционируют системы типа I или III, однако очевидно, что они менее распространены, чем система типа II.

Альтернативные активности ферментов типа I

К ферментам типа I относятся EcoK и EcoV, выделенные из штаммов *E. coli* K и V. Рис. 34.2 дает представление о ферментах рестрикции и модификации типа I, состоящих из трех разных субъединиц. Субъединица R ответственна за рестрикцию, субъединица M – за метилирование, а субъединица S – за узнавание сайта-мишени ДНК. После того как сайт-мишень узнается S-субъединицей, происходит связывание фермента с ДНК, а затем либо ее рестрикция, либо модификация; активности субъединиц R и M являются взаимоисключающими.

Каждая субъединица кодируется одним геном. Они названы *hsdR*, *hsdM* и *hsdS* соответственно. Мутанты *hsdR* имеют фенотип g^-m^+ ; они не способны рестрицировать ДНК, однако могут модифицировать ее. Мутанты *hsdS* являются по фенотипу g^-m^- , т.е. не модифицируют и не рестрицируют ДНК. Мутации в *hsdM* превращают как рестрикцию, так и модификацию. Следовательно, M-субъединица включается в функционирование R-субъединицы. Вероятно, это необходимо для того, чтобы мутанты *hsdM* были по фенотипу g^-m^- , а не g^+m^- , поскольку мутанты с таким фенотипом, по-видимому, нежизнеспособны (рестрикционная активность вызывала бы деградацию собственной немодифицированной ДНК).

Гены, кодирующие ферменты EcoK и EcoV, аллельны. Узнаваемые ими последовательности различны, но путем конструирования диплоидов из соответствующих мутантов можно показать, что S-субъединица одного штамма направляет активности R- и M-субъединицы другого бактериального штамма. Следовательно, стадия узнавания

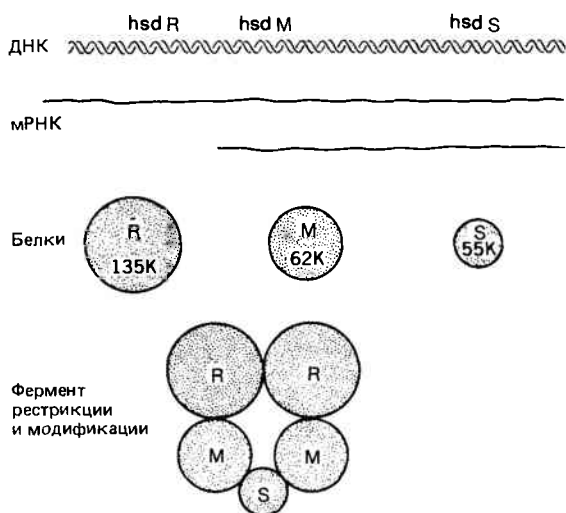


Рис. 34.2. Полифункциональные ферменты типа I содержат различные субъединицы для рестрикции, модификации и узнавания.

не зависит от сменяющих ее событий разрезания или метилирования.

Фермент EcoK имеет мол. массу около 400 000. Он содержит две копии R-субъединицы по 135 000 дальтон каждая, две копии M-субъединицы (62 000 дальтон) и одну копию S-субъединицы с мол. массой 55 000. Фермент EcoV состоит из субъединиц такого же размера, однако они могут быть представлены в других молярных соотношениях. Субъединицы M и S в соотношении 1:1 формируют комплекс, который проявляет функцию метилирования независимо от функции рестрикции. Это естественно, поскольку три гена локализируются в опероне в следующем порядке: $P1-hsdR-P2-hsdM-hsdS$, где $P1$ и $P2$ — независимые промоторы (см. рис. 34.2). Таким образом, гены $hsdM$ и $hsdS$ могут экспрессироваться независимо от $hsdR$.

Сайты узнавания ферментов EcoV и EcoK — это структуры, состоящие из двух частей. В обоих случаях первая часть представлена специфической последовательностью из трех пар оснований, отделенной у фермента EcoV семью парами, а у EcoK — шестью парами оснований от специфической последовательности в четыре нуклеотидные пары. Расстояние между двумя частями сайта узнавания указывает на то, что обе они находятся на одной стороне ДНК; последовательность промежуточной области, по-видимому, важной роли не играет. Заметим, что оба сайта узнавания несимметричны. Однако каждый из них действительно несет на каждой из цепей по одному аденину для метилирования, как показано звездочками в последовательности узнавания ферментов EcoV.



Будет ли данная ДНК разрезана или модифицирована, определяется состоянием сайта-мишени (рис. 34.3). Если сайт-мишень полностью метилирован, фермент может связаться с ним, но затем диссоциировать без какого-то дальнейшего действия. Если мишень частично метилирована, фермент метилирует неметилированную цепь; благодаря этому сохраняется состояние метилирования в ДНК хозяина. Если мишень не метилирована, ее узнавание способствует реакции разрезания.

Группы для метилирования обеспечиваются кофактором S-аденозилметионином (SAM), который в ходе реакции превращается в S-аденозилгомоцистеин (SAH). Связывание SAM с M-субъединицей необходимо для узнавания ДНК субъединицей S. Мутации в $hsdM$ могут привести к неспособности связываться с SAM. На начальной стадии реакции SAM действует как аллостерический эффектор, т.е. изменяет конформацию белка, позволяя ему связаться с ДНК. Конформационное изменение захватывает S-субъединицу, о чем свидетельствуют мутации в $hsdS$, предотвращающие ее переход в активную форму.

После связывания с ДНК следует реакция с АТФ. Если фермент связался с полностью метилированным сайтом, наличие АТФ ведет к отделению его от ДНК. В неметилированном сайте АТФ превращает фермент в форму, способную осуществить разрезание. Обнаружение мутаций в $hsdR$, предотвращающих АТФ-индуцируемое конформационное изменение, свидетельствует о том, что это зависит от R-субъединицы. Для разрезания ДНК необходим гидролиз АТФ. SAM освобождается от фермента до того, как происходит реакция разрезания. Роль АТФ в случае частично метилированного субстрата неясна; метилирование частично метилированной ДНК может происходить и без АТФ, но реакция идет более эффективно в его присутствии. Для сайтов узнавания и разрезания ферментов типа I характерны необычные взаимоотношения. Сайт разрезания расположен на расстоянии более 1000 пар оснований от сайта узнавания. Разрезание производится не в специфической последовательности, но и не в совершенно случайном сайте: имеются области, предпочтительные для событий разрезания. (Различие между сайтами узнавания и разрезания состоит в том, что сайты узнавания нельзя определить по такому характерному признаку, как разорванные концы ДНК; сайты-мишени идентифицируются благодаря их модификации при метилировании, а также при локализации мутаций, препятствующих узнаванию.)

Реакция разрезания осуществляется в две ступени. Сначала разрезается одна цепь ДНК, а затем рядом разрезается другая. В областях, прилегающих с каждой стороны к сайту разрезания, может иметь место экзонуклеотическая деградация. Происходит эффективный гидролиз АТФ, роль которого еще не выяснена.

Каким образом фермент узнает один сайт, а разрезает другой, достаточно удаленный? Важно отметить, что белок никогда не отделяется от молекулы ДНК, с которой он первоначально связался. Если фермент инкубировать со смесью модифицированной и немодифицированной ДНК, он предпочтительно разрезает немодифицированную ДНК. Следовательно, узнавая сайт связывания, белок не отделяется от неметилированной ДНК для того чтобы найти сайт разрезания. Существуют две альтернативные модели, объясняющие взаимосвязь между сайтами узнавания и разрезания: в соответствии с одной из них движется фермент, согласно другой модели, перемещается ДНК. Предполагаемые схемы этих процессов представлены на рис. 34.4. Если движется фермент, то его перемещение вдоль ДНК будет продолжаться до тех пор, пока он не сделает выбор сайта разрезания. Если же движется ДНК, то фермент остается прикрепленным в сайте узнавания, а ДНК протаскивается через второй сайт связывания на ферменте, и это продолжается до тех пор, пока фермент не достигает области разрезания (пока не-

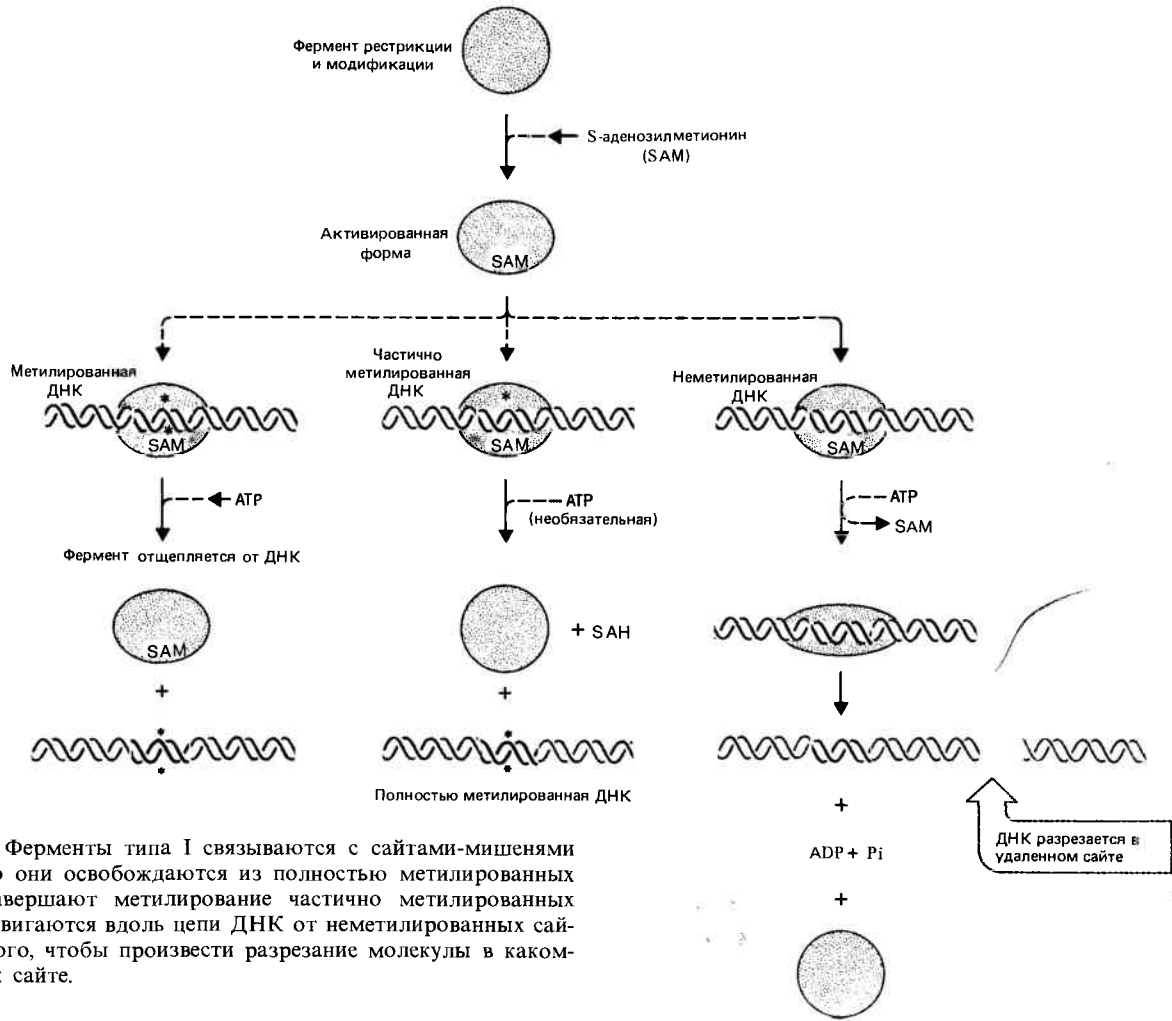


Рис. 34.3. Ферменты типа I связываются с сайтами-мишенями после чего они освобождаются из полностью метилированных сайтов, завершают метилирование частично метилированных сайтов и двигаются вдоль цепи ДНК от неметилированных сайтов для того, чтобы произвести разрезание молекулы в каком-то другом сайте.

охарактеризованной). Получены электронно-микроскопические данные, свидетельствующие, что фермент вызывает образование петли в ДНК и остается, по-видимому, связанным с сайтом узнавания после разрезания; эти данные подтверждают вторую модель.

Два вида активности ферментов типа III

Исследованы три фермента рестрикции и модификации типа III: *EcoP1* и *EcoP15*, кодируемые плазмидами P1 и P15 у *E. coli*, и *Hinf*, обнаруженный у *H. influenzae* серотипа R_f. Каждый фермент состоит из двух типов субъединиц. Субъединица R ответственна за рестрикцию; мутанты *hsdR* имеют фенотип $\gamma^- m^+$. Субъединица MS ответственна и за модификацию, и за узнавание. Мутации в гене *hsdMS* определяют фенотип $\gamma^- m^-$, если они локализованы в узнающей части последовательности; если же такие мутации локализованы в ее модифицирующей части, они обуславливают фенотип $\gamma^+ m^-$ (который, по-видимому, летален). Вероятно, узнающая часть гена лежит в середине его последовательности; к такому выводу приводит сравнение генов *hsdMS* систем P1 и P15. Субъединица R имеет мол. массу 106 000–110 000; полипептид MS – 73 000–80 000.

У ферментов типа III рестрикционная и модифицирующая активности выражаются одновременно. Сначала фермент присоединяется к своему сайту на ДНК; это событие требует ATP. (На последующих стадиях зависимость от ATP у разных ферментов неодинакова.) Затем активности рестрикции и метилирования конкурируют за вступление в реакцию с ДНК. Событие метилирования происходит в сайте связывания, что согласуется с объединением функций метилирования и узнавания в субъедини-

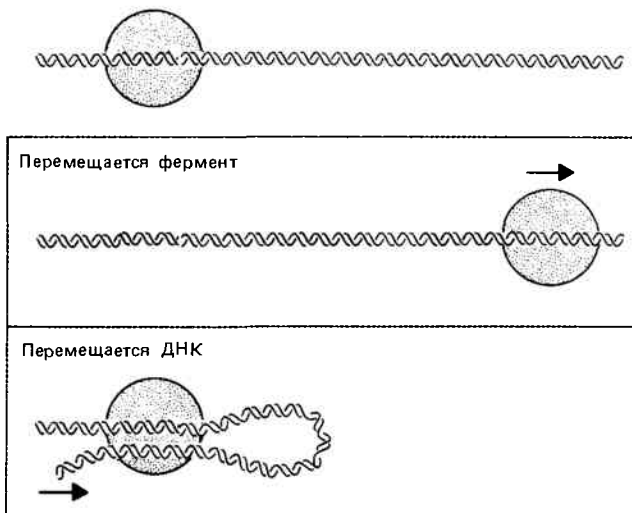


Рис. 34.4. Перемещается ли фермент типа I по ДНК, или он остается в своем сайте-мишени и протаскивает ДНК через белок?

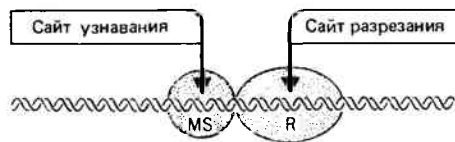


Рис. 34.5. Ферменты типа III содержат две субъединицы. Узнавание и метилирование, осуществляемые субъединицей MS, происходят в сайте-мишени; событие рестрикции может происходить в соседнем сайте, находящемся в контакте с субъединицей R.

це MS. Рестриktionное разрезание происходит на расстоянии 24–26 пар от сайта узнавания с одной его стороны, что, вероятно, обусловлено тем, что достаточно большие размеры фермента не позволяют субъединице рестрикции контактировать с ДНК в этой же точке (рис. 34.5). Рестрикция включает ступенчатые разрезы, отделенные друг от друга 2–4 основаниями. Ферменты метилируют адениновые остатки, однако сайты-мишени P1 и P15 имеют загадочное свойство: они могут быть метилированы *только на одной цепи*. Ниже представлены последовательности сайтов

P1	AGACC TCTGG
P15	CAGCAG GTCGTC

Возникает вопрос: как сохраняется состояние метилирования? На рис. 34.6 показано, что в процессе репликации возникают два типа сайтов. Одна реплика содержит исходную метилированную цепь и фактически неотличима от родительской. Другая реплика полностью не метилирована и поэтому представляет собой мишень либо для рестрикции, либо для модификации. Чем объяснить, что из двух этих возможностей чаще выбирается вторая? Ответ нам неизвестен, но одно из объяснений состоит в том, что реакция метилирования связана с актом репликации.

Механизмы репарации повреждений ДНК

Повреждения в ДНК сводятся к минимуму благодаря существованию систем, которые узнают эти повреждения и исправляют их. Поэтому определяемая частота мутаций – спонтанных и индуцированных – отражает равновесие между числом повреждающих событий, происходящих в ДНК, и количеством повреждений, которые были исправлены (или неправильно исправлены). Значение репарирующих систем в жизни клетки так велико, что, ве-

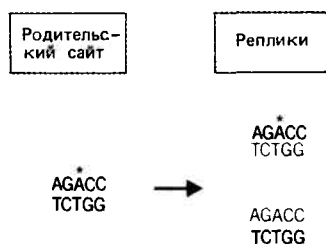


Рис. 34.6. Репликация метилированной последовательности sP1 приводит к образованию метилированной и неметилированной реплик.

роятно, по сложности они не отличаются от репликационного аппарата.

Под «повреждением» понимают любое изменение ДНК, которое вызывает отклонение от обычной двухцепочечной структуры. К повреждающим событиям могут быть отнесены следующие:

- 1) введение одноцепочечных разрывов;
- 2) удаление основания, в результате чего его гомолог остается неспаренным;
- 3) превращение одного основания в другое, которое неправильно спарено с основанием-партнером;
- 4) введение ковалентных связей между основаниями на одной цепи ДНК или между основаниями на противоположных цепях.

Некоторые репарирующие системы не специфичны в отношении каких-то определенных типов повреждений, а узнают искажения в структуре ДНК как сигналы, указывающие на необходимость действия. Перечисленные выше повреждающие изменения можно разделить на два основных класса.

Точковые мутации могут быть узнаны благодаря тому, что они нарушают правильное спаривание оснований. Это не влияет на транскрипцию или репликацию, когда цепи ДНК разделены. Следовательно, точковые мутации проявляют свой повреждающий эффект на будущие поколения вследствие изменений в последовательности ДНК (гл. 2).

Структурные нарушения могут создать физические препятствия для репликации или транскрипции. Разрыв цепи или отсутствие основания может привести к помехам при использовании цепи в качестве матрицы для синтеза РНК или ДНК. Присутствие внутрицепочечных поперечно сцепленных структур может подавить репликацию, вызывая остановку полимеразы в сайте повреждения. В таком случае фермент продолжает синтез в последующей точке, оставляя во вновь синтезированной цепи брешь (пробел). Межцепочечные поперечные сшивки препятствуют разделению цепей ДНК. Наличие любого типа сшивок приводит к преждевременной остановке транскрипции. Наиболее изучен пример образования внутрицепочечных **тиминовых димеров** под действием ультрафиолетового облучения (рис. 34.7).

В исправлении этого повреждения участвуют три типа событий. Связи в тиминовых димерах могут быть удалены при участии фермента, зависящего от света. Такой

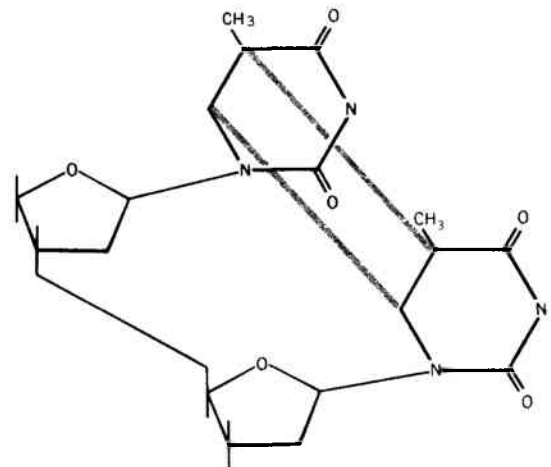


Рис. 34.7. Тиминовый димер образуется в результате возникновения ковалентных связей между смежными основаниями.

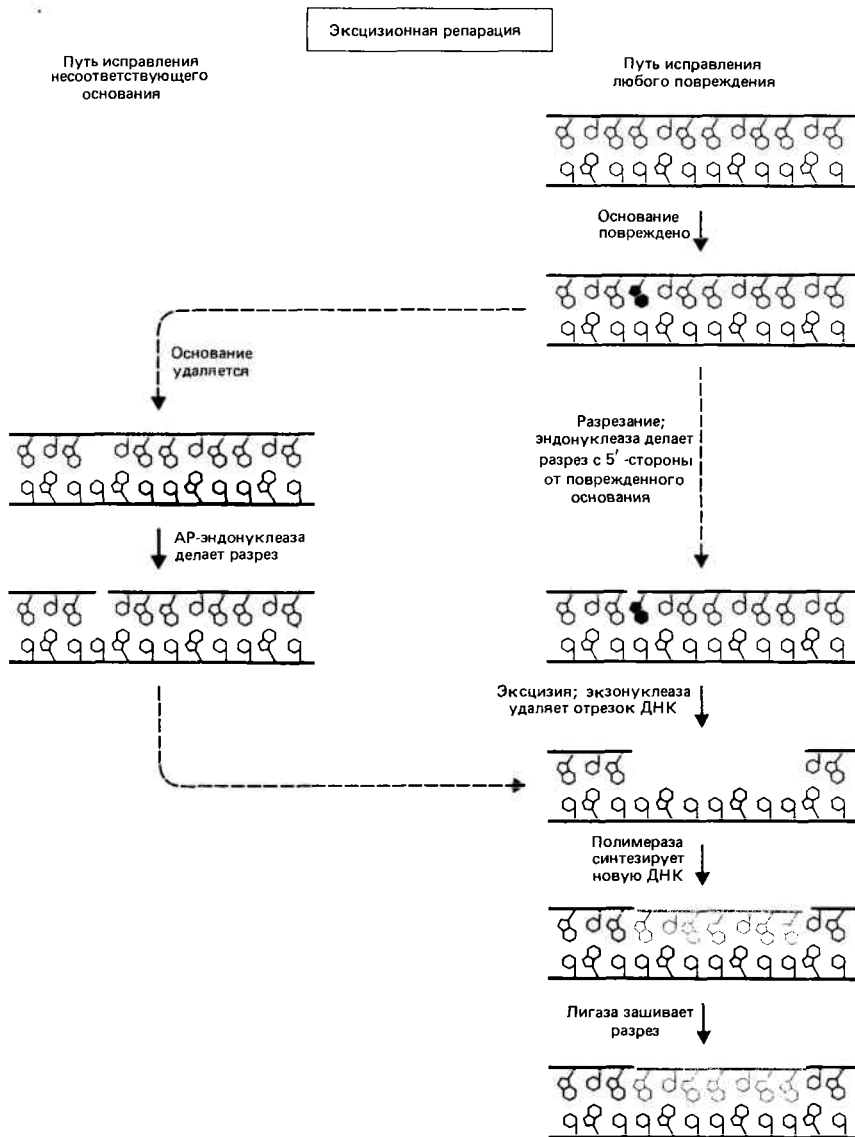


Рис. 34.8. Эксцизионная репарация удаляет и замещает последовательность ДНК, которая включает поврежденное (поврежденные) основание (основания), но имеет большую протяженность.

процесс получил название **фотореактивации**; у *E. coli* он зависит от продукта одного гена (*phr*). Поврежденный материал может быть удален и замещен с помощью **репарирующей системы**. Наконец, повреждение может оставаться в исходном сайте, но дочерние цепи, синтезируемые в процессе репликации, могут быть репарированы с помощью **исправляющей системы**. Клетка может иметь независимые системы репарации и исправления.

Системы **эксцизионной репарации** удаляют неправильно спаренные или поврежденные основания из ДНК и затем синтезируют новую последовательность ДНК, замещающую их. Основной тип эксцизионной репарации схематически изображен на **рис. 34.8**. На первом этапе поврежденная структура узнается эндонуклеазой, которая **разрезает** цепь ДНК с 5'-стороны от повреждения. На стадии **вырезания** 5'—3'-экзонуклеаза удаляет поврежденный участок. Образующийся одноцепочечный участок служит в качестве матрицы для ДНК-полимеразы при **синтезе цепи**, замещающей вырезанную последовательность. Наконец, ДНК-лигаза ковалентно связывает 3'-конец нового материала со старым материалом.

Если повреждение в ДНК представляет собой структурное искажение, например тиминный димер, поврежденные основания удаляются, что ведет к восстановле-

нию последовательности дикого типа. Однако в том случае, когда нарушение заключается в неправильном спаривании оснований, возникающем в результате мутирования одного из них, репарирующая система не может определить, какое именно основание представляет дикий тип, а какое — мутантный. Все это узнается как два неправильно спаренных основания, каждое из которых может служить объектом для эксцизионной репарации. Если вырезается мутантное основание, восстанавливается дикий тип последовательности. Но если случается так, что вырезается исходное основание (дикого типа), то новая (мутантная) последовательность закрепляется. Во многих случаях направление эксцизионной репарации не случайно, а определяется локальной последовательностью. Это может иметь значение в генетической рекомбинации (гл. 35).

В случае ошибок спаривания, которые не узнаются ДНК-полимеразой во время репликации, оказывается возможным отличить исходную цепь ДНК. После репликации метилированной ДНК только исходная родительская цепь несет метильные группы (см. рис. 34.1). В период, когда вновь синтезированная цепь еще ждет введения метильных групп, две цепи можно различить. Вероятно, это важно для системы исправления реплика-

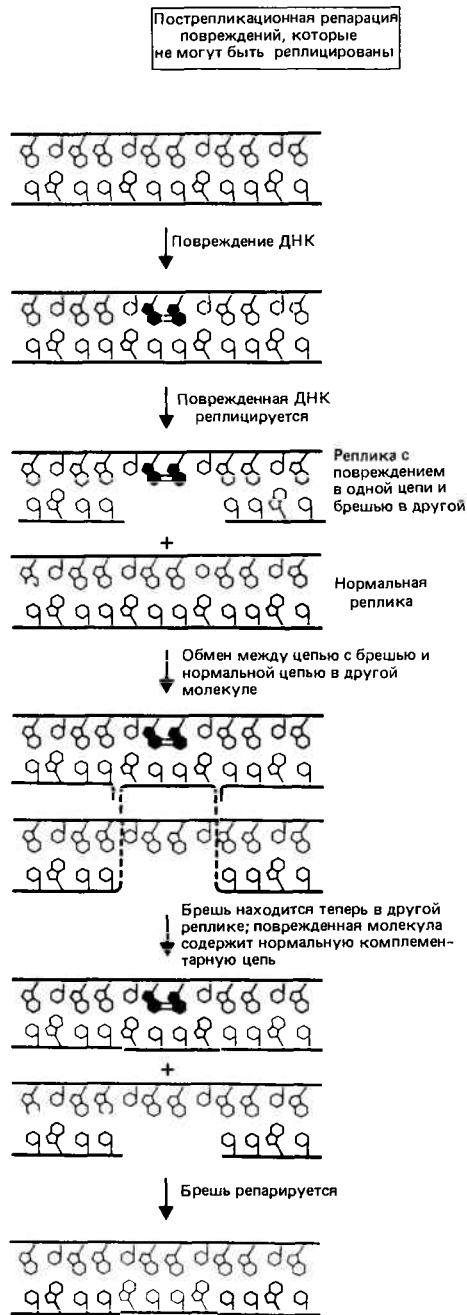


Рис. 34.9. Рекомбинационная репарация восстанавливает нормальную цепь ДНК, замещая брешь (пробел), оставленную во вновь синтезированной цепи напротив сайта нерепарируемого повреждения.

ционных ошибок; мутанты *dam⁻ E. coli* не способны метилировать аденин, и для них характерна увеличенная частота спонтанного мутирования. Свойства таких мутантов подразумевают существование репарирующей системы, которая специфически исправляет цепь ДНК, содержащую неметилированные адениновые основания, чтобы обеспечить ее комплементарность метилированной цепи.

Другой тип специфической репарации, не связанной с необходимостью различать последовательности дикого и мутантного типов, осуществляет урацил-гликозилаза, которая специфически удаляет урациловые основания из

ДНК (гл. 2 и 32). Механизм ее действия в отношении неправильного основания схематически изображен на рис. 34.8. Из фосфодиэфирного остова основание удаляется, в результате чего образуется апуриновый сайт. Этот сайт специфически узнается **АР-эндонуклеазой** (АР символизирует отсутствие пурина или пиримидина в зависимости от конкретного фермента), которая вносит разрыв на 5'-стороне сайта. Затем могут следовать этапы эксцизионной репарации.

Исправление повреждений осуществляется ферментами **рекомбинационной репарации**, которые используют материал одной молекулы ДНК для восстановления другой. Название этого типа репарации свидетельствует о перекрывании с процессами генетической рекомбинации. Поскольку для рекомбинационной репарации необходимо, чтобы ДНК мишени имела реплику, ее иногда называют **пострепликационной репарацией**. Она эффективна в случае дефектов, образуемых в дочерних молекулах при репликации матрицы, содержащей поврежденные основания (рис. 34.9).

Рассмотрим последствия образования тиминового димера в одной цепи ДНК. При репликации поврежденный сайт не может использоваться в качестве матрицы, направляющей включение оснований в дочернюю цепь. Репликация вынуждена пропустить димер, оставляя брешь. Полимераза движется до точки (или почти до точки), в которой находится тиминовый димер, и прекращает синтез соответствующей дочерней цепи. Репликация возобновляется на некотором расстоянии дальше по цепи — например, в следующем сайте инициации фрагмента Оказаки в отстающей цепи. Поэтому брешь, оставляемая во вновь синтезированной цепи, может быть существенной.

Образующиеся дочерние двойные нити различны по природе. Одна из них содержит родительскую цепь, несущую повреждение, и находящуюся напротив вновь синтезированную цепь с длинной брешью. Другая имеет неповрежденную родительскую цепь, на которой была скопирована нормальная комплементарная цепь. Брешь напротив поврежденного сайта первой молекулы застраивается при использовании гомологичной отдельной цепи ДНК другой двойной спирали. Образующаяся в результате такого **одноцепочечного обмена** реципиентная молекула содержит родительскую (поврежденную) цепь, спаренную с нормальной цепью. Донорная молекула содержит нормальную родительскую цепь с находящейся напротив разорванной цепью; брешь может быть застроена обычным путем в результате репарирующего синтеза. Таким образом, повреждение ограничивается исходным нарушением (хотя одни и те же рекомбинационно репарирующие события должны повторяться после каждого репликационного цикла до тех пор, пока повреждение не будет устранено с помощью системы эксцизионной репарации).

Системы эксцизионной репарации у *E. coli*

У *E. coli* мутации в большом числе локусов влияют на ответную реакцию при воздействии агентами, повреждающими ДНК. Многие из этих генов, вероятно, кодируют ферменты, которые участвуют в работе систем, репарирующих ДНК, хотя в большинстве случаев продукты этих генов и их функции еще предстоит идентифицировать. При повреждении репарирующих систем клетки

Таблица 34.2

Гены, участвующие в репарации повреждений ДНК у *E. coli*

Ген	Эффект мутации	Продукт гена	Функция белка		
<i>uvrA</i> <i>uvrB</i> <i>uvrC</i> <i>uvrE</i>	Чувствительность к УФ-облучению	АТРазный компонент эндонуклеазы	Иницирует удаление тиминовых димеров из ДНК Неизвестна		
		Компонент эндонуклеазы			
		Компонент эндонуклеазы			
		Неизвестен			
<i>recA</i>	1. Нарушение рекомбинации 2. Не индуцирует репарацию после повреждения	Белок RecA, 40 000 дальтон	1. Участвует в обмене цепей ДНК в случае рекомбинации и рекомбинационной репарации 2. Иницирует SOS-индукцию путей репарации		
<i>recB</i> <i>recC</i> <i>sbcB</i>	Нарушение рекомбинации Супрессия <i>recBC</i> мутаций	{Субъединицы экзонуклеазы V То же Экзонуклеаза I	Необходимы для рекомбинации и рекомбинационной репарации Неизвестна		
<i>recF</i>				Нарушение рекомбинационной репарации и рекомбинации	Неизвестна
<i>recJ</i> <i>recK</i> <i>recE</i>				в мутантах <i>recBCsbcB</i> Нарушение рекомбинации	» » Экзонуклеаза VIII
<i>umiC</i>	Устраняет мутагенные эффекты репарации, исправляющей ошибки	Неизвестен	Неизвестна		
<i>lon</i>	УФ-чувствительность, нарушение деления	АТР-зависимая протеаза, связывающаяся с ДНК	Одна активность может контролироваться генами полисахаридов оболочки		
<i>mut</i>	По крайней мере 6 генов, мутации в которых вызывают повышенную частоту мутирования	Неизвестен	Неизвестна		
<i>dam</i>	УФ-чувствительность, повышенная частота мутаций	Может быть метилазой	Связана с метилированием аденина		
<i>lexA</i>	Устраняет SOS-ответ	Репрессор (22 000 дальтон)	Действует на многие гены		

становятся чрезвычайно чувствительными к облучению ультрафиолетом. Введение УФ-индуцируемых повреждений использовали в качестве основного теста в отношении репарирующих систем. Следовательно, при оценке их активности и эффективности мы должны помнить, что, если изучается другой тип повреждений, их проявление может быть различным.

В табл. 34.2 суммированы свойства некоторых мутаций, которые влияют на способность клеток *E. coli* осуществлять репарацию ДНК. Нам известны по крайней мере три способа репарации повреждений. Два разных способа эксцизионной репарации идентифицируются по гетерогенности длины сегментов репарируемой ДНК. Эти пути описываются как **репарация короткими и длинными последовательностями ДНК**. Ферменты рекомбинационной репарации в значительной мере перекрываются с ферментами, осуществляющими собственно рекомбинацию (гл. 35).

Система *uvr*, участвующая в осуществлении обоих типов эксцизионной репарации, включает три гена, *uvr A, B, C*, кодирующие компоненты репарационной эндонуклеазы. Первичная характеристика ферментативной активности позволяет предполагать, что эндонуклеаза узнает тиминный димер (или другие искажения) и делает первоначальный разрез на его 5'-стороне; другой разрез производится с 3'-стороны примерно на расстоянии 12 нуклеотидов. Определенные ферменты *E. coli* могут вырезать тиминные димеры *in vitro* из ДНК, в которой уже сделаны разрезы. Такой 5'—3'-экзонуклеазной активностью обладают ДНК-полимераза I и ДНК-полимераза II (см. табл. 32.2) и экзонуклеаза III, специфичная в отношении одноцепочечной ДНК. Однако мутации, по-

вреждающие любую из этих предполагаемых экзонуклеаз, не приводят к измеримому уменьшению способности клеток *E. coli* удалять тиминные димеры. Основываясь на данных о поведении различных двойных мутантов, в настоящее время основная роль в выщеплении повреждений *in vivo* отводится ДНК-полимеразе I.

Средняя протяженность вырезаемых сегментов ДНК составляет примерно 20 нуклеотидов. Такая длина характерна при способе репарации, определяемом как репарация короткими последовательностями. Ферментом, который осуществляет репарирующий синтез, вероятно, также является ДНК-полимераза I (хотя ферменты II и III способны заменить ее).

С помощью мутаций, которые нарушают процесс репарации короткими последовательностями, были идентифицированы другие потенциальные компоненты системы *uvr*, однако до сих пор продукты или функции генов, кодирующих эти компоненты, не были установлены.

Репарация короткими последовательностями используется в 99% случаев репарационных событий, включающих вырезание. Остаточный 1% приходится на замещение протяженных последовательностей ДНК, содержащих в основном около 1500 нуклеотидов; меньшая часть протяженных последовательностей включает приблизительно 9000 нуклеотидов. Этот способ требует участия генов *uvr* и ДНК-полимеразы I. Различие между двумя способами репарации состоит в том, что репарация короткими последовательностями представляет собой конститутивную функцию бактериальной клетки, тогда как репарация длинными последовательностями — это процесс, индуцируемый повреждениями. Мы еще не охарактеризовали различия между этими способами репарации

по тому кругу генов, которые вовлечены в их осуществление.

Существование репарирующих систем, участвующих в синтезе ДНК, ставит вопрос о сравнимости контроля качества этого процесса с контролем процесса репликации ДНК. Насколько было известно до сих пор, большинство систем, включающих *uvr*-контролируемую репарацию с вырезанием, по частоте ошибок значительно не отличаются от репликационных систем. Однако при определенных условиях в клетках *E. coli* система репарации «склонна» ошибаться чаще, чем обычно.

Впервые способность делать ошибки была обнаружена при таких обстоятельствах. Оказалось, что, если фаг введен в клетки, предварительно обработанные УФ-светом, репарация поврежденной фаговой ДНК сопровождается индукцией мутаций. Это позволило предположить, что УФ-облучение клетки-хозяина способствовало активации какого-то белка (белков), который не функционировал в необлученных клетках и активность которого в процессе репарации порождает мутации. Мутагенный эффект имеет место также и для ДНК бактерии-хозяина.

Что представляет собой механизм появления ошибок? Можно предположить, что определенный компонент пути репарации обуславливает продолжение репликации за сайтом повреждения. Когда ДНК-полимераза минует тиминный димер, она включает неправильные основания и это приводит к появлению мутации. Существуют доказательства того, что для индукции ошибок необходимо присутствие ДНК-полимеразы III, обычной репликазы. Следовательно, рассматриваемая функция действует согласованно с нормальным репликационным аппаратом. Мутации в гене, получившем название *umuC*, устраняют УФ-индуцируемый мутагенез, но не нарушают какие-либо известные ферментативные функции. Вероятно, продукт этого гена, *umuC*, служит компонентом системы, продуцирующей ошибки.

Системы репарации *E. coli*, включающие рекомбинацию

Если в клетках с нарушением эксцизионной репарации происходит мутация в гене *recA*, она лишается всех восстановительных способностей. В результате попытки реплицировать ДНК в клетках *uvr⁻recA⁻* ведут к образованию фрагментов ДНК, размер которых соответствует ожидаемому расстоянию между тиминовыми димерами. Эти результаты указывают, что неудаленные димеры в отсутствие RecA-функции служат летальными помехами для репликации. Становится понятным, почему двойной мутант не может допустить наличия в своем геноме более чем 1–2 димеров (по сравнению со способностью бактерии дикого типа выдерживать присутствие около 50 димеров).

Влияние мутаций *recA* на жизнеспособность позволило понять реакцию клеток *E. coli* на повреждение в них ДНК. Первое объяснение сводилось к тому, что ген *recA*, вероятно, вовлечен и в другой путь. Поскольку у мутантов *recA* почти полностью нарушена генетическая рекомбинация, а также восстановление повреждений, казалось вероятным, что таким путем является репарация, включающая рекомбинацию. Данная точка зрения получила подтверждение при выделении других *rec*-мутантов с нарушением (хотя и в меньшей степени) и репарации,

и рекомбинации (см. табл. 34.2). Компоненты, включаемые в рекомбинационную репарацию, перекрываются с компонентами, участвующими в рекомбинации, но не полностью идентичны им (некоторые мутации повреждают только одну из двух активностей).

Белок RecA функционирует при обмене цепями между молекулами ДНК. Это главная активность в процессе рекомбинации (гл. 25), родственная предполагаемому одноцепочечному обмену в рекомбинационной репарации. Свойства двойных мутантов позволяют предполагать существование двух Rec-путей. Чтобы проверить, принадлежат ли два гена с родственными функциями одному и тому же или различным путям, сравнивали фенотипы одиночных мутантов с фенотипами двойных мутантов. Если гены принадлежат к одному и тому же пути, фенотип двойного мутанта не будет отличаться от фенотипа одиночного мутанта. Если гены контролируют разные пути, у двойного мутанта будут утрачены они оба. Согласно этому критерию, один *rec*-путь включает гены *recBC*, другой — *recF*.

Гены *recBC* кодируют две субъединицы экзонуклеазы V, чье действие ограничено другим компонентом пути. (В клетках *recA*-мутантов фермент вызывает чрезмерную деградацию ДНК.) Функция гена *recF* неизвестна. Благодаря эффектам мутаций на рекомбинационную репарацию или рекомбинацию были идентифицированы и другие локусы *rec*. Некоторые из них супрессируют определенные *rec*-мутации. Как и в случае дополнительных локусов *uvr*, их не удалось связать с какими-то продуктами или функциями.

Обозначения этих генов основаны на фенотипах мутантов; однако в некоторых случаях мутация, выделенная при одних условиях и названная как *uvr*, может оказаться идентичной мутации, выделенной при других условиях и отнесенной к локусу *rec*. Это очень важный факт. Мы не можем еще точно определить, сколько функций принадлежит каждому из путей или каким образом они взаимодействуют. Механизмы *uvr* и *rec* не могут быть совершенно независимыми, так как для мутантов *uvr* характерна пониженная эффективность рекомбинационной репарации. Можно ожидать выявления ряда нуклеаз, полимераз и других ферментов, составляющих репарирующие системы, которые могут быть частично перекрывающимися (или в которых один фермент, обычно используемый для обеспечения какой-то функции, может быть заменен ферментом, принадлежащим другому пути).

Прямое участие RecA-белка в рекомбинационной репарации представляет только одну из его активностей. Этот необычный белок имеет также другую, весьма важную функцию. Он является протеазой, активируемой УФ-облучением (или другими воздействиями, блокирующими репликацию). Протеазная активность ответственна за индукцию экспрессии многих генов, чьи продукты имеют и репарирующие функции. Среди них, например, можно назвать репарацию длинными последовательностями; следовательно, белок RecA необходим для индукции этой эксцизионной репарации и для участия в работе рекомбинационной репарации. Такая двойная роль белка RecA затрудняет выяснение вопроса о том, чем обусловлено нарушение репарации в мутантных клетках *recA*. Связано ли это с утратой функции белка RecA в обмене между цепями ДНК или какой-то другой функции, индукция которой зависит от протеазной активности этого белка?

SOS-репарация

Многие воздействия, которые повреждают ДНК или ингибируют ее репликацию у *E. coli*, индуцируют серию фенотипических изменений, получивших название SOS-ответа. Начало такого ответа определяется взаимодействием белка RecA с репрессором LexA. Повреждение может быть вызвано УФ-облучением (наиболее изученный случай), возникновением поперечных сшивок или алкилирующими агентами. Подавление репликации различными способами, включая тиминное голодание, добавление ядов или возникновение мутаций в некоторых из генов *dna* имеют тот же эффект. Ответ клетки, выражающийся в увеличенной способности репарировать поврежденную ДНК, достигается путем индукции синтеза компонентов системы эксцизионной репарации длинных последовательностей и Rec-зависимой репарации. Наряду с индукцией происходит подавление клеточного деления. Лизогенизирующие профаги могут быть индуцированы.

Ответ клетки на повреждающее воздействие начинается с активации протеазной активности белка RecA. О взаимосвязи между повреждающим событием и внешним изменением в RecA-активности нам известно немного. Поскольку SOS-ответ способны индуцировать разные повреждающие агенты, можно предположить, что белок RecA стимулируется каким-то общим промежуточным продуктом в метаболизме ДНК. Индуцирующим сигналом может служить небольшая молекула, освобожденная из ДНК, или какая-то структура, образуемая в самой ДНК. В системе *in vitro* протеазная активность белка RecA требует присутствия одноцепочечной ДНК и АТФ. Таким образом, активирующим сигналом может быть присутствие одноцепочечной области в сайте повреждения.

Независимо от того, какую форму имеет сигнал, его взаимодействие с белком RecA происходит очень быстро: SOS-ответ проявляется через несколько минут после повреждающей обработки. Активируясь, RecA-протеаза разрезает белок-репрессор, кодируемый геном *lexA*. Белок LexA относительно стабилен в необработанных клетках, где он функционирует как репрессор многих оперонов. Протеолитическое разрезание репрессора координированно индуцирует все эти опероны. Гены, представляющие мишени для LexA-репрессии, включают многие репарационные функции, из которых в настоящее время идентифицированы лишь некоторые. Путем слияния случайных оперонов с геном *lacZ* была разработана система, позволяющая определять LexA-зависимые генов на основании увеличения продукции β-галактозидазы после обработки клеток митомицином С. Было идентифицировано по крайней мере пять LexA-зависимых генов; они были названы *din* (от англ. *damage inducible*). Ранее известные гены, которые участвуют в SOS-ответе в результате активации их продуктов, включают *recA*, *lexA*, *uvrA*, *uvrB*, *umuC* и *himA*.

Некоторые из SOS-генов активны только в обработанных клетках; другие активны в необработанных, но уровень их экспрессии увеличивается при разрезании белка LexA. Для гена *uvrB*, который представляет собой компонент системы эксцизионной репарации, имеются два промотора; один функционирует независимо от LexA, другой подвержен его контролю. Вот почему после разрезания репрессора LexA ген может экспрессироваться при использовании обоих промоторов.

Очистка белка LexA позволила охарактеризовать его действие *in vitro*. Он репрессировал гены-мишени, связываясь с последовательностью ДНК длиной около 20 пар оснований, названной SOS-блоком. Эта последовательность симметрична, и ее копия представлена в каждом локусе-мишени. SOS-блоки различных локусов не идентичны, но во всех случаях представлены последовательностями с восемью абсолютно сохраняемыми позициями. Подобно другим операторам, SOS-блоки перекрываются с соответствующими промоторами.

RecA и LexA являются взаимными мишенями в SOS-цикле. RecA разрезает LexA, который в свою очередь репрессировал RecA. SOS-ответ вызывает амплификацию обоих белков. Рис. 34.10 иллюстрирует происходящие события; их результаты не столь противоречивы, как может показаться на первый взгляд.

Увеличение экспрессии белка RecA, по-видимому, необходимо для выполнения его роли в рекомбинационной репарации. Это означает, что в клетке имеется достаточное количество протеазы для разрезания всего дополнительного белка LexA, что необходимо для предотвращения повторной репрессии генов-мишеней. Однако основное значение этого цикла для клетки может заключаться в способности быстро восстанавливать обычное состояние.

При прекращении индуцирующего сигнала белок RecA теряет свою протеазную активность. В этот момент ген *lexA* имеет высокий уровень экспрессии; в отсутствие RecA-протеазы белки LexA быстро накапливаются в неразрезанной форме и выключают SOS-гены. Этим можно объяснить легкую обратимость SOS-ответа.

Протеаза RecA действует и на некоторые другие репрессорные белки, например репрессоры отдельных профагов (именно на примере репрессора профага лямбда и была открыта протеазная активность). Это объясняет, почему фаг лямбда индуцируется ультрафиолетовым облучением: RecA разрезает лизогенный репрессор, и фаг вступает в литический цикл развития. Указанная реакция не является клеточным SOS-ответом: она свидетельствует о способности профага узнавать, что клетка находится в опасности, и обеспечивать выживаемость вступлением в литический цикл. В этом смысле индукция профага зависит от клеточной системы, поскольку она реагирует на тот же самый сигнал (активацию RecA-белка).

Протеаза RecA разрезает все доступные ее действию белки в Ala-Gly-последовательности в середине полипептидной цепи. Мутанты типа *lexA* (Ind^-) индуцировать нельзя; один из них содержит замещающую последовательность Ala-Asp.

С каждой стороны от дипептида имеется лишь ограниченная гомология, что позволяет предполагать, что третичная структура белка может играть важную роль в узнавании мишени.

Две активности белка RecA проявляются относительно независимо. Мутация *recA441* вызывает SOS-ответ без индуцирующей обработки, вероятно, потому, что протеаза спонтанно активна. Другая мутация инактивирует протеазную активность. Ни та, ни другая мутация не влияют на способность RecA-белка взаимодействовать с ДНК. При выяснении прямого и непрямого эффектов RecA-белка в путях репарации были бы полезны мутации, инактивирующие рекомбинационную функцию, но оставляющие интактной способность индуцировать SOS-ответ.

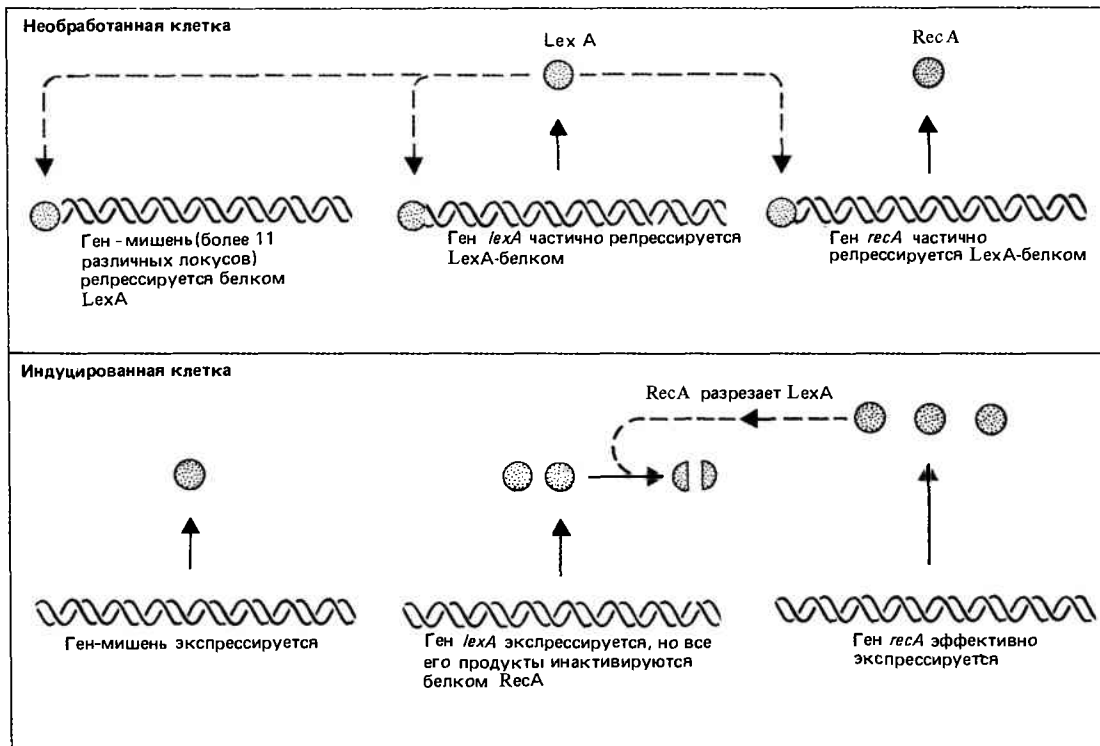


Рис. 34.10. Белок LexA репрессирует многие гены, включая те, которые кодируют функции *recA* и *lexA*. Активация RecA-про-

теазы ведет к разрезанию белка LexA и индукции всех репрессируемых генов.

Репарационные системы млекопитающих

Биохимическая характеристика репарационных систем клеток эукариот находится на начальном этапе и в основном ограничена выделением случайных ферментных препаратов, свойства которых *in vitro* позволяют предполагать, что они могут быть частью репарационной системы. Существование путей эксцизионной репарации установлено в культивируемых клетках благодаря замещению сегментов ДНК в ответ на повреждающие обработки. Один из методов связан с регистрацией незапланированного синтеза ДНК в синхронизированной культуре вне S-фазы. Его наличие может быть полностью приписано работе репарационных систем.

Клетки млекопитающих гетерогенны в отношении количества ДНК, включенной после повреждения. Однако самые длинные сегменты, обнаруживаемые в клетках млекопитающих, сравнимы по длине с короткими последовательностями, репарируемыми бактериальной системой. Такой путь репарации эффективен в отношении повреждений, индуцируемых ультрафиолетовым облучением или действием других агентов, вызывающих в ДНК родственные эффекты. Другой путь репарации способен вводить только 3–4 основания в сайты повреждений, индуцированных X-лучами или обработкой алкилирующими соединениями.

О существовании репарационных систем в клетках млекопитающих свидетельствует и анализ причин определенных наследственных болезней человека. Так, показано, что причиной заболевания *Xeroderma pigmentosum* (гиперчувствительность к солнечным лучам, в особенности к ультрафиолету) является неспособность клеток осуществлять эксцизионную репарацию, а точнее, вырезать

тиминные димеры и другие дефектные сегменты. Были идентифицированы семь генетических групп комплементации, большинство из которых характеризовались нарушением в стадии, связанной с введением разреза при репарации.

Некоторые непрямые данные позволяют предположить, что клетки млекопитающих имеют и системы рекомбинационной репарации. Вероятно, и в клетках эукариот эти системы могут быть связаны с процессом генетической рекомбинации. Например, в случае рецессивной болезни человека, синдрома Блума, возросшая частота хромосомных aberrаций, включая обмены между сестринскими хроматидами, может быть связана с нарушением рекомбинационных систем.

Рекомендуемая литература

Следует обратиться к обзору Модриха (*Modrich, Quart. Rev. Biophys.*, **3**, 315–369, 1979), в котором рассматриваются отдельные активности рестрикции и модификации (ферменты типа II). Ценные сведения о мультифункциональных ферментах (типа I и типа III) приведены в обзоре Юана (*Yuan, Ann. Rev. Biochem.*, **50**, 285–315, 1981), разработавшего модель транслокации фермента типа I (*Yuan, Cell*, **20**, 237–244, 1980). Функции AP-эндонуклеаз рассмотрены Линдэлом (*Lindahl, Prog. Nucleic Acid Res.*, **22**, 135–192, 1979). Подробный обзор систем и механизмов репарации ДНК у прокариот и эукариот представлен Ханаволтом (*Hanawalt, Ann. Rev. Biochem.*, **48**, 783–836, 1979). Систему SOS-контроля у *E. coli* проанализировал Готтесман (*Gottesman, Cell*, **23**, 1–2, 1981) и всесторонне обсудили Литл и Маунт (*Little, Mount, Cell*, **29**, 11–22, 1982).

Глава 35

ВОССТАНОВЛЕНИЕ И РЕКОМБИНАЦИЯ ДНК

В отсутствие рекомбинации генетический материал каждой хромосомы был бы фиксирован в ее аллелях. Единственным источником изменчивости в этом случае служили бы мутации. Протяженность мишени для мутационных повреждений увеличилась бы от одного гена до целой хромосомы. Накопление вредных изменений в отдельной хромосоме приводило бы к ее элиминации вместе с присутствующими в ней полезными мутациями. Однако рекомбинация существует, в результате чего в хромосомах происходит перераспределение генов, полезные мутации отделяются от вредных и проверяются в новых сочетаниях. Таким образом, генетическая рекомбинация способствует спасению и распространению полезных и элиминации вредных аллелей. С эволюционной точки зрения хромосома – это непостоянная структура, образуемая временно связанными определенными аллелями. Такое непостоянство обусловлено рекомбинацией.

Тип рекомбинации, включающий взаимодействие между гомологичными последовательностями ДНК, получил название **общей рекомбинации**. Ферменты, ответственные за ее осуществление, могут использовать в качестве субстратов *любую* пару гомологичных последовательностей (хотя возможность предпочтительного использования определенных последовательностей существует всегда). Данные о поведении кластеров генов (гл. 21) свидетельствуют о том, что гомология может быть неполной; родственные неаллельные гены способны рекомбинировать с частотой, которая тем меньше, чем больше степень дивергенции между последовательностями этих генов (неравная рекомбинация).

Для генома в целом частота рекомбинаций не является постоянной. На нее может влиять структура хромосомы; например, близость гетерохроматина подавляет кроссинговер.

Рекомбинации между *специфическими* парами последовательностей способствует другой тип событий. Такая **сайтспецифическая рекомбинация** обуславливает интеграцию фаговых геномов в бактериальную хромосому. В рекомбинационное событие вовлекаются специфические последовательности фаговой и бактериальной ДНК; в пределах этих последовательностей имеется короткая область гомологии, присутствие которой хотя и необходимо, но недостаточно для осуществления рекомбинации. Ферменты, участвующие в этом процессе, действуют на определенной паре фаговой и бактериальной последовательностей и не способны вести рекомбинацию между другими парами. Такой тип рекомбинации также называется **консервативной рекомбинацией**, так как он всецело определяется реакцией двух предсуществующих молекул ДНК.

Третий тип рекомбинации дает возможность одной последовательности ДНК встраиваться в другую независимо от гомологии между ними. **Репликативная рекомбинация** обуславливает способность определенных элементов перемещаться в геноме. Подвижный элемент копи-

руется, и затем реплика переносится в какой-то другой сайт. Детально такой тип рекомбинации рассматривается в гл. 36.

Для осуществления рекомбинации необходим синапсис гомологичных молекул ДНК

Давно известно, что для рекомбинации между генами необходим физический обмен частей хромосомы (см. рис. 1.10). Образующую в результате такого обмена структуру (хиазму) можно визуальное наблюдать в мейозе (см. рис. 1.9). Она формируется в результате разрыва и воссоединения двух несестринских хроматид (каждая из которых содержит двухцепочечную ДНК). Разрез и последующее соединение происходят между точно соответствующими последовательностями, в результате чего ни одна пара оснований из рекомбинантных хромосом не теряется и ни одна пара к ним не добавляется.

Первая решающая стадия рекомбинации заключается в том, что две гомологичные двухцепочечные молекулы ДНК входят в тесный контакт; благодаря этому становится возможным обмен соответствующими последовательностями. У эукариот и прокариот этот процесс может в какой-то мере отличаться.

В прокариотических системах мы имеем дело главным образом с самой ДНК (хотя и в составе нуклеоида). Механизм узнавания гомологичных участков пока не выявлен. Можно полагать, что в этих событиях участвуют только те области, которые действительно вовлекаются в рекомбинационный обмен. По-видимому, узнавание является неотъемлемой частью рекомбинационного механизма. Оно может основываться на сравнении отдельных цепей ДНК с целью идентификации комплементарных последовательностей (см. ниже).

В клетках эукариот ДНК упакована в дискретные структуры – хромосомы. Каким образом молекулы ДНК приходят в непосредственное соприкосновение? Контакт между парами родительских хромосом осуществляется в раннем мейозе. Процесс назван **синапсисом** или **спариванием хромосом**. Гомологичные хромосомы (каждая из которых состоит из двух сестринских хроматид, образуемых предшествующей репликацией) приближаются друг к другу. Они оказываются латерально связанными в форме синаптонемального комплекса, который у каждого вида имеет характерную структуру. Детали этой структуры могут существенно отличаться.

Пример синаптонемального комплекса представлен на **рис. 35.1**. Каждая хромосома на этой стадии состоит из хроматина, связанного **латеральными элементами** (которые в данном случае имеют исчерченную структуру). Два латеральных элемента отделены друг от друга **центральным элементом**. Три параллельные плотные цепи ле-



Рис. 35.1. Структура синаптонемального комплекса *Neotellia* свидетельствует о том, что спаривание хромосом не вызывает их наложения друг на друга.

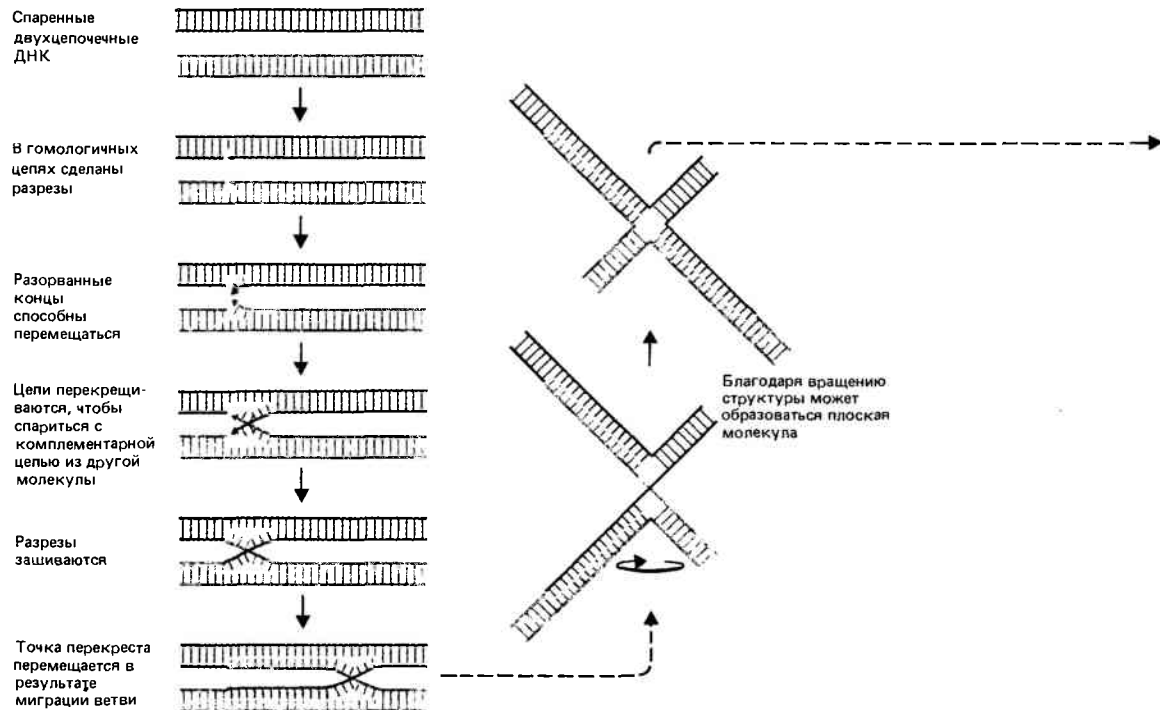
Хр – хроматин, ЛЭ – латеральные элементы, имеющие примерно 50 нм в диаметре; ЦЭ – центральный элемент диаметром около 18 нм. Расстояние между латеральными элементами равно примерно 120 нм. (Фотография любезно предоставлена д-ром M. Westergaard и д-ром D. von Wettstein.)

жат в одной плоскости, извиваясь и закручиваясь вдоль своей оси. Гомологичные хромосомы значительно удалены друг от друга: расстояние между ними достигает более 200 нм (диаметр кольцевой ДНК 2 нм).

Образование синаптонемального комплекса совпадает с предполагаемым временем кроссинговера, хотя прямых доказательств, что рекомбинационное событие происходит именно на стадии синапсиса, нет. Основная проблема заключается в том, что на этой стадии осуществляется контакт между гомологичными хромосомами, а не гомологичными молекулами ДНК. Только у грибов и насекомых компоненты синаптонемального комплекса связаны между собой сферическими или цилиндрическими структурами. Они лежат поперек комплекса и носят на-

звание узелков или **рекомбинационных утолщений**; их частота и распределение соответствуют частоте и распределению хиазм. Название этих узлов отражает предположение, что именно они представляют собой сайты рекомбинации.

На следующей стадии мейоза хромосомы теряют синаптонемальный комплекс; хиазмы становятся видимыми как точки, в которых связаны хромосомы. Вероятно, их наличие указывает на осуществление генетического обмена. На более поздних стадиях мейоза хиазмы могут перемещаться к концам хромосом. Такая подвижность говорит о том, что хиазмы являются скорее следствием рекомбинационного события, чем его промежуточным продуктом.



Разрыв и воссоединение осуществляются через гетеродуплексные ДНК

До сих пор не ясно, каким образом узнают друг друга гомологичные двухцепочечные последовательности ДНК, однако механизм узнавания комплементарности отдельных цепей хорошо известен—это спаривание оснований. Он используется для образования рекомбинационных промежуточных структур. Одна из нескольких существующих моделей этой реакции представлена на рис. 35.2.

Процесс начинается с разрыва в соответствующих точках гомологичных цепей двух спаренных молекул ДНК. Разрыв обуславливает подвижность образовавшихся свободных концов. Каждая цепь оставляет своего партнера; происходит перекрещивание. При этом каждая цепь спаривается с комплементарной ей цепью другой молекулы. Такой реципрокный обмен связывает две двухцепочечные ДНК. Вначале это только водородные связи; затем благодаря сшивки разрезов в сайтах обмена в какой-то точке устанавливается ковалентная связь (сшивка может произойти и позднее, чем показано на рисунке). Пара соединенных двухцепочечных молекул получила название **сцепленной молекулы**.

Каждая двухцепочечная ДНК в сцепленной молекуле имеет область, в которой из разделенных родительских молекул ДНК присутствует по одной цепи. Такая область получила название **гибридной ДНК** или **гетеродуплексной ДНК**. Модель показывает, что существует небольшая стерическая помеха для образования реципрокных гетеродуплексных областей между спаренными молекулами ДНК; так практически все основания могут сохраняться спаренными (либо в родительской молекуле, либо в гетеродуплексных областях). Существование протяженных одноцепочечных областей не обязательно.

Если в клетке произошла инициация генетического обмена, он продолжается, перемещаясь вдоль цепей ДНК. Такое явление получило название **миграции ветви**. Ветви в структуре ДНК могут включать либо отдельные, либо двойные цепи. На рис. 35.3 показан пример миграции одноцепочечной ветви. Точка ответвления может мигри-

ровать в любом направлении, когда одна цепь замещается другой. Эта реакция важна с теоретической и практической точек зрения. Она придает динамичность рекомбинантным структурам. Однако в практическом отношении возможность миграции в любом направлении означает, что точку ответвления нельзя обнаружить *in vitro* (поскольку ветвь может мигрировать уже после того, как молекула была выделена).

Такой тип движения позволяет точке обмена в рекомбинационной промежуточной структуре перемещаться в любом направлении. Чтобы в естественных условиях поддерживалось образование протяженных областей гетеродуплексной ДНК (более 1000 пар оснований), скорость миграции ветви должна быть, по-видимому, достаточно высокой. При вовлечении в реакцию двухцепочечных молекул могут оказаться необходимыми топологические перестройки: либо спираль ДНК должна иметь возможность свободно вращаться, либо с помощью топоизомеразы должны сниматься топологические ограничения.

Сцепленная молекула, образуемая при обмене цепями, должна разделиться на две отдельные двухцепочечные молекулы. Для этого необходимы два дополнительных разреза. Альтернативные способы такой реакции рассмотрены в нижней части рис. 35.2 на примере плоскостной молекулы, образуемой при вращении одной из двойных цепей рекомбинационного интермедиата. Конечный результат зависит от того, какая именно пара цепей разрезается.

В тех случаях, когда разрезы вносятся в ту пару цепей, которая исходно не разрезалась (т. е. в пару, не принимавшую участия в инициации обмена), каждая из четырех исходных цепей окажется разрезанной. Это приведет к разделению рекомбинационных молекул ДНК. Одна двухцепочечная родительская ДНК ковалентно связывается с другой родительской молекулой через отрезок гетеродуплексной ДНК. Между маркерами, расположенными с той или другой стороны от гетеродуплексной области, осуществляется рекомбинационное событие.

Если разрезанию подвергаются цепи, уже разрезанные в начале реакции, две другие цепи остаются интактными. Разрезы ведут к освобождению исходных интактных ро-

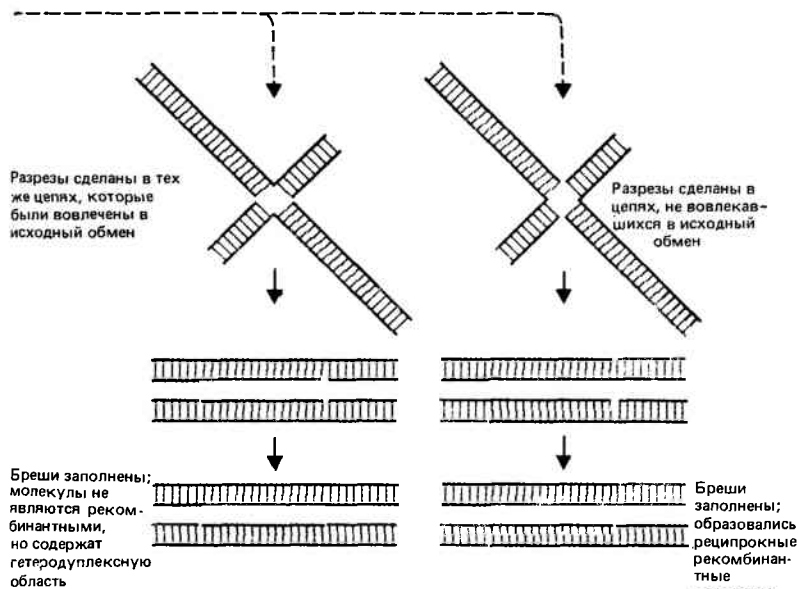


Рис. 35.2 Рекомбинация между двумя спаренными двухцепочечными молекулами ДНК может быть иницирована реципрокным одноцепочечным обменом, расширяющимся в результате миграции ветви и разрешающимся путем разрезания.

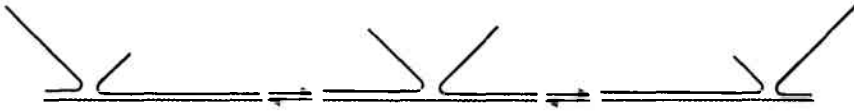


Рис. 35.3. Миграция ветви может происходить в любом направлении, когда неспаренная одиночная цепь вытесняется спаренной цепью.

Такая структура формируется при ренатурации, когда одиночная цепь ДНК отжигается в одном конце с одной комплементарной цепью, а в другом конце — с независимой комплементарной цепью.

дательских молекул, каждая из которых, однако, имеет отрезок гетеродуплексной ДНК.

Таким образом, основной принцип альтернативных способов разделения сцеплений молекулы заключается в образовании области гетеродуплексной ДНК при обмене цепями между молекулами ДНК, причем такой обмен не обязательно сопровождается рекомбинацией фланкирующих областей.

Можно рассмотреть несколько вариантов использования гетеродуплексной ДНК. Обмен цепей может быть инициирован только одной цепью вместо двух; миграция ветви может способствовать расширению реакции, в результате чего произойдет вытеснение другой цепи. Реакция может происходить путем раскручивания и крестообразного спаривания цепей без какого-либо разрезания (хотя это связано с гораздо большими топологическими перестройками). В модели, представленной на рис. 35.2, гетеродуплексные участки в каждой молекуле ДНК имеют равную длину. Генетические исследования, проведенные на грибах, показали, что области гетеродуплексной ДНК в рекомбинантных хромосомах варьируют по длине (см. ниже). Этот факт согласуется с такой модификацией модели, при которой допускается синтез новой последовательности ДНК, замещающей часть последовательности одной из обменивающихся цепей. Следует отметить, что детали механизма и порядок событий могут различаться, однако все модели основаны на одном и том же основном принципе: образовании промежуточной структуры, включающей гетеродуплексную ДНК, которая может быть расширена посредством миграции ветви.

Действительно ли двухцепочечные разрывы инициируют рекомбинацию?

Другая модель рекомбинации основана на предположении, что *двухцепочечный разрыв инициирует генетический обмен*. Одна из образовавшихся одиночных цепей мигрирует к другому дуплексу, в результате чего две двухцепочечные молекулы оказываются связанными с помощью гетеродуплексной последовательности ДНК.

Эндонуклеаза производит двухцепочечный разрыв в одной хроматиде — «реципиенте». Разрыв увеличивается до размера бреши, по-видимому, за счет действия экзонуклеаз. Экзонуклеаза (экзонуклеазы) разрушает одну из цепей с одной стороны от разреза, в результате чего образуется одноцепочечный 3'-конец. Свободный 3'-конец вторгается в гомологичную область другого, «донорного» дуплекса. Образование гетеродуплексной ДНК сопровождается вытеснением D-петли, в которой одна цепь донорного дуплекса вытесняется. Репарирующий синтез расширяет D-петлю; при этом в качестве затравки используется свободный 3'-конец. В конечном счете D-петля достигает размера, соответствующего полной длине бреши реципиентной хроматиды. При достижении вытесняемой отдельной цепью удаленного конца пробела, компле-

ментарные одноцепочечные последовательности «отжигаются». В результате с обеих сторон от бреши находится гетеродуплексная ДНК, а сама брешь представлена одноцепочечной D-петлей.

Дуплексная целостность области бреши может быть восстановлена с помощью репарирующего синтеза, использующего в качестве затравки ее левый 3'-конец. Во всех случаях брешь репарируется в результате двух отдельных циклов синтеза одноцепочечной ДНК.

Миграция ветви превращает эту структуру в молекулу с двумя рекомбинантными соединениями. В тех случаях, когда оба соединения разрешаются одним и тем же способом, например разрезанием внутренних цепей, происходит высвобождение исходных, не подвергавшихся кроссинговеру молекул. Если же два соединения разрешаются противоположным способом, т.е. в одном случае разрез вносится на внутреннюю цепь, а в другом — на внешнюю, происходит генетический кроссинговер.

Модели с двухцепочечным и одноцепочечным разрывами предполагают некоторые различия в структуре молекулы, сочлененной в двух местах перед ее разрешением. Гетеродуплексная ДНК образуется в каждом конце области, вовлекаемой в обмен. Между двумя гетеродуплексными сегментами находится область, соответствующая бреши, которая теперь содержит последовательность донорной ДНК в обеих молекулах. Таким образом, гетеродуплексные последовательности расположены асимметрично и часть одной молекулы превращена в последовательность другой (вот почему инициирующая хроматида названа реципиентом). Модель одноцепочечного разрыва предсказывает противоположный результат, при котором инициирующая хроматида выполняет роль донора генетической информации. Данные, полученные на клетках дрожжей, согласуются с моделью двухцепочечного разрыва: инициация сопряжена с получением генетической информации. Эта модель предусматривает существование родства между сайтами, участвующими в событии генетического обмена, и подтверждается данными, полученными при анализе систем рекомбинации грибов.

О правомочности этой модели свидетельствует тот факт, что у дрожжей двухцепочечные разрывы участвуют в определенных рекомбинационных событиях. При введении плазмиды с двухцепочечным разрывом в клетки дрожжей наблюдается стимуляция рекомбинации и, кроме того, могут образовываться продукты рекомбинации с репарированной брешью.

Модель двухцепочечного разрыва не умаляет важности образования гетеродуплексной ДНК, которая вероятно представляет собой единственный реальный способ взаимодействия двух дуплексных молекул. Однако, возлагая ответственность за инициацию рекомбинации на двухцепочечные разрывы, мы должны изменить наши представления о способности клетки манипулировать с ДНК.

На первый взгляд участие двухцепочечных разрывов в инициации рекомбинации кажется удивительным. Ведь

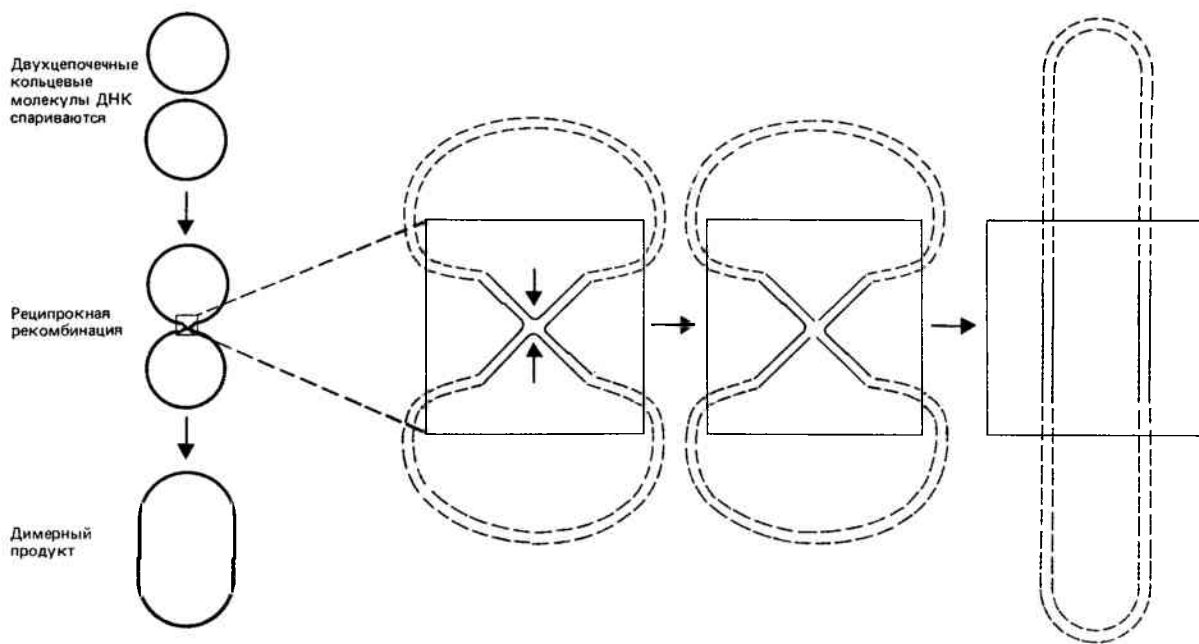


Рис. 35.4. Реципрокная рекомбинация между гомологичными двухцепочечными кольцевыми молекулами приводит к образо-

ванию димерного кольца. Промежуточный продукт имеет структуру, напоминающую по форме цифру восемь.

поперечный разрез молекулы ДНК приводит к необратимым последствиям. Согласно модели одноцепочечного обмена ни на одной из стадий не может быть потеряна никакая информация. В соответствии с моделью двухцепочечного разрыва первый же разрыв немедленно сопровождается потерей информации; при этом фатальной может оказаться любая ошибка при восстановлении. С другой стороны, именно способность восстанавливать утраченную информацию путем ресинтеза ее с использованием другого дуплекса обеспечивает основную систему безопасности для клетки.

Выделение промежуточных продуктов рекомбинации

Свойства кольцевых молекул ДНК могут быть использованы при разработке подходов к выделению промежуточных рекомбинационных структур. До сих пор мы рассматривали рекомбинирующие двухцепочечные ДНК как линейные молекулы, хотя многие геномы (особенно вирусные и плазмидные) имеют кольцевую структуру. На рис. 35.4 изображены последствия реципрокной рекомбинации между гомологичными сайтами на двух кольцевых молекулах ДНК. Увеличенная часть рисунка показывает, что в месте соединения эта структура не отличается от той, которая изображена на рис. 35.2, за исключением того, что концы каждой двухцепочечной родительской ДНК соединены. Этот факт имеет важные топологические последствия. Разделение такой структуры при рекомбинации ведет к образованию димерного кольца, содержащего две исходные родительские последовательности, соединенные по типу «голова к хвосту».

Какое это имеет значение для жизненного цикла кольцевых ДНК? Если они вступают в рекомбинацию, исходные (мономерные) кольцевые геномы могут сохраниться только благодаря парным рекомбинационным событиям. Одно реципрокное рекомбинационное событие всегда порождает димер, изображенный на рис. 35.4. Та-

кой димер может восстановить мономерную структуру благодаря второму рекомбинационному событию, вовлекающему любую пару гомологичных последовательностей (рис. 35.5). Поэтому в любой популяции кольцевых геномов в момент рекомбинации можно обнаружить мультимерные кольца.

В результате рекомбинации кольцевых молекул ДНК образуется структура, имеющая форму восьмерки (рис. 35.4). Димерные молекулы такой формы были обнаружены у нескольких фагов в инфицированных клетках, а также в популяции плазмидной ДНК у дрожжей. Однако внешний вид этих структур не позволяет с определенностью идентифицировать их как рекомбинационные ин-



Рис. 35.5. При реципрокной рекомбинации между гомологичными последовательностями димерных колец образуются мономерные кольца.

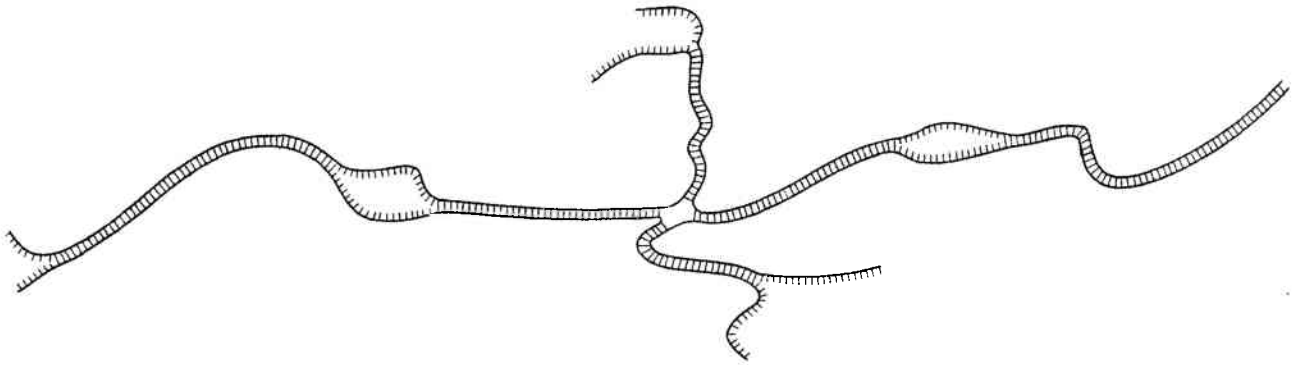


Рис. 35.6. Разрезание промежуточной структуры, имеющей форму восьмерки в гомологичных сайтах на каждом кольце приводит к образованию χ -подобной структуры, у которой

четыре двухцепочечных отростка остаются объединенными в сайте обмена цепи.

(Фотография любезно предоставлена D. Dressler.)

термедиаты; например, вид восьмерки могут иметь взаимосцепленные (катенаны) мономерные кольца.

Природу «восьмерки» можно определить, если подействовать на нее рестрикционным ферментом, разрезающим каждое мономерное кольцо только в одном сайте. Если две части молекулы, имеющей вид восьмерки, не связаны ковалентно, они разделятся. Если же «восьмерка» — результат рекомбинации, ее разрезание приведет к появлению структуры, в которой родительские молекулы удерживаются вместе областью гетеродуплексной ДНК в точке слияния. Разрезанная молекула имеет четыре ответвления, причем каждая пара ответвлений соответствует каждому из исходных мономерных колец. В разрезанной форме молекула получила название *chi*-структуры, поскольку она напоминает греческую букву χ . Ее микрофотография представлена на рис. 35.6 вместе с поясняющей схемой. Четыре ответвления двухцепочечных ДНК соединены областью, в которой гетеродуплексные сегменты были разорваны на составляющие их отдельные цепи (по-видимому, в результате приготовления препарата для электронной микроскопии). Такая структура точно соответствует предсказанной (рис. 35.2 и 35.4); ее выделение является прямым доказательством существования гетеродуплексной ДНК *in vivo*.

Chi-структуры найдены только в клетках дикого типа, но не обнаружены в мутантах *recA*⁻. Следовательно, их существование коррелирует со способностью клетки к генетической рекомбинации. Этот факт еще раз подтверждает, что *chi*-структуры не случайные примеси, а имеют прямое отношение к генетической рекомбинации.

Данные наблюдения можно использовать для разработки методов очистки рекомбинационных ферментов. Бесклеточные экстракты проверяют на способность формировать димерные молекулы из мономерных колец. Экстракты из клеток дикого типа выполняют эту реакцию; в них образуются молекулы типа восьмерки, которые далее могут быть разрезаны с образованием *chi*-структур. Экстракты из клеток *recA*⁻ в этом отношении неактивны. *RecA*-систему можно фракционировать и использовать компоненты в комплементационных экспериментах, подобно тому как это делается для изучения репликации ДНК (гл. 33).

Обмен между цепями, осуществляемый при участии белка *RecA*

К сожалению, о генетическом аппарате, осуществляющем общую рекомбинацию, известно мало. Охарактеризованы ферменты, участвующие в каких-то специализированных рекомбинационных событиях; описаны некоторые связанные с ними топологические изменения. Однако все это лишь доказало сложность изучения гомологичной рекомбинации.

Другой подход к исследованию свойств ферментов, вовлекаемых в генетическую рекомбинацию, — получение *rec*⁻-мутаций в их генах. Малочисленность известных продуктов свидетельствует о том, что до сих пор этот подход использовали главным образом для получения

данных о способности белка RecA участвовать в обмене цепей ДНК.

Точно неизвестно, как ферментативные активности RecA связаны с рекомбинацией *in vivo*, однако показано, что они участвуют в нескольких реакциях, включенных в рекомбинационный механизм. Мы уже упоминали, что для проявления протеазной активности RecA-белка необходимы одноцепочечная ДНК и АТР (гл. 34). Те же субстраты требуются для осуществления его манипуляций с молекулами ДНК.

Белок RecA стимулирует спаривание оснований между отдельной цепью ДНК и комплементарной ей цепью в суперспиральной кольцевой молекуле. Одна цепь атакует кольцо, вытесняет исходную и спаривается с комплементарной цепью, в результате чего образуется D-петля (рис. 35.7, *верхний ряд*). Эта реакция получила название **поглощения** или **ассимиляции одной цепи**. Процесс ассимиляции одной цепи родствен образованию гетеродуплекса, модель которого представлена на рис. 35.2. Разница состоит в том, что рекомбинационная модель предполагает два вторжения. В каждом случае одна цепь реципрокно вытесняет своего гомолога в двухцепочечной молекуле-партнере.

Белок RecA способен связываться как с одноцепочечной ДНК (действие стимулируется АТР), так и с двухцепочечной (действие требует присутствия АТР). Реакция ассимиляции требует гидролиза АТР. Ассимиляция происходит в оптимальном режиме при наличии одного мономера АТР на каждые 5–10 нуклеотидов одноцепочечной ДНК. Итак, реакция осуществляется при связывании RecA-белка с одноцепочечной ДНК. Нам еще непонятна последовательность событий, предшествующих образованию D-петли, однако можно предположить, что RecA-белок сначала связывается с одноцепочечной, а затем с двухцепочечной ДНК в поисках комплементарной одноцепочечной области. Основной вопрос состоит в том, каким образом происходит плавление двухцепочечной ДНК, позволяющее вторгающейся цепи найти комплементарный участок.

В реакции образования D-петли вытеснение гомологичной кольцевой цепи происходит частично; структура ковалентно замкнутого двухцепочечного кольца может служить топологической помехой распространению реакции. Реакция будет более эффективной, если поставить структуры двух- и одноцепочечной молекул в обратном порядке. Кольцевая одноцепочечная молекула может взаимодействовать с гомологичной двухцепочечной, вытесняя полностью одну цепь. Схема реакции представлена на рис. 35.7 (*средний ряд*). Она начинается в одном конце молекулы и распространяется к другому ее концу, образуя гетеродуплекс длиной до нескольких тысяч пар оснований.

Такая пара субстратов была использована для выяснения некоторых вопросов, касающихся механизмов реакции. Происходит ли она в каком-то определенном направлении? На рис. 35.8 показано, что это можно проверить, используя линейную одноцепочечную ДНК, которая комплементарна одной цепи только в одном конце. Такой субстрат получают при добавлении чужеродного фрагмента ДНК к одному концу линейной молекулы. Если фрагмент локализован в 5'-конце комплементарной цепи, т.е. той, которая должна спариваться с вторгающейся цепью, реакция протекает вполне нормально (рис. 35.8, *верхний ряд*). Однако, если чужеродную ДНК содержит 3'-конец комплементарной цепи, реакция не происходит (рис. 35.8, *средний ряд*). Следовательно, ассимиляция одной цепи может происходить только в фиксированном направлении, от 3'-конца комплементарной цепи к 5'-концу, или (относительно вторгающейся цепи) от 5'-конца к 3'-концу.

Что происходит, если две последовательности только частично комплементарны и не соответствуют одна другой полностью? Какая степень дивергенции может допускаться при работе RecA-белка? В том случае, когда чу-

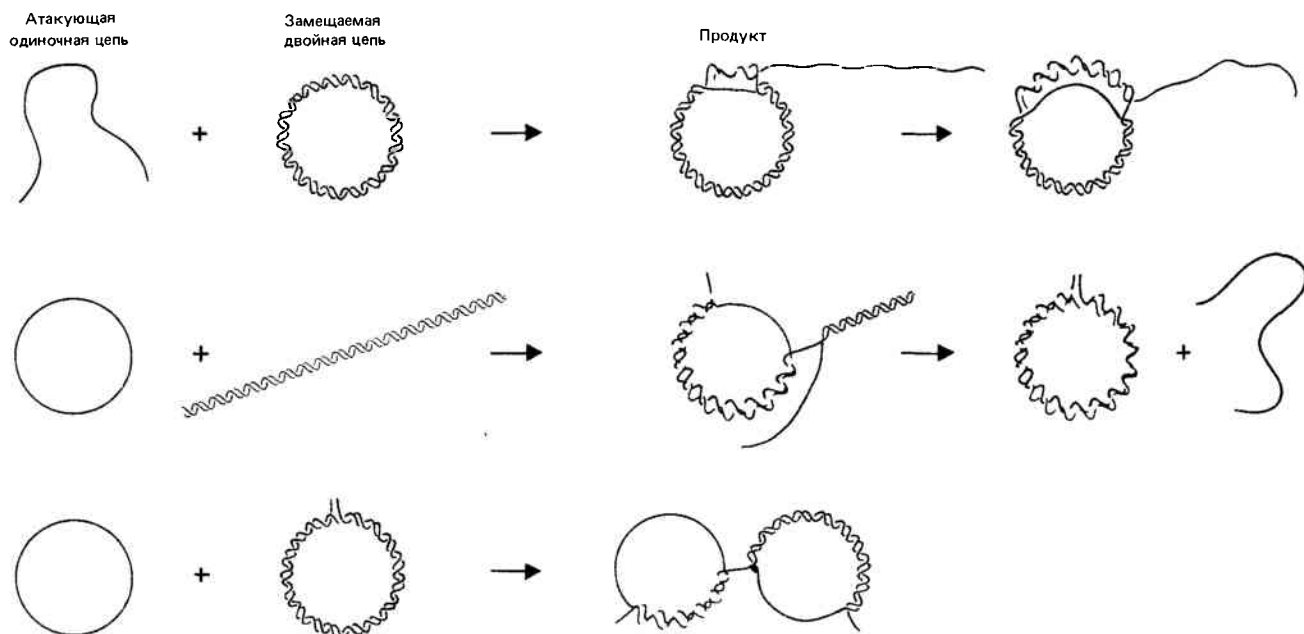


Рис. 35.7. RecA-белок стимулирует ассоциацию атакующих одиночных цепей молекулы ДНК до тех пор, пока одна из взаимодействующих цепей имеет свободный конец.

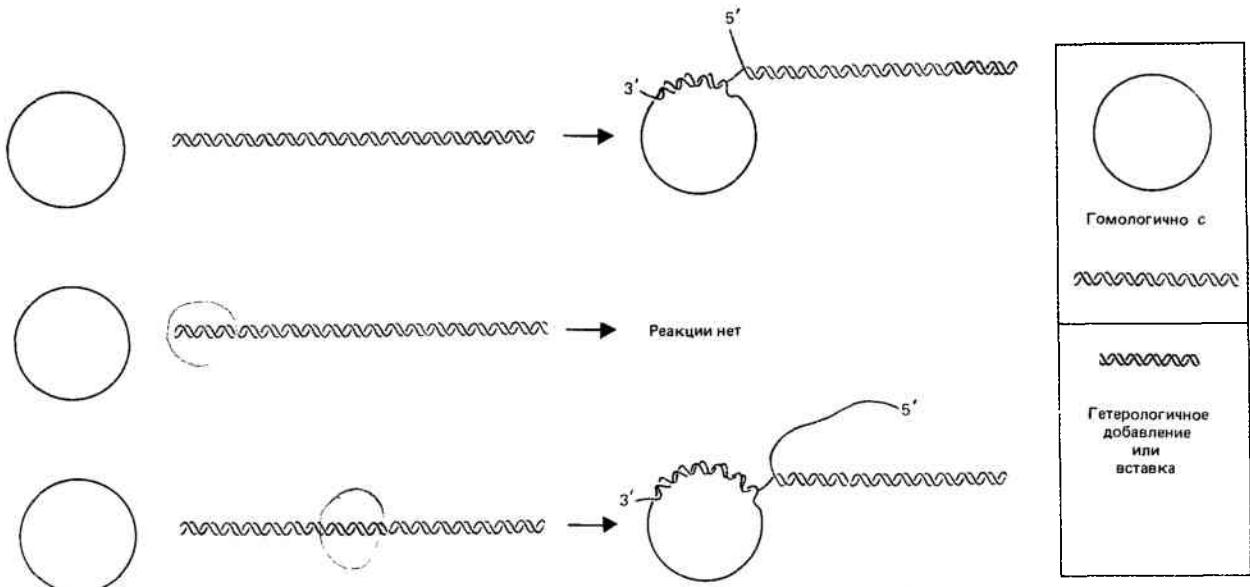


Рис. 35.8. Белок RecA способен стимулировать ассимиляцию кольцевой одиночной цепи молекулы ДНК только в том случае, если существует комплементарный 3'-конец, с которым одиноч-

ная цепь может спариваться. Если в составе ДНК присутствует чужеродная последовательность, реакция ассимиляции останавливается в конце комплементарного участка.

жеродная ДНК внедрена в центральную часть двойной цепи ДНК (рис. 35.8, *нижний ряд*), реакция ассимиляции цепи останавливается на границе между ними. Следовательно, ее продолжение, как и инициация, зависит от спаривания комплементарных оснований. Однако ассимиляция цепи может развиваться при умеренной степени несоответствия цепей. Например, ДНК фагов fd и M13 имеют примерно 97% гомологии и могут вместе подвергаться ассимиляции цепи. В то же время фаги фХ174 и G4, ДНК которых имеет 70% гомологии, не могут вступать в эту реакцию. Максимум допустимой дивергенции еще не установлен.

Белок RecA и условия рекомбинации

Мы знаем, что белок RecA необходим для того, чтобы иницировать ассимиляцию одной цепи, однако существование спонтанной миграции ветвей ставит вопрос о том, требуется ли этот белок впоследствии. В присутствии белка SSB (связывающегося с одной цепью) RecA-белок необходим для продолжения реакции, а также инициации образования гетеродуплексной ДНК. Для расширения реакции требуется также гидролиз АТР.

В основе механизма активации RecA-белка, стимулирующего миграцию ветви, лежит его способность формировать нити. Этот белок может образовывать длинные нити с одноцепочечной ДНК. В присутствии нуклеотида-кофактора он способен полимеризоваться сам по себе или включать в нити двухцепочечные ДНК. Такие нити по длине превышают исходную молекулу ДНК, что, возможно, свидетельствует о ее раскручивании. Нити из RecA-молекул не образуются в присутствии АТР, однако могут сформироваться в присутствии негидролизуемых аналогов.

Вероятно, молекулы RecA-белка способствуют расширению реакции ассимиляции цепи путем кооперативного связывания с ДНК с образованием структуры, родственной нити. Такая структура может быть разрушена при гидролизе АТР. Эта модель объясняет требования реак-

ции в отношении стехиометрии между RecA-белком и одноцепочечной ДНК. Присутствие белка SSB стимулирует как ассимиляцию одной цепи, так и образование нитей, возможно, посредством кооперирования с RecA-белком и создания соответствующего остова для гетеродуплексной ДНК.

Ассимиляция одной цепи требует наличия свободного конца ДНК. Как показывает рис. 35.7, свободный конец может находиться на атакующей цепи (*верхний ряд*) или на двухцепочечной молекуле (*средний ряд*). Если обе молекулы кольцевые, двухцепочечное кольцо должно иметь разрыв цепи, которая должна быть замещена (*нижний ряд*). Необходимость в свободном конце может отражать свойство RecA-белка или обуславливаться топологическими помехами. Отдельная цепь со свободным концом будет закручиваться вокруг комплементарной ей цепи, но, если свободный конец отсутствует, молекулы довольно быстро прекратят дальнейшее спаривание.

Если одноцепочечная молекула взаимодействует с двухцепочечной, происходит раскручивание дуплекса в области рекомбинации. Образующийся участок гетеродуплексной ДНК может даже не иметь традиционную форму двойной спирали. Он может состоять из двух цепей, связанных друг с другом бок о бок. Область такого типа получила название **области паранемного соединения** (по сравнению с классическим **плектономным** взаимодействием цепей в двойной спирали).

Паранемное соединение нестабильно. Дальнейший ход реакции требует его перехода в форму двойной спирали. Это событие эквивалентно удалению отрицательных сверхвитков и требует участия топоизомеразы I. Топоизомераза, по-видимому, осуществляет раскручивание/закручивание путем введения временных разрывов, которые позволяют одной цепи вращаться вокруг другой. Однако рекомбинация без исходных разрывов цепей обнаружена и между неспаренными или перекрестно спаренными цепями (см. рис. 35.2), и это запутывает и без того сложный вопрос об участии топоизомераз в рекомбинационном процессе.

Важность топологии молекул доказывается способностью одноцепочечного кольца атаковать ковалентно замкнутое суперспиральное двухцепочечное кольцо только в присутствии топоизомеразы I. По-видимому, топоизомераза решает проблему раскручивания (повторного закручивания), производя временные разрезы, которые устраняют необходимость в свободном конце. Этот факт свидетельствует о том, что рекомбинация может происходить при неспаренности или перекрестной спаренности без первичных разрывов цепей, показанных на рис. 35.2.

Все обсуждавшиеся до сих пор реакции представляют только часть потенциального рекомбинационного события: атаку одной цепью двухцепочечной молекулы. При участии RecA-белка могут взаимодействовать и две двухцепочечные молекулы при условии, что одна из них содержит одноцепочечную область по крайней мере из 50 оснований. Такая область может иметь форму хвоста на линейной молекуле или бреши в кольцевой молекуле.

Реакция между частично двухцепочечной и полностью двухцепочечной молекулами ведет к обмену цепями (рис. 35.9). Ассимиляция начинается в одном конце линейной молекулы, где атакующая цепь вытесняет своего гомолога в двойной цепи обычным способом. Когда реакция достигает области, в которой обе молекулы двухцепочечные, атакующая цепь не находит пары в своем партнере, который спаривается с другой вытесненной цепью.

На этой стадии молекула имеет структуру, которая неотличима от рекомбинационного места стыковки, изображенного на рис. 35.2. Если взаимодействующие молекулы кольцевые, продукт может выглядеть как *chi*-структура. Выделение таких продуктов реакции, стимулируемых *in vitro* белком RecA, демонстрирует способность этого белка обеспечивать реципрокный перенос цепей.

Для нормального функционирования белка RecA необходимо, чтобы в одной из взаимодействующих молекул ДНК имелась одноцепочечная область. Каким образом происходит ее образование *in vivo*? Одна из гипотез допускает ограниченную деградацию одной цепи ДНК, другая — существование какой-то рекомбинационной активности, раскручивающей ДНК. Вполне вероятно, что

в этот процесс может вовлекаться продукт RecBC. В присутствии SSB-белка этот фермент, исходно идентифицированный как сильно действующая нуклеаза, также способен осуществлять раскручивание ДНК с образованием одноцепочечных областей.

Белок, напоминающий по своим свойствам RecA, был выделен из клеток гриба *Ustilago maydis*. Этот белок (Rec1) способен приводить гомологичные молекулы двухцепочечной ДНК в состояние синапсиса тем же (зависящим от одной цепи) способом, что и белок RecA. В присутствии топоизомеразы I типа молекулы ДНК, находящиеся в состоянии синапса, могут быть связаны.

До сих пор мы обсуждали действие RecA-белка в контексте независимой последовательности ДНК. Однако существуют горячие точки, которые стимулируют рекомбинационную RecA-систему. К их открытию у фага лямбда исследователи пришли до некоторой степени окольным путем.

Созревание ДНК фага лямбда требует образования конкатемерных молекул. Обычно их появление связано с изменением в способе репликации, от первоначального типа с θ -структурой к типу «катящегося кольца». Хвост катящегося кольца и обеспечивает конкатемерную ДНК. Переключение с одного способа репликации на другой требует участия продуктов *red*- и *gam*-генов фага лямбда, вовлекаемых также в общую рекомбинацию фаговой ДНК. Роль Gam-продукта заключается в подавлении действия экзонуклеазы V (продукта RecBC), которая предотвращает переход θ -молекул в катящиеся кольца. В случае инфекции фагами *red*⁻ *gam*⁻ ДНК лямбда не только остается в θ -форме, но и не способна к рекомбинации. Чтобы обеспечить соответствующий субстрат для созревания фага, в результате рекомбинации между кольцами-мономерами должны образовываться кольца-димеры (рис. 35.4). Поскольку рекомбинационная функция, кодируемая фагом, отсутствует, это событие полностью зависит от функции RecA клетки-хозяина. Частота рекомбинации сильно возрастает (об этом свидетельствует увеличение выхода фага) при определенных мутациях лямбда, получивших название *chi*. Оказалось, что *chi*-мутанты несут замены одной пары оснований,

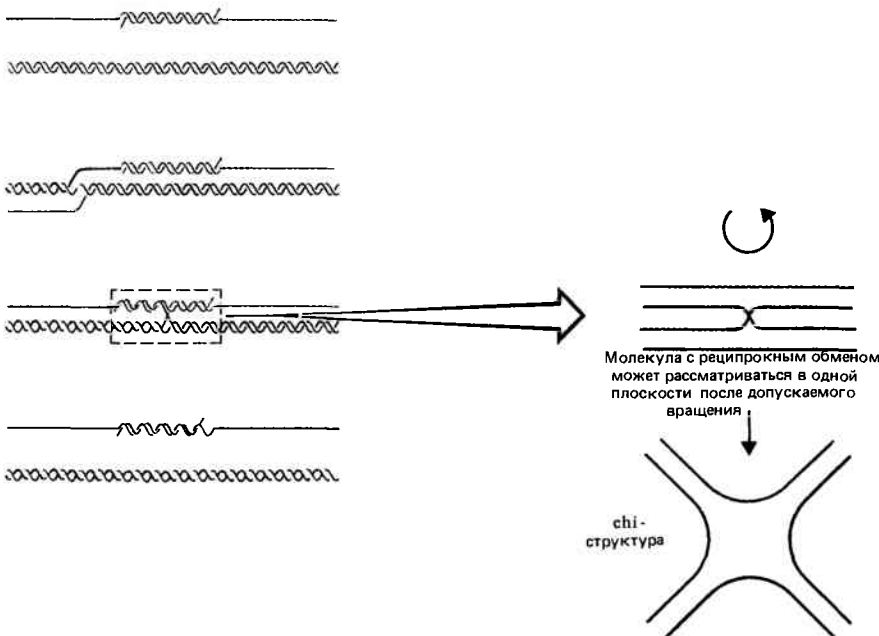


Рис. 35.9. В результате стимулированного RecA-белком обмена цепи между частично и полностью двухцепочечной ДНК образуется сложная молекула, имеющая ту же структуру, что и промежуточный продукт рекомбинации.

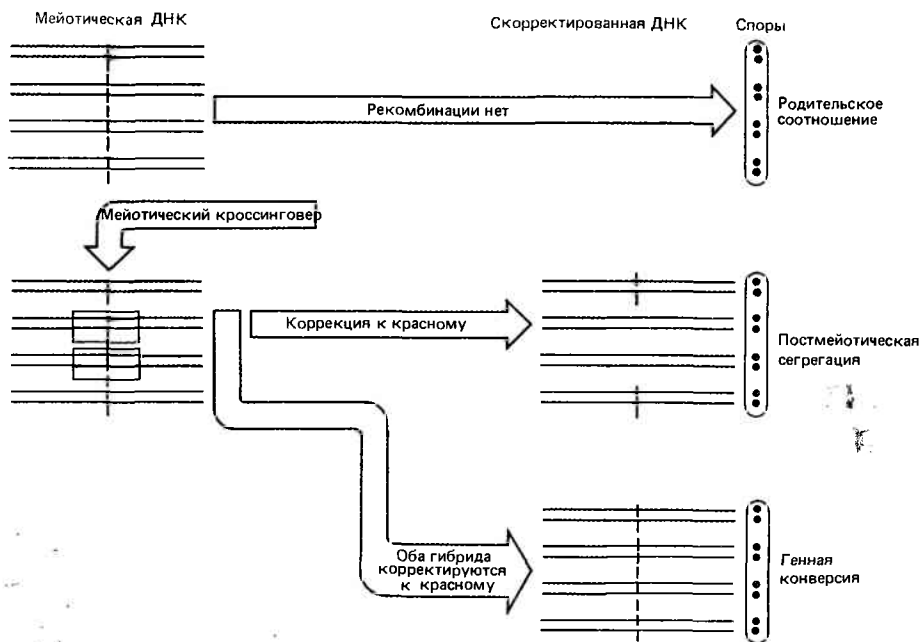


Рис. 35.10. Механизм образования спор у *Ascomycetes* позволяет определить генетический состав каждой из цепей ДНК, участвующих в мейозе.

Рекombинационное событие ведет к образованию гибридной ДНК в области, ограниченной прямоугольником. Последующая генетическая судьба локуса в этой области указана вертикальной линией. Если одна из гибридных молекул ДНК подвергается коррекции, споры имеют соотношение 3:5. Если обе молекулы корректируются одинаковым способом, соотношение будет 2:6. Каждая спора представляет одну цепь ДНК, поскольку дополнительная репликация происходит между мейозом и образованием спор.

в результате чего создаются сайты, стимулирующие рекомбинацию.

Такие сайты разделяют одну и ту же несимметричную последовательность из 8 пар оснований. В присутствии родственных последовательностей стимулирующий эффект утрачивается; следовательно, любая замена основания в такой последовательности приводит к потере ею способности выступать в качестве «горячей точки». Последовательность *chi* встречается обычно в ДНК *E. coli* примерно один раз на каждые 10 тысяч пар оснований. Ее присутствие в ДНК фага лямбда дикого типа, а также в других генетических элементах свидетельствует, что она не существенна для *RecA*-зависимости рекомбинации. Последовательность *chi* способна стимулировать рекомбинацию в соседних областях, находящихся от нее на расстоянии до 10 тысяч пар оснований. Мы не знаем, какую именно роль она играет в рекомбинации, но возможно, что *chi*-последовательность необходима для инициации или разрешения рекомбинационных интермедиатов.

Конверсия гена ответственна за межаллельную рекомбинацию

Участие гетеродуплексной ДНК в рекомбинации объясняет многие особенности рекомбинации между аллелями; эти данные послужили основой для разработки модели рекомбинации с участием гетеродуплекса. Когда было открыто явление межаллельной рекомбинации, предполагали, что она осуществляется с помощью того же механизма реципрокной рекомбинации, который используется для более удаленных локусов, однако в этом случае при образовании реципрокной пары рекомбинантных хромосом между локусами происходит отдельное событие разрыва и воссоединения. Следует отметить, что образование гетеродуплексной ДНК само по себе может привести к рекомбинационному событию.

Удобным объектом для изучения последовательных стадий рекомбинаций служат грибы-аскомицеты, так как у них продукты одного мейотического деления удерживаются вместе в одном аске. Более того, четыре гаплоидных ядра, образуемые в результате мейоза, распо-

ложены в линейном порядке. Следующий за мейозом митоз дает линейную серию из восьми гаплоидных ядер. На рис. 35.10 показано, что каждое из этих ядер соответствует одной из восьми цепей четырех хромосом, образуемых при мейозе, и представляет ее генетические особенности.

Мейоз в гетерозиготе приводит к образованию четырех копий каждого аллеля. Анализ спор в большинстве случаев свидетельствует об этом. Однако встречаются аски с ненормальным соотношением спор. Этот факт можно объяснить образованием и коррекцией гетеродуплексной ДНК в области, в которой аллели различаются.

Допустим, что два аллеля отличаются одной точковой мутацией. Если произойдет обмен и образуется гетеродуплексная ДНК, две цепи этого гетеродуплекса будут ошибочно спарены в сайте мутации. В результате окажется, что каждая цепь ДНК несет различную генетическую информацию. Если не происходит изменения в последовательности, цепи разделяются в последующей репликации, причем каждая дает начало двухцепочечной молекуле, которая сохраняет свою информацию. Это ведет к соотношению 4:4, при котором порядок спор изменен из-за одноцепочечных обменов. Такое соотношение названо **постмейотической сегрегацией**, так как ее наличие зависит от разделения цепей ДНК при репликации, которая следует за мейозом.

В некоторых асках наблюдается соотношение 3:5 или 2:6, т.е. одна или две споры вместо одного аллельного типа имеют другой тип. Это может быть обусловлено коррекцией неправильно подобранных пар в гетеродуплексной ДНК. Если репарирующая система узнает неправильно спаренные основания в гетеродуплексной ДНК, она может вырезать и изменить одну из цепей, восстановив комплементарность. Такое событие приводит к изменению цепи ДНК, представляющей один из аллелей, превращению ее в последовательность другого аллеля. Если это происходит только в одном из реципрокных гетеродуплексов, возникает соотношение 3:5 или 5:3 (в зависимости от направления конверсии). Если оба гетеродуплекса подвергаются коррекции одинаковым способом, возникает соотношение 2:6 или 6:2. Процесс коррекции

получил название **конверсии гена**; о нем свидетельствуют отклоняющиеся от нормы соотношения. (Равновероятно любое направление конверсии, но могут существовать аллельспецифичные эффекты, обуславливающие предпочтительность какого-то одного направления.)

Конверсия гена не зависит от кроссинговера, но коррелирует с ним. Большая часть аномальных асков появляется в результате генетической рекомбинации между двумя маркерами с каждой стороны от сайта межallelльной геной конверсии. Именно так и должно быть, если ненормальные соотношения возникают вследствие образования гетеродуплексной ДНК с примерно равной вероятностью разрешения структуры с участием рекомбинации или без нее (как показано на рис. 35.2). Предполагается, что хромосомы грибов инициируют кроссинговер примерно в два раза чаще, чем можно ожидать для отдаленных генов.

В пределах гена рекомбинация между двумя сайтами происходит, когда образуется гетеродуплексная ДНК или встречается коррекция в одном, но не в другом сайте. Таким образом, межallelльная рекомбинация проявляет то же свойство, что и конверсия гена: она не обязательно связана с реципрокной рекомбинацией между фланкирующими маркерами.

Почему же этот процесс позволяет построить линейную генетическую карту при рекомбинации в пределах гена? По-видимому, чем ближе лежат два сайта друг к другу, тем больше вероятность того, что они будут претерпевать **совместную конверсию** в том же отрезке гетеродуплексной ДНК. В этом случае оба сайта подвергаются коррекции в одном и том же направлении, которая нарушает соотношения генотипов, появляющихся при скрещивании, но не приводит к рекомбинации между маркерами.

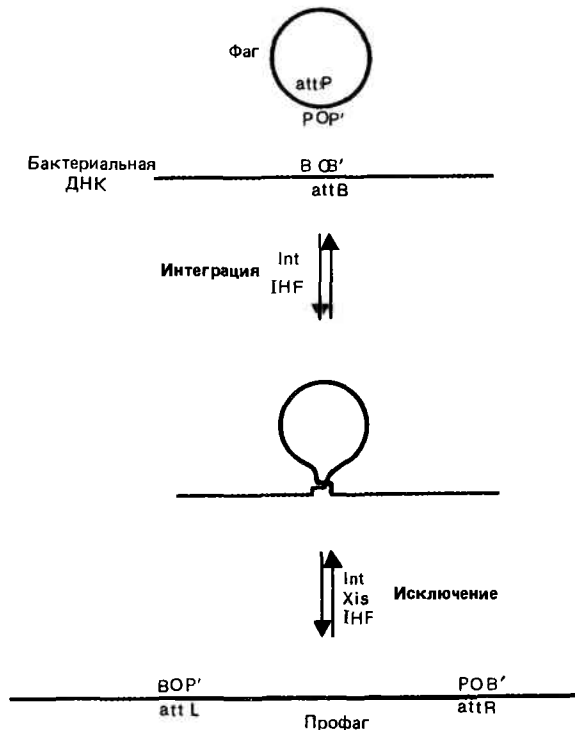


Рис. 35.11. Кольцевая ДНК фага превращается в интегрированный профаг в результате реципрокной рекомбинации между сайтами *attP* и *attB*; профаг исключается путем реципрокной рекомбинации между сайтами *attL* и *attR*.

Гетеродуплексная ДНК может иметь длину, поддающуюся оценке. Некоторую информацию о размере участка ДНК, на которой происходит коррекция, можно получить при анализе последовательностей, принадлежащих одному кластеру. Обычно продукты рекомбинационного события разделяются и становятся недоступными для определения последовательности их ДНК. Однако в том случае, если неравный обмен произошел между двумя членами одного кластера, как было показано на рис. 21.3, гетеродуплекс может сформироваться между двумя неallelльными генами. Такой гетеродуплекс может подвергнуться геной конверсии в направлении последовательности одной из его цепей. В результате произойдет превращение последовательности одного из неallelльных генов в последовательность другого. Подобный гетеродуплекс может образоваться посредством прямой реакции между двумя копиями гена на одной и той же хромосоме.

Присутствие более одной копии гена на хромосоме дает возможность проследить за этими событиями. Например, если образование гетеродуплекса и геной конверсия происходят и при этом захватывается часть гена, эта часть может иметь последовательность, идентичную или родственную другому гену, в то время как остающаяся часть проявляет большую дивергенцию. Установлено, что геной конверсия может захватывать до нескольких тысяч оснований.

Специализированная рекомбинация узнает специфические сайты

Переход ДНК фага лямбда из одного жизненного цикла в другой включает два типа событий. Во-первых, способ выражения генов регулируется, как описано в гл. 16. В период лизогении устанавливается и поддерживается репрессия операторов; во время литического цикла репрессия отсутствует. Во-вторых, физическое состояние ДНК различно в лизогенном и литическом состояниях; взаимопревращение этих состояний обеспечивается сайтспецифической рекомбинацией.

В литическом цикле ДНК фага лямбда существует в инфицированной бактерии в виде независимой кольцевой молекулы. Когда бактерия находится в лизогенном состоянии, фаговая ДНК представляет собой составную часть бактериальной хромосомы (названную профагом). Чтобы лизогенизировать бактерию, свободная ДНК фага лямбда должна **внедриться** в ДНК клетки-хозяина. Исключение профага из хромосомы бактериальной клетки выводит ее из лизогенного состояния и обуславливает лизис.

Внедрение и исключение осуществляются посредством рекомбинации в специфических локусах бактериальной и фаговой ДНК, получивших название сайтов присоединения или *att*-сайтов (от англ. *attachment*). В бактериальной генетике сайт присоединения ДНК фага лямбда называется *att^L*; этот локус был определен по наличию в нем профага λ в лизогенных штаммах и по локализации мутаций, предотвращающих интеграцию фага лямбда. Реакция рекомбинации изображена на рис. 35.11. Сайт присоединения на бактериальной хромосоме, обозначенный как *attB*, представлен компонентами последовательности *BOB'*. Сайт присоединения на фаговой ДНК, *attP*, состоит из компонентов *POP'*. Последовательность *O* является общей для *attB* и *attP* и представляет собой сайт реципрокной рекомбинации между ними. Поскольку

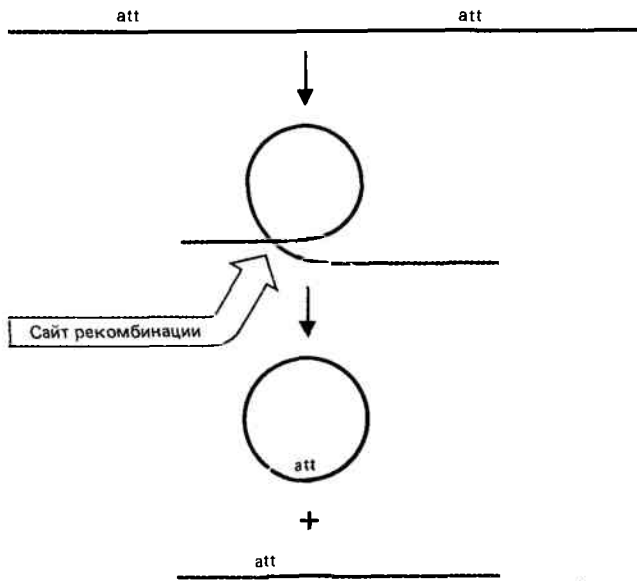


Рис. 35.12. Незаконная рекомбинация между сайтом в фаге и сайтом в бактериальной ДНК ведет к исключению трансдуцирующего фага, в котором бактериальные гены заменили некоторые фаговые гены.

фаговая ДНК имеет форму кольца, рекомбинационное событие ведет к ее внедрению в бактериальную хромосому в виде линейной последовательности. Профаг ограничен двумя новыми *att*-сайтами, продуктами рекомбинации. Обычно при изображении карты слева указывается *attL*, состоящий из *BOP'*, и справа — *attR*, состоящий из *POB'*.

В более ранних описаниях рекомбинационной модели два сайта, *attB* и *attP*, представляли в виде двухполовинчатых структур, *B.V'* и *P.P'*. Точка между ними означала точку кроссинговера. Считали, что реакция происходит между сайтами, полностью лишенными гомологии. Позднее генетические данные показали, что должна существовать и некоторая гомология между *att*-сайтами. В результате был введен символ *O*, обозначающий общую кор (сердцевинную)-последовательность, в пределах которой происходит кроссинговер. Фланкирующие области *B.V'* и *P.P'* рассматриваются по отношению к ней как отличающиеся друг от друга «плечи».

Важным следствием структуры *att*-сайтов является то, что реакции внедрения и исключения не вовлекают одну и ту же пару взаимодействующих последовательностей. Внедрение требует узнавания *attP* и *attB*, тогда как при исключении должны узнаваться сайты *attL* и *attR*. Таким образом, несмотря на обратимость рекомбинационного события, для каждого направления реакции необходимы определенные условия. Это может иметь большое значение в жизни фага, поскольку такая требовательность гарантирует, что событие внедрения не будет немедленно сопровождаться исключением и наоборот.

Различие в парах взаимодействующих сайтов при интеграции и исключении отражено различием белков, осуществляющих две реакции. Интеграция требует продукта фагового гена *int* и бактериального белка, названного фактором интеграции хозяина (ИНФ). Фактор ИНФ состоит из двух субъединиц, одну из которых кодирует SOS-регулируемый локус *himA*. В дополнение к *Int* и ИНФ реакция исключения требует продукта фагового гена *xis*. (При

определенных условиях исключение может происходить и в отсутствие *Xis*, правда, довольно неэффективно, что, по-видимому, свидетельствует об остаточной способности *Int* и ИНФ обращать реакцию интеграции непосредственно.)

Все четыре типа *att*-сайтов доступны для изучения. Они обеспечиваются в соответствующей форме при «незаконной рекомбинации», когда фаг лямбда освобождается благодаря реципрокной рекомбинации между последовательностью в профаге и последовательностью во фланкирующей бактериальной ДНК. Реакция происходит посредством генерализованного (*RecA*-зависимого) пути и, вероятно, обусловлена случайной гомологией между рекомбинирующими сайтами. Последствие такой реакции показано на рис. 35.12.

Незаконная рекомбинация приводит к образованию дефектных фаговых частиц, у которых часть фаговой ДНК замещена фланкирующей бактериальной ДНК. Исключенный геном несет *attL*, если рекомбинация включила бактериальный сайт слева; если же событие произошло с другой стороны, он содержит *attR*. Присутствие бактериальной ДНК в фаговом геноме оказывается очень полезным качеством. Процесс переноса бактериальных генов с помощью таких фагов получил название **трансдукции**: до разработки методов получения рекомбинантных ДНК это свойство использовали для выделения бактериальных генов, находящихся с той или другой стороны от *att*-сайта.

Используя фаги такого типа, можно получить сходные молекулы ДНК, каждая из которых несет определенный *att*-сайт. Эти субстраты используются для проверки роли всех возможных комбинаций *att*-сайтов в сайтспецифической рекомбинации. Белок *Int* необходим для всех комбинаций; ему помогает белок *Xis*, причем наиболее выражен этот эффект в реакции *attL* × *attR*. Белок *Xis* оказывает стимулирующее действие и на некоторые другие комбинации. Различия в реакциях, требующих участия *Xis*-белка, нам пока непонятны.

Ступенчатый разрез и воссоединение в коре

Природа кор-последовательности стала ясна, как только удалось расшифровать последовательности *att*-сайтов. Все они представляют собой АТ-богатые области, включающие общую последовательность из 15 пар оснований. То, что рекомбинация происходит в пределах именно этой последовательности, показано обменом «плечей» в *attL* и *attR* в сравнении с сайтами *attP* и *attB*.

Происходит ли рекомбинация в любом участке внутри последовательности кора или в специфической точке? Проследивание судьбы радиоактивной метки в фосфатных группах *att*-сайта показало, что материал перемещается непосредственно при рекомбинации: точный разрыв и воссоединение имеют место в отсутствие какого-либо синтеза ДНК. Способ переноса метки указывает на существование специфических точек рекомбинации. Поскольку точки обмена различны на каждой цепи ДНК, можно предположить, что разрыв и воссоединение имеют ступенчатый характер. Согласно модели, представленной на рис. 35.13, при внесении одинакового ступенчатого разреза в *attP*- и *attB*-сайты, комплементарные одноцепочечные концы могут быть использованы в перекрестной гибридизации. Это напоминает реакцию между липкими концами, образующимися при действии некоторых ре-

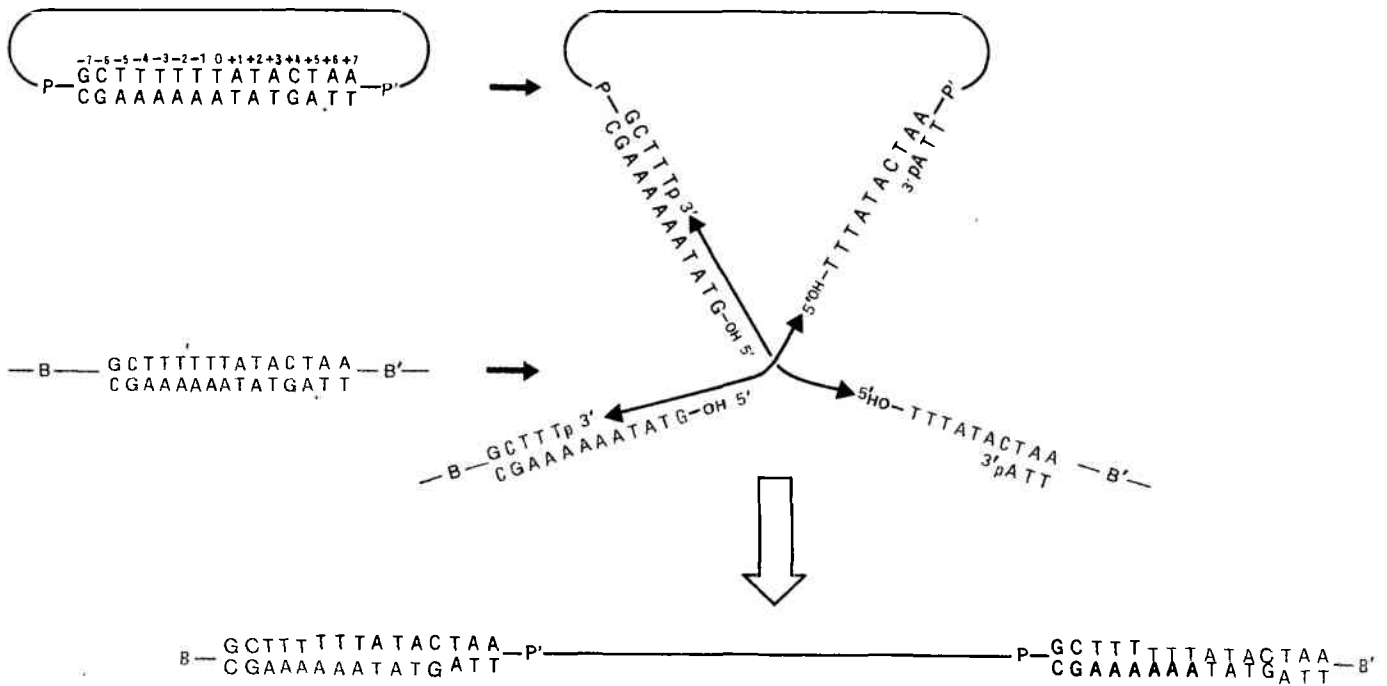


Рис. 35.13. Ступенчатые разрезы в одинаковых коровых последовательностях сайтов *attP* и *attB* обеспечивают перекрестное воссоединение, которое приводит к реципрокным рекомбина-

ционным обменам.

Последовательность пронумерована относительно центра кора.

стрикционных ферментов (гл. 19). Поскольку фосфатные группы были мечеными, точно установить расстояние между ступенчатыми разрезами невозможно, если неизвестно, с какой стороны разрезана фосфоэфирная связь. Промежуточные продукты этой реакции с одноцепочечными концами выделить не удалось. По-видимому, протяженность образующегося уступа должна соответствовать 5 или 7 парам оснований. Образование уступа из 7 пар оснований подразумевает появление 3'-фосфатного и 5'-ОН концов.

Характер рекомбинационного события обуславливает сохранение мутации в «плече», в котором она локализована с любой стороны от перекрывающихся 7 пар оснований. Следовательно, мутация никогда не может перейти от *attB* к *attP* (или от *attL* к *attR*). Однако в пределах 7 перекрывающихся пар оснований она может быть передана путем генной конверсии между *attP* и *attB* (и, следовательно, может передаваться между всеми *att*-сайтами). Открытие такой передачи послужило основой для концепции, что все *att*-сайты должны иметь общий кор, если даже общая последовательность (15 пар оснований) оказывается длиннее, чем область, вовлекаемая в образование уступа.

Если в хромосоме *E. coli* сайт *att^λ* делетирован, инфицирующий фаг лямбда способен лизогенизировать бактерию, интегрируясь в каком-то другом сайте, однако эффективность такой реакции составляет менее 0,1% от частоты интеграции в сайт *att^λ*. Такая неэффективная интеграция происходит во **вторичных сайтах присоединения**. Эти сайты имеют последовательности, напоминающие те, которые находятся внутри и вокруг общего кора из 15 нуклеотидных пар. Рекомбинанты, образуемые путем вставки в эти сайты, имеют последовательности, согласующиеся с данными о том, что сайты перекреста на двух цепях разделяют 7 пар оснований. В таких вторичных реакциях важную роль играет гомология между

определенной парой рекомбинирующих сайтов в 7 парах оснований коровой области; кроме того, большое значение имеет гомология с последовательностями, окружающими сайты разрезания, которые узнает белок Int (см. ниже).

Сайтспецифическую рекомбинацию можно проанализировать в отношении ее ферментативного механизма, поскольку рекомбинация этого типа воспроизводима в экспериментах *in vitro* с помощью белков Int и IHF. Роль плечевых последовательностей изучают с помощью делеций, полученных с каждой стороны от кора. Делеции, уменьшающие размер *attP*-сайта за пределы от -152 (в P) до +82 (в P'), приводят к исчезновению его функции. Для выяснения роли отдельных областей в пределах отрезка из 240 пар оснований необходимо ввести изменения в каждый из сегментов. Сайт *attB* имеет меньшую длину; его функция может быть выполнена сегментом из 23 пар оснований, занимающим положение от -11 до +11. В этот сегмент входит лишь по 4 пары оснований с каждой стороны от кора. Фрагмент *attP* подобного размера ведет себя в таком эксперименте как *attB*-сайт. Несоразмерность в их величине позволяет предполагать, что сайты *attP* и *attB* играют различную роль в рекомбинации. Вероятно, сайт *attP* обеспечивает дополнительную информацию, необходимую для того, чтобы его можно было отличить от сайта *attB*.

Если реакция осуществляется между двумя суперспиральными молекулами ДНК *in vitro*, почти все супервитки сохраняются. Следовательно, не существует каких-либо свободных промежуточных структур, в которых могло бы происходить вращение цепи. Любой свободный конец, образующийся при разрыве, сохраняется в фиксированном положении с помощью рекомбинационного фермента (или ферментов), или один конец переносится непосредственно к другому концу с образованием четырехце-

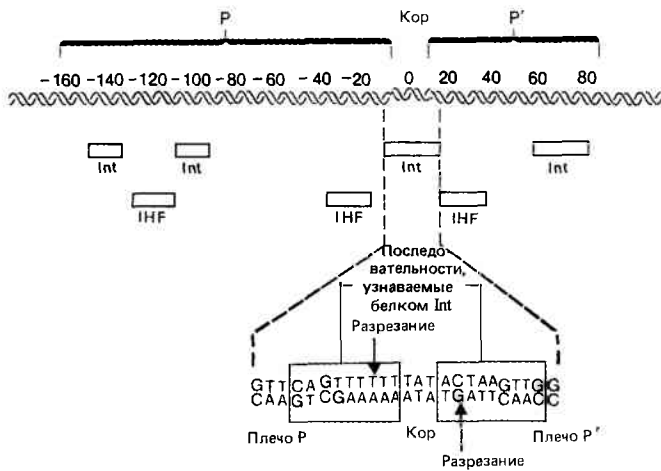


Рис. 35.14. Белки Int и IHF связываются с различными сайтами в *attP*. Последовательности, узнаваемые белком Int в коре, включают сайты разрезания.

почечной промежуточной структуры. Разрыв и воссоединение могут быть осуществлены ферментом, напоминающим по активности топоизомеразу I, с тем отличием, что в этом случае разрезанная цепь спlicingается вместе с разрезанной цепью другой, а не той же молекулы. В такой процесс может быть вовлечен белок Int, так как он проявляет активность (довольно слабую) топоизомеразы I, что приводит к релаксации отрицательно суперспиральной ДНК.

Как белок Int, так и белок IHF связываются со специфическими сайтами в *att*-ДНК *in vitro*. Для осуществления рекомбинации на одну рекомбинантную ДНК необходимо около 20–40 молекул Int-белка с мол. массой 40 000. Белка IHF требуется еще больше, примерно 70 молекул с мол. массой 20 000 дальтон на рекомбинант. (Белок IHF содержит субъединицы с мол. массой 11 000 и 9 500.) Высокая стехиометрия свидетельствует о том, что белки не функционируют каталитически, а образуют какую-то структуру, которая поддерживает только одно рекомбинационное событие.

Белки Int и IHF связываются со специфическими сайтами в *att*-области. Связывающие сайты представлены на рис. 35.14. Белок Int связывается с сегментом из 30 пар оснований, окружающим область кора, с последовательностью той же длины в P' и двумя отдельными, более короткими чем 15 пар оснований, последовательностями в P. Различия в длинах отдельных сайтов, вероятно, отражают число связывающих мономеров белка Int. (Связывание Int с *attB* напоминает связывание в кор-области *attP*.) Белок IHF связывается с последовательностями примерно из 20 пар оснований в «плечах» *attP*;

сайты связывания белка IHF расположены почти рядом с Int-сайтами. Белки Int и IHF покрывают большую часть последовательности *attP*. Характер связывания белка Int в плечевых областях и коре различен. Во-первых, в плечевых областях связывающие сайты родственны обобщенной последовательности из 8 пар оснований, которая отсутствует в области кора. Во-вторых, белок Int способен связываться с фрагментами ДНК, содержащими последовательность СААСТТНТТ. Родственные с ней последовательности найдены в местах сочленения кор—плечо *attP* и *attB* в инвертированной ориентации. Эти последние сайты окружают сайты разрезания кора, как показано на рис. 35.14, что ставит белок Int в правильное положение для выполнения рекомбинации. Если определить на двойной спирали локализацию участков кора, связывающихся с белком Int, то окажется, что практически все они располагаются на одной стороне ДНК.

При связывании белка Int с *attP* образуется комплекс, в котором все сайты связывания могут находиться на поверхности белкового олигомера. Реакция включает обертывание ДНК протяженностью около 230 пар оснований вокруг 4–6 мономеров белка Int. Весьма вероятно, что этот комплекс промежуточный, что далее он вступает в реакцию с *attB*, в котором единственной областью, узнаваемой белком Int является последовательность кора.

Ассиметрия реакций внедрения и исключения доказана тем фактом, что белок Int способен образовывать подобный комплекс с сайтом *attR* только в присутствии белка Xis. Этот комплекс способен спариваться с комплексом, который образует белок Int в сайте *attL*. В данном случае участие фактора IHF не требуется.

Сложность сайтспецифической рекомбинации может быть обусловлена необходимостью регулировать реакцию таким образом, чтобы при переходе вируса в состояние профага происходила интеграция, а при вступлении профага в литический цикл развития — исключение. Нужная в данной ситуации реакция осуществляется благодаря контролю количества белков Int и Xis.

Рекомендуемая литература

Рекомбинации посвящены многие обзоры последних лет. Среди них особого внимания заслуживает обзор Дресслера и Поттера (*Dressler, Potter, Ann. Rev. Biochem.*, **51**, 727–761, 1982), рассматривающий механизмы рекомбинации и роль RecA-белка. Сталь в своем обзоре связывает известные механизмы рекомбинации с фактами, описанными ранее на грибах и, кроме того, обсуждает роль горячих точек (*Stahl, Ann. Rev. Genet.*, **13**, 7–24, 1979). Механизм сайтспецифической рекомбинации бактериофага лямбда обсуждается в обзоре Нэша (*Nash, Ann. Rev. Genet.*, **15**, 143–167, 1981).

Часть X

ДИНАМИЧНОСТЬ ГЕНОМА: ПОСТОЯННОЕ ИЗМЕНЕНИЕ ДНК

Элементы, содержащиеся в хромосомах кукурузы ... предназначены для контроля действия генов и индукции в них наследственных модификаций, влияющих на это действие. Эти элементы были открыты благодаря их способности менять свое положение в геноме: исчезать и появляться в новых местах хромосом. Присутствие такого элемента в локусе известного гена или вблизи него влияет на проявление данного гена. Эти элементы получили название контролирующих... Можно думать, что они представляют собой некую внехромосомную субстанцию, способную присоединяться к хромосоме и воздействовать на активность генетического субстрата в этом участке. Однако способы действия контролирующих элементов не подтверждают это предположение. Скорее они свидетельствуют о том, что контролирующие элементы являются компонентами хромосом и обладают таким большим количеством специфических активностей и способов их проявления, каким должны обладать гены... Транспозиции контролирующих элементов либо являются результатом какого-то еще неизвестного механизма, либо происходят во время процесса редупликации хромосомы и представляют собой его следствие.

Барбара Мак-Клинтон, 1958

ТРАНСПОЗИРУЮЩИЕСЯ ЭЛЕМЕНТЫ БАКТЕРИЙ

До недавнего времени было принято считать, что геномы про- и эукариот статичны, что последовательности, образующие их, подвергаются только медленным эволюционным изменениям. Мы привыкли к мысли, что генетическая карта отражает порядок расположения известных генов; подразумевается, что неидентифицированные последовательности также сохраняют постоянное место в геноме. На стабильность генетической организации указывает наличие родственных последовательностей у представителей дивергировавших видов, например у человека и обезьяны. Различие во времени генераций про- и эукариот свидетельствует о том, что они эволюционируют с разной скоростью, но даже у прокариот организация генома меняется относительно медленно. Например, очень сходные генетические карты имеют разные бактериальные виды *E. coli* и *S. typhimurium*. Эволюция генов происходит как в результате приобретения новых последовательностей, так и в результате перераспределения уже имеющихся. Новые последовательности могут быть введены с помощью векторов или появляться при мутировании существующих генов. Возникновение новых последовательностей возможно также в результате перестроек генетического материала. Такие перестройки могут изменить и функции имеющихся генов путем создания для них новых условий регуляции.

«Традиционная» эволюция генома связана с механизмами, которые хорошо известны в течение многих лет. У прокариот за осуществление генетических обменов ответственны внехромосомные элементы. Плазмиды способствуют конъюгации бактерий, в то время как фаги осуществляют инфекционный процесс. И те и другие иногда переносят вместе с автономным репликоном гены клетки-хозяина. У эукариот в каждой генерации происходит реципрокная рекомбинация между соответствующими сайтами на гомологичных хромосомах; очень редко она сопровождается дупликацией или перестройкой локусов. Такая реорганизация, по существу, представляет собой побочный эффект обычных механизмов, включаемых в генетическую рекомбинацию и (или) синтез ДНК (таких, как неравный кроссинговер или конверсия гена). Механизмы, ответственные за транслокацию между негомологичными хромосомами, остаются неизвестными.

Потенциальная возможность для изменения прокариотических и эукариотических геномов обеспечивается способностью определенных последовательностей перемещаться из одного сайта в другой. Эти последовательности получили название транспозирующихся элементов, или транспозонов. В ДНК прокариот было найдено множество элементов этого типа. Бактериальные транспозиции включают дупликацию элемента: одна копия сохраняется в исходном донорном сайте, в то время как другая появляется в новом реципиентном сайте, или сайте-мишени. Транспозиция не основана на каком-либо родстве между последовательностями в донорных и реципиентных сайтах, и это отличает ее от всех других механизмов, участвующих в перестройках ДНК. Каждый транспозон несет гены, требуемые для его собственной транспозиции,

хотя для этого необходимы также функции генома, в котором находится транспозон (в частности, ДНК-полимераза и ДНК-гираза хозяина). О распространенности подобных событий среди эукариот можно судить на основании общего структурного сходства и разнообразия в локализации аналогичных элементов (гл. 37 и 38).

Каковы последствия присутствия в геноме транспозирующихся последовательностей, способных к воспроизведению и перемещению? Возможные преимущества, создаваемые для организма транспозонами, связаны с их способностью индуцировать прямым или косвенным путем перестройки. Событие транспозиции само по себе может вызывать делеции, инверсии или способствовать переносу последовательности хозяина в новое место на хромосоме. Транспозоны могут создавать в клеточных системах частичные области гомологии, поскольку их копии в разных местах (даже на разных хромосомах) обеспечивают возможность реципрокной рекомбинации. Такие обмены могут приводить к делециям, инсерциям, инверсиям или транслокациям.

Транспозирующиеся элементы играют активную роль в эволюции некоторых бактериальных плазмид. Компоненты плазмид могут объединяться друг с другом с помощью реакций, включающих узнавание последовательностей транспозона. Интеграция F-плазмиды (полового фактора) *E. coli* в бактериальную хромосому часто осуществляется с помощью RecA-зависимой реципрокной рекомбинации между транспозоном в плазмидной ДНК и гомологичным транспозоном в бактериальной ДНК.

Геномы клетки или фага иногда используют транспозон в собственных целях. В этом случае элемент становится статичным и играет роль в регуляции, участвуя в специфических событиях, подобных транспозиции. Примеры с *Salmonella* и фагом Mu показывают, что регулятор может включать инверсию специфического сегмента ДНК, осуществляемую посредством рекомбинационного механизма, связанного с транспозицией. Эти процессы проявляют некоторое сходство с интеграцией фага лямбда. Таким образом транспозиционные механизмы принимают участие в целом ряде событий, от соединения негомологичных последовательностей ДНК до обеспечения специфических рекомбинационных процессов.

Следует отметить, что значение транспозонов для природной селекции еще неясно. Существует предположение, что (по крайней мере некоторые) транспозирующиеся элементы не дают ни преимуществ, ни вреда клетке, а представляют собой «эгоистическую ДНК», занятую только своим собственным размножением. Согласно этой концепции, взаимосвязь транспозона с геномом напоминает отношения между паразитом и его хозяином. По-видимому, распространение элементов путем дупликации и транспозиции уравнивается теми вредными последствиями, которые оказывает внедрение транспозона, приводящее к инактивации необходимого гена. Кроме того, распространение транспозонов может быть ограничено, если их число в клетке становится для нее грузом. Мы должны иметь в виду, что любое транспозиционное

событие, обеспечивающее селективные преимущества, например генетические перестройки, обуславливает предпочтительную выживаемость генома, несущего активный транспозон.

Открытие транспозиции у бактерий

Транспозирующиеся элементы были открыты при обнаружении вставок (инсерций) нового материала в пределах бактериальных оперонов. Такие вставки локализируются внутри гена и предотвращают его транскрипцию и (или) трансляцию. Кроме того, они могут проявлять полярный эффект, выражающийся в уменьшении степени экспрессии следующих за ними генов оперона. Первоначально определенные как негативные мутации, вставки были идентифицированы благодаря тому, что в отличие от точечных мутаций единственной формой реверсии для них является делетирование встроенного материала. При сравнении последовательностей, присутствующих в различных инсерционных мутантах, оказалось возможным классифицировать несколько дискретных элементов. Каждый тип элемента мог быть встроен в любой ориентации в определенный сайт.

Поведение многих транспозонов было прослежено в процессе нескольких циклов транспозиции. На рис. 36.1 приведена схема, показывающая состояние как донорной, так и реципиентной молекул. Плазмида, несущая транспозон, введена в клетку, которая содержит другую плазмиду без транспозона. Производится отбор реципиентных молекул, получивших транспозон (например, по инактивации маркера резистентности к антибиотику в реципиентной плазмиде). Затем выделяются молекулы плазмидной ДНК донора и реципиента. Их исследование подтверждает, что реципиентная ДНК действительно приобрела вставку транспозона, а донорная ДНК не утратила его копию. Таким образом, транспозиция связана с образованием *дополнительной копии* транспозона в реципиентном сайте.

Транспозицию часто изучают путем выделения многих независимых вставок в селективируемую область мишени. Дальнейший анализ тонкой структуры показывает, существуют ли какие-либо закономерности в селекции сайтов при внедрении. Большинство транспозонов имеют несколько сайтов внедрения. Некоторые проявляют предпочтительность в отношении «горячих» точек. Транспозоны, у которых в пределах небольшой области сайты внедрения выбираются случайно, могут тем не менее проявлять предпочтительность в отношении одного района по сравнению с другим. Например, в пределах какой-то определенной последовательности длиной приблизительно в 3000 пар оснований внедрения могут быть более частыми, но при этом сайты внутри этой последовательности могут выбираться случайно. Причина такой предпочтительности неизвестна; в таблицах 36.1 и 36.3 она описана как «региональная» специфичность. Наиболее важно, что на молекулярном уровне выбор мишени не зависит от нуклеотидной последовательности (с одним исключением).

Частота транспозиции варьирует для различных элементов в пределах, сравнимых с частотой спонтанных мутаций (от 10^{-5} до 10^{-7} на поколение). Реверсия (путем делеции) происходит приблизительно в 1000 раз ре-

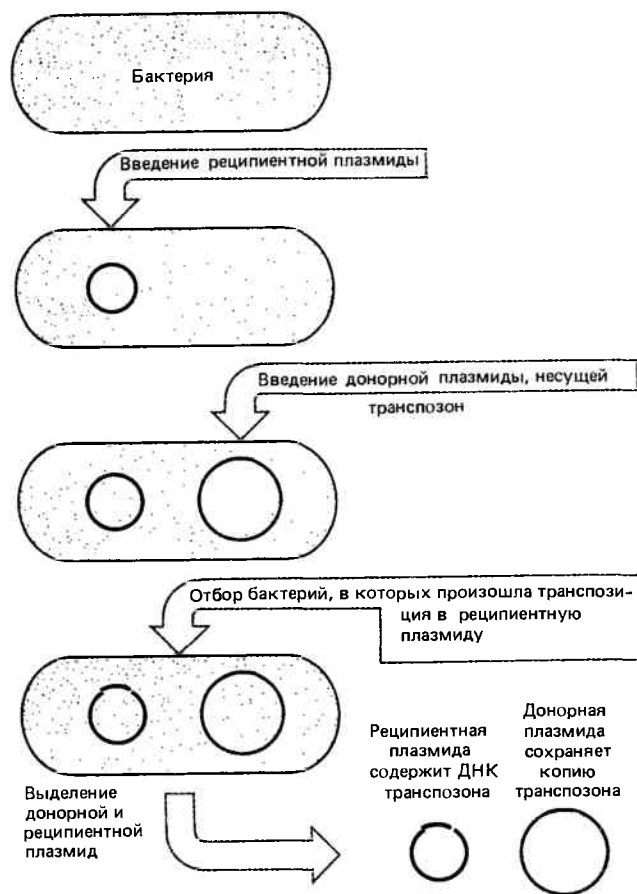


Рис. 36.1. Транспозиция ведет к появлению новой копии транспозона в сайте-мишени, в то время как исходная копия сохраняется в донорном сайте.

же, ее частота колеблется от 10^{-6} до 10^{-10} на поколение.

Существует несколько типов транспозонов, хотя определенные черты являются общими для всех элементов. Каждый из них представляет собой автономную единицу, которая кодирует белки, необходимые для транспозиции. Реакция включает узнавание концов транспозирующегося элемента. Самые короткие транспозоны кодируют только белки, участвующие в транспозиции. (К ним подходит определение «эгоистичной» ДНК.) Более крупные транспозоны несут дополнительные генетические маркеры.

Чем объясняется полярный эффект транспозонов в отношении генов, расположенных дистальнее сайта внедрения транспозона? Внедренная последовательность становится частью полицистронной мРНК оперона. В результате он может препятствовать транскрипции или трансляции. В некоторых случаях проявление полярности объясняется *rho*-зависимой терминацией транскрипции: в *rho*⁻штаммах полярность не наблюдается (гл. 13). В других случаях полярные эффекты возникают в результате терминации трансляции в нонсенс-кодоне транспозона. Так как транспозоны имеют регуляторные сигналы, относящиеся к их собственным генам, эти сигналы иногда могут влиять на события в оперонах. Некоторые транспозоны вблизи их границ содержат промоторы, в которых инициируется транскрипция фланкирующего материала; активируются при этом гены, смежные с элементом.



Рис. 36.2. Транспозоны имеют инвертированные концевые повторы. Внедрение транспозонов вызывает образование в сайте-мишени прямых повторов фланкирующей ДНК.

В этом примере сайт-мишень представлен последовательностью из 5 п.н. Концы транспозона содержат инвертированные повторы из 9 п.н. Цифры 1–9 указывают последовательность пар оснований.

Инсерционные последовательности – это простейшие транспозоны

Первая группа выделенных транспозонов получила название **инсерционных последовательностей**. Каждый тип инсерционной последовательности обозначается префиксом IS, далее следует номер, соответствующий именно этому типу. (Первые классы получили номера IS 1–4; номера классов, выделенных позднее, отражают историю и не соответствуют общему числу известных в настоящее время элементов.) IS-элементы являются нормальными компонентами бактериальных хромосом и плазмид. Например, один стандартный штамм *E. coli* содержит восемь копий IS1 и пять копий IS2. Для обозначения инсерции в определенном сайте используется двоеточие; так, обозначение лямбда::IS1 показывает, что IS1-элемент внедрен в геном фага лямбда.

IS-элементы представляют собой простейший класс транспозонов; их генетические функции связаны только

со способностью к транспозиции. Несмотря на то что каждый IS-элемент имеет присущую только ему последовательность, форма организации у них общая. Некоторые параметры IS-элементов суммированы в верхней части табл. 36.1.

Каждый элемент обладает короткими инвертированными концевыми повторами. Обычно их протяженность колеблется от 15 до 25 нуклеотидных пар. Две копии повторов очень близки, но не идентичны. На рис. 36.2 показано, что по направлению к элементу с каждой стороны от него идет одна и та же последовательность.

При транспозиции IS-элемента последовательность ДНК хозяина в сайте внедрения дублируется. Природа дубликации установлена при сравнении последовательности сайта-мишени до и после внедрения. ДНК IS-элемента всегда фланкирована очень короткими прямыми повторами. (В этом контексте определение «прямые» означает, что две копии последовательности повторены в одной и той же ориентации, а не то, что они смежные.) Сайт-мишень содержит последовательность только одного из этих повторов. В случае гипотетической последовательности, изображенной на рис. 36.2, сайт-мишень содержит последовательность $\frac{ATGCA}{TACGT}$. После транспозиции по одной копии этой последовательности представлено с каждой стороны транспозона.

Таблица 36.1
Некоторые транспозоны состоят из отдельных модулей или содержат их в своем составе

	Длина, п. н.	Инвертированные концевые повторы, п. н.	Прямые повторы, п. н.	Возможное число белков	Выбор мишени
IS-элементы					
IS1	768	23	9	2	Региональный
IS2	1327	41	5	2	Горячие точки
IS4	1428	18	11 или 12	2	
IS5	1195	16	4	4	
IS-подобные элементы					
IS10R	1329	22	9	3	Горячие точки
IS50R	1531	9	9	3	
IS903	1057	18	9	2	Неизвестен

Для большинства транспозонов последовательность прямого повтора в каждом отдельном событии транспозиции отличается, но длина ее остается постоянной (что связано с механизмом транспозиции). Наиболее часто встречаются прямые повторы длиной 5 и 9 пар оснований.

Из рис. 36.2 следует, что транспозон характеризуется определенной структурой; его концы представляют собой инвертированные концевые повторы, а примыкающие концы фланкирующих ДНК хозяина являются короткими прямыми повторами. При анализе последовательности ДНК факт обнаружения такого типа структуры указывает на присутствие транспозона.

Транспозиция всегда связана с внедрением непермутированной линейной последовательности. В результате все копии данного транспозона имеют одинаковые инвертированные терминальные повторы в местах соединения с хозяйской ДНК. Важность концевых структур подтверждается терминальной локализацией мутаций, которые предотвращают транспозицию элемента. Эти мутации действуют в *цис*-положении, следовательно концы узнаются белками, ответственными за транспозицию.

Самый короткий по длине транспозон, IS1, содержит две открытые рамки считывания в одном и том же направлении; они транслируются в белки InsA и InsB (их функции еще не установлены). Все другие IS-элементы содержат одну длинную кодирующую область, начинающуюся внутри инвертированного повтора в одном конце и заканчивающуюся перед инвертированным повтором (или внутри его) в другом конце. Такая длинная рамка считывания фактически транслируется всегда; в некоторых случаях продукт считывания был выделен. Обычно существуют некоторые потенциально кодирующие области на другой цепи ДНК (т.е. та же последовательность считывается в противоположном направлении). Используются ли такие перекрывающиеся рамки считывания для синтеза белков, не известно, но они включены в табл. 36.1 при подсчете возможных продуктов.

Сложные транспозоны содержат IS-модули

Некоторые транспозоны (помимо генов, связанных с транспозицией) несут маркеры лекарственной устойчивости (или какие-либо другие). Такие транспозоны обозначаются как Tn с добавлением соответствующего номера. Один класс более крупных транспозонов представлен **сложными элементами**, состоящими из центральной области, несущей маркер лекарственной устойчивости, и фланкирующих с каждой стороны родственных последовательностей (плеч). Их иногда называют **длинными концевыми повторами (LTR)**. (Не следует путать с короткими инвертированными концевыми повторами у IS-элементов.) Плечи могут либо иметь одну и ту же или (более часто) инвертированную ориентацию. Таким образом, сложный транспозон с плечами, представленными прямыми повторами, имеет структуру:

Плечо L Центральная область Плечо R

Если плечи представлены инвертированными повторами, транспозон имеет структуру:

Плечо L Центральная область Плечо R

Стрелки указывают ориентацию плеч, которые идентифицированы как L и R в соответствии с ориентацией генетической карты транспозона слева направо. Детали структуры сложного транспозона представлены на **рис. 36.3**. Иногда плечи состоят из идентичных IS-элементов. В случае Tn9 на каждом конце транспозона присутствует прямой повтор IS1-элемента.

В других случаях плечи напоминают IS-элементы. Каждое включает одну непрерывную открытую рамку считывания, фланкированную короткими инвертированными концевыми повторами. Однако, поскольку такие элементы в виде отдельных компонентов в составе бактериальной ДНК не найдены, они получили название **IS-подобных элементов**. IS-подобные элементы так же, как и IS-элементы, обычно получают номер того транспозона, в котором они обнаружены. Их свойства сравниваются со свойствами IS-элементов в нижней части табл. 36.1.

Особенности некоторых сложных транспозонов суммированы в **табл. 36.2**. Поскольку плечи представляют собой часть сложного транспозона и в то же время на-

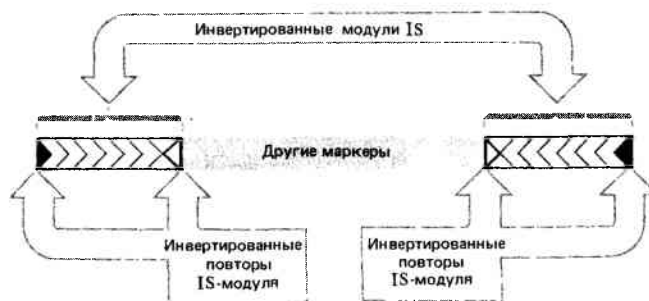


Рис. 36.3. Сложные транспозоны имеют центральную область, несущую маркеры, не связанные с транспозицией (например, ген лекарственной устойчивости). Центральная область фланкирована IS-модулями. Модули имеют короткие инвертированные концевые повторы. Если модули находятся в инвертированной ориентации относительно друг друга (как показано на рисунке), короткие инвертированные повторы на концах транспозона являются идентичными.

поминают IS-элемент или даже ими являются, иногда их описывают как IS или IS-подобные модули. В некоторых случаях модули сложных транспозонов идентичны, например у Tn9 (прямые повторы IS1) или у Tn903 (инвертированные повторы IS903). В других случаях модули очень близки, но не идентичны. Мы можем отличить модули IS10L и IS10R. (В этом случае L и R терминологически идентифицируют различные модули; если же модули идентичны, та же терминология указывает на их положение относительно центральной области транспозона.)

Все модули имеют концевые инвертированные повторы, поэтому сложный транспозон также заканчивается короткими инвертированными повторами. (Если плечи имеют инвертированную ориентацию относительно друг друга, короткие инвертированные повторы на концах транспозона, вероятно, идентичны, а не родственны.)

Несмотря на то что IS-подобные модули обычно не существуют в виде независимых элементов в бактериальной ДНК (отсюда и происходит их название), функциональный модуль может транспозироваться, как это показано в случае IS10R транспозона Tn10 или IS50R Tn5. Если модули сложных транспозонов идентичны, по-видимому, любой из них может обеспечивать перемещение. Если же модули различны, транспозиция, вероятнее всего, будет зависеть от одного из них.

Мы предполагаем, что сложные транспозоны образуются путем объединения двух первоначально независимых модулей с центральной областью. Такая ситуация может возникнуть в случае, когда IS-элемент транспозируется в реципиентный сайт, находящийся вблизи донорного сайта. Будучи исходно идентичными, эти модули могут сохранять свою идентичность или дивергировать. Способность одного модуля обеспечивать транспозицию всего сложного элемента можно объяснить отсутствием селективного давления в отношении сохранения активности обоих модулей.

Представляет интерес вопрос о том, как кооперируются два модуля при транспозиции сложного транспозона в тех случаях, когда функционально активны оба. Например, у Tn9, по-видимому, каждый модуль (IS-элемент) активен как в отношении своей собственной транспозиции, так и в отношении сложного транспозона. Если один из модулей транспозируется независимо, существованию сложного транспозона это не грозит, поскольку транспозиция ведет к появлению новой копии в сайте-мишени

Таблица 36.2
Некоторые сложные транспозоны фланкированы отдельными модулями

Элемент	Длина, п.н.	Генетические маркеры	Концевые модули	Ориентация модулей	Родство модулей	Функция модулей
Tn10	9300	tet^R	IS10R IS10L	Инвертированная	2,5% дивергенции	Полностью функциональный Ограниченно функциональный
Tn5	5700	kan^R	IS50R IS50L	Инвертированная	1 п.н.	Функциональный Нефункциональный
Tn903	3100	kan^R	IS903	Инвертированная	Идентичны	Оба функциональных
Tn9	2500	cam^R	IS1	Прямая	Предположительно идентичны	Предположительно функциональный.

(хотя такое событие и не служит цели дальнейшего распространения всего транспозона).

В действительности два IS-элемента способны транспозировать любую последовательность, заключенную между ними, а также свои собственные последовательности. Схема эксперимента, который демонстрирует такую способность, представлена на рис. 36.4. Если Tn10 находится в составе кольцевого репликона, два его модуля могут рассматриваться фланкирующими либо ген tet^R исходного транспозона, либо последовательность в другой части кольца. В результате в событие транспозиции могут вовлекаться исходный транспозон Tn10 или транспозон «наизнанку» с противоположающей центральной областью. Оба они имеют инвертированные модули, но эти модули, очевидно, способны функционировать в любой ориентации относительно центральной области. В таких случаях частота транспозиции тем меньше, чем больше расстояние между модулями. Зависимость от длины, вероятно, свидетельствует о проблемах, которые возникают во время транспозиции. Используя эту зависимость, можно определять размеры сложных транспозонов.

Суммируя свойства, определяющие способность к транспозиции IS- или IS-подобных элементов, следует отметить, что все они, кроме IS1, кодируют один главный продукт гена, соответствующий рамке считывания, простирающейся на длину элемента. Этот единственный продукт, вероятно, ответствен за нахождение сайта-мишени и узнавание концов транспозона. Для

транспозона, используемого в качестве субстрата при транспозиции, необходимы только концы. Функции хозяина могут быть нужны для репликации транспозона. О деталях процессов и порядке событий, включаемых в эти реакции, нам известно немного, однако можно предположить их сходство для всех IS-элементов и сложных транспозонов (пожалуй, за исключением IS1).

Только один модуль транспозона Tn10 функциональный

Об эволюции сложного транспозона можно судить при сравнении модулей уже не идентичных, но еще сохраняющих родство. В этом отношении больше всего известно о транспозонах Tn10 и Tn5.

Для узнавания элемента в качестве субстрата транспозиции необходимы несколько оснований в каждом конце транспозона Tn10. Инвертированные концевые повторы имеют длину, равную 22 парам оснований; из них, вероятно, только последние 13 важны для реакции транспозиции. Мутации в этой области предотвращают транспозицию и являются *cis*-активными.

Tn10 — необычный элемент, имеющий специфическую последовательность в сайте-мишени. Прямые фланкирующие повторы из 9 пар оснований, которые образуются при транспозиции, обнаруживают обобщенную последовательность из 6 нуклеотидных пар, симметрично расположенную в сайте-мишени. Так, повторы с каждой

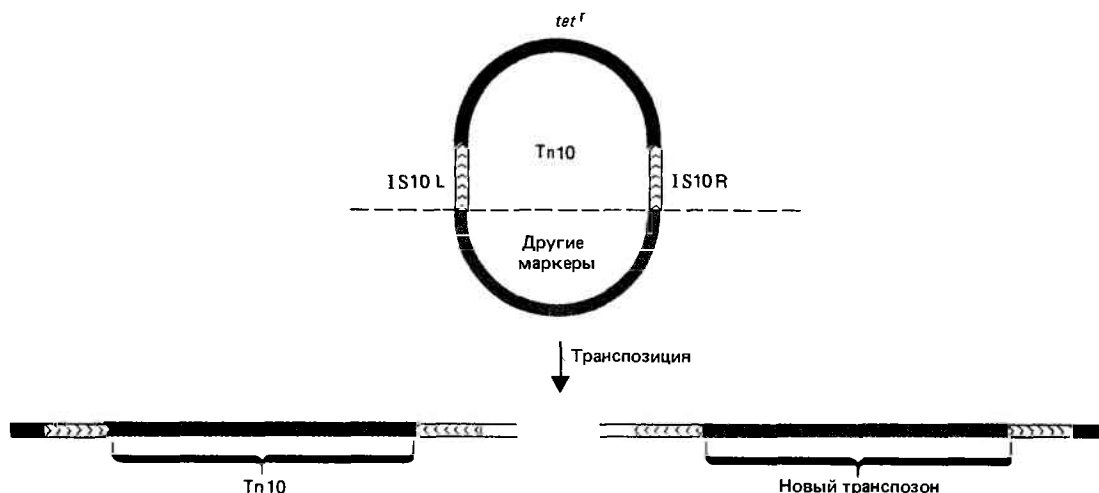


Рис. 36.4. Модули IS10 способны кооперироваться и транспозировать любую область ДНК, которая лежит между ними. Если Tn10 входит в состав кольцевой молекулы, повторы IS10 могут

обеспечить транспозицию любой стороны кольца. Транспозиция tet^R соответствует перемещению Tn10. Транспозиция маркеров, находящихся на другой стороне, создает новый «вывернутый» транспозон.

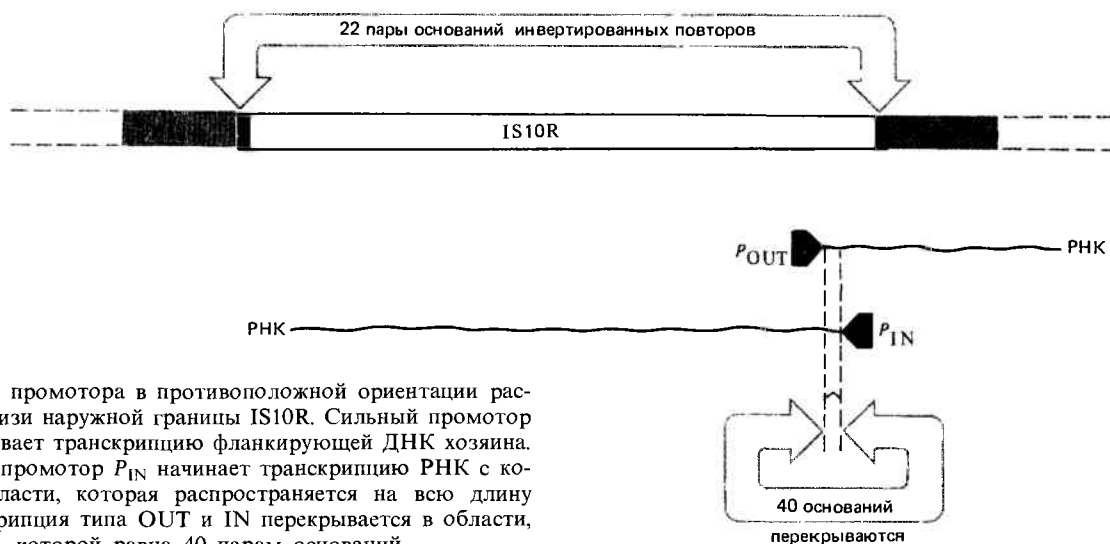


Рис. 36.5. Два промотора в противоположной ориентации расположены вблизи наружной границы IS10R. Сильный промотор P_{OUT} обеспечивает транскрипцию фланкирующей ДНК хозяина. Более слабый промотор P_{IN} начинает транскрипцию РНК с кодирующей области, которая распространяется на всю длину IS10R. Транскрипция типа OUT и IN перекрывается в области, протяженность которой равна 40 парам оснований.

стороны Tn10 часто представлены последовательностями NGCTNAGCN NCGANTCGN, где N соответствует любой паре оснований. Чем «сильнее выражена» горячая точка, тем больше она соответствует канонической последовательности. Не исключено, что одна и та же Tn10-кодируемая функция узнает последовательность мишени и концы транспозона. Активным модулем Tn10 служит элемент IS10R. Модуль IS10L функционально дефектен и обеспечивает только 1–10% транспозиционной активности IS10R. Накопление мутаций в IS10L по сравнению с IS10R и обуславливает их примерно 2,5%-ную дивергенцию.

На рис. 36.5 показана организация IS10R. Он имеет непрерывную рамку считывания, которая кодирует существенный белок с мол. массой 47 000 дальтон; значение другой перекрывающейся рамки считывания неизвестно. Уровень продукции существенного белка повышается благодаря мутациям, увеличивающим степень выражения гена IS10R. Следовательно, количество этого белка может ограничивать скорость транспозиции.

Мутации в этом гене могут быть комплементированы в *транс*-положении другим элементом дикого типа, однако эффективность комплементации выражена слабо. Это свидетельствует о предпочтительном *цис*-действии белка. По непонятной причине белок эффективно функционирует только на той ДНК-матрице, с которой он был транскрибирован и транслирован. Это общее свойство белков, участвующих в IS-опосредованной транспозиции. (Мы уже говорили о *цис*-предпочтительности белка А, участвующего в репликации фХ174, гл. 33.)

При увеличении расстояния между геном транспозазы и концами транспозона IS10 эффективность действия транспозазы падает. Причина этого явления, по-видимому, заключается в предпочтительном узнавании транспозазой последовательностей ДНК мишени, расположенных вблизи того сайта, в котором она синтезируется. Можно предположить, что транспозаза настолько быстро связывается с ДНК после (или во время) белкового синтеза, что не успевает диффундировать. Реальная и другая ситуация: транспозаза может отличаться такой нестабильностью, что белковые молекулы, неспособные быстро связываться, никогда не смогут стать активными. Вблизи внешней границы IS10R обнаружено два про-

мотора. Один из них P_{IN} ответствен за транскрипцию IS10R, другой — P_{OUT} вызывает транскрипцию, протекающую в смежной фланкирующей ДНК. Возможно, этот эффект связан с активацией бактериальных генов, сцепленных с Tn10.

Мультикопийное ингибирование свидетельствует о регуляции экспрессии гена IS10R. Транспозиция Tn10-элемента на бактериальной хромосоме уменьшается, если в составе мультикопийной плазмиды ввести в клетки дополнительные копии IS10R. Ингибирование требует присутствия последовательности промотора P_{OUT} и эффективно только в отношении гена IS10R, экспрессируемого при участии P_{IN} -промотора (если ген находится под контролем другого промотора, эффект увеличивается). Однако мультикопийное ингибирование действует на уровне трансляции, а не транскрипции. Этот эффект может объясняться наличием перекрывания на участке в 40 пар оснований 5'-концевых областей транскриптов, инициируемых в P_{IN} и P_{OUT} . Так как P_{OUT} является более сильным промотором, может образовываться большее количество OUT-РНК, и эта РНК может в свою очередь спариваться с IN-РНК, предотвращая трансляцию белка, необходимого для транспозиции.

Модули транспозона Tn5 почти идентичны, однако функционально очень различаются

Инвертированные модули Tn5 служат удивительным примером того, как небольшие мутационные изменения вызывают важные функциональные эффекты. Модуль IS50R функциональный; модуль IS50L нефункциональный в отношении транспозиции. Единственное различие между ними обусловлено заменой одной пары оснований. Результаты этой замены демонстрируются на рис. 36.6.

В модуле IS50R при использовании одной и той же рамки считывания образуются два белка. Единственное различие между ними состоит в наличии у белка 1 примерно 40 дополнительных N-концевых аминокислот, отсутствующих у белка 2. Точные события, включаемые в транскрипцию и трансляцию этих белков, еще не установлены; известно только, что белок 2 образуется в больших количествах, чем белок 1.

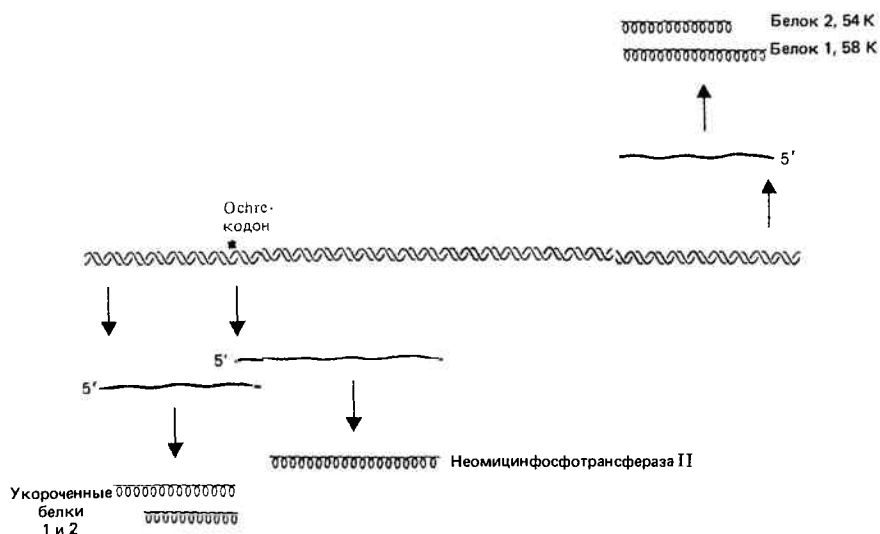


Рис. 36.6. Замена одной пары оснований в IS50L преждевременно терминирует синтез белков 1 и 2 и создает промотор для транскрипции гена неомидин- и канамицин-резистентности.

Замена одной пары оснований в IS50L одновременно влияет на трансляцию этих белков и контролирует транскрипцию центральной области. Замещение ведет к образованию *ochre*-кодона, который преждевременно терминирует трансляцию как белка 1, так и белка 2. Урезанные белки (иногда называемые белками 3 и 4) утрачивают транспозиционную активность. То же самое замещение создает промотор для транскрипции гена центральной области, который кодирует неомидинфосфотрансферазу II, фермент, ответственный за устойчивость к таким антибиотикам, как неомидин и канамицин. Следовательно, для проявления этой функции транспозона такая замена необходима.

Функции белков 1 и 2 родственны, но различны. Белок 1 существен для транспозиции любого модуля IS50 или интактного транспозона Tn5. Белок дикого типа комплементирует дефектный модуль в *транс*-положении крайне неэффективно. Следовательно, белок 1 может быть *цис*-активной функцией транспозиции.

Белок 2 служит ингибитором транспозиции. Его действие осуществляется на какой-то еще неизвестной стадии самого процесса транспозиции (а не путем регуляции экспрессии гена). Белок 2 функционирует в *транс*-положении; механизм его действия может заключаться в связывании с теми же сайтами, которые должны узнаваться белком 1 для осуществления его транспозиционной функции. Таким образом, эта функция будет подавлена. Другая возможность — образование белками 1 и 2 олигомерного комплекса, в котором активность белка 1 также окажется подавленной.

Поведение белка 2 может объяснить свойство транспозона Tn5. При введении в клетки нового бактериального хозяина Tn5 транспозируется с высокой частотой.

Внедренный элемент способен также подавлять транспозицию входящего элемента. Такая *транс*-действующая функция может отражать способность белка 2 препятствовать транспозиции; при вхождении нового элемента в клетку в ней уже присутствует такое количество белка 2, которое достаточно, чтобы ингибировать *цис*-действующую функцию белка 1 как нового, так и старого Tn5-элемента. Вполне возможно, что укороченные белки IS50L тоже вовлечены в этот эффект.

Взаимосвязь между белками 2 и 1 транспозона Tn5 отличается по природе от взаимосвязи между транскриптами P_{OUT} и P_{IN} транспозона Tn10. Однако их объединяет общее свойство: они ограничивают частоту транспо-

зиции. Механизмы, которые ограничивают транспозоны в размножении, могут играть важную роль в предотвращении слишком сильного вмешательства транспозонов в клеточные функции.

Транспозиция включает репликативную рекомбинацию

Термин «транспозиция» является, по существу, неправильным, поскольку транспозон не покидает своего исходного сайта. Процесс транспозиции включает дупликацию транспозона; при этом одна копия сохраняется в исходном сайте, а другая обнаруживается в новом. Следовательно, транспозиция сопровождается увеличением числа копий транспозона.

Природа транспозиционного события, по-видимому, одинакова для всех транспозонов; об этом свидетельствует наличие повторов коротких последовательностей ДНК мишени. Транспозиция включает два типа реакций, которые в принципе могут происходить в любом порядке. Одна из них связана с репликацией транспозона без репликации прилегающих хромосомных последовательностей. Другая требует разрыва последовательности ДНК мишени, приводящего к образованию сайта, в который и внедряется транспозон. Ни на одной из стадий в этом процессе ДНК транспозона не существует в виде свободной молекулы; все события транспозиции происходят на ДНК, в пределах которой находится транспозон.

Основные этапы реакции внедрения показаны на рис. 36.7. Они включают образование ступенчатого разреза в ДНК мишени, соединение транспозона с образующимися одноцепочечными концами и заполнение брешей. Образование и достройка ступенчатых концов делают понятным наличие прямых повторов ДНК мишени в сайте внедрения. Расстояние между разрезами на двух цепях определяет длину прямых повторов; следовательно, повтор последовательности мишени, характеризующий каждый транспозон, отражает способ действия фермента, вовлекаемого в процесс разрезания ДНК мишени.

Общий конечный результат транспозиции маскирует наличие более чем одного механизма дупликации транспозона и его внедрения. Существование по крайней мере двух основных типов транспозиции показано как различиями в самих транспозиционных событиях, так и различиями между продуктами реакции.

Наряду с «простой» межмолекулярной транспозицией,

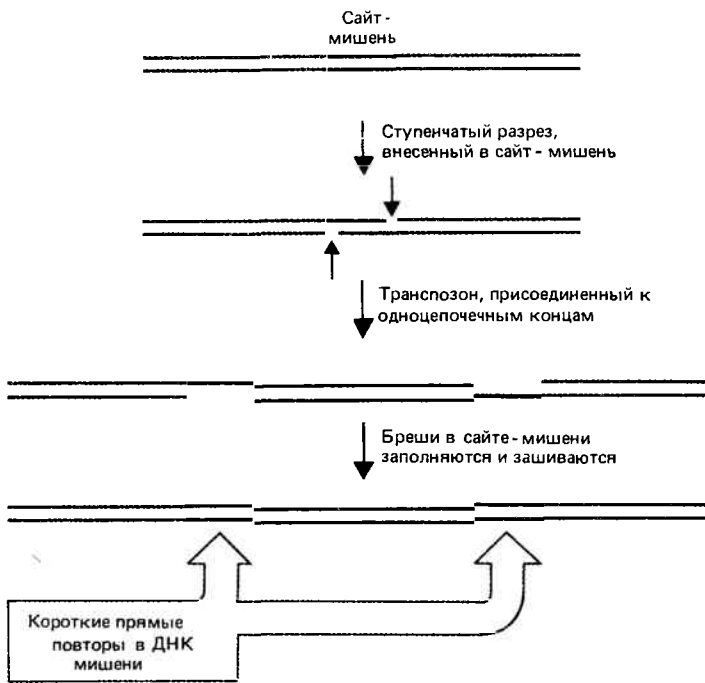


Рис. 36.7. Прямые повторы ДНК мишени, фланкирующей транспозон, формируются в результате введения ступенчатых разрезов; образовавшиеся выступающие концы соединяются с транспозоном.

которая ведет к внедрению дополнительной копии в новый сайт, транспозоны способствуют другим типам перестроек в ДНК. Некоторые из этих событий могут представлять альтернативные результаты работы транспозиционного механизма; они дают ключ к решению вопроса о природе этого механизма. Другие события являются следствием взаимоотношений между множеством копий транспозона.

Перестройки в хозяйской ДНК могут возникать в тех случаях, когда вторая копия транспозона внедряется в другой сайт вблизи исходного. Системы клетки-хозяина могут осуществить реципрокную рекомбинацию между копиями транспозона; последствия этих событий зависят от того, находятся повторы в прямой или инвертированной ориентации.

На рис. 36.8 показано, что рекомбинация между любой парой прямых повторов будет приводить к делетированию материала между ними. Промежуточная область вырезается в виде кольцевой ДНК (которая утрачивается клеткой); хромосома сохраняет одну копию прямого повтора. Следует отметить, что рекомбинация между прямо повторенными модулями IS1 сложного транспозона Tn9 будет поэтому заменять транспозон на один модуль IS1. Делетирование последовательности, прилегающей к транспозону, может происходить в результате двухступенчатого процесса; сначала в результате транспозиции возникает прямой повтор последовательности транспозона, а затем между повторами происходит рекомбинация. Однако большинство делеций, которые возникают вблизи транспозонов, являются результатом вариации в процессе транспозиции.

На рис. 36.9 изображены последствия реципрокной рекомбинации между парой инвертированных повторов. Область между повторами становится инвертированной; сами повторы сохраняют способность стимулировать

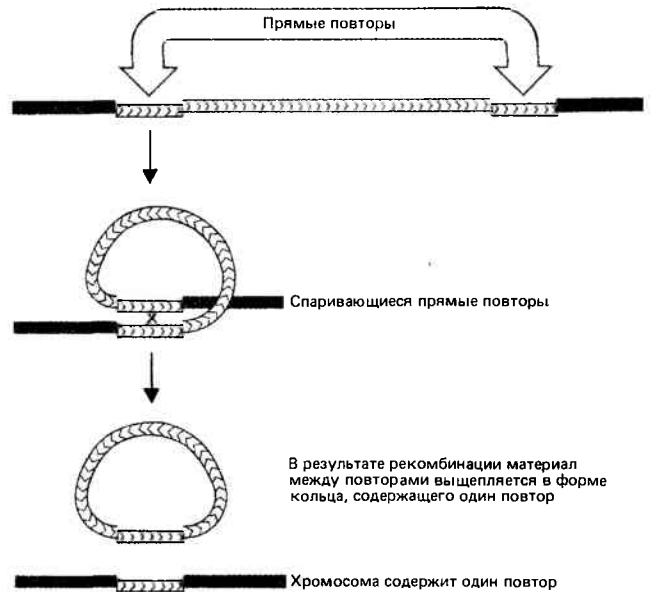


Рис. 36.8. Реципрокная рекомбинация между прямыми повторами ведет к вырезанию материала, находящегося между ними; каждый продукт рекомбинации содержит одну копию прямого повтора.

дальнейшие инверсии. **Двойные инверсии** связаны с транспозицией и принимают форму транспозонов, находящихся в инвертированной ориентации с каждой стороны от центральной области, которая тоже была инвертирована по отношению к ее обычной ориентации. Заметим, что сложный транспозон, чьи модули инвертированы, представляет собой стабильный компонент генома, хотя направление центральной области по отношению к модулям может быть инвертировано в результате рекомбинации.

В списке событий, происходящих с транспозоном важное место занимает процесс его **исключения**. Утрату транспозона можно определить по восстановлению функ-

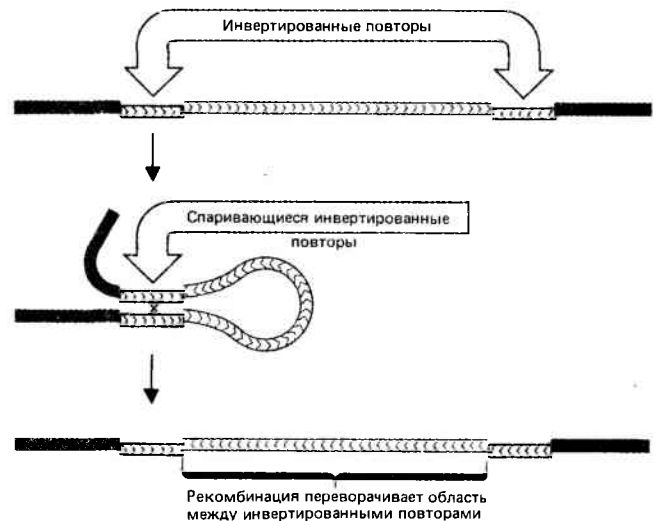


Рис. 36.9. Реципрокная рекомбинация между инвертированными повторами переворачивает область между ними.

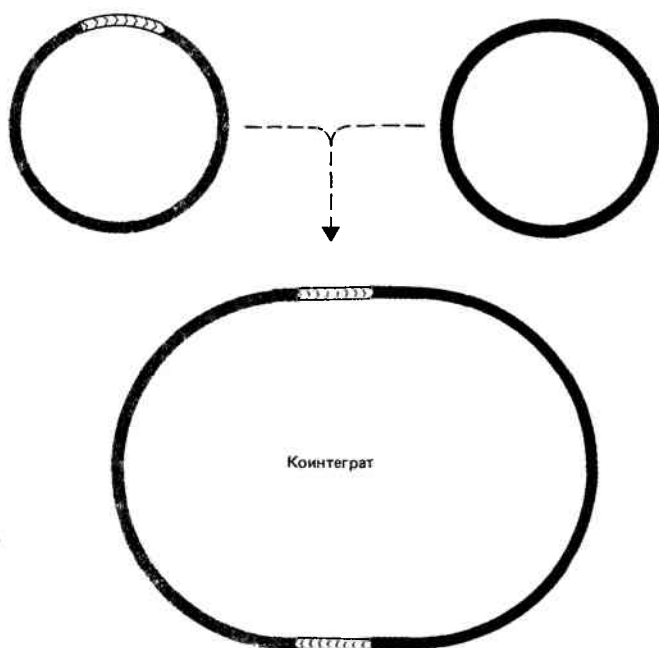


Рис. 36.10. Транспозиция может вести к слиянию донорного и реципиентного репликонов с образованием коинтеграта.

пии в сайте внедрения. Такое восстановление происходит только при **точном исключении**; частота этого события примерно 10^{-6} для транспозона Tn5. Кроме того, потеря транспозона может быть определена по снятию полярного эффекта в смежных генах; это бывает и в результате **неточного исключения** (которое оставляет часть транспозона, достаточную для предотвращения реактивации гена, служащего мишенью, но недостаточную для того, чтобы обуславливать полярный эффект). Неточное исключение встречается с частотой примерно 10^{-5} для Tn5. Ни один тип исключения не связан с транспозицией в новые сайты. Процесс не зависит от функций, кодируемых транспозоном; он является RecA-независимым и может осуществляться с помощью какого-то клеточного механизма, порождающего спонтанные делеции между близко расположенными повторяющимися последовательностями.

Одной из важнейших реакций, опосредованных транспозонами, является **слияние репликонов** с образованием **коинтегрированной структуры**. Репликон, содержащий транспозон, может слиться с репликоном, не несущим элемент, как это показано на **рис. 36.10**. Образующийся коинтеграт имеет две копии транспозона, по одной в каждой точке соединения между исходными репликонами, ориентированными в виде прямых повторов. Коинтеграты, образуемые с помощью сложных транспозонов, могут иметь дублированным либо весь транспозон (Tn-стимулируемые коинтеграты), либо модуль IS (IS-стимулируемые коинтеграты).

Частоты, с которыми различные транспозоны образуют коинтеграты, свидетельствуют о существовании различных путей транспозиции. Например, Tn903 и IS1 могут примерно с равной частотой осуществлять транспозицию и формировать коинтеграты. Однако у Tn5 и Tn10 транспозиция происходит примерно в 50–200 раз чаще, чем образование коинтегрированных структур. Коинтеграты, образуемые при участии Tn5, отличаются стабильностью. Это позволяет предположить, что их

образование – результат транспозиции, а не промежуточный продукт этого процесса.

Модель транспозиции, по которой коинтеграт представляет собой обязательный промежуточный продукт, иллюстрирована на **рис. 36.11**. Четыре одноцепочечных разреза инициируют процесс. Донорная молекула разрезается в каждом конце транспозона с помощью сайтспецифического фермента, который узнает его конец. Молекула мишени разрезается в сайтах, расположенных ступенчато, обычно на расстоянии 5 или 9 оснований друг от друга. Порядок и время, в которое производится разрез во время транспозиции, могут быть изменены. Разрезы в цепях донора и мишени сшиваются. Каждый конец последовательности транспозона присоединяется к одному из одноцепочечных концов, образовавшихся в сайте мишени. В результате образуется chi-образная структура, удерживаемая двухцепочечной ДНК транспозона. Образование такой структуры приводит к перемещению транспозона; ферментативная активность, ответственная за такое перемещение, получила название **транспозазы**.

Chi-образная структура содержит одноцепочечную область в каждом из ступенчато разрезанных концов. Эти области используются в качестве псевдорепликационных вилок, обеспечивающих матрицу для синтеза ДНК. (Использование концов в качестве затравок для репликации означает, что должен иметь место разрыв цепи с полярностью, в результате чего в этой точке образуется 3'-конец.) Если репликация продолжается от обеих репликационных вилок, она идет через транспозон, разделяет его цепи и терминируется в его концах. Репликация выполняется, по-видимому, с помощью ферментов, кодируемых клеткой-хозяином. По окончании репликации структура становится коинтегратом, обладающим прямыми повторами транспозона в местах соединения между репликонами (в чем можно убедиться при анализе последовательности, составляющей коинтеграт).

Сайтспецифическая рекомбинация между двумя копиями транспозона может восстановить исходный донорный репликон, освободив при этом репликон мишени, который приобрел транспозон. Такое событие получило название **разрешения**; фермент, ответственный за эту активность, называется **резолвазой**.

Транспозиция Tn3 происходит путем разрешения коинтеграта

Существуют большие (~ 5 т.п.н.) элементы, транспозиция которых не зависит от модулей IS-типа. Они представляют собой независимые единицы, несущие как гены, кодирующие функции транспозиции, так и гены, кодирующие резистентность к антибиотикам. К наиболее изученным транспозонам такого типа относятся представители семейства TnA. Первоначально полагали, что оно включает только один транспозон, однако их обнаружено много (**табл. 36.3**).

Транспозиция этих элементов осуществляется в две ступени: сначала образуются коинтеграты, а затем эти коинтеграты разрешаются. Каждая стадия проходит при участии различных продуктов, кодируемых транспозоном (в отличие от единственного белка транспозиции IS-элемента). Так же как у элементов типа IS, в этом случае в стадии транспозиции участвуют концы элемента. Уникальным свойством транспозонов TnA-семейства являет-

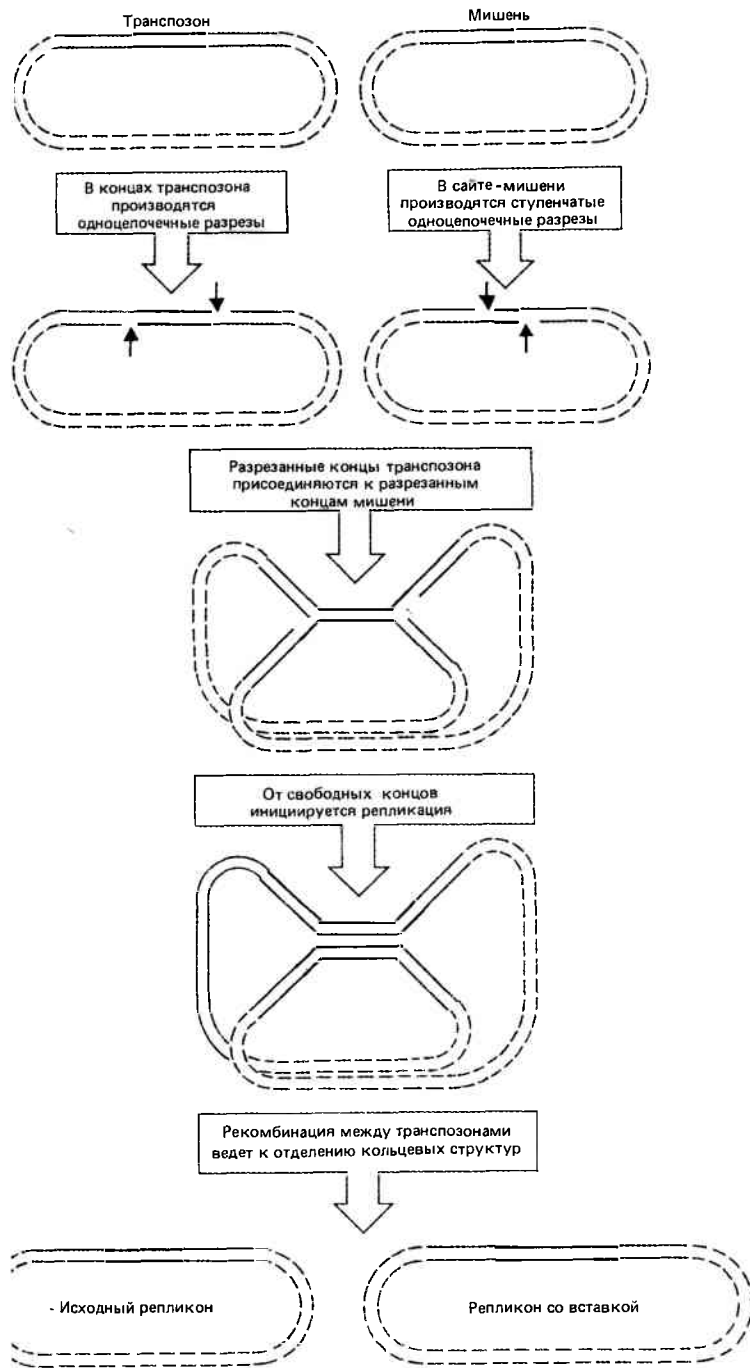


Рис. 36.11. Один из путей транспозиции ведет к образованию коинтегратов, которые разрешаются рекомбинацией.

Таблица 36.3
Семейство транспозонов TnA представлено большими, но тем не менее индивидуальными единицами

Элемент	Длина, п. н.	Генетические маркеры	Инvertированные концевые повторы, п. н.	Фланкирующие повторы мишени, п. н.	Выбор мишени
Tn3	4257	<i>amp^R</i>	38	5	Региональный
Tn1	5000	<i>amp^R</i>	?	5	»
$\gamma\delta$	5800	?	37	5	»
Tn501	8200	<i>Hg^R</i>	38	5	Неизвестный

ся то, что их стадия разрешения требует специфического внутреннего сайта.

Транспозоны TnA обычно на концах имеют близкородственные инvertированные повторы длиной около 38 пар оснований. *Цис*-активные делеции в любом из них предотвращают транспозицию элемента. В сайте-мишени образуются прямые повторы из 5 пар оснований, выбираемых случайно, хотя и с региональной предпочтительностью.

С помощью рецессивных мутаций идентифицированы два гена, продукты которых необходимы для транспозиции. Так как мутантные гены могут быть комплементированы в *транс*-положении генами дикого типа другой копии транспозона, их действие не проявляет *цис*-активной предпочтительности, характерной для транспозиционного процесса IS-типа.

Мутанты по гену *tnpA* не способны осуществлять транспозицию. Продукт этого гена – транспозаза – состоит из 1021 аминокислоты (примерно 120 000 дальтон). Очищенный белок способен связываться с одноцепочечной ДНК. О его функции известно немного, хотя показано, что он необходим для узнавания концов элемента и для осуществления ступенчатого разреза с промежутком между разрезами на отдельных цепях в 5 пар оснований. В месте ступенчатого разреза происходит соединение концов транспозона с ДНК мишени. Транспозазы Tn3 и родственного элемента $\gamma\delta$ эффективно работают только на свой собственный тип транспозона, хотя концы элементов одинаковы на участке в 27 из 38 пар оснований инvertированных повторов. Возможно, что некоторые из отличающихся позиций представляют собой критические точки при узнавании концов.

Ген *tnpR* кодирует белок из 185 аминокислот (23 000 дальтон). Различный эффект мутаций в этом гене свидетельствует о двойной функции кодируемого им белка. Этот белок действует как репрессор экспрессии гена и обеспечивает функцию резолвазы.

Мутации в гене *tnpR* увеличивают частоту транспозиции. Причина такого эффекта заключается в том, что TnpR репрессирует транскрипцию как *tnpA*, так и своего собственного гена. В результате инактивация белка TnpR ведет к увеличению синтеза белка TnpA, что в свою очередь обуславливает возрастание частоты транспозиции. Следовательно, количество транспозазы TnpA служит ограничивающим фактором в транспозиции.

Гены *tnpA* и *tnpR* транскрибируются дивергентно от А–Т-богатой межцистронной контролирующей области длиной в 163 пары оснований, как показано на карте транспозона Tn3, представленной на рис. 36.12. (Другие транспозоны TnA-типа имеют такую же организацию.) Оба эффекта белка TnpR проявляются при его связывании в этой области.

Белок TnpR, функционируя в качестве резолвазы, вовлекается в рекомбинацию между прямыми повторами Tn3 в структуре коинтеграта. Коинтеграт представляет собой обязательный промежуточный продукт при транспозиции; разрешение коинтегратов блокируется в мутантах *tnpR⁻*, что ведет к образованию только коинтегрированных структур.

В принципе коинтеграты могут быть разрешены с помощью гомологичной рекомбинации между любой соответствующей парой точек в двух копиях транспозона (рис. 36.11). Однако на деле реакция разрешения происходит только в специфическом сайте. Этот сайт первона-

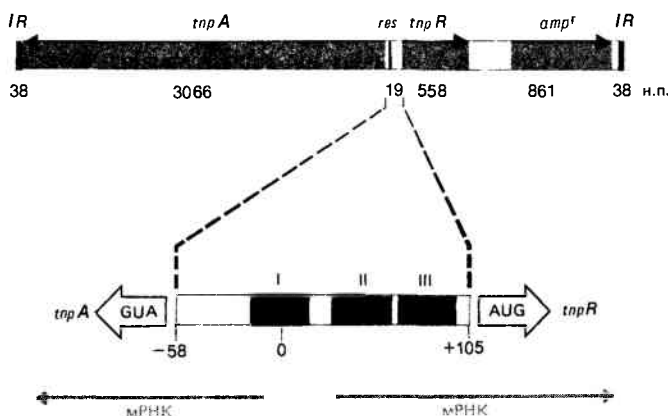


Рис. 36.12. Транспозоны, подобные Tn3, имеют инвертированные концевые повторы, внутренний сайт *res* и три известных гена.

Стрелками указывается направление выражения кодирующих областей. Цифры в верхней части диаграммы указывают длину каждой области. Цифры в нижней части диаграммы показывают расстояния от точки перекреста (0).

начально был идентифицирован *cis*-действующими делециями во внутренней области, которые блокируют завершение транспозиции, вызывая накопление коинтегратов. Сначала сайт был обозначен как *IRS* (от англ. *internal resolution sequence*); теперь он известен как *res*-сайт. В отсутствие *res* реакция разрешения может быть замещена *RecA*-зависимой общей рекомбинацией, но она происходит со значительно меньшей эффективностью.

Сайты связывания резолвазы TnpR удалось определить с помощью метода отпечатков комплекса ДНК—белок. Данные об их локализации суммированы в нижней части рис. 36.12. Связывание происходит независимо в каждом из трех сайтов длиной в 30–40 пар оснований. Сайт I включает область, генетически идентифицированную как *res*-сайт; в его отсутствие реакция разрешения не происходит вообще. Однако разрешение включает также связывание в сайтах II и III, поскольку при делеции любого из них реакция разрешения идет с низкой эффективностью. По-видимому, сайт I перекрывается со стартовой точкой транскрипции *tnpA*. Сайт II перекрывается со стартовой точкой транскрипции *tnpR*; мутация оператора картируется в самом левом конце сайта.

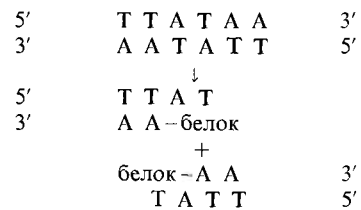
Три сайта связывания в Tn3 несут гомологичную последовательность, которая определяет каноническую (усредненную) последовательность; имеется двойная симметрия в сайтах. Как они взаимодействуют между собой? Одна из возможностей сводится к тому, что связывание во всех трех сайтах необходимо для удержания определенной топологической конформации в ДНК; реакция разрешения может начаться, когда комплекс ДНК—белок одного транспозона будет взаимодействовать с таким же комплексом другого транспозона. Связывание только в одном наборе сайтов может репрессировать транскрипцию генов *tnpA* и *tnpR*, не вызывая каких-либо изменений в ДНК.

Область, в которой происходит сайтспецифическая рекомбинация, впервые была идентифицирована при сравнении реакций транспозонов Tn3 и $\gamma\delta$, у которых гены *tnpR* проявляют высокую гомологию и могут замещать

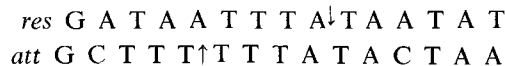
друг друга в комплементарных исследованиях. Последовательности рекомбинантов свидетельствуют о том, что реакция разрешения происходит в пределах 19 пар оснований сайта I, идентичных в обоих транспозонах.

Значительно больше информации о реакции разрешения было получено в опытах по разрешению коинтегратов *in vitro*. В качестве субстрата в этих экспериментах использовали коинтеграт-подобную молекулу ДНК. (Субстрат был сконструирован путем клонирования с образованием прямых повторов в двух областях плазмиды pBR322.) Субстрат должен иметь суперспиральную структуру; его разрешение ведет к образованию сцепленных колец, каждое из которых содержит один *res*-сайт. Реакция требует присутствия больших количеств резолвазы TnpR (которую получают в нужном количестве благодаря применению стандартных методов клонирования). Разрешение происходит наиболее эффективно при соотношении фермент/ДНК, равном примерно 20/1. Факторы хозяина в этой реакции не требуются (в противоположность сайтспецифической рекомбинации фага лямбда).

Разрешение—консервативный процесс; связи при этом разрываются и восстанавливаются без поглощения энергии. Для завершения реакций необходимы ионы магния; если они отсутствуют, происходит лишь частичная реакция с низкой эффективностью. В результате такой реакции получают продукты, соответствующие промежуточной стадии в разрешении коинтегратов; они состоят из резолвазы, ковалентно связанной с обоими концами, образовавшимися при двухцепочечном разрезе, произведенном в *res*-сайте. Разрезание производится симметрично в коротких палиндромных областях; в результате образуются выступы из двух оснований. Развивая представление об области перекреста, расположенной в сайте I, мы можем описать реакцию разрезания следующим образом:



На основании родства сайта *res* с кором *att*-сайта фага лямбда (гл. 35) можно предположить наличие сходства между этой системой и сайтспецифической рекомбинацией фага лямбда. Между двумя сайтами существует гомология в 10 из 15 пар оснований, непосредственно окружающих точку перекреста. Ниже представлены эти последовательности:



Стрелками показаны сайты, в которых белок TnpR разрезает локус *res*, а белок Int—*att*. Сами по себе реакции аналогичны, хотя здесь разрешение происходит только между внутримолекулярными сайтами, в то время как рекомбинация между *att*-сайтами является межмолекулярной и направленной (что обусловлено различиями *att*-сайтов). В обоих случаях очевидно сходство с реакциями

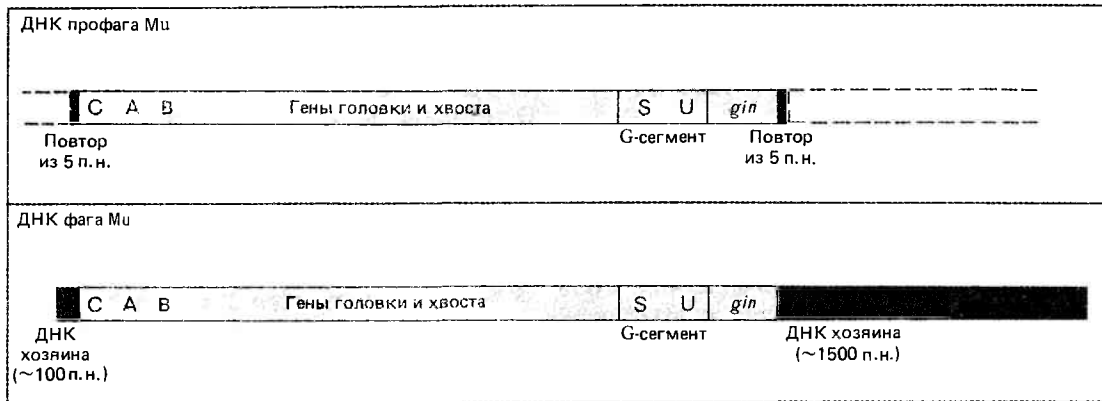


Рис. 36.13. Порядок генов в профаге и свободной ДНК фага Mu одинаков, но концы различны. ДНК профага фланкирована прямыми повторами из 5 пар оснований ДНК хозяина. Свобод-

ная фаговая ДНК содержит различные дополнительные последовательности в левом конце длиной в 100 пар оснований и в правом конце (1500 пар оснований).

разрезания и воссоединения, осуществляемыми топоизомеразами.

Наконец мы должны упомянуть поразительный эффект транспозона Tn3, получивший название **иммунитета транспозиции**. Плазмида, несущая транспозон Tn3, иммунна к внедрениям любого Tn3-элемента, вносимого на другой молекуле ДНК (хотя Tn3, по-видимому, способен транспозироваться в пределах плазмиды). Существует определенный лаг-период между внедрением Tn3 в новую плазмиду и установлением иммунитета. Эффект специфичен; он относится только к Tn3 и не предотвращает внедрения других транспозонов в плазмиду. Природа такого *cis*-действия на далеком расстоянии остается загадочной.

Некоторые необычные свойства транспозирующего фага Mu

Фаг Mu может находиться в клетке в альтернативных состояниях, осуществляя литический цикл или лизогенизируя клетку. Он был открыт благодаря его способности вызывать мутации в *E. coli*. Мутации возникают при внедрении фагового генома, содержащего 37 т.п.н., в случайные сайты (с региональной специфичностью) бактериальной ДНК. При этом реверсий с доступной определению частотой не происходит.

Карта генома фага Mu показана на рис. 36.13. Внедренный геном Mu имеет тот же порядок генов, что и свободная фаговая ДНК, которая имеет линейную форму. (Следует отметить существующее отличие от фага лямбда, свободная линейная ДНК которого при инфекции образует кольцевую молекулу; в результате интеграции образуется линейный профаг, порядок генов которого является пермутированным относительно порядка генов свободной ДНК фага.) Линейные геномы фага Mu не содержат ни липких концов, ни повторов; поэтому не ясно, каким образом осуществляется согласованное действие концов во время интеграции. Возможно, что реакция зависит от гомологии, которая встречается на расстоянии примерно 100 пар оснований от каждого конца. Существование механизма, узнающего специфические концы интегрированной последовательности фага Mu, было обнаружено при открытии способности этого фага вырезаться. Реакция происходит только в том случае, если профаг Mu приобретает IS1-элемент. Механизм вырезания неизвестен, однако установлено, что в него вклю-

чены как функции, кодируемые Mu, так и функции, контролируемые IS1-элементом.

Концы свободной ДНК фага Mu обладают дополнительным материалом, который не является частью фагового генома. Фактически концы представлены последовательностями генома хозяина. Левый конец содержит около 100, а правый — около 1500 пар оснований (они обозначены соответственно как *s*- и *S*-концы). В отдельных молекулах ДНК фага Mu эти последовательности различны; они приобретаются при упаковке фага из предшественников, в которых фаговый геном фланкирован последовательностями хозяина. При внедрении фагового генома в бактериальную хромосому дополнительные области теряются.

Инфицирующая родительская ДНК фага Mu внедряется в геном хозяина неэффективно. Большинство молекул ДНК во время литического цикла сохраняет свои свободные концы. Менее 10% инфицирующих фаговых геномов внедряются. Однако во время репродуктивного цикла последовательности ДНК фага Mu внедряются эффективно. Этот парадокс показывает, что копии ДНК Mu, образуемые при репликации, способны внедряться в геном хозяина. В родственной реакции индукция профага Mu, ведущая к вступлению в литический цикл, не сопровождается вырезанием фаговой ДНК из хромосомы хозяина. Снова ведущую роль в этом цикле играет продукт репликации.

Связь между интеграцией и репликацией обнаружена при изучении мутантов фага Mu, неспособных к интеграции. Оказалось, что и репликация у этих мутантов нарушена. Мутации, определяющие такую дефектность фага, локализованы в одном из двух генов, A или B, необходимых для репликации фага. Таким образом, интеграция, по-видимому, зависит от функции репликации, а не от специальных рекомбинационных ферментов (как у фага лямбда).

Существуют два пути внедрения фага Mu. В результате простых внедрений Mu образуются лизогенные бактерии, содержащие профаги. Обнаружено присутствие прямых повторов из 5 пар оснований с каждой стороны от профага, что указывает на сходство этого события с транспозицией. Кроме того, в процессе литического развития почти все продукты Mu-опосредованной транспозиции имеют структуру коинтегратов. Нам не известно, являются ли простые внедрения и коинтеграты ре-

зультатом альтернативных взаимодействий ДНК фага Mu с ДНК мишени или они представляют собой разные конечные продукты одного общего предшественника. Основные механизмы транспозиции фага Mu пока не изучены.

Другое необычное свойство фага Mu связано с наличием в его геноме, вблизи правого конца, сегмента G, или **инвертируемого сегмента**. Он имеет протяженность около 3 т.п.н. и в различных молекулах ДНК фага Mu представлен в различной ориентации. У фага, полученного в процессе литической инфекции штамма *E. coli* K12, G-сегмент всегда находится в ориентации, получившей название G(+). Если же фаг получают в результате индукции, его ДНК может содержать G-сегмент либо в этой же, либо в противоположной, G(-), ориентации. Для осуществления инверсии требуется ген фага Mu, названный *gip* и локализованный сразу же за пределами G-сегмента.

Сегмент G содержит в своем составе гены, которые кодируют белки, вовлекаемые в процесс адсорбции фага. Ориентация сегмента влияет на экспрессию этих генов. В результате такого контроля фаги G(+) и G(-) обладают различной специфичностью в отношении бактериальных штаммов. Зависимость штаммовой специфичности от G-сегмента схематически изображена на **рис. 36.14**.

Сегмент G несет альтернативные наборы генов. В ориентации G(+) экспрессируются гены S и U. Их продукты позволяют фагу адсорбироваться на клетках *E. coli* K12, но не *E. coli* C. В ориентации G(-) выражаются гены S' и U'. Это дает возможность фагу адсорбироваться на клетках *E. coli* C, но не K.

Установлено, что белок S имеет мол. массу 56 000, U—21 000, S'—48 000 и U'—26 000. Характерно, что их общая протяженность требует кодирующей последовательности 4 т.п.н., которая, по существу, длиннее, чем сам G-сегмент. Это свидетельствует о том, что гены должны перекрываться.

Пока не завершено определение последовательности

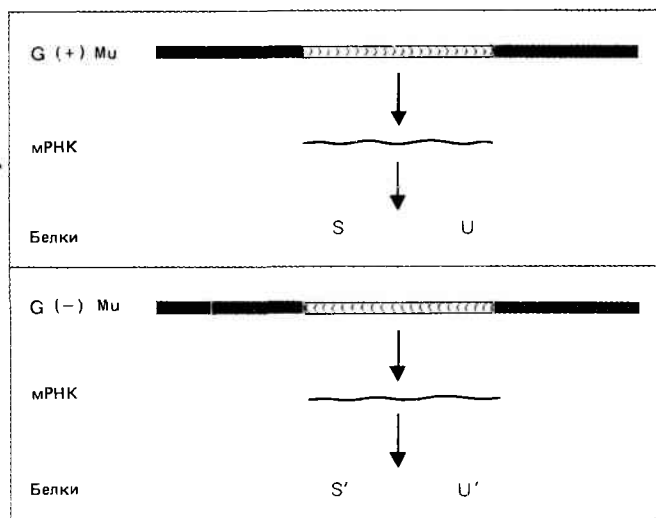


Рис. 36.14. Инверсия G-сегмента определяет, будут ли синтезироваться белки S и U или S' и U'. Одна из моделей предполагает, что их выражение инициируется слева от инвертируемого сегмента. Белки S и U должны кодироваться одной цепью ДНК инвертируемого сегмента, в то время как белки S' и U' кодируются комплементарной цепью.

инвертируемого сегмента, остается нерешенным вопрос и об организации двух наборов генов. Модель, представленная на рис. 36.14, предполагает, что альтернативные наборы находятся на комплементарных цепях ДНК. Ориентация сегмента определяет, какому набору экспрессироваться, поскольку промотор находится вне сегмента и обеспечивает транскрипцию только одной цепи в пределах сегмента.

Фазовая вариация у сальмонелл определяется инверсией

Движение бактерий осуществляется благодаря колебаниям их жгутиков. Многие виды сальмонелл являются **дифазными**, так как они обладают двумя неаллельными генами, контролирующими синтез флагеллина (белковой субъединицы жгутика). Определенный клон бактерий может относиться либо к H1-типу (бактерии находятся в фазе 1), либо к H2-типу (бактерии в фазе 2). Переход от одной фазы к другой происходит примерно один раз на каждую 1000 бактериальных делений, и такая смена фаз получила название **фазовой вариации**.

Гены, кодирующие флагеллины двух типов, локализованы в различных участках хромосомы. Цикл контроля синтеза флагеллина изображен на **рис. 36.15**. Ген H2 тесно сцеплен с другим геном (*rh1*), который кодирует репрессор синтеза H1. Эти два гена экспрессируются координированно. Если функционирует ген H2, экспрессируется и ген репрессора, который предотвращает выражение гена H1. В тех случаях, когда не выражаются ни H2, ни ген репрессора, происходит синтез H1. Таким образом, фаза бактерии определяется активностью транскрипционной единицы H2—*rh1*.

Экспрессия этой транскрипционной единицы зависит от ориентации сегмента ДНК, сцепленного с ней. Сегмент имеет протяженность, равную 995 п.н., и ограничен повторами длиной 14 п.н. (*IRL* и *IRR*). Иницирующий кодон гена H2 лежит на расстоянии 16 п.н. вправо от прилегающего инвертированного повтора. Сегмент ДНК между *IRL* и *IRR* содержит ген *hin*, продукт которого представляет собой **транс-действующий белок**, способствующий инверсии всего сегмента путем реципрокной рекомбинации между инвертированными повторами (рис. 36.9). Мутации в *hin*-гене уменьшают частоту инверсии в 10^4 раз.

Последствия реакции инверсии демонстрируются на **рис. 36.16**. Промотор транскрипционной единицы H2—*rh1* лежит в пределах инвертируемого сегмента. В одной ориентации транскрипция инициируется в промоторе и продолжается в области H2—*rh1*; при этом экспрессируется фаза 2. В другой ориентации промотор обращен в обратную сторону и транскрипционная единица не выражается (хотя транскрипция, по-видимому, происходит только в другом направлении и с неизвестными последствиями). Отсутствие транскрипции H2—*rh1* обуславливает экспрессию фазы 1.

Таким образом, как в случае фазовой вариации, так и штаммовой специфичности Mu инверсия используется для регуляции экспрессии генов. При фазовой вариации процесс включения—выключения обеспечивается перемещением промотора относительно транскрипционной единицы. В случае фага Mu штаммовая специфичность обусловлена инверсией кодирующего сегмента относительно промотора, что приводит к экспрессии альтернативных генов.

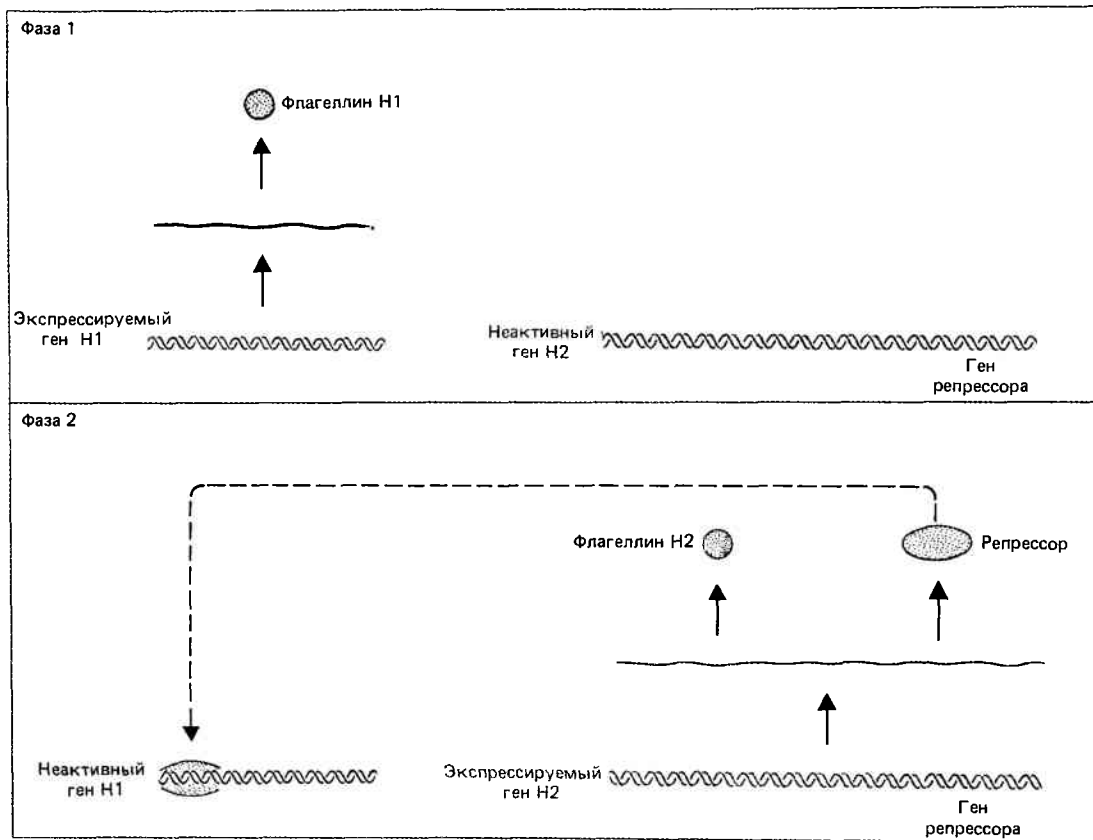


Рис. 36.15. Фаза у сальмонелл определяется активностью транскрипционной единицы Н2. В фазе 1 единица Н2 неактивна, в результате чего синтезируется флагел-

лин Н1. В фазе 2 единица Н2 активна, синтезируются флагеллин Н2 и репрессор. Он и предотвращает выражение гена Н1.

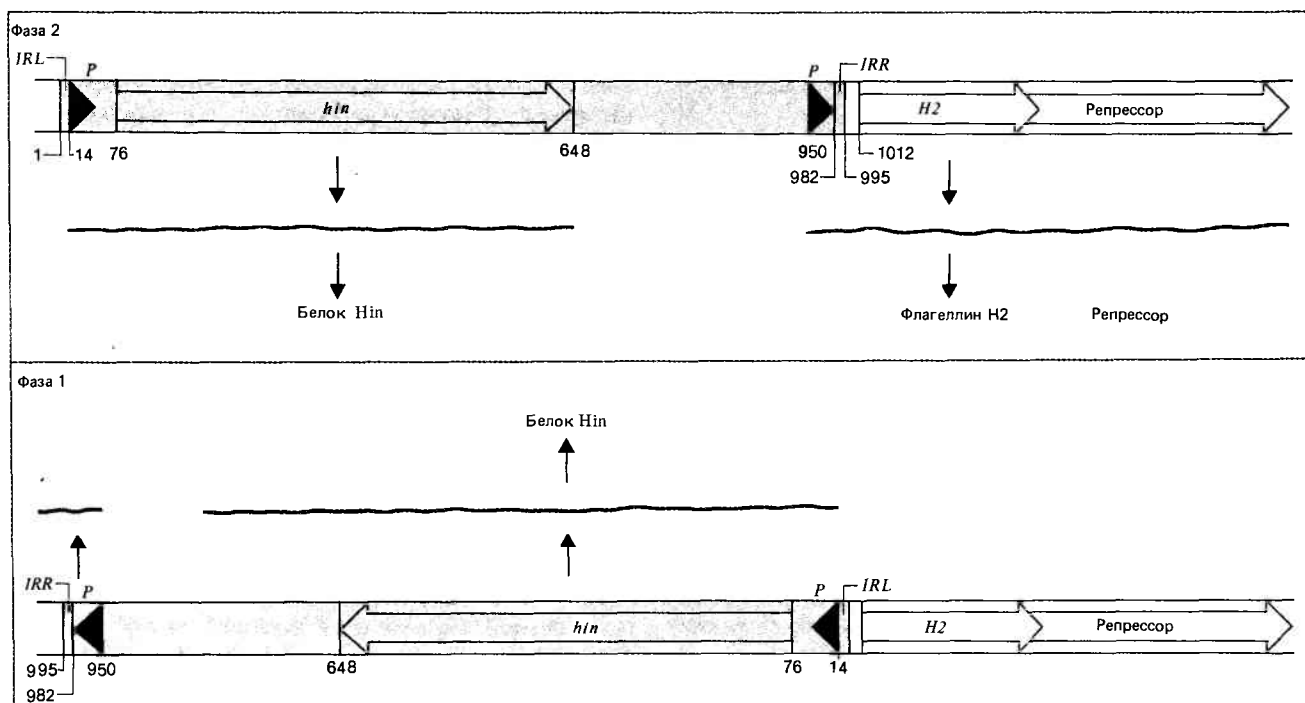


Рис. 36.16. Экспрессия флагеллинового гена Н2 и репрессора фазы 1 контролируется ориентацией инвертируемой области из 995 пар оснований. Фаза 2 проявляется, когда промотор, находящийся в положении 950,

обеспечивает транскрипцию, идущую через последовательность *IRR* в *H2* единицу. Фаза 1 имеет место, когда промотор обращен в противоположную сторону. В обеих фазах ген *hin* экспрессируется со своего собственного промотора (местоположение которого не определено). Цифры указывают пары оснований в пределах инвертируемого сегмента.

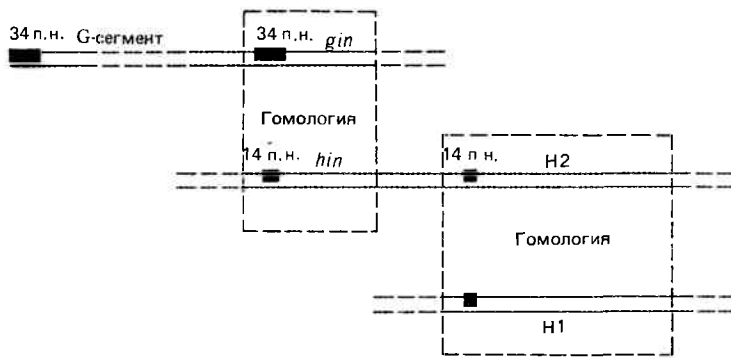


Рис. 36.17. Ген флагеллина H2 гомологичен гену H1, и регуляторная область *hin*, расположенная перед геном H2, гомологична гену *gin* бактериофага Mu.

При сравнении этих реакций друг с другом и с транспозицией Tn3 возникают некоторые параллели. Функция *hin* *Salmonella* и *gin*-функция фага Mu могут заменять одна другую в комплементационных экспериментах. (Эти функции заменяет функция гена *cin*, ответственного за инверсию C-сегмента фага P1). Несмотря на то что ген *tnpR* транспозона Tn3 не способен заместить какую-либо из этих последовательностей, аминокислотная последовательность резолвазы демонстрирует значительную гомологию с последовательностью аминокислот белка Hin.

Область перекреста для действия резолвазы проявляет гомологию с последовательностями *IRL* и *IRR*. Последовательность длиной в 34 пары оснований, включающая *IRL*, имеет только семь несоответствующих пар по сравнению с инвертированными повторами, лежащими в концах G-сегмента. Отсюда следует, что инверсия *hin*-сегмента *Salmonella* и G-сегмента фага Mu происходит с помощью сайтспецифической рекомбинации между инвертированными повторами, механизм которой имеет общие черты с реакцией, участвующей в разрешении коинтегратов Tn3.

Таким образом, эти стационарные инверсионные события могли эволюционировать при взаимобмене между транспозонами и бактериальными локусами, возмож-

но, даже путем наследственной иммобилизации транспозирующихся элементов, функции которых стали ограниченными. Многообразие механизма рекомбинационной инверсии можно видеть при сравнении последовательностей вблизи генов H2, H1, а также последовательности G-сегмента фага Mu. Данные суммированы на рис. 36.17. Флагеллиновые гены H1 и H2 весьма близки, и их гомология распространяется на области, расположенные против направления транскрипции от каждой кодирующей последовательности. Далее в этом же направлении лежит область, включающая ген *hin*, которая имеет тесную гомологию с областью локализации гена *gin* в геноме фага Mu. Таким образом, можно предположить, что между флагеллиновым геном H2 и геном, кодирующим инверсионную функцию фага Mu, произошло наследственное слияние, приведшее к появлению системы контроля фазовой вариации.

Рекомендуемая литература

Развитие представлений о транспозонах может быть прослежено в обзорах Клекнер (Kleckner, *Cell*, **11**, 11–23, 1977; *Ann. Rev. Genet.*, **15**, 341–404, 1981), Калоса и Миллера (Calos, Miller, **20**, 579–595, 1980). Их роль в эволюции была рассмотрена в обзоре Кэмпбелла (Campbell, *Ann. Rev. Microbiol.*, **35**, 55–83, 1981). Оригинальная модель транспозиции, включающая образование коинтегратов, была представлена Шапиро (Shapiro, *Proc. Nat. Acad. Sci. USA*, **76**, 1933–1937, 1979); последовательность и функции Tn3 проанализированы Хеффроном (Heffron *et al.*, *Cell*, **18**, 1153–1163). Две главы в книге “Mobile Genetic Elements” (ed. Shapiro, Academic Press, New York, 1983) посвящены отдельным транспозонам: Tn3 (написано Хеффроном) и Tn10 (автор Клекнер); глава, написанная Сильверманом и Симоном, связана с фазовой вариацией и инверсией сегмента фага Mu. Эти вопросы ранее рассматривались Симоном и соавт. (Simon, *Science*, **209**, 1370–1374, 1980). Зиег и Симон расшифровали последовательность, вовлекаемую в фазовую вариацию (Zieg, Simon, *Proc. Nat. Acad. Sci. USA*, **74**, 4196–4200, 1980). Данные об инвертируемом сегменте G кратко обсуждала в своей работе Хау (Howe, *Cell*, **21**, 605–606, 1980).

Глава 37

МОБИЛЬНЫЕ ЭЛЕМЕНТЫ ЭУКАРИОТ

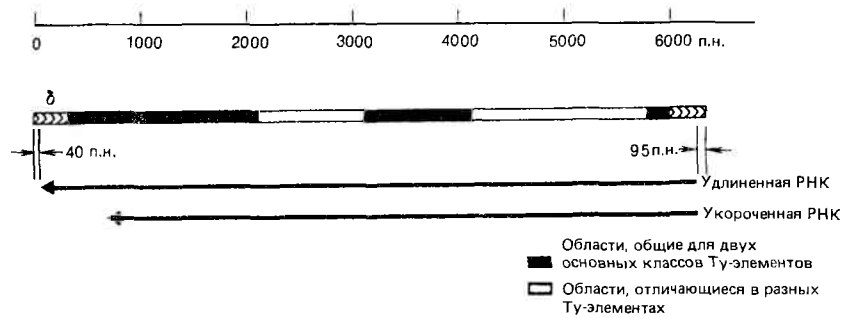
Вездесущность транспозирующихся элементов может рассматриваться либо как доказательство их важности для эукариот и прокариот, либо как указание на то, что они одинаково успешно размножаются в составе геномов любого типа. Вместе с плазмидами, способными переносить генетическую информацию между бактериями, транспозоны прокариот обеспечивают подвижность генов хозяина (с успехом компенсируя отсутствие истинного полового процесса). В некоторых случаях механизмы, подобные транспозициям, вовлекаются в регуляцию генов. У эукариот, с другой стороны, уже известны отдельные случаи, когда перемещение последовательностей из сайта или в специфические сайты играет роль в регуляции генов. В их геномах найдены также элементы,

сравнимые с бактериальными транспозонами, способные перемещаться в случайно выбираемые места генома.

Мобильные элементы, наиболее близкие по своим свойствам к бактериальным транспозонам, были обнаружены у *S. cerevisiae* и *D. melanogaster*. Они имеют небольшой размер, содержат концевые повторы и найдены в различных участках геномов. Очевидно, что в состав этих элементов входят гены, продукты которых необходимы для транспозиции; включены ли в них гены, кодирующие второстепенные функции, не связанные с транспозиционным событием, не известно.

Каждый из этих элементов аналогичен бактериальным транспозонам, которые представляют собой последовательности, способные перемещаться в пределах

Рис. 37.1. Ту-элементы содержат на концах короткие прямые повторы и транскрибируются с образованием двух перекрывающихся РНК.



хозяйского генома. Транспозоны про- и эукариот, по-видимому, не могут существовать вне генома.

Функции, связанные с осуществлением транспозиции у эукариот, пока не установлены; нам еще предстоит идентифицировать продукты эукариотических генов, участвующих в этом процессе.

Явление транспозиции было описано много лет назад Мак-Клинток, проводившей генетические исследования на кукурузе. До совсем недавнего открытия этого эффекта у бактерий полученные ею результаты оставались «в вакууме», однако сейчас события, связанные с транспозицией, могут быть изучены на молекулярном уровне и у эукариот. На кукурузе получена значительная информация о генетических взаимодействиях между членами семейства транспозонов, об эффектах их внедрения в определенные локусы и различных результатах транспозиционного события.

В составе геномов животных мобильные элементы еще не найдены, но обнаружены следы транспозиционных событий в форме прямых повторов мишени, фланкирующих диспергированные повторяющиеся последовательности. Механизмы, подобные транспозиции, по-видимому, используются ретровирусами, чтобы внедрять ДНК-копии своего РНК-генома в хромосомы клеток хозяина. Эти случаи обсуждаются вместе с другими примерами вариабельности эукариотической ДНК в гл. 38.

Молекулярные механизмы транспозиции в клетках эукариот еще не изучены так тщательно, как у бактерий: в частности, еще не идентифицировали продукты генов эукариотического транспозона. Однако присутствие коротких прямых повторов ДНК мишени позволяет предполагать использование сходного механизма, при котором такая «генерализованная» транспозиция выбирает случайные сайты для внедрения. Нам бы хотелось знать, какова взаимосвязь между транспозирующимися элементами различных видов и могут ли функции транспозиции сохраняться при перемещении элемента из одних видов в другие.

Другой тип событий представляют «направленные» транспозиции, при которых перемещение последовательностей из одного специфического сайта в другой имеет регуляторный эффект. У бактерий существует очевидная взаимосвязь между механизмами сайтспецифической инверсии и транспозиции. У эукариот направленное перемещение ДНК может осуществляться с помощью иных механизмов. Например, у дрожжей наблюдается односторонний перенос копий генов, определяющих тип скрещиваемости, из «молчащих» локусов в «активные». Этот процесс напоминает бактериальную транспозицию, когда донорный локус остается интактным, а реципиентный изменяется; однако у дрожжей существует только один реципиентный локус, который сам инициирует перенос, вместо того чтобы остаться «пассивной» ми-

шенью. Иная ситуация имеет место в случае иммуноглобулиновых генов. Здесь в результате рекомбинации меняется содержание генома, создаются активные гены в соответствующих соматических клетках (гл. 39). Диапазон изменений в геноме, таким образом, простирается от случайных транспозиций к тканеспецифичным перестройкам.

Дрожжевые элементы Ty напоминают бактериальные транспозоны

Элементы Ty дрожжей представлены семейством разобщенных повторяющихся последовательностей ДНК. Каждый элемент состоит из 6,3 т.п.н.; последние 330 пар в его концах представлены прямыми повторами, названными δ . В различных штаммах *S. cerevisiae* число Ty-элементов варьирует, но в обычных лабораторных культурах оно колеблется в пределах от 30 до 35. Наряду с интактными Ty-элементами встречается примерно 100 независимых элементов δ , называемых одиночными. Название Ty представляет собой сокращение от английских слов *transposon yeast*, что означает «транспозон дрожжей»; наличие функций транспозиции у этого элемента предполагают на основании данных, показавших его присутствие в различных сайтах разных штаммов. До сих пор частота транспозиции Ty-элементов еще не определена, но, по-видимому, она ниже, чем частота перемещения бактериальных транспозонов. Перемещение Ty-элементов было прослежено благодаря отбору мутаций, вызываемых их внедрениями в регуляторные локусы, поэтому большая часть имеющейся информации об эффектах, вызываемых Ty-элементами, касается генов, смежных с сайтами внедрений.

Наблюдается значительная дивергенция между отдельными Ty-элементами. Большинство из них попадает в один из двух основных классов, получивших название Ty1 и Ty917. Их взаимосвязь показана на рис. 37.1. Элементы обоих классов фланкированы дельта-повторами. Общим в их структуре является также присутствие длинных областей в левом конце, в середине и короткой области, прилегающей к правой дельта-последовательности. Судя по результатам рестрикционного картирования и гетеродуплексного анализа, две основные области замещения не проявляют гомологии. Отдельные Ty-элементы каждого класса могут отличаться от своего прототипа одним или несколькими основаниями, вставками или делециями. Значительная гетерогенность характерна и для δ -последовательностей, хотя более вероятно, что оба повтора одного Ty-элемента идентичны или по крайней мере очень близки. (Самый выраженный случай дивергенции соответствует шести различиям в 330 парах оснований). Не известно, чем обусловлена идентичность терминирующих последовательностей δ -элементов: конверсией,

происходящей до или после транспозиции, или тем, что оба повтора «нового» Ту-элемента образуются из одного повтора донора. Дельта-последовательности, связанные с Ту-элементами, проявляют большую степень консерватизма, чем отдельные дельта-элементы, что позволяет предполагать, что в процесс транспозиции включается узнавание повторов.

Транспозиция Ту-элемента, так же как и бактериальная транспозиция, сопровождается образованием повторов из 5 пар оснований с каждой стороны от внедренного элемента. Мы не знаем происхождения фермента (ферментов), необходимого для транспозиции Ту-элементов; широкое разнообразие последовательностей этих элементов вызывает сомнения в транспозонном происхождении ферментов.

Присутствие Ту-элементов обеспечивает наличие областей частичной гомологии, в которых могут происходить рекомбинационные события, контролируемые системами хозяина. Обычно такие события сводятся к индукции делеций или инверсий в хромосоме (путем рекомбинации между двумя находящимися в ней Ту-элементами) или более существенных изменений, вызываемых рекомбинацией между Ту-элементами из разных хромосом. По-видимому, рекомбинация между Ту-элементами происходит по типу «взрывов»: осуществление одного рекомбинационного события с большой вероятностью влечет за собой последующие акты рекомбинации. Следует отметить, что для Ту-элементов характерен и такой тип взаимодействия, как конверсия гена; она происходит между элементами, локализующимися в различных местоположениях.

Ту-элементы способны вырезаться при участии гомологичной рекомбинации между прямыми повторами δ -последовательностей, что приводит к появлению большого числа отдельных δ -элементов. Вырезание такого типа может быть связано с реверсией мутаций, вызванных внедрением Ту; частота реверсий зависит от точки вырезания.

Ту-элемент транскрибируется в два вида poly(A)⁺-РНК. Как показано на рис. 37.1, оба вида РНК иницируются приблизительно в области 95-й пары оснований от правого конца элемента. Терминация синтеза одной из них происходит на расстоянии 5 т.п.н. от точки инициации в пределах общей последовательности левого конца; синтез другой РНК терминируется на расстоянии 5,7 т.п.н. в пределах 40 пар оснований левой пограничной области. Следовательно, более длинная РНК начинается и кончается в пределах δ -последовательностей и поэтому имеет повторы на своих концах. Парадокс состоит в том, что оба элемента δ содержат одинаковые последовательности, однако промотор активен в правом элементе, а терминатор — в левом. Аналогичная особенность обнаружена и в других транспозирующихся элементах (включая ретровирусы, описанные в гл. 38). Мы не знаем, осуществляется ли трансляция Ту-РНК с образованием белка.

В геноме *D. melanogaster* содержится несколько типов мобильных элементов

Вывод о существовании мобильных элементов у *D. melanogaster* был сделан на основании наблюдений, аналогичных тем, которые позволили идентифициро-

вать первые инсерционные последовательности у *E. coli*. Были обнаружены нестабильные мутации, ревертирующие к дикому типу или вызывающие делеции в областях, прилегающих к исходному сайту мутации. Прямые доказательства наличия нескольких типов транспозирующихся последовательностей удалось получить при изучении локализации определенных компонентов повторяющихся последовательностей ДНК.

Наиболее изучено семейство элементов, получивших название *copia*. Название свидетельствует о присутствии большого числа близкородственных последовательностей, кодирующих множество мРНК. Семейство *copia* берется за образец для нескольких других типов элементов, структура и общие свойства которых, по-видимому, очень сходны, однако последовательности не являются родственными. Они описываются как *copia*-подобные элементы¹⁾. Список таких элементов приведен в табл. 37.1. Вместе взятые, элементы *copia*-подобных семейств составляют более 1% ДНК *D. melanogaster*.

Число копий этих элементов зависит от линии мухи; обычно оно колеблется между 20 и 60. Члены семейства сильно отличаются друг от друга. При сравнении геномов мух различных линий с помощью рестрикционного анализа фрагменты, гибридизующиеся с *copia*-специфичной пробой, обнаруживают в каждой линии различный, хотя и перекрывающийся спектр последовательностей. С помощью гибридизации *in situ* с политенными хромосомами установлено, что последовательности *copia* присутствуют в разных (и в этом случае перекрывающихся) сериях сайтов. О вариабельности распределения *copia* свидетельствует тот факт, что в каждой из нескольких исследованных линий последовательности *copia* обнаружены в разных сайтах. Эти различия выработаны в процессе эволюции. Сравнение линий, которые дивергировали в процессе их размножения в лаборатории, в течение последних 40 лет обнаружило мало изменений. В настоящее время мы не можем оценить скорость изменчивости. Однако о природе лежащих в ее основе событий свидетельствуют результаты наблюдений за ростом клеток в культуре. Число элементов *copia* на геном в этих условиях увеличивается, по существу, в 2–3 раза. Дополнительные копии элементов представлены внедрениями последовательностей элементов *copia* в новые сайты. Процесс адаптации клеток в культуре каким-то непонятным способом временно увеличивает частоту транспозиции до 10^{-3} – 10^{-4} событий на поколение.

Схематическое изображение элемента *copia* представлено на рис. 37.2; данные, характеризующие его структуру, приведены в табл. 37.1. Элемент имеет протяженность, равную 5000 пар оснований, и содержит идентичные прямые концевые повторы из 276 пар оснований. Каждый из прямых повторов заканчивается сходными инвертированными повторами. В сайте интеграции образуется прямой повтор из 5 пар оснований ДНК мишени. Среди известных сайтов внедрения общая последовательность не выявлена, хотя региональная предпочтительность, свойственная некоторым бактериальным транспозонам, вероятно, может существовать. Дивергенция между отдельными членами семейства *copia*

¹⁾ В советской литературе обычно используется термин *мгд*-элементы (мобильные диспергированные генетические элементы). Именно для одного из них (*мгд 1*) была впервые показана подвижность в геноме.—Прим. ред.

Таблица 37.1

Типы транспозирующихся элементов у *D. melanogaster*

Элемент	Число копий на геноме	Длина, п.н.	Длина терминального прямого повтора, п.н.	Длина терминального инвертированного повтора, п.н.	Длина прямого повтора мишени
<i> copia </i>	20–60	5000	276	17	5
<i> FB </i>	~ 30	500–5000	Нет	250–1250	9
<i> P </i>	~ 50 или 0	500–2900	Нет	31	Неизвестно

Транспозирующиеся элементы *D. melanogaster*, напоминающие по организации *copia* : 412, 297, *mgd1* , *mgd3* , B104 и *gypsy* . Число копий этих элементов варьирует от 10 до 1000 на геноме; их длина колеблется от 5500 до 8500 п.н., длина прямых повторов — от 269 до 571. Длина повторов мишени составляет обычно 4 или 5 оснований.

незначительная (менее 5%), варианты часто содержат небольшие делеции. Все эти особенности касаются и других *copia* -подобных семейств, хотя их отдельные члены могут проявлять большую степень дивергенции. Например, максимальная дивергенция, обнаруженная в случае элемента 297, достигает ~20%.

Идентичность двух прямых повторов каждого элемента свидетельствует, что они либо взаимодействуют между собой, что ведет к исправлению отклонений, либо образуются из одного прямого повтора предшествующего элемента во время транспозиции. (Мы уже рассматривали аналогичные особенности *Tu* -элементов дрожжей; это также характерно и для ретровирусов; см. гл. 38). Элементы *copia* в геноме всегда интактны; отдельные копии концевых повторов не были обнаружены (хотя можно было бы ожидать их возникновения в результате рекомбинационных событий, делетирующих промежуточный материал). С другой стороны, элементы *copia* иногда существуют в форме свободных кольцевых молекул ДНК, что отличает их от бактериальных транспозонов, не образующих свободных форм. Два главных типа кольцевых молекул в популяции элементов *copia* представлены структурами протяженностью в 5000 и 4700 пар оснований. Последние имеют делецию, захватывающую один из концевых повторов (которая была бы естественной, если бы кольцевые молекулы возникали в результате вырезания из хромосомы с помощью реципрокной рекомбинации).

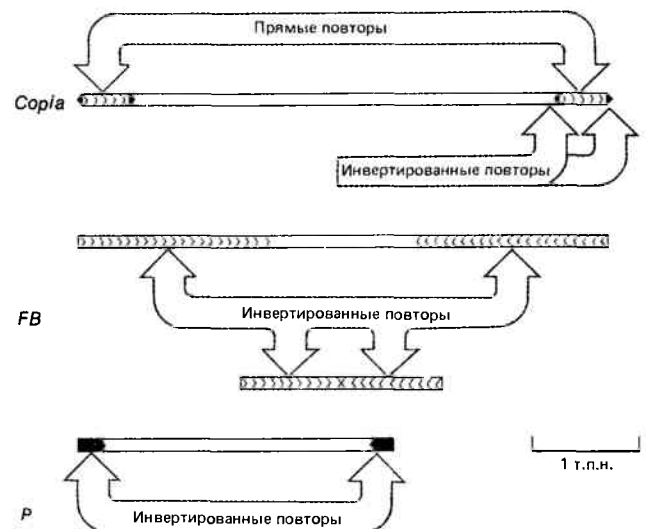
Транскрипты *copia* присутствуют в виде многочисленных *poly(A)* ⁺-мРНК, представляющих как полные, так и частичные транскрипты. мРНК имеют одинаковый 5'-конец, образующийся в результате инициации в середине одного из концевых повторов. Укороченная мРНК (две т.п.н.) содержит информацию для белка с мол. массой 51 000, функции которого неизвестны. Так как прямые повторы идентичны, не исключено, что оба они функционируют в качестве промоторов, один из которых может использоваться для транскрибирования последовательности самого элемента с образованием *copia* -специфичных РНК, а другой — для транскрибирования генов в областях ДНК, прилегающих к сайту внедрения (такие же свойства характерны для ретровирусов).

Члены другого семейства мобильных элементов называемых *FB* (от англ. *faldback*), содержат инвертированные

концевые повторы разной длины. Некоторые *FB* -элементы состоят исключительно из прилежащих друг к другу повторов; у других они отделены неповторяющейся последовательностью ДНК. Несмотря на вариации в длине, инвертированные повторы всех членов *FB* -семейства гомологичны. Это свойство обусловлено их структурой, которая представлена тандемными копиями простых последовательностей ДНК, разделенных длинными участками различающихся последовательностей. В направлении от конца к центру элемента длина простых повторяющихся единиц увеличивается: начинается с 10 п.н., затем увеличивается до 20 п.н. и, наконец, достигает 31 п.н. Две копии инвертированного повтора в одном элементе не идентичны. Инвертированные повторы внешне напоминают сателлитную ДНК. Представляет интерес их взаимосвязь у различных членов семейства. Структура концевых последовательностей *FB* -элементов необычна, поэтому не понятно, чем обусловлена их способность к транспозициям. Иногда два (неидентичных) *FB* -элемента кооперируются и переносят большой сегмент ДНК, расположенный между ними. В этом случае они напоминают сложные бактериальные транспозоны (хотя длина участка ДНК между *FB* -элементами может быть значительно больше). Известны два примера таких структур: одна из них имеет длину 200 т.п.н. и названа *TE* (транспозирующийся элемент). Возможно, что любая последовательность ДНК, фланкированная на концах *FB* -элементами, способна вести себя как такая единица.

Другой тип транспозирующихся последовательностей представлен значительно меньшими по размеру *P* -элементами; их максимальная длина достигает 2,9 т.п.н. Они содержат на концах короткие инвертированные повторы и транспозируются только в определенных условиях (см. ниже).

По структуре генов рРНК можно судить о возможности предшествующей локализации в этом локусе мобильного элемента. Примерно две трети генов, кодирующих 28S рРНК, у *D. melanogaster* имеют прерывистое строение. Промежуточные области можно отнести к двум классам, локализованным в пределах гена, в участках, отделенных 50–80 п.н. Ни один из типов раз-

Рис. 37.2. Три типа мобильных элементов у *D. melanogaster* имеют различную структуру.

единающих последовательностей не может быть классифицирован как интрон, поскольку прерывистые гены не транскрибируются. Вполне вероятно, что существует связь между присутствием разъединяющих последовательностей и неспособностью генов инициировать транскрипцию в промоторе (который, по-видимому, идентичен промотору непрерывной последовательности генов), однако природа этой связи неясна.

В генах рРНК X-хромосомы обнаружены внедрения последовательностей I типа. Последовательности гомологичны, но варьируют в длине от 0,5 до 5,0 т.п.н.; внедрения более коротких последовательностей могут быть результатом внутренних делеций в пределах исходной инсерционной последовательности протяженностью 5 т.п.н. Сравнение их структурной организации дано на рис. 37.3.

Инсерционные последовательности, содержащие 0,5 или 1,0 т.п.н., несут дубликации фланкирующей кодирующей последовательности из 14 или 11 пар оснований соответственно. Дубликации того же сайта мишени из 14 пар оснований присутствуют в прерывистых генах типа I у *D. virilis*. Возникают ли вставки I типа в результате транспозиции? По-видимому, дубликация сайта мишени была утрачена последовательностью из 5 т.п.н., поскольку вместо нее эта последовательность имеет делецию кодирующей последовательности с левой стороны вставки протяженностью в 9 пар оснований.

Последовательности, гомологичные инсерциям I типа, встречаются помимо генов рРНК в других положениях, что согласуется с представлением о них как о бывших транспозонах. Они часто встречаются в виде тандемных повторов, перемежающихся с негомологичной ДНК.

Инсерционные последовательности II типа обнаружены в генах рРНК как X-, так и Y-хромосом. Их длина варьирует от 1,5 до 4,0 т.п.н. Несмотря на сложную структуру, они не родственны инсерционным последовательностям I типа и не проявляют каких-либо признаков, свидетельствующих об их транспозонной природе.

Сложные локусы и «прогулка по хромосоме»

Рассмотрим некоторые особенности генома *Drosophila*, анализ которого стал возможным благодаря преимуществам, создаваемым мобильными элементами. Поведение нескольких локусов у *D. melanogaster* не согласуется с представлением о них как о генах, каждый из которых кодирует единственный продукт. Они получили название **сложных локусов**. Традиционный ген, даже прерывистый, идентифицируется на уровне генетической карты с помощью группы тесно сцепленных некомплементирующих между собой мутаций. Отличительным признаком сложных локусов является присутствие довольно разбросанных групп мутаций, занимающих относительно большой отрезок карты и проявляющих сложное поведение при комплементации. Например, мутации могут попадать в несколько перекрывающихся комплементационных групп, а их влияние на фенотип может быть сложным и неодинаковым у разных мутаций.

Примером локуса такого типа служит *rudimentary*, мутации в котором обуславливают глубокие изменения в морфологии крыла. Локусу *rudimentary* на карте соответствует участок размером около 0,07 единиц (примерно 35 т.п.н., если учитывать среднюю частоту рекомбинации). Другим примером может служить локус *decapentaplegic*, который нарушает развитие имаги-

нальных дисков (личиночных тканей, из которых после метаморфоза формируются различные части тела взрослой мухи). Классическим локусом этого типа можно назвать локус *bithorax*. К нему относятся несколько групп мутаций, которые нарушают развитие грудного отдела, вызывая удвоение грудных сегментов и образование второй пары крыльев. Существующая в настоящее время модель предполагает наличие нескольких генов в локусе *bithorax* (раньше их называли позиционными псевдоаллелями), кодирующих те продукты, которые участвуют в регуляции развития грудного отдела.

Локус *w* (*white*) был идентифицирован с помощью самой первой мутации, выделенной у плодовой мушки. Такая мутация определяла появление мужской особи с белыми или полупрозрачными глазами вместо обычных глаз красного цвета у особи дикого типа. Этот локус оказался интересным сам по себе; кроме того, его изучение дало информацию о событиях, включаемых в транспозицию разных элементов.

Мутации локуса *w* распределяются в области, имеющей размер около 0,04 единиц генетической карты, и составляют только одну группу комплементации. Вот почему локус *w* вряд ли можно рассматривать как удачный пример сложного локуса. В то же время, поскольку у мутантных особей *w* глаза не окрашены, возникает вопрос о функции этого локуса. Дело в том, что у мух дикого типа цвет глаз обусловлен присутствием красного и коричневого пигментов, неродственных по структуре и синтезируемых разными путями. Были идентифицированы отдельные гены, нарушающие тот или иной путь (мутанты *vermilion* утрачивают коричневый пигмент, мутанты *brown* — красный и т.п.). Отсутствие продуктов обоих путей у мутантов *w* позволяет предполагать, что локус *w*⁺ не кодирует фермент, участвующий в образовании пигмента, а регулирует их продукцию. Мутанты *w* классифицируются в фенотипические группы согласно значимости эффекта мутаций на пигментобразование. Аллели, подобные исходному *w*¹, ведут к полной утрате красного и коричневого пигментов. Другие аллели вызывают промежуточный фенотип, что свя-

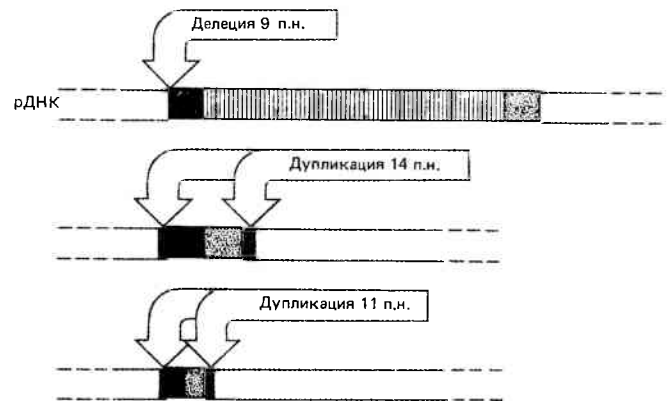
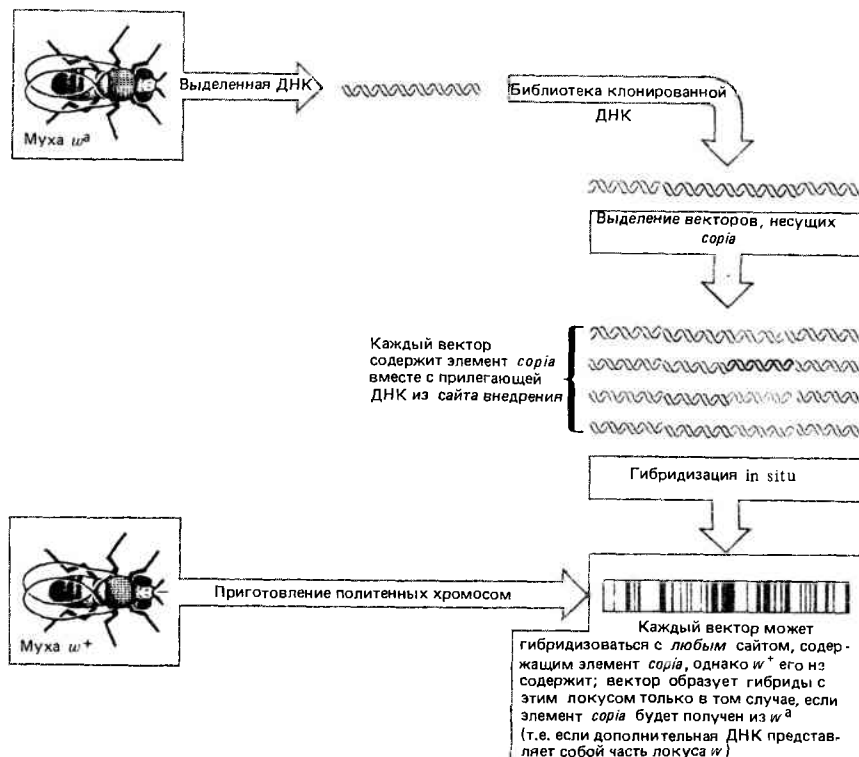


Рис. 37.3. Вставки рДНК типа I (красные сегменты) представляют собой гомологичные последовательности различной длины, вероятно образовавшиеся в результате внутренних делеций. В точке прерывания кодирующей последовательности вставка из 5 т.п.н. имеет делецию протяженностью в 9 п.н. рДНК; вставка из 1 т.п.н. содержит дубликацию рДНК протяженностью 14 п.н., а вставка из 500 п.н. — дубликацию рДНК длиной 11 п.н.

Рис. 37.4. Мобильные элементы могут быть использованы для выделения областей генома, фланкирующих сайт внедрения.



зано с уменьшенной продукцией одного или обоих пигментов у этих линий. Обычно аллели различают по индексам, отражающим фенотипы, которые они обуславливают, например w^a -аллель, определяющий абрикосовый цвет глаз, w^i -цвет слоновой кости, w^e -эозиновый и т.д. Большинство аллелей влияет на качество пигмента, однако группа w^{sp} влияет также на его распределение внутри глаза. Это согласуется с точкой зрения, что w контролирует распределение пигмента.

ДНК локуса w была клонирована с помощью метода, использующего преимущества, создаваемые присутствием элемента *soria* в локусе w^a (рис. 37.4). Геномную библиотеку линии w^a подвергли скринингу с целью выделения клонов, несущих элементы *soria*. Выделенные химерные фаговые ДНК содержали все локусы, в которых присутствовал элемент *soria*. Отдельные клонированные молекулы гибридизовали затем с полигенными хромосомами мухи w^+ . Поскольку локус w такой мухи не содержит элемент *soria*, гибридизация в w происходит только с ДНК определенного клона, содержащего аллель w^a , поскольку именно он несет последовательности w , фланкирующие сайт внедрения элемента *soria*. Примерно 1 из 50 клонов попадает в эту категорию.

Любой сегмент локуса w , полученный в таком клоне, может быть использован для выделения всего локуса с помощью метода «прогулка по хромосоме», который представляет собой усовершенствованный вариант рестрикционного картирования. Он основан на использовании перекрывающихся фрагментов, полученных в результате разрывов генома в одной и той же области. Принцип метода иллюстрирует рис. 37.5.

Анализ начинается с клона, несущего фрагмент с интересующей областью или участком вблизи нее. Затем идентифицируются другие клоны, несущие последовательности, перекрывающиеся с последовательностью первого клона. Такие клоны будут содержать более протяженные последовательности с одной или дру-

гой стороны от фрагмента, содержащегося в первом клоне. Проводя в каждом случае рестрикционное картирование, можно определить, в каком именно направлении увеличен фрагмент нового клона. Фрагменты могут иметь довольно большую протяженность, поэтому удобно для выделения каждого следующего клона использовать субфрагмент неповторяющейся ДНК из конца предыдущего. Процесс может быть повторен неограниченно. Таким образом, оказывается возможным проанализировать сотни тысяч пар оснований. Следует отметить, однако, что, если область содержит многократно повторенную ДНК, многие гибридизующиеся клоны могут представлять совершенно различные области хромосом.

Внедрения в локус w отражают сложность мишени

На рис. 37.6 представлена карта области, содержащей примерно 15 т.п. оснований, которая включает все известные сайты мутаций. Мутации классифицированы в соответствии с изменениями, которые они вносят в рестрикционную карту. Тип мутаций и их расположение соответствуют двум зонам с потенциально разными функциями.

Доказано, что многие мутации возникают в результате вставок генетического материала. В некоторых случаях такими вставками могут быть известные транспозирующие элементы, в других внедренная ДНК представляет семейство повторяющихся последовательностей и может оказаться неидентифицированным транспозирующим элементом. Единственными поддающимися обнаружению делециями являются делеции участков, соответствующих аллелям w^{sp3} и w^{sp4} , которые локализируются вблизи или в самом сайте внедрения w^{sp1} .

В локусе w идентифицированы потенциальные точечные мутации, которые могут не вызывать изменений в рестрикционной карте. Генетическое картирование по-

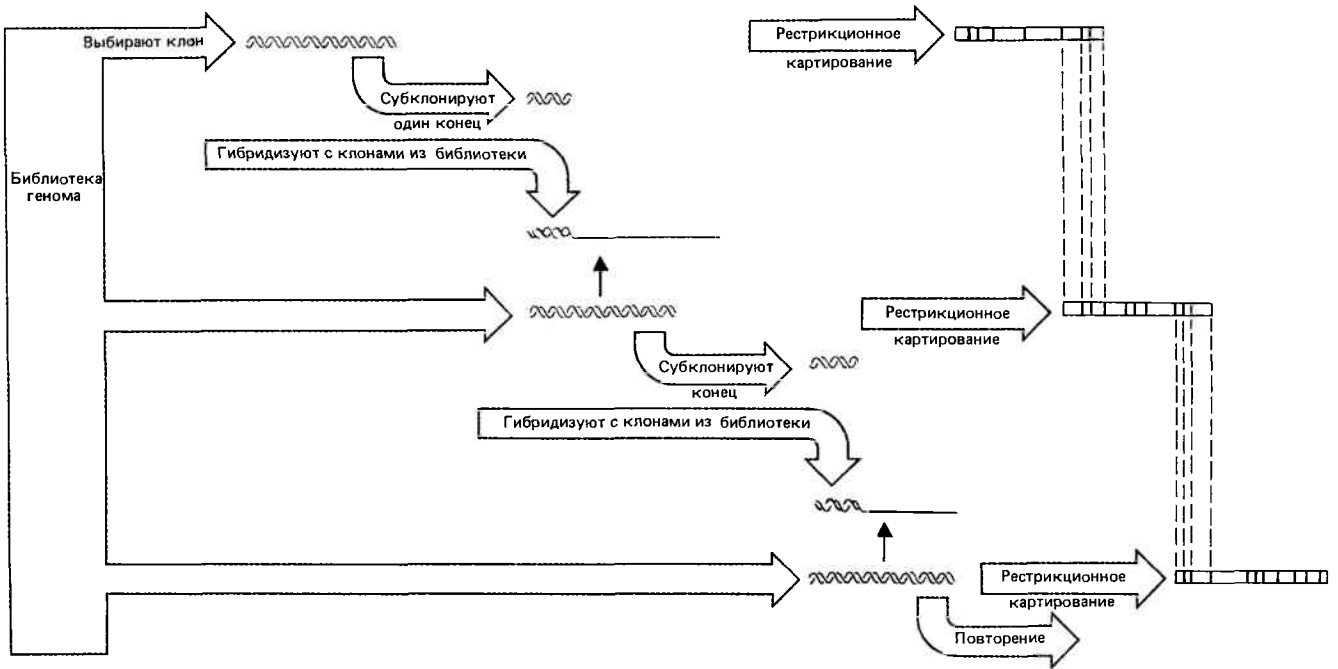
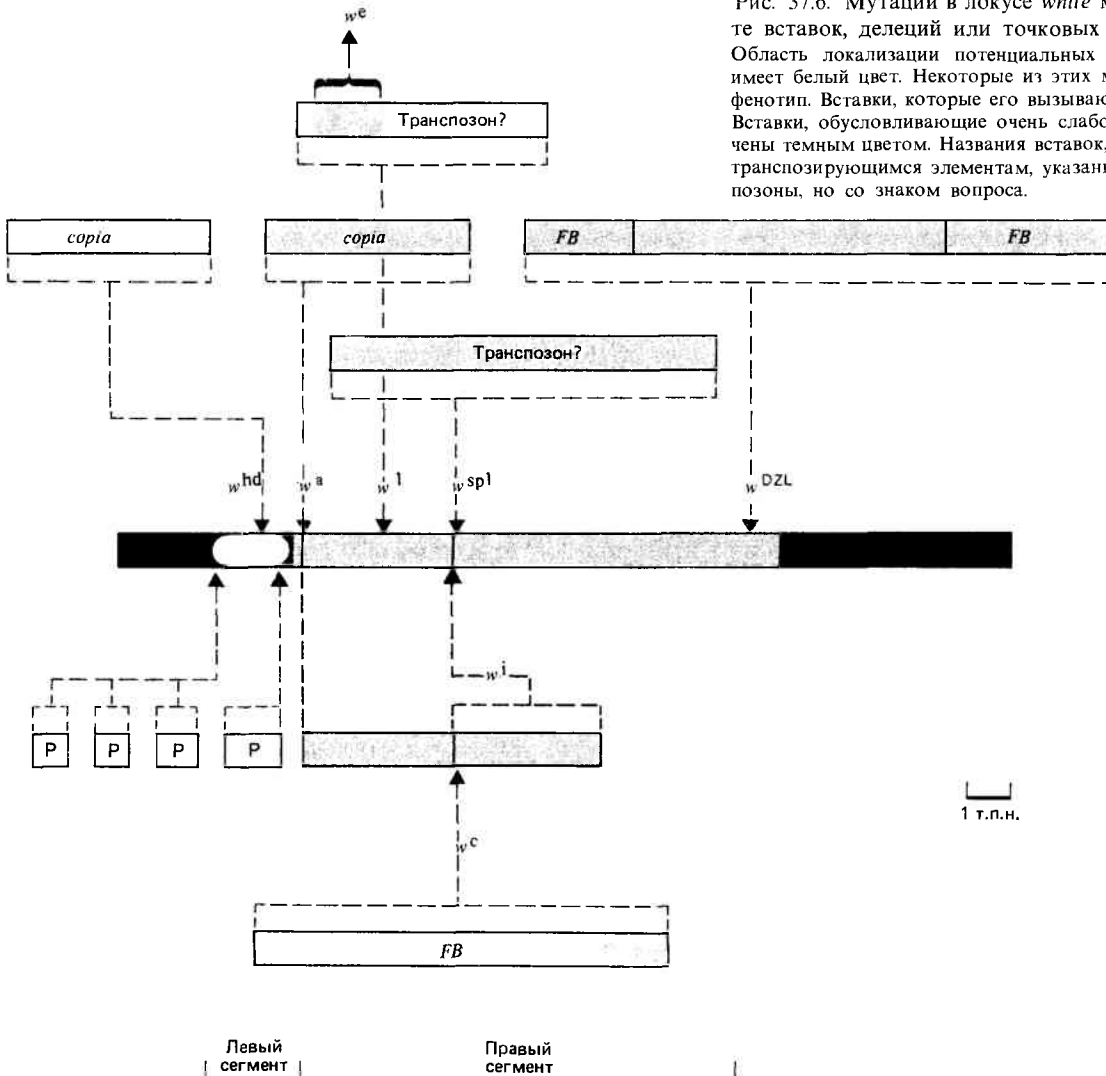


Рис. 37.5. «Прогулка по хромосоме» осуществляется путем последовательных актов гибридизации между перекрывающимися клонами генома.

Рис. 37.6. Мутации в локусе *white* могут возникнуть в результате вставок, делеций или точковых замен.

Область локализации потенциальных точковых мутаций на рисунке имеет белый цвет. Некоторые из этих мутаций обуславливают нулевой фенотип. Вставки, которые его вызывают, отмечены красным контуром. Вставки, обуславливающие очень слабое пигментообразование, обозначены темным цветом. Названия вставок, которые относятся к известным транспонирующимся элементам, указаны, другие обозначены как транспозоны, но со знаком вопроса.



казало, что все мутации такого типа локализованы слева от внедрения w^a (этот участок на карте имеет белый цвет). Чтобы подтвердить, что эти мутации вызваны заменой оснований, необходимо определить последовательность оснований в данной области. Многие мутации обуславливают либо полную потерю функции локуса (глаза приобретают белый цвет), либо образование пигмента в мизерном количестве. Левый сегмент представляет потенциальную кодирующую область локуса. Об этом свидетельствует тот факт, что все инсерции в нем вызывают появление фенотипа «белых глаз», чего следует ожидать, если нарушенная область кодирует белок.

Мутации, картирующиеся справа от w^a (включая этот сайт), почти всегда отличаются нестрогим (leaky) эффектом и обуславливают остаточную пигментацию. В этой области не найдено потенциально точечных мутаций; все мутации ведут к изменениям в рестрикционной карте. Подавляющее большинство мутаций представляют собой внедрения известных или предполагаемых подвижных элементов. (Атипичным аллелем является w^1 , исходный мутант с фенотипом «белые глаза»). Эти мутации и определяют правый сегмент, простирающийся от w^a до w^{DZL} .

Поскольку внедрения в правом сегменте полностью не блокируют его функции, можно предположить, что вставки не прерывают кодирующий участок, а вмешиваются в осуществление какой-то другой функции, контролируемой смежной регуляторной областью. Вставки картируются в нескольких дискретных сайтах. Один из них w^a , который находится на границе потенциально кодирующей области. Следующим является сайт w^1 , включающий серию вставок определенных последовательностей и их производных. В сайте w^{sp} локализуется несколько вставок, делеций. Вставка w^{DZL} — самая удаленная в правом конце, возможно, именно она служит границей локуса. Весьма вероятно, что такие сайты внедрений идентифицируют отдельные регуляторные области; внедрения в другие сайты правого сегмента не вызывают появления мутантного фенотипа, доступного определению. Все мутации, которые повреждают потенциально регуляторные функции, такие, как синхронность исчезновения пигментов или их исчезновение в определенном порядке, дозовую компенсацию, картируются в правом сегменте.

Вывод о том, что вставки в правом сегменте оказывают непрямой эффект на активность продукта гена, подтверждается природой ревертантов. Изменения во внедренном сегменте часто приводят к изменению функции локуса. Например, делетирование небольшой части вставки w^1 дает начало аллелю w^c , который в какой-то мере восстанавливает образование пигмента. Многие примеры эффектов такого типа позволяют заключить, что утрата функции не обусловлена только самим фактом внедрения вставки, а может зависеть и от природы, и от протяженности внедренного сегмента.

Реверсия не обязательно обусловлена утратой части внедренного материала, как это было показано на производных мутации w^1 . Эта мутация в отличие от других изменений, увеличивающих размер локуса, не связана с внедрением постороннего материала, а представляет собой тандемную дупликацию области из 3 т.п. н. внутри локуса w . Механизм реверсии здесь заключается в точной утрате дупликации, происходящей, вероятно, за счет внутривнутрихромосомной рекомбинации с относительно высокой частотой (около 10^{-5}).

Исключительным ревертантом оказался мутант w^c , отличающийся малиновой окраской глаз; он возникал

в результате внедрения в дупликацию w^1 элемента FB протяженностью 11 т.п. н. Аллель w^c нестабилен, ревертирует с частотой $\sim 10^{-3}$. При этом восстанавливается тип w^1 или образуются стабильные производные с делециями, оканчивающимися в локусе w ; представляется вероятным, что такие различные перестройки стимулируются FB -элементом. Удивительно, что такая большая вставка, как FB -элемент, способна восстановить фенотип, который был утрачен в результате внутренней дупликации. Возможно, что FB -элемент привносит новый промотор для поврежденного сайта, хотя имеются основания считать, что увеличение расстояния до других сайтов правого сегмента должно вызывать отрицательный эффект.

В сложных локусах обнаружен загадочный регуляторный эффект благодаря феномену так называемой **трансквекции**. Этим термином обозначают способность хромосомных перестроек, предотвращающих спаривание гомологичных хромосом (как это показано для политенных хромосом), изменять фенотип, оставляя генотип в целом неизменным. В основе эффекта лежит свойство мутаций из одного аллеля влиять на активность другого аллеля, пока эти аллели спарены. Следовательно, трансквекция может зависеть от какого-то структурного изменения, определяемого физическим контактом. Ее существование свидетельствует о возможности хромосомного спаривания в нормальных диплоидных клетках.

Причина явления трансквекции несколько прояснилась после выделения мутации w^{DZL} (DZL — от англ. *dominant zest-like*). Вставка DZL в *cis*-положении подавляет функцию левой части области w . Этот эффект демонстрируется на **рис. 37.7**; аллель w^{DZL} подавляет также экспрессию аллеля w^+ , находящегося в состоянии синапсиса с ним. Это не связано с доминантным эффектом в *транс*-положении, поскольку тот же самый аллель w^+ экспрессируется нормально, если он транслоцирован в другой участок, и не попадает под влияние спаренного с ним аллеля w^{DZL} .

Аллель w^{DZL} содержит внедренную последовательность длиной 13 т.п. н., состоящую из двух неидентичных FB -элементов, которые фланкируют сегмент негомологичной ДНК протяженностью 6 т.п. н. (неповторяющаяся последовательность хромосомы II). Подобно другим случаям FB -вставок (в w^c), этот аллель нестабилен, ревертирует с частотой от 10^{-2} до 10^{-3} . Ревертанты к дикому типу возникают в результате утраты части DZL -вставки; обязательным является потеря центральной области длиной 6 т.п. н. Такая частичная эксцизия происходит, вероятно, в результате рекомбинации между FB -элементами, находящимися по концам вставки; ревертанты остаются нестабильными из-за присутствия FB -концов. Однако один FB -элемент в этом сайте не вызывает появления фенотипа DZL ; поэтому наиболее вероятно, что за эффект трансквекции ответственна именно центральная последовательность протяженностью 6 т.п. н. Можно предположить, что эта область кодирует продукт, действующий в *cis*-положении. Весьма вероятно, что роль этого продукта заключается в узнавании контакта между ДНК.

Эффект трансквекции проявляется на значительных расстояниях. В качестве примера можно привести эффект мутации z на локус w , находящийся на расстоянии 0,75 единиц генетической карты правее этой мутации, что соответствует 300–400 т.п. н. Мутация z рецессивна. Она не проявляется у самцов (которые имеют только одну X-хромосому), но у самок, гомозиготных по z , репрессируется пара аллелей w^+ , находящихся в синапсисе, что

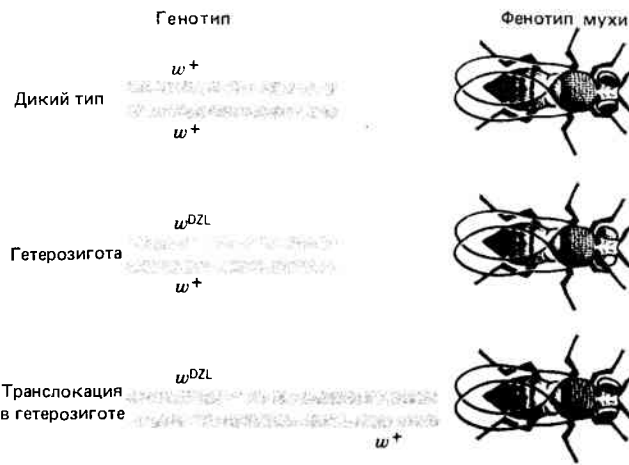


Рис. 37.7. Мутация w^{DZL} предотвращает экспрессию аллеля w^+ , находящегося в состоянии синapsиса с ней, но не влияет на этот аллель, если он транслоцирован.

обуславливает появление бледноокрашенных (желтых) глаз. Эффект исчезает, если один из аллелей w^+ транслоцирован в другой участок или даже делетирован. Он может быть изменен определенными аллелями w , создаваемыми мутациями в правом сегменте локуса.

Характер взаимодействия между z и w позволяет предположить, что синapsис необходим для проявления эффекта, а не только, чтобы эффект распространялся дальше по аллелям, находящимся в состоянии синapsиса. Передача такого эффекта на большое расстояние показывает также, что он может быть обусловлен функциями различных локусов (z и w); следовательно, мутация DZL не обязательно должна являться составной частью самого локуса w , а может быть сцепленной с ним.

Роль мобильных элементов в гибридном дисгенезе

Скращивание определенных линий *D. melanogaster* приводит к образованию потомства с «дисгенетическими признаками». Это выражается в появлении у них серии генетических дефектов, таких, как мутации, хромосомные aberrации, нарушение расхождения хромосом в мейозе и стерильность. Комплекс этих генетических аномалий характеризует явление, получившее название **гибридного дисгенеза**.

У *D. melanogaster* были идентифицированы две системы, обуславливающие гибридный дисгенез. Мухи, у которых функционирует первая система, разделены на типы *I* (*inducer*) и *R* (*reactive*). При скращивании самцов *I* с самками *R* наблюдается уменьшение плодовитости, однако реципрокное скращивание проходит нормально. Мухи со второй системой дисгенеза подразделяются на два типа, *P* (*paternal contributing*) и *M* (*maternal contributing*). На рис. 37.8 продемонстрирована асимметрия системы: скращивание между самцом *P* и самкой *M* вызывает дисгенез, а в реципрокном скращивании этого эффекта нет.

Дисгенез преимущественно проявляется в зародышевых клетках. В скращиваниях $P \times M$ гибридные мухи F_1 имеют нормальные соматические ткани, однако их гонады не развиваются. Морфологический дефект в развитии гамет ведет начало от стадии, на которой в зародышевой линии начинается быстрое клеточное деление.

При скращивании самца *P* с самкой *M* индукцию дисгенеза способна вызвать любая из хромосом самца. Конструкция рекомбинантных хромосом показывает, что в пределах каждой *P*-хромосомы есть области, способные вызывать дисгенез. Отсюда следует, что именно у самцов *P* имеются **факторы *P***, расположенные в разных местах их хромосом. Локализация таких последовательностей в различных *P*-линиях разная. Факторы *P* отсутствуют в хромосомах мух *M*-типа.

Мухи *M*-типа вносят в явление дисгенеза другой вклад. Показано, что гибридный дисгенез проявляется только в том случае, если факторы *P* подвергаются действию цитоплазмы, унаследованной по материнской *M*-линии. Материнский вклад назван *M*-цитотипом; его механизм непонятен, но, по-видимому, это не просто материнский эффект, поскольку он наблюдается в течение нескольких поколений. В тех редких линиях, которые удалось получить от скращиваний самцов *P* с самками *M*, факторы *P* вызывали состояние, при котором они перестают быть активными. В процессе картирования ДНК мутантов w , обнаруженных среди дисгенетичных гибридов, были выявлены те события, которые отвечают за индукцию мутаций при дисгенезе. Все мутации были обусловлены внедрениями ДНК в локус w . Три из четырех охарактеризованных внедрений *P*-элемента обнаружены в одном и том же сайте (рис. 37.6). Внедрения *P* стабильны, если хромосома находится в *P*-цитотипе; если же она попадает в *M*-цитотип, эти последовательности вырезаются. Наряду с внедрениями *P* в составе хромосом дис-

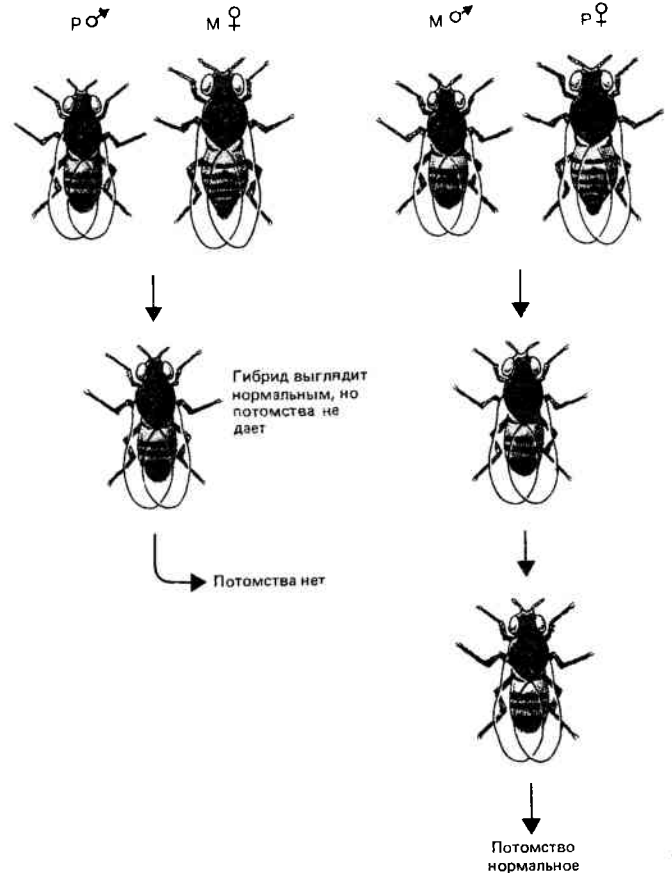


Рис. 37.8. Гибридный дисгенез асимметричен. Он индуцируется в скращиваниях: самец *P* × самка *M*; реципрокное скращивание к дисгенезу не приводит.

геничных мух были обнаружены внедрения элементов *copia* . Одно из них обозначено на рисунке символом w^{hd} .

P-элементы варьируют по длине, но последовательности нуклеотидов во всех случаях гомологичны. Разная длина обусловлена внутренними делециями в пределах P-фактора, которые, по-видимому, происходят довольно часто. Самые длинные P-элементы содержат около 2,9 т.п.н. и имеют прямые концевые повторы (см. рис. 37.2 и табл. 37.1). Существующая модель событий, приводящих к дисгенезу, основана на предположении, что интактные P-факторы кодируют транспозазу, которая неактивна в P-цитотипе, но становится активной в M-цитотипе. Некоторые короткие P-элементы могут утратить способность синтезировать транспозазу, однако их активация в *транс*-положении возможна за счет активности фермента, кодируемого P-элементом нормальной длины. Концы P-фактора могут указывать, что они служат субстратом для ферментов транспозиции. Дисгеничные взаимодействия P-M способны также активировать подвижность элементов *copia* (и других транспозирующих элементов), однако эти элементы независимо от цитотипов, в которые попадает хромосома, далее остаются стабильными.

Гибридизация *in situ* позволила идентифицировать в каждом из P-штаммов 30–50 копий P-фактора; в штаммах M такие факторы не обнаружены. Хромосомы мух, полученных от скрещивания P × M, содержат P-факторы, внедренные во множество новых сайтов. Типичные для гибридного дисгенеза хромосомные разрывы происходят в горячих точках, которые и являются сайтами внедрения P-фактора. Хотя частота транспозиции P-факторов в хромосомы M у самцов и самок в потомстве дисгеничного скрещивания различна, в среднем она достигает порядка одного события на поколение. Хромосома, которая приобрела P-элементы, становится дисгеничной, если она попадает в M-цитотип; ее способность индуцировать стерильность связана с числом приобретенных P-элементов. С другой стороны, существуют данные, свидетельствующие о независимости процессов транспозиции и индукции стерильности: обнаружена линия Q, которая во всех отношениях аналогична линии P, но не индуцирует в гибридах стерильности.

Линии *D. melanogaster*, которые были введены в лабораторную культуру более 30 лет назад, принадлежат к M-типу. Линии, ведущие свое начало от мух, пойманных в минувшие 10 лет, почти всегда относятся к типу P. Значит ли это, что семейство P-элементов распространилось в природных популяциях дрозофилы в последние годы? P-элементы действительно отличаются высокой инвазивностью при введении в новую популяцию. Источником элемента должны служить другие виды мух. С другой стороны, показано, что P-элементы могут утрачиваться популяцией, сохранявшейся в лабораторных условиях в течение большего промежутка времени.

Так как гибридный дисгенез подавляет «срещиваемость», его можно считать ступенью на пути видообразования. Допустим, что в какой-то географической зоне благодаря транспозирующему элементу возникла дисгеничная система. Другой элемент в другой зоне может привести к появлению другой системы. Если это определяет их стерильность при взаимных скрещиваниях, популяция становится генетически изолированной и происходит дальнейшее отдаление. Следовательно, множественные дисгеничные системы могут привести к неспособности скрещиваться и к видообразованию.

Контролирующие элементы кукурузы способны транспозироваться

Генетические исследования кукурузы, начатые в 1940 г., показали, что во время деления соматических клеток в геноме происходят определенные изменения. Эти изменения стимулируются **контролирующими элементами**, которые были обнаружены благодаря их способности перемещаться из одного сайта в другой. Внедрение контролирующего элемента в определенный сайт влияет и на активность смежных генов. Частота и синхронность событий, связанных с контролирующими элементами, регулируются в процессе развития; действия контролирующих элементов могут проявлять тканеспецифичность.

В настоящее время можно рассматривать контролирующие элементы как транспозоны. Их внедрение способно вызывать нестабильность аллеля в этом локусе. Такие аллели раньше называли *изменчивыми (mutable)*; эта терминология еще сохранилась в виде символа *m*, используемого при описании таких аллелей. Утрата самого контролирующего элемента или его способности к транспозиции превращает изменчивый аллель в стабильный. В сайтах, где присутствуют контролирующие элементы, могут происходить делеции, дупликации, инверсии и транслокации, а также разрывы хромосом. Геном кукурузы содержит несколько семейств контролирующих элементов. Разные линии отличаются по количеству, типу элементов и сайтам их присутствия. Члены каждого семейства могут быть подразделены на два класса. **Автономные элементы** способны вырезаться и транспозироваться; их внедрение ведет к появлению нестабильных (*m*) аллелей. **Неавтономные элементы** теряют свою стабильность только в том случае, если в какой-то области генома присутствует автономный член того же семейства. Неавтономный элемент может комплементировать в *транс*-положении автономным и осуществлять свои функции. Весьма вероятно, что неавтономные элементы произошли от автономных в результате утраты последними способности к транспозиции.

Автономные элементы подвержены «изменениям состояния», наследственным, но относительно нестабильным изменениям в их свойствах. Эти изменения способны влиять на периодичность и частоту транспозиции. Некоторые изменения приводят, по-видимому, к утрате автономным элементом его автономности. Некоторые повреждения, действующие в *цис*-положении неавтономных элементов, могут сделать их недоступными воздействию автономных. Таким образом, неавтономный элемент может стать надолго стабильным из-за того, что он не способен быть активированным к транспозиции.

Нуклеотидные последовательности контролирующих элементов еще не установлены. Однако уже полученные данные о характере взаимоотношений между автономными и неавтономными элементами. Так, известно, что неавтономные элементы могут быть активированы в *транс*-положении только определенными автономными элементами (вероятно, членами того же семейства). Данные о трех наиболее изученных семействах контролирующих элементов, *Ac-Ds*, *Spm* и *Dt*, суммированы на рис. 37.9.

Функции контролирующих элементов могут быть идентифицированы с помощью мутаций. Например, два свойства представителей семейства *Spm* отражены в их названии: супрессор—мутатор. Их действие показано при введении *Spm* в определенный локус генома для актива-

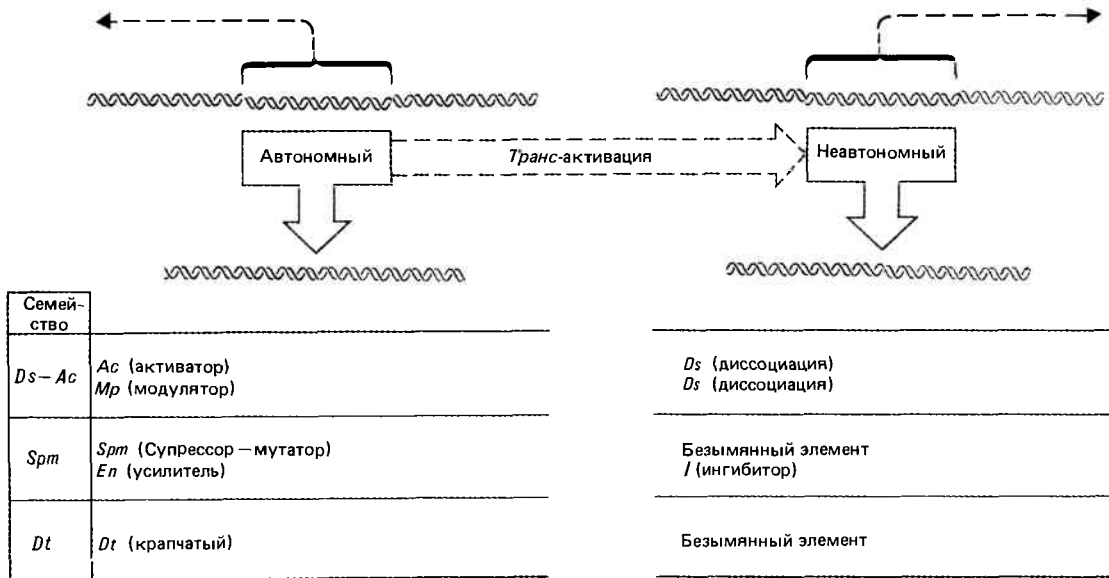


Рис. 37.9. Каждое семейство контролирующих элементов содержит как автономные, так и неавтономные члены. Автономные элементы способны к транспозициям и, кроме того, проявляют другие активности (например, влияют на экспрессию генов). Неавтономные элементы дефектны, по крайней мере в способности к транс-

позиции. Названия, первоначально присвоенные парам автономных и неавтономных элементов, классифицированы как семейства. Каждое семейство может включать более одного элемента; например, *Spm-s* и *Spm-w* представляют собой различные модификации *Spm*-элемента.

пции неавтономного элемента. «Супрессор» подавляет выражение структурного гена в этом локусе. «Мутатор» проявляет способность транспозировать, в результате чего ген восстанавливает полную активность. Оба свойства принадлежат элементу, так как они переносятся вместе; однако обнаружены мутанты, фенотип которых свидетельствует о возможности их разделения. Поскольку способность неавтономного элемента супрессировать смежные функции контролируется автономным элементом в *транс*-положении, вряд ли этот эффект обусловлен только внедрением неавтономного элемента; вероятно, он должен иметь специфическую регуляторную основу.

Результаты транспозиционных событий удалось проследить благодаря двум свойствам кукурузы. Контролирующие элементы часто внедряются вблизи генов, которые проявляют видимый (но не летальный) эффект на фенотип. И, поскольку для кукурузы характерно клональное развитие, место и время транспозиционного события можно проследить (рис. 37.10). Природа события в данном случае не имеет значения: это может быть вставка, вырезание или (как показано на рисунке) хромосомный разрыв. Важно, что это событие происходит в гетерозиготе и изменяет экспрессию одного аллеля. (В следующем разделе рассматривается пример, когда разрыв приводит к утрате доминантных аллелей, что обуславливает рецессивный фенотип.) В результате потомки клетки, в которой произошло событие, определяемое присутствием элемента, приобретают новый фенотип, в то время как потомки клетки, которая не подвергалась изменениям, продолжают проявлять исходный фенотип.

Клетки, образующиеся в результате последовательных митотических делений одной исходной, расположены рядом и образуют сектор ткани. Изменения в фенотипе во время соматического развития получили название **мозаицизма**; он проявляется в виде секторов нового фенотипа, находящихся внутри ткани исходного фенотипа. Размер сектора зависит от числа делений клетки, дающей начало сектору; следовательно, размер зоны нового фе-

нотипа определяется временем появления изменения в генотипе. Чем раньше оно произошло в клеточной родословной, тем большее число потомков образуется и, следовательно, больше размер пятна в зрелой ткани.

Элемент *Ds* способен транспозироваться или вызывать хромосомные разрывы

Присутствие неавтономного элемента *Ds* может иметь разные последствия для смежного локуса. При активации автономным *Ac*-элементом *Ds*-элемент может либо переместиться в новый сайт, либо вызвать хромосомный разрыв. Транспозиция сопровождается вырезанием из донорного сайта, в результате чего нестабильный аллель превращается в стабильный (но не обязательно дикого типа). Обратная связь между реверсией к стабильности и наличием хромосомного разрыва позволяет предполагать, что транспозиция и разрыв представляют собой альтернативные результаты активации *Ds*-элемента.

Пока *Ds*-элемент присутствует в локусе, он может влиять на уровень экспрессии гена; больше того, он способен изменять время экспрессии определенных генов в ходе развития. Кроме того, взаимодействие *Ds*-элемента со структурным геном влияет на структуру продукта этого гена.

Наиболее доступны для структурного анализа *Ds*-элементы, внедренные вблизи локуса *Sh*. Аллели *Sh* кодируют сингтазу сахарозы, отсутствие которой ведет к появлению фенотипа «морщинистого зерна», от которого и произошло название локуса. Была выделена мРНК сингтазы сахарозы протяженностью 3 т.п.н. С ее помощью удалось изолировать соответствующую геномную ДНК. Локусы *shm*, в которых присутствует *Ds*-элемент, отличаются ретриктионной картой в 5'-области гена, соответствующей внедрениям или перестройкам ДНК. Точ-

ная локализация изменений в последовательности неодинакова у различных аллелей *sh-m*.

Присутствие *Ds*-элемента изменяет экспрессию *sh-m*. У некоторых мутантов он почти не выражается. У других этот локус транскрибируется с образованием большого количества aberrантной РНК, которая короче, чем мРНК гена *Sh*, и состоит частично из последовательности *Sh*, а частично из новых последовательностей. В результате трансляции образуется белок, иммунологически родственный синтетазе сахарозы, но имеющий измененную электрофоретическую подвижность. После того как станет известна структура локусов *Sh* и *sh-m*, появится возможность и ответить на вопрос: какие именно сайты в пределах структурного гена или вблизи него подвержены влиянию *Ds*-элемента и как это сказывается на выражении гена?

Локус *Waxy* (*wx*) определяет содержание амилозы в пыльце и ткани эндосперма. Он кодирует фермент UDP глюкозу-крахмал—глюкозилтрансферазу с мол. массой 58 000 дальтон, который связывается гранулами крахмала. Обнаружены внедрения в этот локус *Ac*- и *Ds*-элементов. Сравнивая последовательности этих вставок с после-

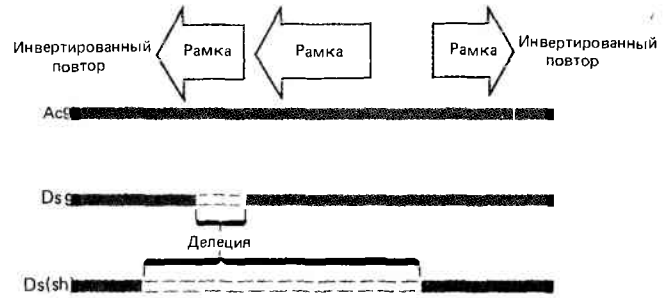


Рис. 37.11. *Ac*-элементы содержат три открытые рамки считывания; *Ds*-элементы имеют внутренние делеции.

довательностями вставок в локус *Sh*, нам удалось определить структуру *Ac*- и *Ds*-элементов. Она представлена на рис. 37.11.

Автономный *Ac*-элемент содержит 4563 пары оснований и включает три основные открытые рамки считывания. Рамки 1 и 2 считываются в одном и том же направлении и, судя по составу триплетов, кодируют белки. Функция рамки 3, которая считывается в противоположном направлении, неясна.

Рамки считывания составляют большую часть последовательности *Ac*-элемента. На его концах имеются инвертированные повторы из 11 п.н.; в сайте внедрения дуплицируется последовательность мишени из 8 п.н.

Последовательности *Ds*-элементов варьируют как по длине, так и по составу. Однако между ними и *Ac*-элементами есть родство: они содержат на концах те же инвертированные повторы из 11 п.н. Элемент *Ds9* имеет в рамке считывания 1 делецию протяженностью 194 п.н. Элемент *Ds6* имеет протяженность, равную 2 т.п.н., соответствующую 1 т.п.н. с каждого конца *Ac*-элемента, поэтому основная часть рамок 1 и 2 отсутствует. Сложный двойной *Ds*-элемент содержит последовательность, представленную одним *Ds6*-элементом, внедренным в обратной ориентации в другой.

Неавтономные элементы по-видимому могут иметь разную форму, однако во всех случаях для них характерна утрата некоторых внутренних последовательностей при сохранении терминальных инвертированных повторов (и возможно каких-то других особенностей организации). Структура неавтономных элементов согласуется с общим представлением об их происхождении из автономных в результате делеций. По-видимому, такие делеции обуславливают инактивацию *trans*-активной транспозазы; при этом сайты, на которые действует транспозаза, сохраняются. Таким образом, *Ds*-элементы в чем-то напоминают дефектные *P*-элементы *D. melanogaster*.

Первоначально *Ds*-элемент был идентифицирован благодаря его способности обеспечивать сайт хромосомного разрыва при активации регулятором *Ac*. Последствия такой активации показаны на рис. 37.12. Рассмотрим гетерозиготу, в которой *Ds*-элемент расположен на одном из гомологов между центромерой и серией доминантных маркеров (*C1*, *Bz*, *Wx*), эффекты которых могут быть прослежены по цвету клеток или при соответствующем окрашивании. (*C1* представляет собой регулятор, который вызывает обесцвечивание алейрона; *Bz* гликозилирует антоцианиновые пигменты; *Wx* обуславливает образование крахмала, окрашивающегося синим при добавлении I_2-KI). Другой гомолог утрачивает *Ds* и содержит рецессивные маркеры (*C*, *bz*, *wx*). Разрыв в локусе *Ds*

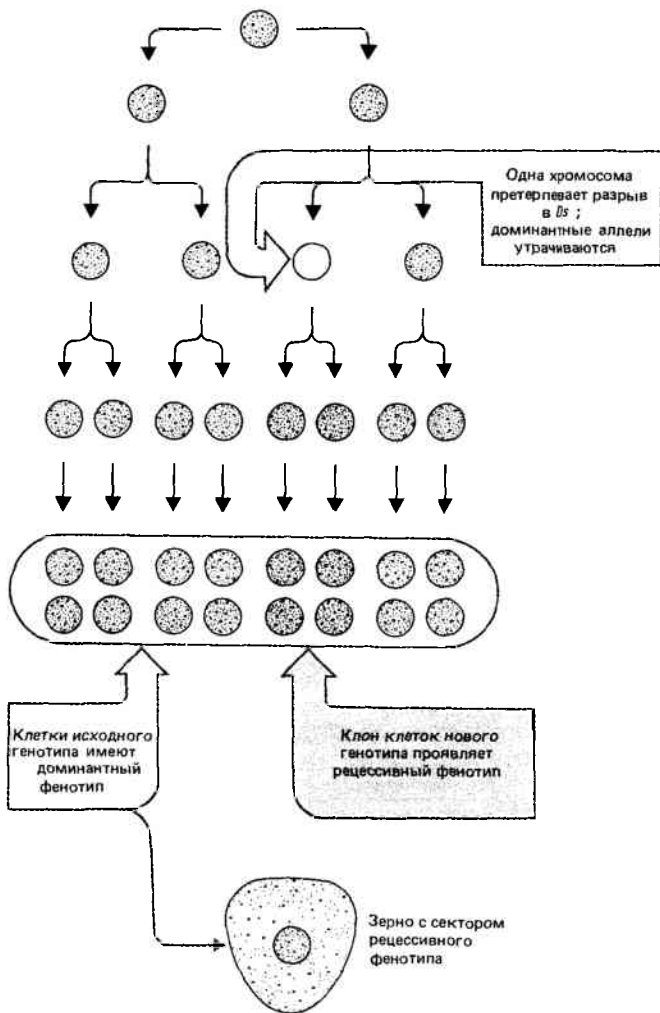


Рис. 37.10. С помощью клонального анализа можно идентифицировать группу клеток, ведущих происхождение от одной родительской клетки, у которой в результате транспозиционного события изменился фенотип.

По числу измененных клеток судят, на какой стадии развития произошла транспозиция; локализация таких клеток указывает на тканевую специфичность эффекта.

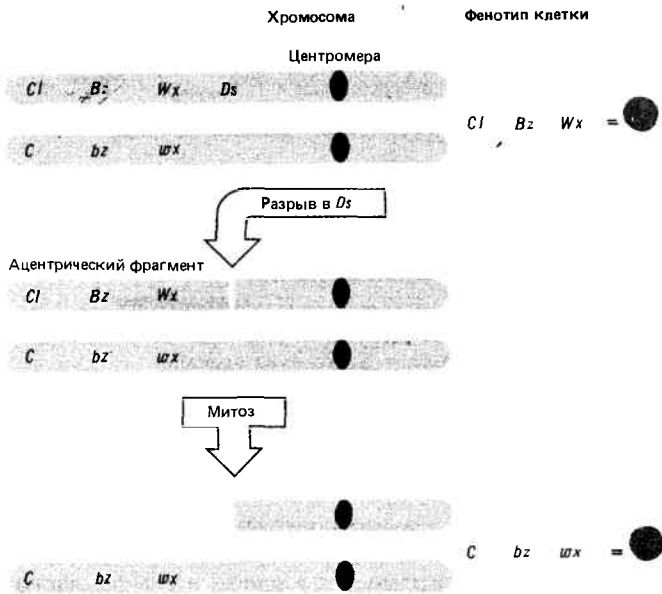


Рис. 37.12. Разрыв в контролирующем элементе ведет к утрате ацентрического фрагмента. Если фрагмент несет доминантные маркеры гетерозиготы, его утрата обуславливает изменение фенотипа.

ведет к появлению **ацентрического фрагмента**, несущего доминантные маркеры. Из-за отсутствия центромеры этот фрагмент теряется во время митоза. В результате клетки потомков имеют только рецессивные маркеры, находящиеся на интактной хромосоме. Возникает ситуация, изображенная на рис. 37.11.

Разрыв в локусе *Ds* ведет к образованию двух необычных хромосом. Они возникают при соединении разорванных концов продуктов репликации. Одна из них представлена U-образным ацентричным фрагментом, состоящим из соединенных сестринских хроматид, в области, расположенной дистальнее *Ds*-элемента (на рис. 37.12 изображены *слева*). Другая является U-образной **дигентричной хромосомой**, состоящей из сестринских хроматид, проксимальных по отношению к *Ds* (на рисунке изображены *справа*). Последняя структура ведет к возникновению классического цикла **разрыв—слияние—мост**, изображенного на рис. 37.13.

Рассмотрим судьбу хромосомы-дигентрика в митозе. Наличие двух центромер приводит к тому, что она растягивается к противоположным полюсам. Напряжение разрывает эту хромосому в случайном сайте между центромерами. На примере, приведенном на рис. 37.13, разрыв происходит между локусами *A* и *B*, в результате чего одна из дочерних хромосом содержит дупликацию *A*, а другая — делецию. В случае доминантности маркера *A* клетки с дупликацией будут сохранять фенотип *A*, а клетки с делецией — проявлять рецессивный фенотип *a*.

Цикл разрыв—слияние—мост продолжается в течение дальнейших клеточных поколений, что обеспечивает сохранение генетических изменений у потомков. Для примера рассмотрим хромосому, утратившую маркер *A* в результате деления. В следующем цикле разрыв может произойти между *B* и *C*; часть потомства в этом случае будет нести дупликацию маркера *B*, а другая — делецию. Последовательные потери доминантных маркеров могут быть прослежены по появлению в пределах секторов субсекторов.

Транспозиция *Ds*-элемента связана с репликацией

Мы уже упоминали, что транспозиция *Ds*-элемента обычно сопровождается его исчезновением из донорного сайта, что отличает этот эффект от транспозиции у бактерий. Результаты клонального анализа свидетельствуют о том, что транспозиция *Ds*-элемента почти всегда происходит после того, как донорный элемент реплицировался. Больше всего данных получено при изучении транспозиций из донорного в реципиентный сайт в пределах одной и той же хромосомы. Такие исследования были проведены с элементом *Mr*, относящимся к *Ac*-типу.

В зависимости от того, произошла ли репликация реципиентного сайта, различают два возможных результата транспозиции (рис. 37.14). Если реципиентный сайт не реплицировался перед транспозицией, элемент будет реплицирован в своем новом сайте. В этом случае он присутствует в реципиентном сайте на обеих дочерних хромосомах. Если же реципиентный сайт был предварительно реплицирован, то изменена в реципиентном локусе будет только та хроматида, которая физически приобрела элемент. В любом случае элемент утрачивается из донорного сайта одной хроматиды и остается в донорном сайте другой.

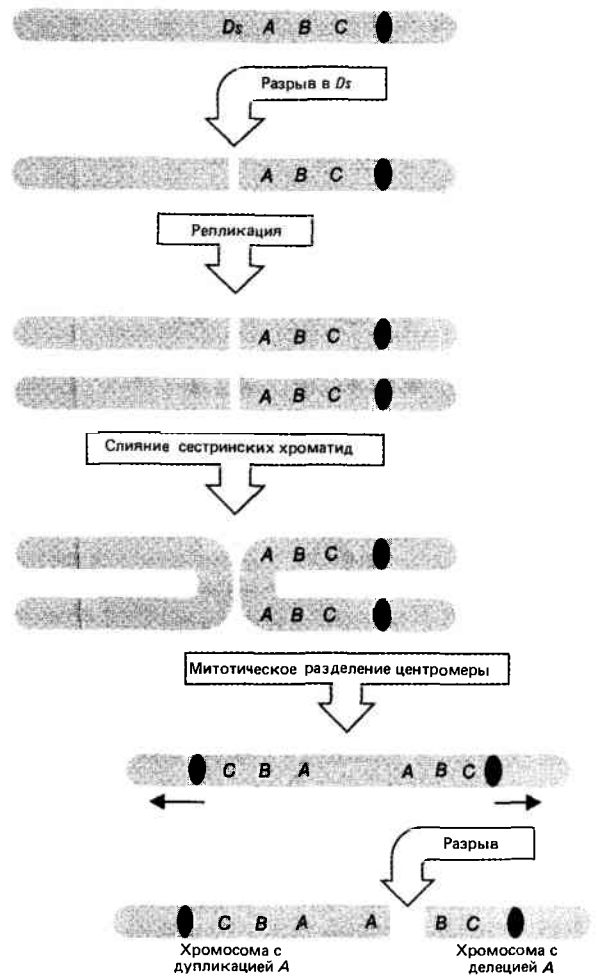


Рис. 37.13. *Ds*-элемент обеспечивает сайт, в котором инициируется цикл разрыв—слияние—мост (при активации *Ac*-элементом). Клональный анализ дает возможность анализировать последствия этого цикла.

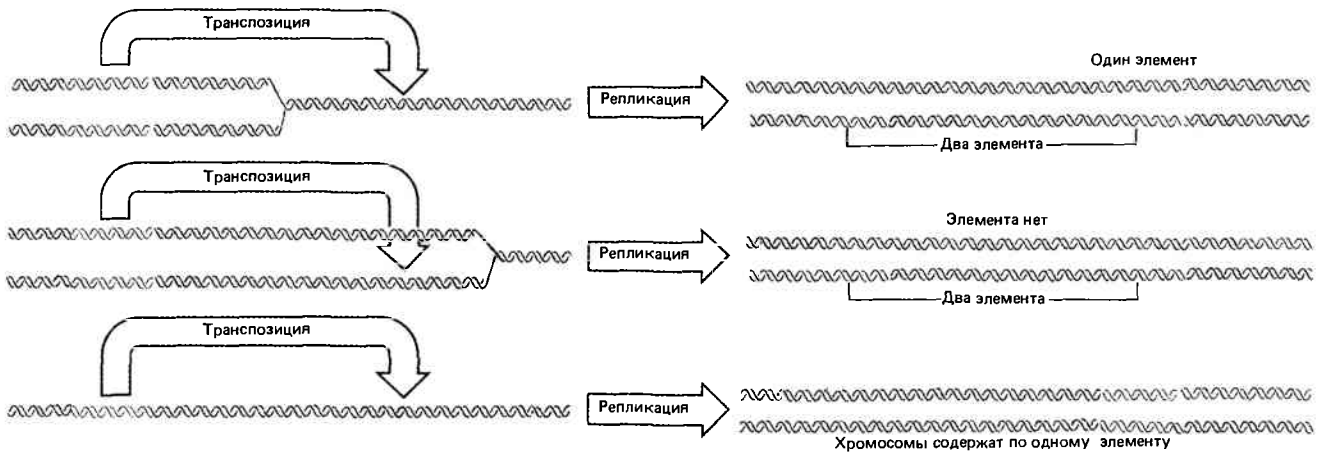


Рис. 37.14. Транспозиция у кукурузы заключается в перемещении элемента из донорного сайта после его репликации; природа продуктов зависит от того, произошла ли до транспозиции репликация реципиентного сайта.

Вверху: если реципиентный сайт не реплицирован, элемент дублируется после транспозиции, в результате чего каждая дочерняя хромосома приобретает элемент. Только одна из дочерних хромосом не содержит его в исходном сайте. У этой хромосомы элемент локализуется в реципиентном сайте, в то время как другая дочерняя хромосома содержит его как в донорном, так и в реципиентном сайтах.

В центре: если реципиентный сайт был реплицирован, только одна дочерняя хромосома приобретает элемент в этом сайте. Обычно элемент перемещается из донорного сайта, находящегося на одной хроматиде, в реципиентный сайт, находящийся на другой хроматиде, в результате в генотипе обеих хроматид происходят изменения. (Если элемент перемещается в реципиентный сайт, находящийся на той же хроматиде, генетические последствия будут аналогичны представленным в нижней части рисунка.)

Внизу: если бы транспозиция происходила перед репликацией донорного сайта, обе дочерние хромосомы содержали бы элемент в реципиентном сайте.

На рис. 37.14 (два верхних ряда) представлена ситуация, когда деление клетки ведет к образованию дочерних особей, генотипы которых отличаются друг от друга и от генотипа родительской клетки. Клональный анализ обнаруживает потомков этих клеток в виде **сдвоенных секторов**: появление сектора одного нового генотипа сопровождается появлением другого сектора, представляющего другой генотип. Заметим, что сдвоенные сектора представляют собой неизбежный результат транспозиции в нереплицировавшийся реципиентный сайт. Если реципиентный сайт уже реплицировался, они возникают только в том случае, когда мишень находится в противоположной по сравнению с донорной хроматиде. Преобладающие именно сдвоенные сектора (80% от всех транспозиций у кукурузы) свидетельствуют, что у этого растения такие события можно считать обычными. Сравнительно небольшое число транспозиций, ведущих к появлению единичных секторов, может быть обусловлено либо расположением и донорных и реципиентных сайтов на одной хроматиде, либо тем, что такие транспозиции не связаны с репликацией (рис. 37.14, нижний ряд). Возникновение единичного сектора может быть и результатом артефакта, например при потере одного из продуктов транспозиции. Хотя транспозиция обычно все-таки связана с репликацией донорного сайта, мы не знаем, насколько эта связь обязательна.

Молчание и активные локусы дрожжей, контролируемые тип спаривания

Дрожжевые клетки *S. cerevisiae* могут успешно размножаться как в гаплоидном, так и в диплоидном состоянии. Переход из одного состояния в другое имеет место при скрещивании (слияние гаплоидных спор с образованием диплоидов) и при споруляции (мейоз в диплоидах дает гаплоидные споры). Способность осуществлять эти действия определяется **типом спаривания**

штамма. Свойства штаммов, относящихся к разным типам спаривания, суммированы в **табл. 37.2**. Следует иметь в виду, что скрещивание возможно, если гаплоиды принадлежат к различным генетическим типам спаривания, а споруляция происходит, если диплоиды представляют собой гетерозиготы и, следовательно, способны образовывать рекомбинанты.

Два типа спаривания определяются аллелями *MATa* и *MATα* локуса типа спаривания. Гаплоидные клетки должны иметь генотип *MATa*, или *MATα*. Спариваться могут только клетки противоположного типа. Клетки дрожжей выделяют феромоны; благодаря этому происходит узнавание клеток противоположного типа спаривания. Клетки *MATα* выделяют небольшой полипептид-фактор α ; клетки *MATa* выделяют фактор *a*. Каждый фактор действует на клетки противоположного типа спаривания, чтобы координировать слияние клеток и ядер. Хотя обычно мы имеем дело со спариванием гаплоидных клеток, гомозиготные диплоиды (*MATa/MATa* или *MATα/MATα*) ведут себя подобно гаплоидам и способны спариваться с клетками противоположного типа.

Гомозиготные диплоиды не способны спорулировать. С другой стороны, диплоиды типа *MATa/MATα*, образующиеся при спаривании гаплоидов, способны спорулировать. На **рис. 37.15** показано, каким образом сохраняется у дрожжей нормальный жизненный цикл, включающий чередование стадий гаплоид/диплоид. Существуют и другие различия между гетерозиготами и гомозиготами в харак-

Таблица 37.2

Тип спаривания контролирует несколько активностей

Свойство	<i>MATa</i> или <i>MATa/MATa</i>	<i>MATα</i> или <i>MATα/MATα</i>	<i>MATa/MATα</i>
Спаривание	Есть	Есть	Нет
Споруляция	Нет	Нет	Есть
Секреция	Фактор <i>a</i>	Фактор α	Нет

тере почкования (отделения дочерних клеток), чувствительности к облучению и некоторых функциях клеточного цикла. Так как гетерозигота фенотипически отличается от гомозиготы, *MATa* и *MATα* можно считать кодоминантными аллелями: каждый из них детерминирует свой собственный продукт, и ни один из этих продуктов не является дефектным вариантом другого.

Некоторые штаммы дрожжей способны менять (переключать) тип спаривания. Благодаря этому удалось расшифровать механизм генетического контроля типов спаривания у дрожжей. Такие **гомоталлические** штаммы несут доминантный аллель *HO* и меняют свой тип спаривания часто (один раз на каждое поколение). **Гетероталлические** штаммы с рецессивным аллелем *ho* имеют стабильный тип спаривания, подвергающийся изменению с частотой примерно 10^{-6} .

Экспрессия гена *HO* находится под контролем типа спаривания, так как он не функционирует в диплоидах *MATa/MATα*. По-видимому, в этом случае переключения не требуется, поскольку и так оба аллеля *MAT* экспрессируются. В присутствии гена *HO* в популяции дрожжей (независимо от первоначального типа спаривания) после нескольких поколений появляется большое число клеток обоих типов, в результате чего образуются диплоиды *MATa/MATα*, которые начинают преобладать в популяции. Образование стабильных диплоидов в гаплоидной популяции может рассматриваться как смысл переключения.

Существование переключения позволяет предполагать, что все клетки содержат информацию, необходимую для экспрессии обоих типов спаривания, но выражают только один. Это заключение подтверждается анализом мутаций в каждом типе *MAT*-аллеля. Переключение в штаммах *HO* обуславливает замену мутантного аллеля либо локусом типа *MATa*, либо *MATα*.

Откуда появляется генетическая информация для экспрессии дикого типа? Переключение требует два допол-

нительных локуса. Чтобы образовался *MATα*-тип, необходим локус *HMLα*; для образования *MATa*-типа при переключении необходим локус *HMRa*. Эти локусы локализованы в одной хромосоме с геном *MAT*, однако удалены от него на значительное расстояние; *HML* находится слева, а *HMR* — справа от него.

Кассетная модель переключения типа спаривания основана на предположении, что локус *MAT* содержит **активную кассету** α- или a-типа. *HML* и *HMR* содержат **молчащие кассеты**, которые в процессе переключения замещают кассету в *MAT*-локусе. Обычно *HML* содержит кассету α, а *HMR* — кассету a. В кассетах заключена генетическая информация, способная транспозироваться.

Процесс переключения иллюстрирован на **рис. 37.16**. Событие переключения не является реципрокным, так как копия *HML* или *HMR* замещает аллель в локусе *MAT*. Это стало известно благодаря тому, что при переключении мутация в *MAT* утрачивается навсегда, она не обменивается с копией, которая замещает ее.

Копии, присутствующие в молчащих кассетах *HML* или *HMR*, также могут быть мутантными. В этом случае переключение перебрасывает мутантный аллель в локус *MAT*. Мутантная копия сохраняется в *HML* или *HMR* в течение неопределенного числа переключений. Эти события очень напоминают транспозицию, при которой донорный элемент копируется, копия встраивается в реципиентный сайт, в то время как сам элемент сохраняется неповрежденным. Переключение типа спаривания соответствует направленной транспозиции, так как существует только один реципиент (*MAT*) и два донора (*HML* и *HMR*).

Как показано на **рис. 37.16**, при переключении происходит замещение *MATa* копией *HMLα* или замещение *MATα* копией *HMRa*. Однако данные, полученные на мутантных локусах, показывают, что иногда *HMLα* замещает *MATα*, а *HMRa* — *MATa*. Если локусы не маркированы, такие переключения проходят незамеченными. В 80–90% переключений аллель *MAT* замещается аллелем противоположного типа. Эффект замещения определяется фенотипом клетки. Клетки фенотипа a в качестве донора предпочтительно выбирают *HML*; клетки фенотипа α предпочитают *HMR*. Можно получить дрожжевые штаммы, у которых обычная ориентация молчащих кассет изменена на обратную. Если их генотипы представлены локусами *HMLa* и *HMRα*, 92% переключений будут гомологичными, так как выбор доноров остается неизменным.

Интересной особенностью события переключения является его зависимость от происхождения клетки. Переключение определяется только в продуктах деления: **обе дочерние клетки имеют один и тот же тип спаривания**, отличный от типа спаривания родительской клетки. Можно предположить, что переключение происходит перед тем, как локус *MAT* реплицируется. (Сравните эту ситуацию с существующей зависимостью между транспозицией и репликацией у кукурузы, показанной на **рис. 37.14**).

Молчащие и активные кассеты имеют одинаковые последовательности

Как молчащие, так и активные локусы типов спаривания были клонированы в плаزمиды, реплицирующиеся в дрожжевых клетках. Структура генов представлена на

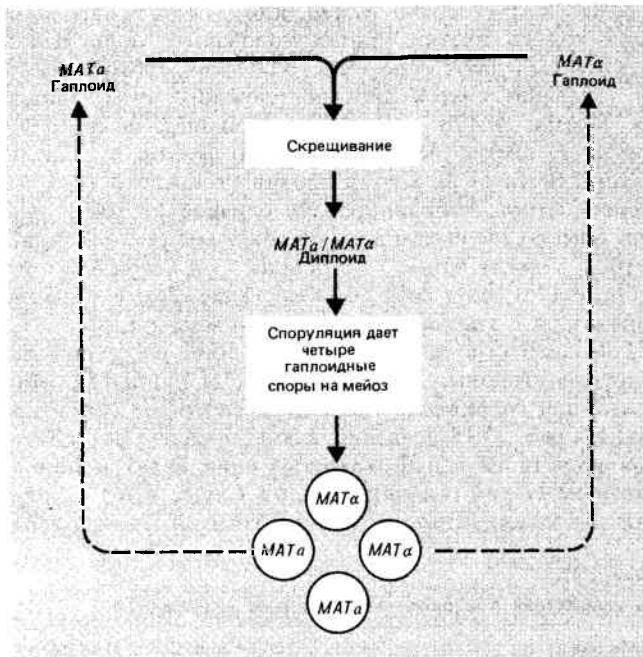


Рис. 37.15. Жизненный цикл дрожжей включает спаривание гаплоидов *MATa* и *MATα*, ведущее к образованию гетерозиготных диплоидов, которые при споруляции образуют гаплоидные споры.

рис. 37.17. Каждый тип аллеля содержит отличающуюся область, обозначенную $Y\alpha$ или $Y\alpha$. Эта область с двух сторон фланкирована участками X и $Z1$, которые практически идентичны во всех кассетах. (Идентифицированы только четыре различия, каждое из которых обусловлено точечной мутацией.) Более отдаленные фланкирующие области W и $Z2$ одинаковы в HML и MAT , однако в HMR они отсутствуют. (Организация одинаковая и не зависит от ориентации аллелей a и α .)

Мутации в $MAT\alpha$ можно отнести к двум группам комплементации (α^1 и α^2), в то время как мутации в $MATa$ попадают в одну группу ($a1$). Способность плазмид, несущих кассеты, комплементировать мутантные локусы в MAT позволяет приравнять отдельные области каждого локуса к комплементационным группам. Взаимосвязь между комплементационными группами и транскриптами, соответствующими MAT -аллелям, показана на рис. 37.17. Следует указать, что белки еще не охарактеризованы и данные о них получены только благодаря анализу кодирующих последовательностей.

Комплементационная группа $a1$ соответствует транскрипту $a1$, который кодируется полностью в пределах области $Y\alpha$, и поэтому является уникальным для локуса a -типа. Последовательность $a1$ кодирует два потенциальных продукта. Более короткий терминирует в кодоне UGA; однако нам известно, что может происходить считывание информации и за терминирующим кодоном, поскольку мутации, следующие за ним предотвращают проявление функции $a1$. Мы не знаем, выполняет ли более короткий белок функцию, отличающуюся от функции более длинного белка. Для $a2$ -транскрипта функция не была определена. Его транскрипция осуществляется с промотора, который, по крайней мере частично, находится в области $Y\alpha$. Транскрипт расположен в пределах общей для $MATa$ и $MAT\alpha$ области X и может быть транслирован с образованием любого из двух коротких полипептидов. Каждый из транскриптов, образуемых в локусе $MAT\alpha$, соответствует одной из известных групп комплементации. Транскрипты инициируются дивергентно в пределах $Y\alpha$, и каждый распространяется в прилегающую общую область (X и $Z1$). Транскрипт $\alpha1$ кодирует белок, основная часть которого кодируется областью $Y\alpha$. Транскрипт $\alpha2$ может транслироваться областью $Y\alpha$. Транскрипт $\alpha2$ может транслироваться с инициирующего кодона, расположенного в области $Y\alpha$, но большая часть последовательности соответствует области X .

Вывод о функциях продуктов сделан на основе результатов изучения мутантов из каждой группы комплементации. Эти данные представлены на рис. 37.18. У гаплоидов функции спаривания a выражаются конститутивно. Для индуцирования споруляции необходимы функции $a1$ и $\alpha2$. Группа $\alpha2$ также ответственна за репрессирование функций спаривания типа a у гаплоидов. Продукт $\alpha1$ индуцирует у гаплоидов функции α -типа спаривания.

Рассматривая карту транскрипции, представленную на рис. 37.17, можно обнаружить интересный факт. Транскрипция как $MATa$, так и $MAT\alpha$ инициируется в пределах области Y . Экспрессируется только MAT -локус; в то же время область Y представлена в соответствующих нетранскрибируемых кассетах, HML или HMR . Следовательно, регуляция экспрессии не осуществляется посредством прямого узнавания какого-то сайта, перекрывающегося с промотором. Отличать HML и HMR от MAT должен сайт, находящийся вне кассет.

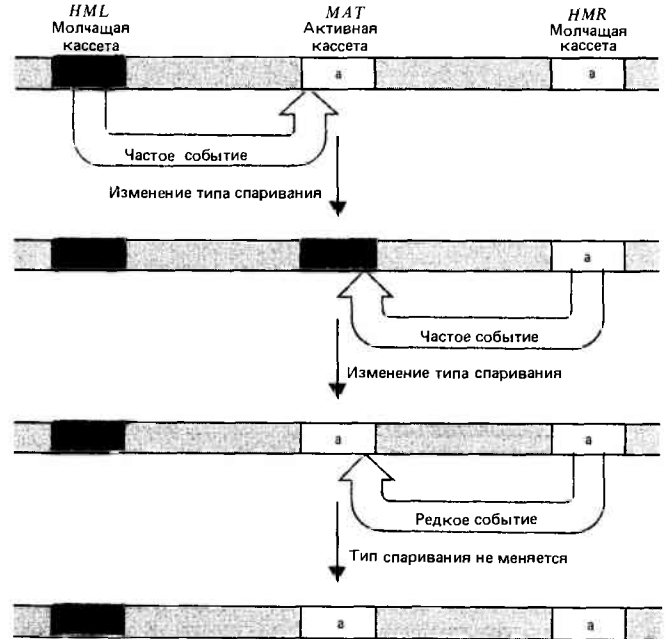


Рис. 37.16. Изменение типа спаривания происходит в результате замещения активных кассет молчащими кассетами противоположного генотипа; если транспозиция осуществляется между кассетами одного и того же типа, тип спаривания не меняется.

Можно ли определить способ контроля активности кассет путем идентификации других локусов, которые влияют на кассеты? Такие локусы были обнаружены при изучении мутаций, предотвращающих выражение аллеля, занимающего MAT . Другие мутации обуславливают экспрессию обычно молчащих кассет в HML и HMR . (Если экспрессируются обе функции, a и α , клетки ведут себя как гетерозиготные диплоиды.) Были идентифицированы четыре группы комплементации, мутации в которых ведут к выражению HML и HMR . Они названы SIR (от англ. *silent information regulator*). Для сохранения HML и HMR в репрессированном состоянии необходимы четыре локуса SIR дикого типа; мутация в любом из них, ведущая к появлению аллеля sir^- , имеет два эффекта: транскрибируются могут как HML , так и HMR -локусы; обе молчащие кассеты становятся мишенями для замещения при переключении. Таким образом, одно и то же регуляторное событие участвует как в репрессии молчащих кассет, так и в предотвращении их использования в качестве реципиентов при транспозиции.

Различные активности молчащих и активных кассет могут зависеть от структуры хроматина. Использование зондов, представляющих области, прилегающие к каждой из кассет, позволяет определить структуру хроматина с помощью метода концевой метки. Все кассеты содержат сайты, сверхчувствительные к ДНКазе I, однако существуют значительные различия по этому признаку между молчащими и активными кассетами. Локализация чувствительных сайтов показана на рис. 37.19.

Общий характер сверхчувствительности одинаков у $MATa$ и $MAT\alpha$. В каждом случае имеется сайт в 3'-конце левостороннего транскрипта ($a2$ или $\alpha2$). В области между двумя транскриптами имеется либо единичный (a), либо сдвоенный (α) сайт. Кроме того, пара сверхчувствительных сайтов есть вблизи границы $Y-Z$.

Все эти сайты отсутствуют в HML и HMR , за исклю-

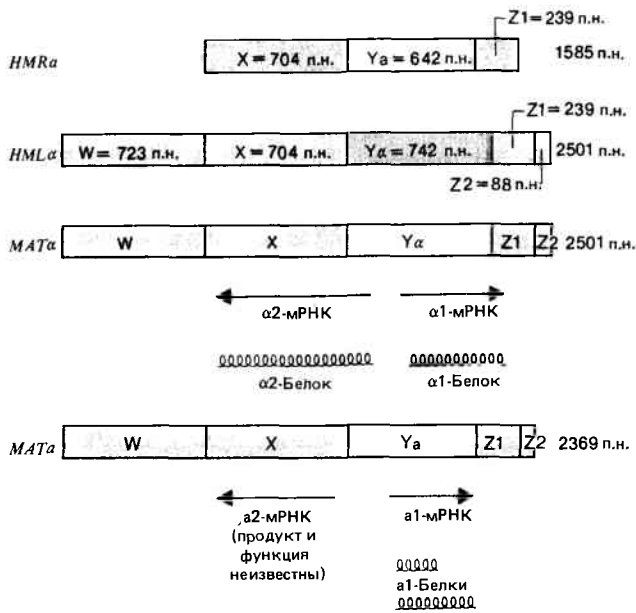


Рис. 37.17. Молчащие кассеты имеют те же последовательности, что и соответствующие активные кассеты. В α - и α -типах изменена только область Y. В *HMR* отсутствуют общие для *HML* и *MAT* последовательности W и Z2.

чением пары между $\alpha 1$ и $\alpha 2$. Однако введение мутации *sir*⁻ восстанавливает все утраченные сайты сверхчувствительности. Следовательно, активация кассет связана с присутствием сверхчувствительных сайтов. Можно предположить, что разные сайты имеют различные функции. Например, сайты в пределах Y могут быть необходимым (хотя и недостаточным) фактором для инициации транскрипции. Функция сайтов в пределах X неизвестна. Пара сайтов на границе Y—Z может быть важна для инициации события транспозиции (см. следующий раздел).

Некоторые сайты обнаружены во фланкирующей ДНК вблизи *HML* и *HMR*. Их присутствие не зависит от функции *SIR*; именно эти сайты помогли идентифициро-

вать молчащие кассеты как мишени для действия *SIR*. Делеционное картирование показывает, что фланкирующий сайт слева от *HML* очень важен для ответа на действие *SIR*; фланкирующий сайт справа также имеет значение. Узнавание этих сайтов должно каким-то неизвестным способом быть передано в пределы кассет для предотвращения образования сверхчувствительных сайтов. Вместо этого различные сверхчувствительные сайты позволяют локусу экспрессироваться и быть подвергнутым переключению.

Каким образом функция локуса *sir* проявляется на расстоянии? Одной из возможностей является передача структурного эффекта по длине ДНК. Например, степень суперспирализации может оказывать влияние на сайты внутри кассет или же доступность сайта может определяться состоянием нуклеосомы.

Однонаправленная транспозиция инициируется реципиентным локусом MAT

В популяциях, подвергающихся переключениям, рекомбинация между *MAT* и *HML* (или *HMR*) происходит с низкой частотой. Можно предположить, что процесс переключения включает спаривание между локусом *MAT* и одной из молчащих кассет. Обычным следствием этого взаимодействия является переключение типа спаривания, однако иногда вместо этого происходит реципрокная рекомбинация. Рекомбинационное событие, по существу, сливает рекомбинирующие локусы и вырезает материал между ними. Это указывает на то, что все три кассеты ориентированы на хромосоме одинаково.

Сайты, необходимые для транспозиции, могут быть идентифицированы с помощью мутаций, предотвращающих переключение. Делеции в правом конце области Y не проявляют какого-либо эффекта до тех пор, пока они не переходят границу области Z1; делеции в этом участке приводят к утрате способности переключения. Действующие в *cis*-положении точечные мутации, названные *inc* (от англ. *inconvertible*), были найдены как в *MAT α* , так и в *MAT α* . Они возникают в результате замещения пар оснований в области Z1, вблизи от границы с областью Y. В редких случаях, когда переключение все же происходит, мутация в *MAT* утрачивается, что свидетельствует о ее локализации в замещаемой области. Мутации, относящиеся к другому классу (*stk*), сохраняются после события переключения, так что, по-видимому, они лежат вне области, которая подвергается замещению. Следовательно, замещение локуса *MAT* включает сайт вблизи границы Y—Z1 и для переключения совершенно необходимы последовательности как внутри, так и вне замещаемой области. Доказательством однонаправленности процесса можно считать отсутствие таких мутаций в *HML* и *HMR*, которые предотвращают переключение; кроме того, если мутации *inc* перенесены в молчащие кассеты, они не предотвращают переключение. В действительности они эффективно транспозируются в локус *MAT*, после чего мутация *inc* снова проявляет свой эффект.

Граница области Y—Z представляет собой сайт изменения ДНК, которое может быть связано с транспозицией. В клетках, подвергающихся переключению, 1–3% ДНК локуса *MAT* содержат двухцепочечный разрез именно в этом сайте. Разрез находится в пределах 10 пар оснований пограничной области и совпадает с сайтом сверхчувствительности к ДНКазе.

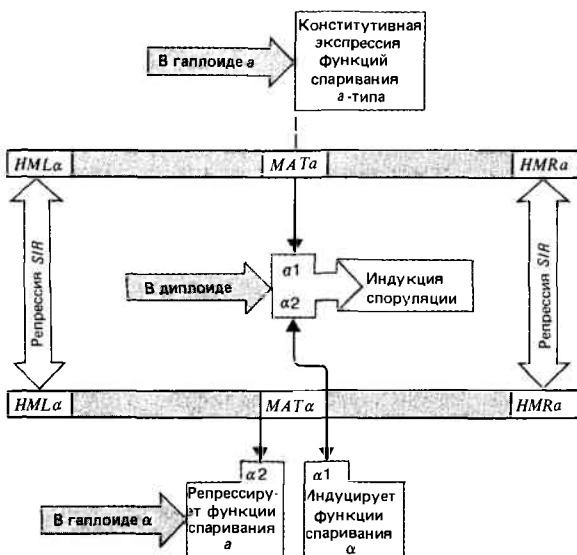


Рис. 37.18. Для индукции споруляции в диплоидах функции $\alpha 1$ и $\alpha 2$ кооперируются. У гаплоидов функция $\alpha 2$ репрессирует функции спаривания; $\alpha 1$ индуцирует функции спаривания α .

	W	X	Y	Z1	Z2
<i>MATa</i>		←*	*→		
<i>MATα</i>		←*	*→		
<i>HMLα</i>		*	* * * *		
<i>HMRα</i>	* *	*	* * *		
		X	Y	Z1	

* Сверхчувствительные сайты всегда присутствуют
 * Сверхчувствительные сайты присутствуют только в штаммах *sir*

Рис. 37.19. Большинство сайтов, сверхчувствительных к ДНКазе I, которые обнаружены в локусе *MAT*, не присутствуют в *HML* или *HMR*; однако в мутантах *sir*⁻ они активируются. Некоторые постоянные сверхчувствительные сайты найдены в областях, фланкирующих *HML* и *HMR*.

Рассмотрим предполагаемую серию событий, происходящих при переключении. Для этой цели может быть пригодным сверхчувствительный сайт на границе Y—Z1 области *MAT*-локуса, так как он отсутствует в нуклеосоме. Его узнает эндонуклеаза, контролируемая локусом *HO*. Возможно, сверхчувствительность к ДНКазе I *in vitro* отражает природную чувствительность к эндонуклеазе *HO in vivo*!

Эндонуклеаза *HO*, называемая также YZ-эндонуклеазой, вносит ступенчатый разрез в области Z1, находящейся на расстоянии трех пар оснований от сайта связывания YZ. Она не разрезает мутантные локусы *MAT*, которые не способны переключаться. Последовательности, содержащие менее 11 пар оснований мишени после делетирования, также не разрезаются, что свидетельствует об относительно большой протяженности сайта узнавания нуклеазой. Вероятно он не встречается ни в одном другом участке генома.

Синтез эндонуклеазы *HO* находится под контролем клеточного цикла. Время ее образования может объяснять характер переключения. Ген выражается только в течение фазы G1 материнской клетки, и это объясняет почему оба ее потомка имеют новый тип спаривания.

На рис. 37.20 схематически изображена реакция в отношении одиночных цепей. (Естественно, что в случае двухцепочечной ДНК события значительно сложнее.) Допустим, что свободный конец Z локуса *MAT* атакует либо *HML*, либо *HMR*-локус и спаривается с областью Z. Область Y локуса *MAT* деградирует до тех пор, пока не открывается участок, гомологичный области X. Затем локус *MAT* спаривается с *HML* или *HMR* как в левой (X), так и правой (Z1) части. Область Y локусов *HML* или *HMR* копируется, чтобы заместить утраченный участок *MAT*-локуса (который может простирается за пределы области Y). Спаренные локусы расходятся. (Порядок событий может быть другим.) Стадии, следующие за первоначальным разрезом, требуют участия ферментов, осуществляющих общую рекомбинацию. Отметим две особенности этого процесса. В отличие от других событий транспозиции переключение предполагает спаривание в обоих концах транспозируемого сегмента, деградацию и замещение промежуточного материала. Процесс инициируется в *MAT*-локусе, который должен быть замещен. В этом смысле описание *HML* и *HMR* как донорных локусов относится к их основной роли, но не к механизму

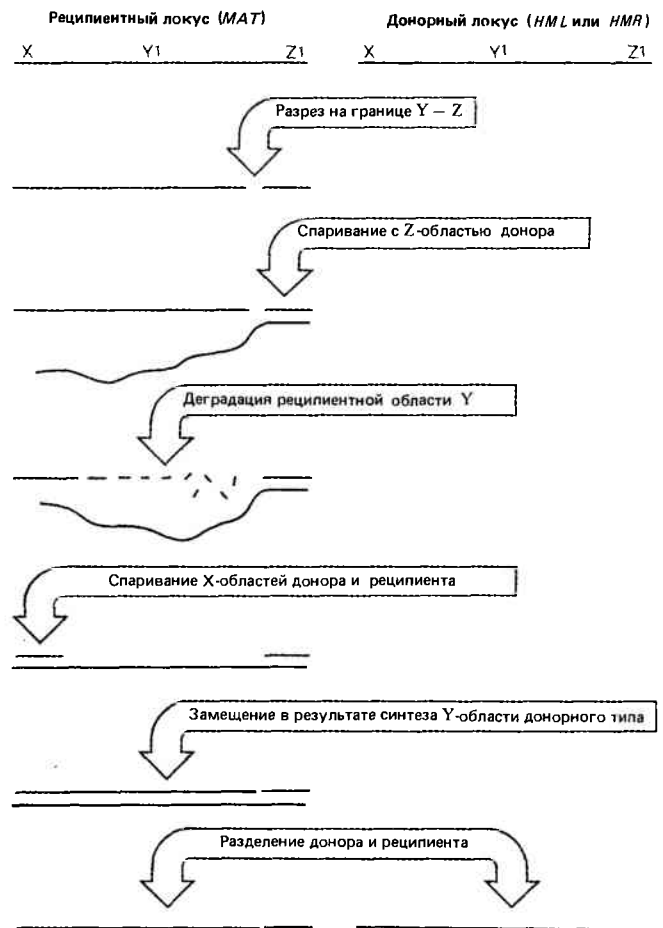


Рис. 37.20. Замещение кассеты инициируется реципиентным локусом и заключается в спаривании с донорным локусом областей, расположенных по обе стороны от области Y. Последовательность, которая удаляется и замещается, может захватывать область X. На рисунке показана реакция, происходящая на отдельных цепях реципиентной и донорной молекул; на самом деле процесс происходит в двухцепочечных областях ДНК.

процесса, который является обратным по отношению к обычной инициации транспозиции донором. В последовательностях этих локусов не идентифицированы какие-либо короткие прямые повторы. По-видимому, в данном случае механизм объединения не включает использование ступенчатых разрезов, обнаруженных для других транспозиционных событий.

Рекомендуемая литература

Эукариотические транспозирующиеся элементы детально рассмотрены в сб. «Mobile Genetic Elements» (ed. Shapiro, Academic Press, New York, 1983). Особый интерес представляют главы, в которых обсуждаются Tu-элементы дрожжей, транспозоны *D. melanogaster*, рассматриваются данные картирования локуса *white*. Заслуживают внимания сведения о молекулярной структуре контролирующих элементов кукурузы. В сборнике представлен исчерпывающий обзор о типах спаривания у дрожжей.

Глава 38

ЭЛЕМЕНТЫ, СПОСОБНЫЕ К ПЕРЕМЕЩЕНИЮ
В ПРЕДЕЛАХ ГЕНОМА И ВНЕ ЕГО

По всей вероятности, транспозоны не способны к существованию вне генома. Независимо от того, являются ли они компонентами, играющими важную роль в жизни клетки, или ее паразитами, занятыми только своим собственным воспроизведением, свободных молекул ДНК они не образуют. В этой главе мы познакомимся с элементами, способными существовать в составе генома и вне его; некоторые из них являются природными, другие образовались в результате рекомбинации в лабораторных условиях.

В отдельных случаях последовательность РНК служит предшественником геномной последовательности (ДНК). Вероятно, РНК должна была превратиться в двухцепочечную ДНК, которая затем внедрялась в геном с помощью события, напоминающего транспозицию. Прямое указание на существование такого пути получено при изучении ретровирусов, у которых такой путь представляет собой часть обычного инфекционного цикла. Структура некоторых псевдогенов и других последовательностей (включая членов семейства *Alu*) свидетельствует, что аналогичный путь может быть пройден клеточной последовательностью (достаточно редко). В обоих случаях внехромосомный элемент дает начало геномной последовательности.

Изменения в относительном соотношении компонентов генома иногда происходят во время соматического развития. Хорошо известно, например, увеличение числа копий определенных генов у личинок насекомых. Благодаря возможности отбирать варианты клеток с увеличенным числом копий определенного гена показана случайная амплификация генов в культуре клеток млекопитающих. Иницируемое внутри генома событие амплификации способствует созданию дополнительных копий гена, которые существуют либо в составе хромосомы, либо в форме внехромосомных элементов.

Открытие транспозиции и других перестроек в ДНК в какой-то мере способствовали нашим знаниям об удивительной способности последовательностей ДНК приспосабливаться. Чужеродную ДНК можно ввести в эукариотические клетки, где она выживает и может экспрессироваться. В некоторых случаях чужеродная ДНК остается внехромосомной, иногда она интегрируется в геном. Взаимоотношения между внехромосомными и геномными формами необычны и скорее зависят от случайных и в некоторой степени непредсказуемых событий, чем напоминают взаимный обмен между свободными и интегрированными формами бактериальных плазмид. Тем не менее процесс может привести к стабильному изменению в геноме. Например, ДНК, инъецированная в яйцеклетки животных, способна включаться в геном, наследоваться и функционировать. Благодаря этому оказывается возможным осуществлять весьма сложные манипуляции с последовательностями ДНК как в природных, так и в экспериментальных условиях.

Жизненный цикл ретровирусов связан
с событиями,
напоминающими транспозицию

Геном ретровирусов представлен одноцепочечной РНК, которая реплицируется в виде промежуточной двухцепочечной ДНК. Класс ретровирусов, о котором мы имеем самую большую информацию, характеризуется способностью образовывать в клетках птиц и млекопитающих частицы *С-типа*. В каждом вирионе упакованы две копии генома РНК, в результате чего каждая отдельная вирусная частица фактически диплоидна. Если клетка одновременно инфицирована двумя различными вирусами *С-типа*, можно получить гетерозиготные вирусные частицы, несущие по одному геному каждого типа. Диплоидность может влиять на возможность приобретения вирусом клеточных последовательностей. Типичная последовательность ретровируса содержит три «гена». Под этим термином в данном случае мы понимаем кодирующие области, каждая из которых на самом деле дает начало множеству белков в ходе процессинга. Порядок генов, *gag—pol—env*, приведен в верхней части рис. 38.1. Вирусная РНК транслируется как обычная информационная РНК (она содержит кэп на 5'-конце и poly(A) на 3'-конце) в полибелок Gag.

Одна из информационных РНК, в результате сплайсинга утратившая терминирующий кодон в гене *gag*, транслируется с образованием белка Gag—Pol. Сплайсинг также способствует образованию более коротких субгеномных информационных РНК, которые при трансляции образуют полипротеин Env, предшественник двух белков.

Ген *gag* дает начало белковым компонентам нуклеопротеинового ядра вириона. Ген *env* кодирует компоненты вирусной оболочки, которая захватывает также компоненты клеточной цитоплазматической мембраны. Ген *pol* кодирует фермент, получивший название обратной транскриптазы, который состоит из двух субъединиц. Одна из субъединиц представляет собой образующийся при процессинге фрагмент другой субъединицы. Соответственно названию фермент осуществляет превращение РНК-генома в комплементарную цепь ДНК. Он катализирует также последующие стадии при образовании двухцепочечной ДНК, обладает ДНК-полимеразной активностью, ведет себя подобно РНКазе H (способен деградировать РНК, входящую в состав гибрида РНК—ДНК) и обладает эндонуклеазной активностью.

Обратная транскриптаза упаковывается вместе с геномом в вирусную частицу. Фермент превращает РНК в линейную двухцепочечную молекулу в цитоплазме инфицированной клетки. Линейная ДНК самостоятельно проникает (неизвестным способом) в ядро, где она превращается в кольцевую молекулу. Одна или несколько копий ДНК интегрируют в геном хозяина. Интегриро-

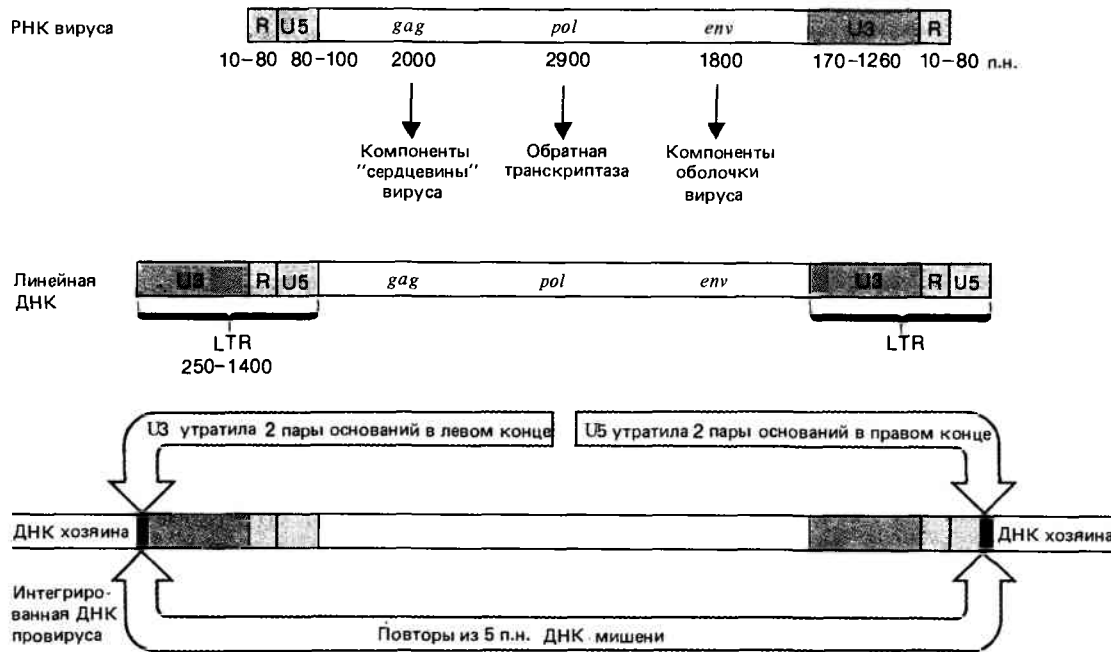


Рис. 38.1. Ретровирусные РНК кончаются прямыми повторами (R), а свободные линейные ДНК — последовательностями LTR;

концы провирусов представляют собой LTR-последовательности, укороченные на два основания.

ванная провирусная ДНК транскрибируется аппаратом хозяина с образованием вирусных РНК, которые затем используются как в качестве мРНК, так и в качестве геномов для упаковки в вирионы. Интеграция представляет собой нормальный этап жизненного цикла вируса и необходима для транскрипции.

Структуры различных форм ДНК вируса сравниваются с РНК на рис. 38.1. Вирусная РНК содержит прямые повторы на концах. В различных штаммах вируса длина таких R-сегментов варьирует от 10 до 80 нуклеотидов. Рядом с R-сегментом в 5'-конце вирусного генома находится область U5, состоящая из 80–100 нуклеотидов. Ее название свидетельствует о локализации этой последовательности на 5'-конце. В 3'-конце R-сегменту предшествует уникальный сегмент U3, содержащий 170–1250 нуклеотидов.

Концы линейной ДНК содержат дополнительные последовательности. Сегмент U3 добавляется к 5'-концу; сегмент U5 — к 3'-концу. В результате каждый конец ДНК имеет последовательность U5—R—U3; она получила название **длинного концевого повтора (LTR)**. Ее образование связано с реакцией, схема которой представлена на рис. 38.2. Обратная транскриптаза переключает матрицы, перенося образующуюся ДНК на новую матрицу. На рисунке показано образование одного LTR; для образования повтора на другом конце требуется подобное событие. Характерная черта вирусного генома — идентичность LTR, находящихся на его концах. Это объясняется их происхождением. 3'-конец U5 состоит из короткого инвертированного повтора, родственного 5'-концу последовательности U3, в результате чего сама последовательность LTR фланкирована короткими инвертированными повторами. Следовательно, организация ДНК ретровируса напоминает организацию транспозирующихся элементов, подобных *copia* (см. рис. 37.2).

Жизненный цикл ретровирусов схематически изображен на рис. 38.3. События, происходящие в период между

образованием линейной ДНК и ее интеграцией, еще недостаточно хорошо изучены. После внедрения в ядро линейная ДНК дает начало двум формам кольцевых ДНК. (Число кольцевых молекул в разных штаммах колеблется от 0,1 до 5 копий на клетку.) Обнаружены две основные формы мономерных замкнутых кольцевых ДНК. Наиболее часто встречаемая форма содержит одну копию LTR, что может объясняться ее происхождением в результате гомологичной рекомбинации между двумя LTR. (Мы не знаем, действительно ли она образуется благодаря такому механизму, но одиночные LTR, которые могут быть последствиями вырезания провирусов, существуют в некоторых клеточных геномах.) Менее часто встречающаяся форма содержит две tandemные копии LTR; иногда они соединены непосредственно, как в случае лигирования тупых концов; иногда в одном или обоих концах LTR в местах соединения утрачены нуклеотиды. Роль таких кольцевых молекул в процессе интеграции еще не выяснена. Необходимо вновь подчеркнуть их сходство с кольцевыми молекулами элементов *copia* , найденными у *D. melanogaster* (гл. 37).

Вирусная ДНК встраивается в случайные сайты генома клетки-хозяина. Инфицированная клетка содержит от одной до десяти копий провируса. В каждом сайте внедрения образуются короткие прямые повторы ДНК мишени. Их длина у разных вирусов может быть равна 4, 5 или 6 парам оснований. Наличие прямых повторов свидетельствует о том, что механизм интеграции включает образование ступенчатых разрезов в ДНК хозяина, аналогичных тем, которые образуются при бактериальной транспозиции. Сайт интеграции специфичен в отношении вируса, и интегрированная ДНК отличается от неинтегрированной двумя парами оснований в каждом конце. Таким образом, интегрированная вирусная ДНК утрачивает две пары оснований в левом конце 5'-концевой последовательности U3 и две пары оснований в правом конце 3'-концевой последовательности U5.

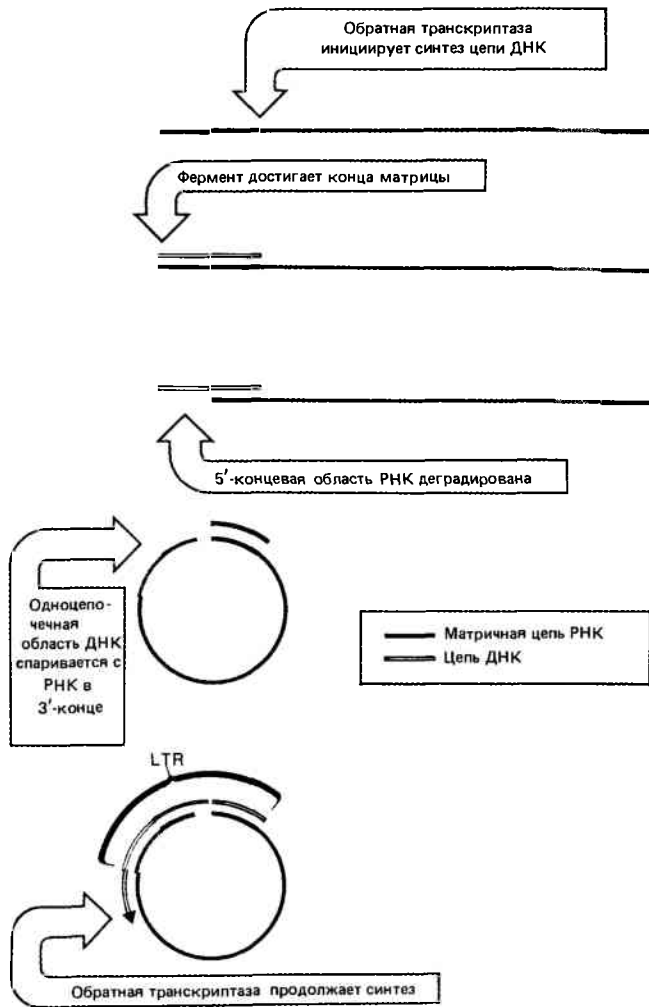


Рис. 38.2. Последовательности LTR образуются при переключении матриц во время обратной транскрипции. Новая матрица может быть другим концом той же молекулы (как показано на рисунке) или другой молекулой.

Область U3 каждого LTR несет промотор. Промотор в левом LTR ответствен за инициацию транскрипции провируса. Иногда (довольно редко) промотор в правом LTR способствует транскрипции последовательности хозяина, прилегающей к сайту внедрения. Интеграция генома ретровируса может обуславливать опухолевую трансформацию клеток путем активации определенных типов клеточных генов.

Способность «включать» фланкирующие последовательности хозяина аналогична свойству некоторых бактериальных транспозонов, хотя способность к включению может быть иногда обусловлена скорее наличием усилителя (*enhancer*), чем промоторными последовательностями. Активность провирусного генома, вероятно, зависит и от сайта, в котором произошла интеграция вируса в геном хозяина.

До сих пор мы имели дело с ретровирусами в условиях инфекционного цикла, когда для образования новых копий РНК необходим процесс интеграции. Однако при интеграции вируса в клетку зародышевой линии его эффективная экспрессия прекращается, он способен стать наследственным эндогенным вирусом организма. Такое состояние вируса более всего изучено у мышей и цыплят, чьи геномы несут неактивные эндогенные вирусы. Иногда

эндогенные вирусы активируются внешними факторами. Одним из таких факторов является инфекция клетки другим вирусом.

Ретровирусы способны транслировать клеточные последовательности

Обнаружение транслирующих вирусов, геномы которых приобрели клеточные последовательности, позволило взглянуть на жизненный цикл вируса с несколько другой точки зрения. На рис. 38.4 показан геном одного из таких вирусов. Часть вирусной последовательности – в данном случае ген *env* – замещена геном *v-onc*. Образовавшийся вирус **дефектен по репликации** и не способен поддерживать самостоятельно инфекционный цикл. Такой вирус может сохраняться вместе с **вирусом-помощником**, который восполняет утраченные вирусные функции.

Название гена *onc* (от англ. *oncogenesis*) связано с названием процесса, который он вызывает: онкогенез или опухолевая трансформация культивируемых клеток. Она выражается в нарушении регуляции роста и вследствие этого в неограниченном делении клеток. За появление опухолеродных клеток могут быть ответственны как вирусные, так и клеточные *onc*-гены.

Гены *onc*, входящие в состав генома ретровирусов, получили название *v-onc*, и именно они придают вирусу способность трансформировать определенный тип клеток хозяина. Лocus с гомологичными последовательностями, найденные в геноме хозяина, получили название *c-onc*-генов. В большинстве случаев о функциях клеточных генов ничего не известно. Однако в некоторых экспериментах по трансфекции были идентифицированы активные *c-onc* гены. Их добавление к реципиентным клеткам вызывало трансформацию последних. Такие гены могут быть активированы и при интеграции вблизи них ге-

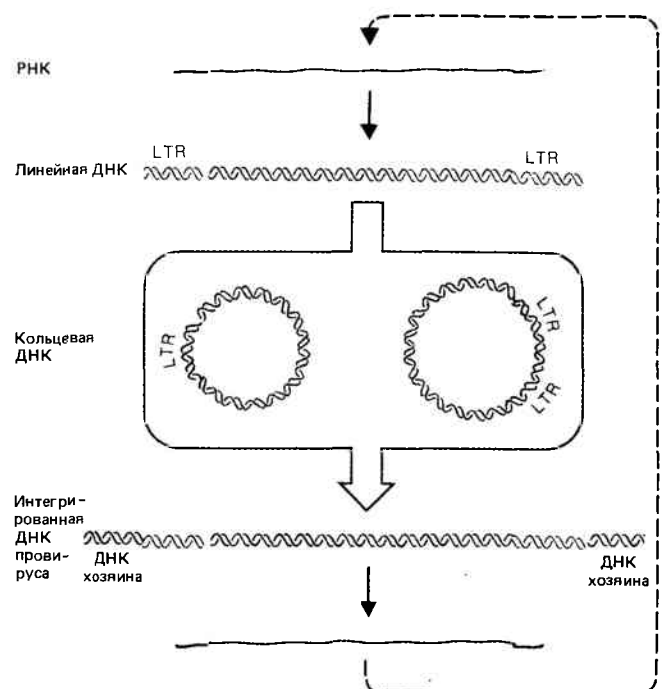


Рис. 38.3. Интеграция ретровируса сопровождается образованием двух типов кольцевых молекул ДНК.



Рис. 38.4. У дефектных в отношении репликации трансформирующих вирусов часть вирусной последовательности замещена последовательностью клетки. Длина замещенной последовательности характерна для каждого вируса.

нома ретровируса, как это описано в предыдущем разделе. Существует большая группа *onc*-генов. Отдельным членам этой группы даны названия, соответствующие названиям вирусов, в которых они впервые были обнаружены (например, *src*-ген выделен из вируса саркомы Рауса, *fes* – из вируса саркомы кошек и т.д.).

Каким образом ретровирусы приобрели гены *onc*? Некоторое представление об этом дает различие, обнаруженное в структуре генов *c-onc* и *v-onc*. Гены *c-onc* обычно прерываются интронами, и их экспрессия в клетке происходит на низком уровне. Гены *v-onc* непрерывны и экспрессируются эффективно как часть вирусной единицы транскрипции. Сравнение их структур позволило предположить, что гены *v-onc* образовались в результате сплайсинга копий РНК генов *c-onc*.

Первые гипотезы о происхождении трансформирующих вирусов основывались на допущении возможности случайной упаковки клеточных мРНК в вирионе вместе с геномом вируса; затем должна была иметь место гомологичная рекомбинация между РНК, в результате которой могли возникнуть трансформирующие вирусные геномы. Однако данную модель пришлось пересмотреть. Для этого были две причины. Во-первых, *c-onc*-последовательности обнаружены в виде редко встречающихся мРНК, в которых отсутствуют сигналы упаковки. Во-вторых, на уровне РНК осуществление рекомбинационных событий кажется маловероятным.

Существующая в настоящее время модель образования трансформирующих вирусов представлена на рис. 38.5. Предполагается, что ретровирус встроился вблизи *c-onc*-гена. В результате делеции происходит слияние генома провируса с геном *c-onc*, затем транскрипция ведет к образованию объединенной РНК, содержащей вирусные последовательности в одном конце и клеточные последовательности *onc* в другом. В результате сплайсинга удаляются интроны как в вирусной, так и в клеточной частях молекулы. РНК имеет соответствующие сигналы для упаковки в вирионы. Вирионы могут образовываться, если в клетке содержится другая, интактная, копия провируса. В этих случаях некоторые диплоидные частицы вирусов содержат одну слившуюся и одну вирусную РНК.

Рекомбинация между этими последовательностями может привести к образованию трансформирующего генома, в котором повторы вируса присутствуют в обоих концах. (Рекомбинация происходит с высокой частотой во время инфекционного цикла ретровируса; по-видимому, в ней используются промежуточные продукты ДНК или дополнительные реакции переноса цепи. Можно предположить, что между вирусным геномом и клеточной частью слившихся РНК происходит негомологичная рекомбинация при участии того же механизма, который ответствен за рекомбинацию ретровирусов.) Какова связь между ними и транспозируемыми элементами? Обнаружена существенная гомология в последовательностях LTR вируса лейко-саркомы птиц AL-SV и прямых концевых повторов *coria*-подобных элементов 297 и 17.6 у *Drosophila*. Возможно *coria*-подобные элементы возникли в результате иммобилизации ретровирусов?

Могли ли ретровирусы произойти от клеточных элементов? Основные свойства, характеризующие всю группу ретровирусов, позволяют высказать предположение о наличии у них общего предшественника. Их

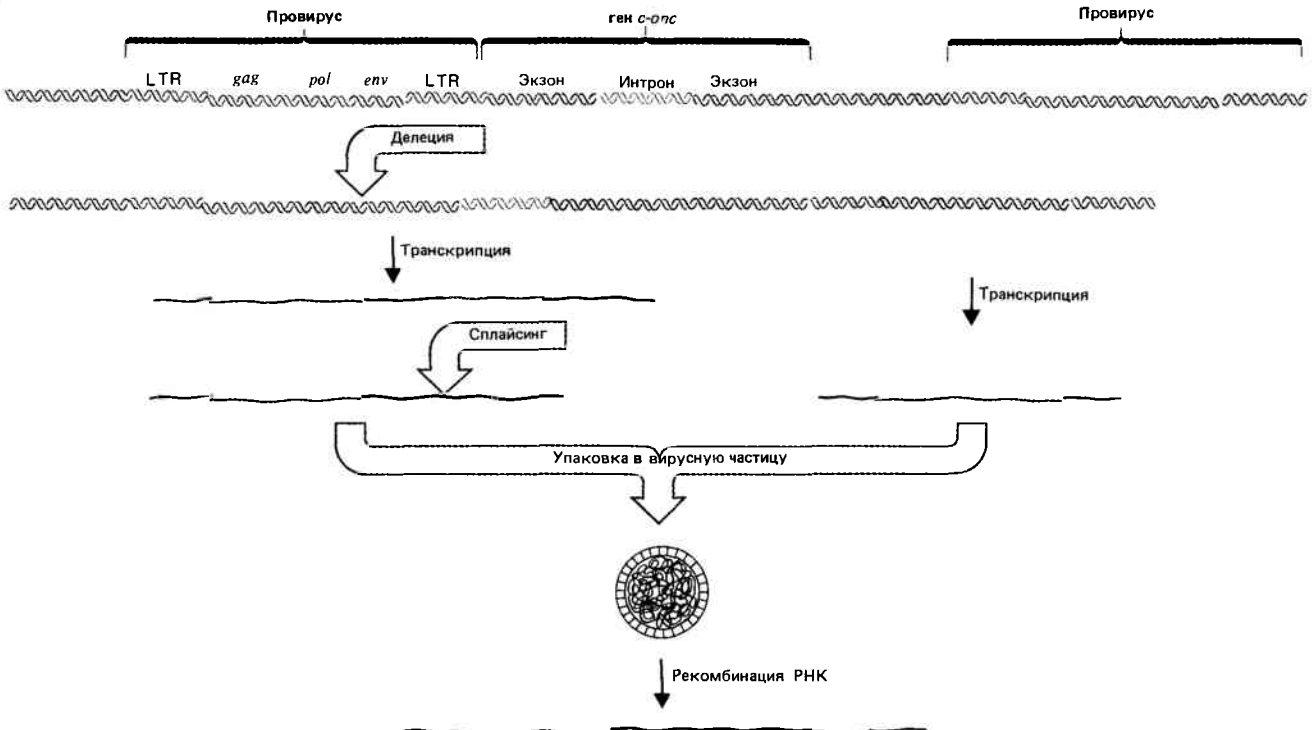


Рис. 38.5. Дефектные в отношении репликации вирусы могут возникать благодаря интеграции и делеции вирусного генома, в результате чего образуются слившиеся вирусно-клеточные транскрипты, которые упаковываются вместе с нормальным

РНК-геномом. В образовании дефектного по репликации трансформирующего генома участвует процесс негомологичной рекомбинации.

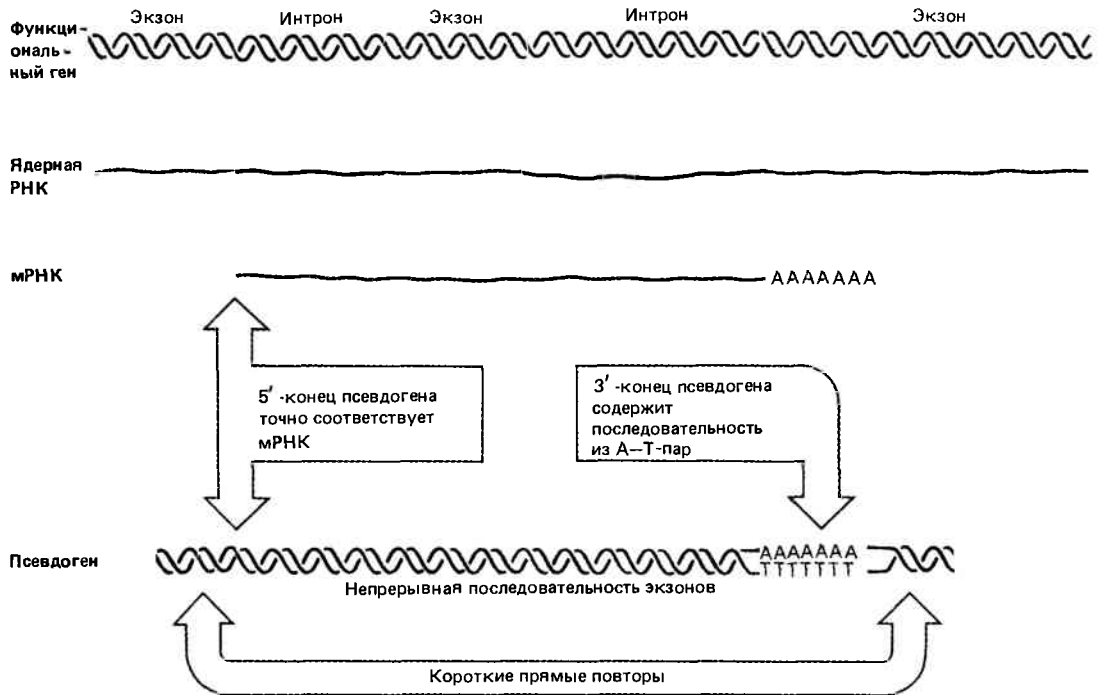


Рис. 38.6. Псевдогены могли возникнуть в результате обратной транскрипции РНК, ведущей к образованию двухцепочечных ДНК, которые интегрировали в геном.

сходство с транспозирующимися элементами свидетельствует о том, что изначально с обеих сторон от гена клетки-хозяина, кодирующего полимеразу нуклеиновой кислоты, могли оказаться IS-подобные элементы; в этом случае образовавшаяся единица должна иметь структуру *LTR-pol-LTR*. Эта структура затем могла приобрести полимеризующую активность, позволяющую манипулировать с ДНК- и РНК-субстратами, и таким образом эволюционировать в ретровирус. Трансформирующие гены могли быть включены позднее. (Нет оснований полагать, что механизм, участвующий в приобретении клеточных функций, является уникальным для трансформирующих генов; однако такие вирусы могли иметь селективное преимущество благодаря их влиянию на рост инфицированных клеток.)

В клетке могли происходить РНК-зависимые транспозиции

Некоторые псевдогены проявляют такие внешние и внутренние признаки, которые свидетельствуют о возможности их происхождения из последовательности РНК, хотя относительно образования ДНК-копий таких генов можно высказать лишь гипотезы. Характерные признаки таких псевдогенов сравниваются на рис. 38.6 с признаками исходного гена и мРНК. На рисунке представлены все диагностические признаки, однако в каждом конкретном случае обнаружить можно лишь некоторые из них.

Начало псевдогена соответствует точке, эквивалентной 5'-концу РНК, и именно это свидетельствует о происхождении его ДНК из РНК. Несколько псевдогенов состоят из точно сочлененных последовательностей экзонов; поскольку известно, что не существует механизма, позволяющего узнавать последовательности интронов

в ДНК, такая особенность доказывает посредничество РНК при образовании псевдогенов. Псевдоген может оканчиваться короткой областью из А—Т-пар оснований, образовавшейся, по-видимому, из poly(A)-хвостовой части РНК. С обеих сторон от псевдогена имеются короткие прямые повторы. Возможно, что их присутствие — результат события, напоминающего транспозицию. Подтверждением гипотезы, что эти псевдогены произошли с помощью иного механизма, нежели тот, который ответствен за образование псевдогенов, найденных вблизи активных членов генных кластеров, служит факт обнаружения псевдогенов в местах, не связанных с их предполагаемыми исходными сайтами локализации.

Если псевдоген произошел из последовательности мРНК, его гомология в 5'-конце не может распространяться на последовательность, расположенную выше сайта инициации, как это показано на рис. 38.6. Однако α -глобиновый псевдоген мыши в действительности имеет некоторую гомологию с активным геном в области, расположенной выше точки инициации. Модель процессинга можно было объяснить, допустив, что инициация происходит в более отдаленном участке, расположенном в направлении против хода транскрипции, в результате чего образуется более длинная РНК (либо вследствие инициационного события, отклоняющегося от нормы, либо вследствие локализации промотора в другой области).

Другие последовательности, напоминающие по своим свойствам транспозоны, — это члены семейства Alu. Они фланкированы короткими прямыми повторами, однако длина повторов у отдельных членов семейства различна. Вставка в Alu-последовательность, ответственная за образование малой ядерной РНК, также, по-видимому, является результатом транспозиционного события. Свидетельства давно происшедших транспозиционных событий можно обнаружить и в генах нескольких малых ядерных РНК.

Все рассмотренные примеры отличаются от ретровирусов: предполагаемые транспозируемые последовательности сами по себе не несут какой-либо информации для транспозиционного события (или обратной транскрипции РНК). Мог ли этот процесс осуществляться с помощью ретровирусов? Происходило ли это с участием клеточной системы? Не исключено, что концы транспозируемых последовательностей случайно напоминали последовательности концов транспозона. Происходит ли в этих генах транспозиция в настоящее время, или мы видим последствия давно прошедших событий?

Тканеспецифичные вариации в геноме дрозофилы

Состав генома в большинстве соматических клеток сохраняется постоянно. Однако теперь мы знаем о некоторых ситуациях, при которых происходят изменения в соотношении отдельных последовательностей. Наиболее известны изменения во время развития личинок у насекомых. Регуляция ведет к образованию недореплицировавшихся или избыточно реплицировавшихся специфических последовательностей.

Недорепликация встречается при формировании политенных тканей мухи. В таких тканях образуются гигантские хромосомы (гл. 28); их возникновение обусловлено множественными последовательными дупликациями исходного (синаптированного) диплоидного набора хромосом.

Примерно четверть генома *D. melanogaster* приходится на гетерохроматиновую область хромосом, которые агрегируют с образованием хромоцентра (большая его часть состоит из последовательностей сателлитной ДНК). Относительное количество гетерохроматина в политенных клетках значительно меньше, чем в клетках диплоидов. Это объясняется тем, что часть генома не способна дублироваться во время политенизации последовательностей ДНК эухроматина. Измерение содержания сателлитной ДНК показывает, что эти последовательности не дублируются или дублируются лишь несколько раз по сравнению с девятью дупликациями эухроматиновой ДНК в слюнной железе.

Гены рРНК также относятся к недореплицированным в политенных тканях. Область рДНК дублируется только 6–7 раз (достигая одного и того же конечного уровня независимо от того, присутствует только один ядрышковый организатор или оба). Такой тип дозового контроля специфичен для рДНК.

Соответственно контролю репликации в политенных тканях геном *D. melanogaster* может быть разделен на три типа областей. Естественно предположить, что точки начала репликации в каждом типе контролируют число инициационных событий. Точки начала репликаций в сателлитной ДНК могут не узнаваться; точки начала репликации в рДНК прекращают функционировать в ответ на действие некоторых конечных продуктов ряда генов рРНК; в то же время инициация в общих точках начала репликации эухроматина продолжается вплоть до самого конца.

У насекомых также наблюдается дифференциальная амплификация определенных последовательностей, кодирующих белок. В слюнных железах *Rhyncosciara* и других сциарид образуются **пуфы ДНК**. Внешне они напоминают пуфы, которые представляют активные области политенных хромосом дрозофил, однако пуфы сциарид содер-

жат локально амплифицированную ДНК, а также РНК. Амплифицированные последовательности могли четыре раза подвергаться дополнительным инициационным событиям репликации, в результате чего число копий увеличивалось в 16 раз.

Пониманию возможных механизмов дифференциальной амплификации способствует состояние генов хориона при развитии *D. melanogaster*. Белки, составляющие хорион (оболочку яйца), синтезируются и секретируются полиплоидными клетками яйцевого фолликула. Гены хориона насекомых имеют тенденцию к группировке; у *D. melanogaster* идентифицированы две группы таких генов. До экспрессии в клетках фолликула гены хориона в X-хромосоме амплифицируются шестнадцатикратно (четыре дополнительных удвоения), тогда как гены в хромосоме III амплифицируются до 60 раз (шесть дополнительных удвоений).

Амплификация всегда распространяется на расстояние, равное примерно 45–50 т.п.н. с каждой стороны от генов хориона. Уровень максимальной амплификации представляет плато, образованное примерно 20 т.п.н., окружающими гены хориона. Степень амплификации показывает градиент уменьшения с каждой стороны от плато.

Чем обусловлена амплификация? Градиент уменьшения показывает, что концевые точки амплифицируемых областей в отдельных молекулах гетерогенны. Для объяснения происходящих событий предложена модель, представленная на **рис. 38.7**. В соответствии с ней в точке начала репликации в центре области происходят множественные инициационные события, ведущие к двунаправленной репликации. Репликационные вилки продвигаются вперед на расстояния от 10 до 50 т.п.н. Эта модель рассматривает всю амплифицируемую область как репликон, представленный множеством только частично реплицировавшихся копий. (Взаимосвязь между частично реплицировавшимися и многократно реплицировавшимися областями в хромосомах слюнных желез могла бы быть объяснена подобной моделью, в которой места соединения между областями представлены неподвижными репликационными вилками.)

Содержит ли амплифицированная область точку, ответственную за начало репликации и активную в *цис*-положении? Если да, то амплификация не происходила бы в хромосоме с мутировавшей точкой начала репликации. В гетерозиготе такая дефектность па репликации должна проявляться только в мутантной хромосоме. В случае транслокации точки начала репликации в другую область она должна была бы обеспечивать амплификацию любого репликона, в котором она оказалась.

Свойства мутанта с дефектным строением глазных фасеток (*ocelliless*) подтверждают такое предсказание. Эта мутация вызывает сложные изменения в фенотипе; она представляет собой инверсию примерно трех дисков, содержащих область генов хориона в X-хромосоме. Гены хориона локализируются в пределах 3 т.п.н. левого конца инверсии.

Инверсия определяет основное изменение, проявляющееся в характере амплификации. На **рис. 38.8** сравниваются области, амплифицированные в клетках мух дико-го и мутантного типов. Область длиной 40 т.п.н. слева от точки разрыва инверсии вообще не способна амплифицироваться. 50 т.п.н., находящихся с другой стороны, амплифицированы, хотя уровень амплификации уменьшен и составляет примерно половину от обычного макси-

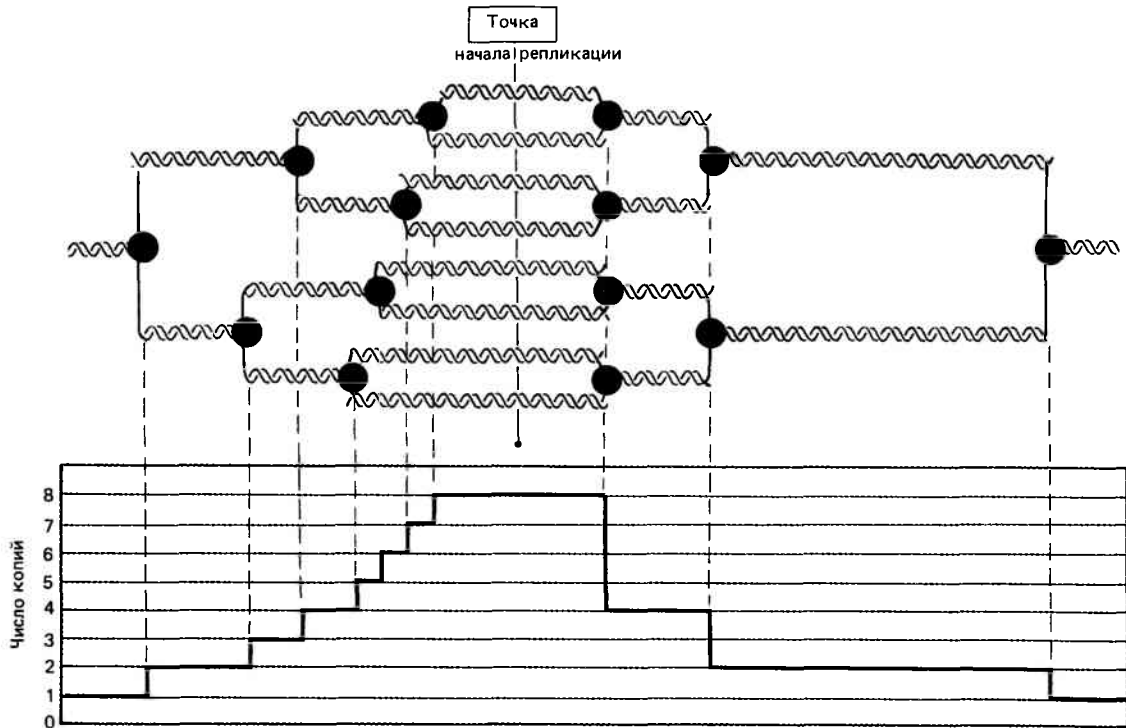


Рис. 38.7. Амплификация локальной области могла происходить путем множественных событий инициации в пределах одного репликона. Каждая реплицированная область соединена с фланкирующими нереплицированными участками неподвижными репликационными вилками. Если

вилки прекращают движение в случайных точках, они образуют постепенный градиент (левая часть рисунка). Если отдельные вилки прекращают движение координированно в фиксированном сайте, они вызывают одновременное уменьшение в уровне амплификации (правая часть рисунка).

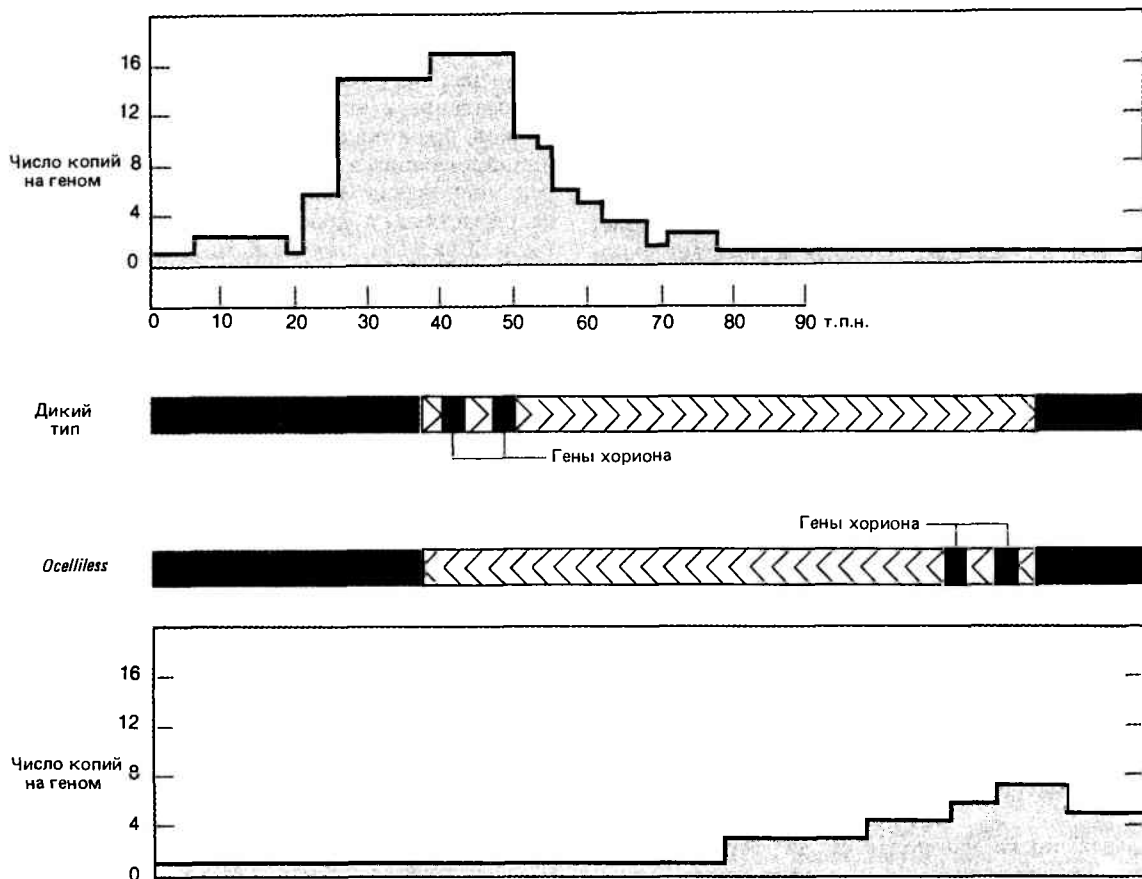


Рис. 38.8. Инверсия, вызывающая нарушение строения глаз у дрозофилы (*ocelliless*) (отсутствие глазков), изменяет область ДНК, которая амплифицируется. Левый конец инверсии находится на расстоянии, не превышающем

3 т.п.н. от левой пограничной области генов хориона; положение правого конца неизвестно, однако на рисунке длина инверсии равна примерно 85 т.п.н.

мального уровня. (Причина уменьшения уровня амплификации неизвестна.) Более того, амплификация теперь распространяется в новые области, что видно по увеличению числа копий последовательностей справа; эти последовательности не амплифицировались в клетках дикого типа. Полученные результаты показывают, что точка начала репликации, способная обеспечить двунаправленную амплификацию, находится по соседству от генов хориона. Протяженность амплифицированной области остается одинаковой независимо от того, находится ли точка начала в обычном месте в штамме дикого типа или перемещена в новое положение в результате инверсии у мух линии *ocelliless*, характеризующейся нарушением строения глаз.

Классический тип эффекта, вызываемого инверсией, — феномен **эффекта положения**. Под этим термином понимают зависимость экспрессии гена от его положения в геноме. Хорошо известно, например, что при транслокации гена из эухроматинового района в гетерохроматиновый активность утрачивается. Хотя ген сам по себе остается неизменным, его активность под влиянием окружения меняется. В случае мутации *ocelliless* инверсия определяет неспособность некоторых генов амплифицироваться, в то время как другие гены начинают амплифицироваться вместо них. Дозовые эффекты вызывают сложные изменения в фенотипе без какого-либо изменения самих генов.

Селекция амплифицированных последовательностей генома

Геном эукариот обладает способностью накапливать дополнительные последовательности экзогенного или эндогенного происхождения. Такие последовательности независимо от того, добавлены ли они извне или имеют хромосомное происхождение, могут давать начало множественным копиям, которые существуют в виде тандемных последовательностей либо вне хромосомы, либо в составе хромосом. Находясь вне хромосом, дополнительный материал наследуется неравномерно (он не сегрегует поровну при делении, как это свойственно истинным плазмидам). Материал, включенный в геном хозяина, является его компонентом.

Способ введения экзогенной донорной ДНК в клетки реципиента получил название **трансфекции** (см. следующий раздел). Селекция клеток на устойчивость к определенным агентам приводит к амплификации эндогенных последовательностей. Механизмы, обеспечивающие способность приспосабливаться при трансфекции и при амплификации, по-видимому, сходны. Нам ничего не известно о ферментативных активностях, участвующих в этом процессе, однако вероятно, что они выполняют родственные функции в обычном процессе воспроизведения ДНК.

Наиболее изучена амплификация генов при добавлении метотрексата (*mtx*) к какой-либо культивируемой клеточной линии. Реагент блокирует метаболизм фолата. Устойчивость к нему может обеспечиваться мутациями, которые изменяют активность фермента дигидрофолатредуктазы (DHFR). Наряду с изменением самого фермента может возрастать его количество в клетке. Причиной такого увеличения служит амплификация структурных генов *dhfr*.

Хотя амплификация относительно редко используется для приобретения резистентности к токсическому реагенту,

она не уникальна для генов *dhfr*. Резистентность к ингибитору фермента аспартат-карбоилтрансферазы также приобретается благодаря амплификации соответствующего структурного генома. Клетки, устойчивые к тяжелому металлу кадмию, содержат увеличенное число копий гена, который кодирует белок металлотионеин, осуществляющий обезвреживание яда. Общей чертой всех трех систем является не единовременное (одноступенчатое) возникновение высокоустойчивых клеток, а постепенное в процессе адаптации к возрастающим дозам токсического реагента. Следовательно, амплификация генов может происходить в несколько стадий.

Число генов *dhfr* в линии клеток, устойчивых к метотрексату, варьирует от 40 до 400 в зависимости от строгости селекции и индивидуальной клеточной линии. Линии клеток *mtx^r* можно отнести к двум классам, отличающимся своим ответом на отсутствие метотрексата в среде, когда потребность в высоких уровнях активности DHFR снимается. В **стабильных** линиях амплифицированные гены сохраняются. В **нестабильных** линиях амплифицированные гены по крайней мере частично утрачиваются при снятии селективного давления. Причиной таких различий, несомненно, служат условия, в которых находятся в них амплифицированные гены *dhfr* (рис. 38.9).

В стабильной линии клеток *mtx^r* амплифицированные гены являются хромосомными и занимают обычный сайт *dhfr* на одной из хромосом. Как правило, другая хромосома сохраняет свою единственную копию *dhfr*. Таким образом, амплификация происходит только в одном из двух аллелей *dhfr* и возросшая устойчивость к метотрексату сопровождается дальнейшим увеличением степени амплификации этого локуса. (Однако иногда хромосома с амплифицированным локусом *dhfr* сама дублируется, по-видимому, в результате нерасхождения при митозе.)

Амплификация гена оказывает видимый эффект на хромосому; локус можно идентифицировать визуально как гомогенно окрашенную область (HSR). Это показано на рис. 38.10. HSR получила свое название из-за присутствия дополнительной области, в которой хромосомные полосы при соответствующих обработках, например при С-окрашивании, выявить не удается (см. рис. 28.9). Это позволяет предполагать, что определенная область между полосами расширилась.

В тех хромосомах нестабильных клеточных линий, которые несут ген *dhfr*, изменений не обнаружено. Однако в этом случае установлено появление большого числа элементов, названных **двойными микрохромосомами**. Такие хромосомы видны на рис. 38.11. В типичной клеточной линии каждая двойная микрохромосома несет 2–4 гена *dhfr*. Двойные микрохромосомы, по-видимому, представляют собой самореплицирующиеся структуры, утратившие центромеры. В результате они не способны прикрепляться к митотическому веретену, сегрегируют нерегулярно и поэтому часто утрачиваются дочерними клетками. Несмотря на данное им название, в действительности двойные микрохромосомы представляют собой внехромосомные элементы.

Нерегулярное наследование двойных микрохромосом объясняет нестабильность признака резистентности к метотрексату в этих линиях. Двойные микрохромосомы непрерывно утрачиваются во время клеточных делений; в присутствии же метотрексата клетки с уменьшенным числом генов *dhfr* погибают. В популяции выживших клеток сохраняются только те клетки, которые сохранили

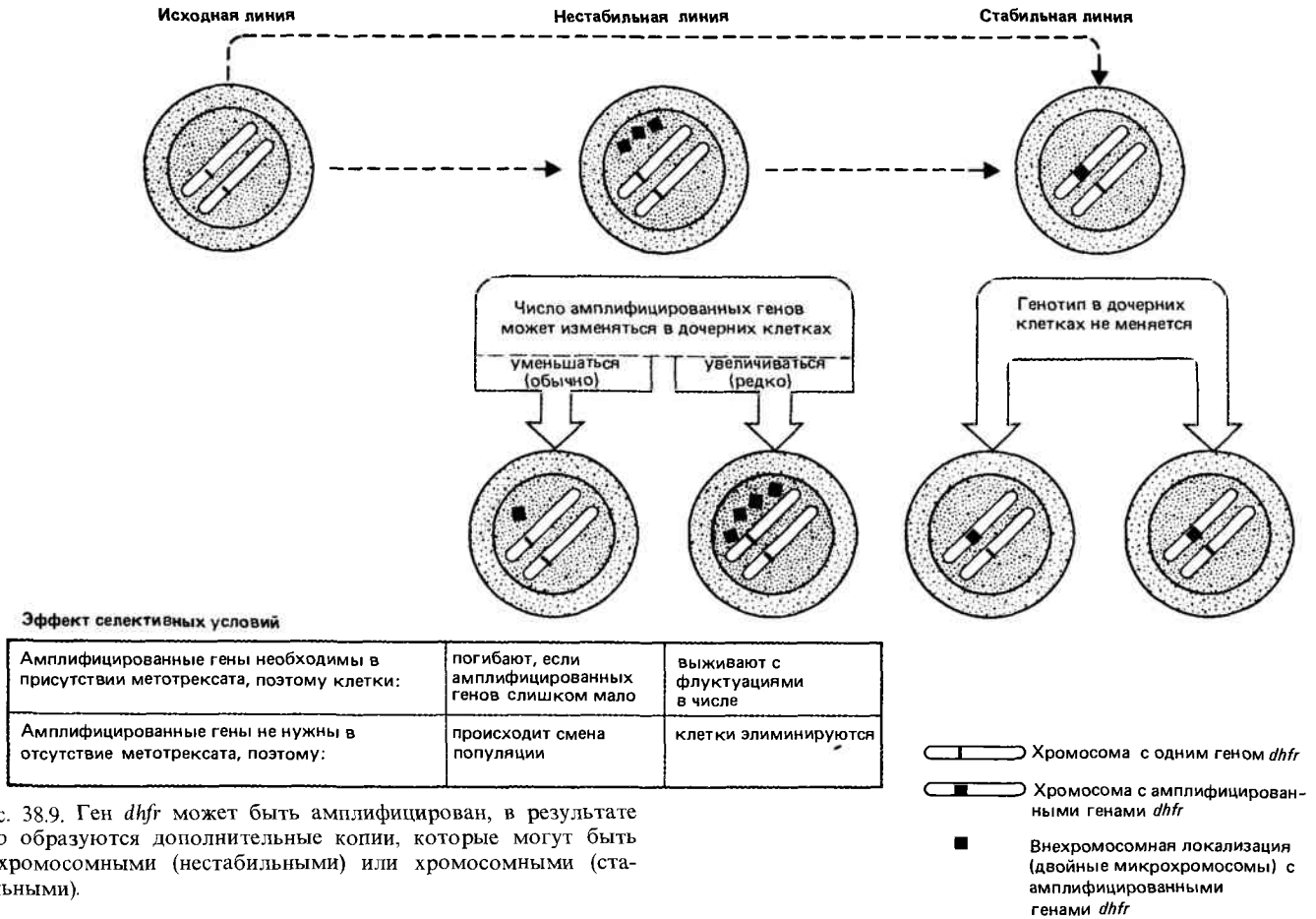


Рис. 38.9. Ген *dhfr* может быть амплифицирован, в результате чего образуются дополнительные копии, которые могут быть внехромосомными (нестабильными) или хромосомными (стабильными).

достаточное количество двойных микрохромосом. Присутствие этих структур уменьшает скорость деления клеток. В результате при снятии селективного давления клетки, утратившие амплифицированные гены, имеют селективное преимущество: они быстрее образуют потомство и вскоре оказываются в популяции в большинстве. Именно этим можно объяснить, почему амплифицированное состояние сохраняется, только пока клетки растут в присутствии метотрексата.

Продолжительная селекция на устойчивость к метотрексату позволила выделить как стабильные, так и нестабильные по этому признаку линии клеток. Какова первоначальная ступень в амплификации гена? На первых ступенях отбора большинство или все резистентные клетки проявляют нестабильность. Образование внехромосомных копий явно представляет собой более частое событие, чем амплификация внутри хромосомы. Мы не знаем, происходит ли внутрехромосомная амплификация менее часто как процесс, происходящий *de novo*, или она связана с внехромосомной амплификацией, являющейся промежуточной ступенью процесса. Из-за неправильной сегрегации двойных микрохромосом увеличение числа копий может происходить относительно быстро в тех случаях, когда при каждом делении отбираются клетки, содержащие больше, чем положено, копий генов *dhfr*. Такие клетки возникают и отбираются в ответ на возросшие уровни метотрексата. Поведение двойных микрохромосом объясняет ступенчатую эволюцию признака *mtx^r* и нерегулярные колебания в уровне генов в нестабильных линиях.

Каким образом возникают внехромосомные копии?

Мы знаем, что их образование не связано с утратой исходной копии хромосомы. Одна из возможностей представлена на рис. 38.12. Показано, что вблизи генов *dhfr* иницируются дополнительные циклы репликации. Благодаря какому-то рекомбинационному событию из хромосомы высвобождаются внехромосомные копии. В зависимости от природы такого события может образовываться внехромосомная молекула ДНК, содержащая одну или несколько копий. Если двойные микрохромосомы содержат кольцевые ДНК, рекомбинация между ними в любом случае будет приводить к образованию мультимерных молекул.

Происхождение амплифицированных хромосомных копий объяснить значительно сложнее. HSR состоит из большого числа тандемно повторенных единиц. Одна из трудностей в объяснении их возникновения путем интеграции последовательностей двойных микрохромосом связана с тем, что число хромосомных повторов во много раз больше, чем число повторов в отдельных двойных микрохромосомах. Большое число интегрированных копий является также препятствием при создании моделей для строго внутрехромосомного события амплификации.

Амплифицирующаяся область значительно протяженнее, чем сам ген *dhfr*. Ген имеет длину, равную примерно 31 т.п.н., в то время как средняя длина повторяющейся единицы составляет $5 \cdot 10^5 - 1 \cdot 10^6$ пар оснований в хромосомной области HSR. Протяженность ДНК двойных микрохромосом, по-видимому, колеблется в пределах от $1 \cdot 10^5$ до $1 \cdot 10^6$ пар оснований. Хромосомная область, дающая начало амплифицированным генам, может со-

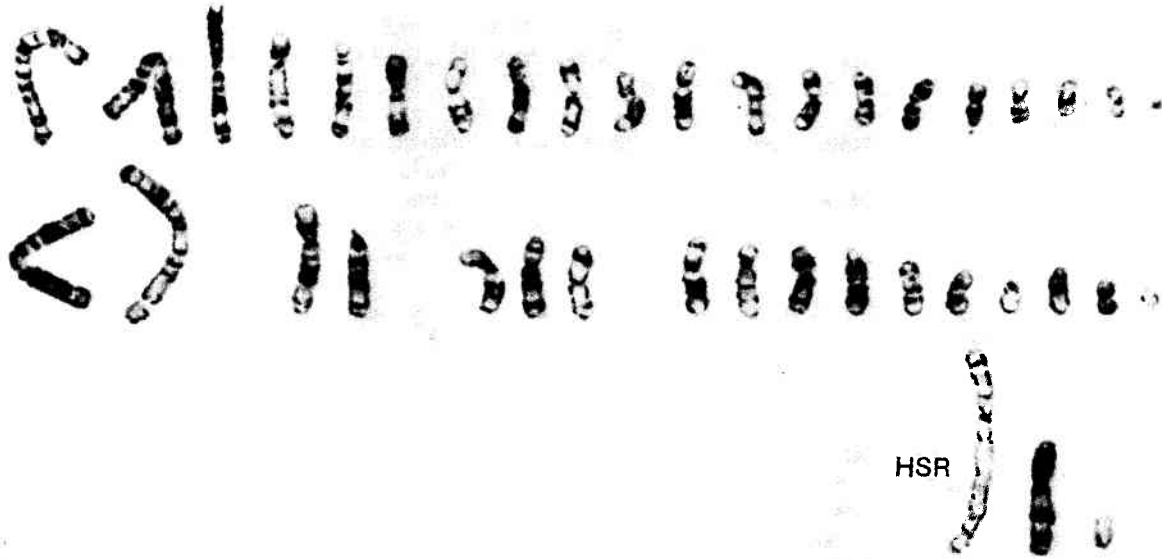


Рис. 38.10. Амплифицированные копии гена *dhfr* образуют гомогенно окрашивающуюся область (HSR) в хромосоме.

ответствовать репликону, в котором она локализована.

Амплификация происходит и в отсутствие метотрексата, однако, если нет какого-либо селективного давления, клеточная популяция утрачивает ее продукты. Почему такое явление характерно лишь для нескольких генов? Возможно, амплификация затрагивает гены, расположенные в соответствующих участках репликона, или происходит в тех случаях, когда в этих участках находятся соответствующие повторяющиеся последовательности, позволяющие этой области выщепляться с помощью рекомбинации.

Некоторая информация о событиях, вовлекаемых в процесс сохранения двойных микрохромосом, получена благодаря изучению нестабильных клеточных линий, в которых амплифицированные гены кодируют мутантный фермент DHFR. Мутантный фермент не присутствует в исходной (диплоидной) линии клеток; следовательно, мутация должна была возникнуть в какой-то момент процесса амплификации. Несмотря на вариации в числе амплифицированных генов, эти клетки обладают только мутантным ферментом. Следовательно, хромо-

сомные гены дикого типа не могут служить постоянным источником больших количеств новых двойных микрохромосом, так как их амплифицированные копии образовывали бы нормальный фермент.

Следовательно, изменения в состоянии клетки осуществляются благодаря однажды возникшим амплифицированным внехромосомным генам, а не благодаря исходным хромосомным копиям этих генов. При удалении метотрексата клеточная линия утрачивает свои двойные микрохромосомы. При повторной обработке реагентом нормальные гены амплифицируются, давая новую попу-



Рис. 38.11. Амплифицированные внехромосомные гены *dhfr* приобретают форму двойных микрохромосом, которые видны в виде мелких белых точек. Фотография любезно предоставлена Р. Шимке.

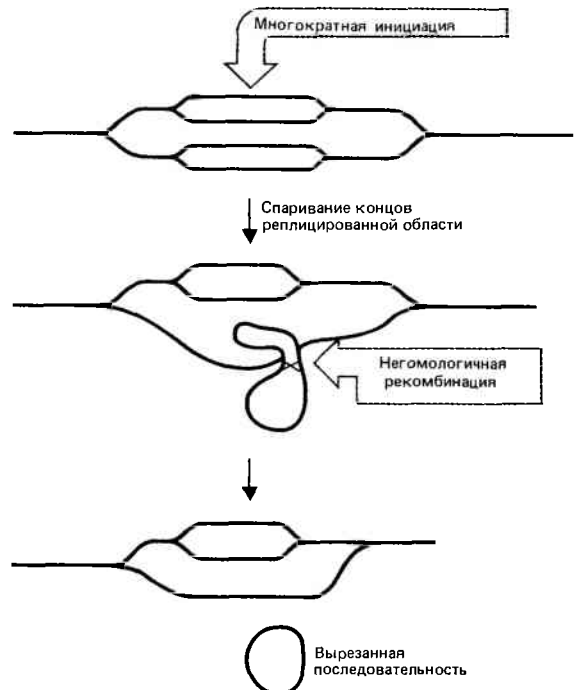


Рис. 38.12. Внехромосомные копии ДНК могли возникать путем реинициации репликации, сопровождаемой негомологичной рекомбинацией между репликациями.

ляцию двойных микрохромосом. Следовательно, ни одна из внехромосомных копий мутантного гена не встраивается в хромосому. Другой важный вывод, следующий из этих результатов, сводится к тому, что двойные микрохромосомы мутантной линии несут только мутантные гены, так что в случае наличия в двойных микрохромосомах более одного гена *dhfr* все они должны быть мутантного типа. Это показывает, что мультикопийные двойные микрохромосомы могут образовываться из отдельных внехромосомных генов.

Введение экзогенных последовательностей посредством трансфекции

Эксперименты с трансфекцией были начаты с добавления препаратов метафазных хромосом к суспензиям клеток. Хромосомы включались клетками довольно неэффективно, и с низкой частотой возникали нестабильные варианты. Интактные хромосомы редко сохранялись во время процедуры; реципиентные клетки обычно приобретали фрагмент донорной хромосомы (который отличался нестабильностью из-за отсутствия центромеры). В редких случаях возникали стабильные линии в результате интеграции материала донора с хромосомой хозяина.

Такие же результаты получены при добавлении к реципиентным клеткам очищенной ДНК. В этом случае оказалось возможным добавлять определенные последовательности, а не полагаться на случайную фрагментацию хромосом. В результате трансфекции с помощью ДНК получают как стабильные, так и нестабильные линии. Эти эксперименты аналогичны бактериальной трансформации, однако их называют трансфекцией, поскольку исторически термин «трансформация» используют при описании изменений, позволяющих эукариотическим клеткам неограниченно размножаться.

Низкая частота трансфекции делает необходимым использование донорных маркеров, присутствие которых в реципиентных клетках способствует отбору. Следует отметить, что трансфицированные последовательности экспрессируются. В большинстве экспериментов по трансфекции использовали маркеры, связанные с легко определяемыми ферментативными функциями, но в принципе может быть использован любой селективный маркер. Это позволяет выделить гены, ответственные за морфологические изменения. Например, могут быть отобраны трансфицированные клетки, которые приобрели трансформированный фенотип (способность неограниченно размножаться с образованием опухоли). Затем можно идентифицировать последовательность ДНК, ответственную за изменение фенотипа. Таким способом были выделены некоторые клеточные *onc*-гены.

Котрансфекция двух и более маркеров дала дополнительные сведения о механизме трансфекции и расширила круг вопросов, которые могут быть решены с помощью этого метода. В случае трансфекции клеток *tk*⁻ препаратами ДНК, содержащими очищенный ген *tk*⁺ и геном Φ X174, все трансформанты содержат обе донорные последовательности. Это наблюдение оказалось полезным, поскольку позволило вводить неселективные маркеры посредством котрансфекции с селективным маркером.

Расположение последовательностей *tk* и Φ X174 различно в каждой трансфицированной линии, но во время размножения одной линии оно сохраняется неизменным.

Часто обнаруживается множество копий донорных последовательностей, причем число их в отдельных линиях варьирует. Ревертанты утрачивают последовательности Φ X174 вместе с последовательностями *tk*. Амплификация трансфицированных последовательностей в условиях селективного давления ведет к равному и одновременному увеличению числа копий всех донорных последовательностей. Следовательно, два типа донорных последовательностей оказываются во время трансфекции связанными физически и имеют общую судьбу.

В опытах по трансфекции массу добавленной к реципиентным клеткам ДНК обычно увеличивают за счет избытка ДНК-носителя, препарата какой-то другой ДНК (например, ДНК спермы лосося). Доказано, что трансфицируемые клетки получают ДНК-носитель в виде последовательностей, фланкирующих селективируемые с каждой стороны. Следовательно, трансфекция осуществляется структурой ДНК, состоящей из ряда сцепленных последовательностей всех типов, присутствующих в препарате донора. Поскольку ревертанты по селективируемому маркеру утрачивают весь этот материал, кажется вероятным, что трансфицируемые клетки приобретают только одну такую структуру ДНК. Данная структурная единица может образовываться благодаря конкатемерному сцеплению донорных последовательностей в ходе реакции, которая проходит очень быстро по сравнению с другими событиями, вовлекаемыми в процесс трансфекции. Такая трансфицируемая единица может иметь протяженность около $1 \cdot 10^6$ п.н. Мы не можем установить с помощью метода гибридизации, сцеплена ли донорная единица с хромосомной ДНК реципиента (интересующие нас концевые фрагменты представлены в слишком незначительных количествах). Вероятно, что первая стадия процесса заключается в образовании нестабильных внехромосомных единиц, которые впоследствии стабилизируются в результате интеграции. В некоторых клеточных линиях методом гибридизации *in situ* было показано, что трансфицированные клетки содержат донорный материал, интегрированный в хромосомы хозяина. Любая определенная клеточная линия имеет только один сайт интеграции; однако сайты в каждой линии различны. Вероятно, выбор сайта для интеграции — случайное событие; иногда оно связано с большими хромосомными перестройками.

Некоторые интересные результаты, полученные при котрансфекции двух селективных маркеров, суммированы на рис. 38.13. Линия реципиентных клеток $APRT^- TK^-$ была трансфицирована с помощью ДНК, несущей гены *aprt* и *tk*. Ген *aprt* был представлен активным аллелем *aprt*⁺, кодирующим фермент аденозинфосфорилтрансферазу. Ген *tk* представлял неактивный аллель, сохраняющий кодирующую область фермента тимидинкиназы, но утративший промотор. Трансфицированные клетки отбирали, используя маркер *aprt*⁺, однако при этом приобретались обе донорные последовательности. Затем клетки $APRT^+ TK^-$ подвергали селекции по способности обеспечивать функцию TK^+ . Осуществление этой функции зависело от двух событий. TK должен был транскрибироваться с какого-то сайта, замещающего его промотор; поскольку это обуславливает только низкий уровень экспрессии, необходимо увеличение числа копий данного гена, т.е. его амплификация.

Отобранные клетки имели фенотип $APRT^+ TK^+$ и содержали около 40 копий трансфицируемых единиц *aprt-tk*, интегрированных в один хромосомный сайт. Наконец,

среди этих клеток отбирали клетки с фенотипом $APRT^-$. Их появление связано с элиминацией экспрессии копий гена $aprt^+$ без повреждения экспрессии гена tk . Интересно, что все копии гена $aprt$ в этих клетках приобретали одну и ту же негативную мутацию. Этот результат свидетельствует о существовании какого-то механизма, который позволяет всей группе генов быть регенерированной из одного гена или подвергаться коррекции в соответствии с последовательностью одного гена; по-видимому, это происходит в течение одной клеточной генерации.

Трансфицированная ДНК способна включаться в геном клеток зародышевой линии

Метод трансфекции успешно применяется для введения генов в геном животных (рис. 38.14). Плазмиды, несущие интересный исследователем ген, вводятся в зародышевый пузырек (ядро) ооцита мыши или в пронуклеус оплодотворенного яйца. Затем это яйцо имплантируется самке. Таким образом, были введены гены глобина и гибридные гены, содержащие промотор для металлотионеина, который сцеплен с кодирующей последовательностью тимидинкиназы. Промотор МТ ведет происхождение от природного гена мыши; его присутствие можно определить по способности отвечать на обычную индукцию тяжелыми металлами или глюкокортикоидами (ответ определяют по активности тимидинкиназы).

В этих опытах ДНК плазмид ввели в 69 яиц. Десять из развившихся мышей несли инъектированную последовательность, что было показано при анализе ДНК из клеток почек. Обычно в один хромосомный сайт множество копий плазмид внедрялось, по-видимому, тандемно. Число копий варьировало от 1 до 150. Они

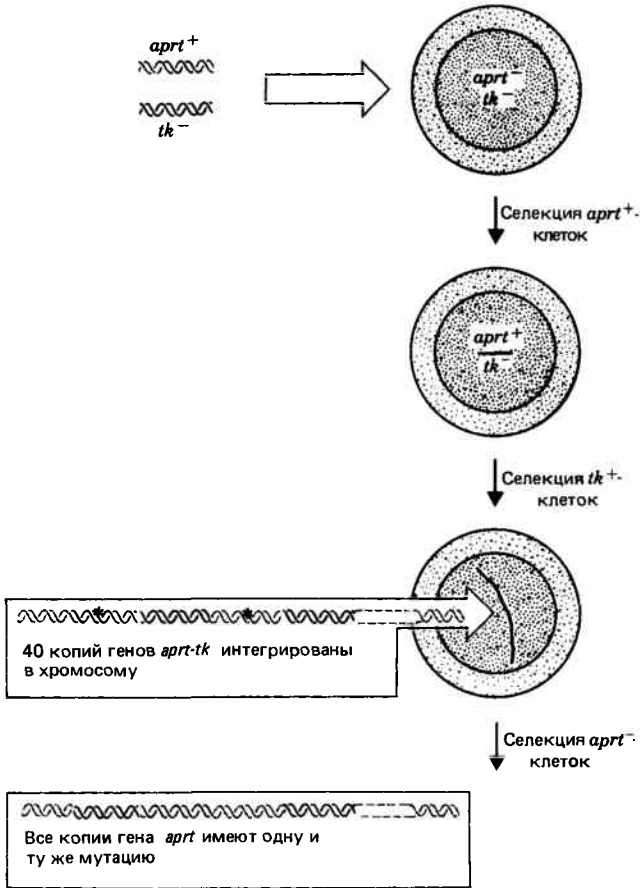


Рис. 38.13. В случае амплификации трансфицированных генов весь кластер координированно реагирует на дальнейшее селективное давление.

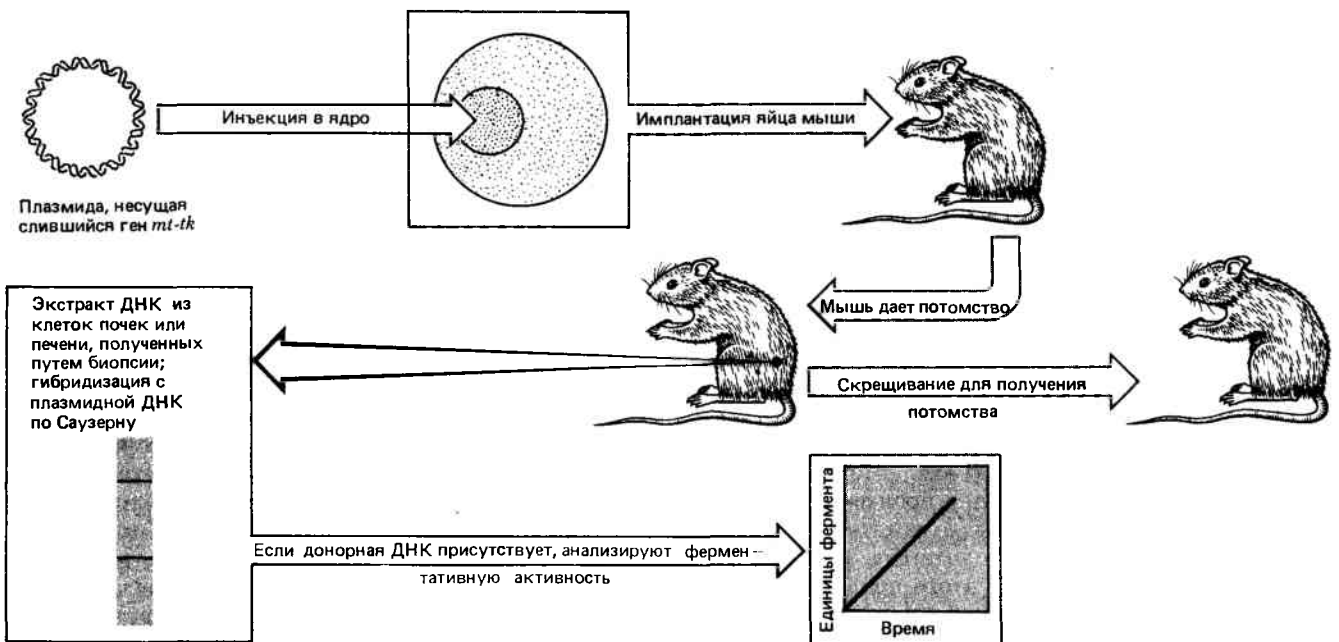


Рис. 38.14. С помощью трансфекции можно вводить ДНК в клетки зародышевых линий животных.

наследовались потомством инъекционной мыши, подобно тому как происходит наследование менделевского локуса. У семи мышей в почках и печени происходила экспрессия гена тимидинкиназы (ткани, в которых обычно ген МТ проявляет активность). Выражения не наблюдали в клетках мозга, ткани, в которой обычно ген не функционирует. Экспрессия индуцировалась при добавлении тяжелых металлов, но не глюкокортикоидов.

В потомстве инъекционной мыши экспрессия донорного гена варьировала. Уровень экспрессии гена не коррелировал с числом tandemно интегрированных копий даже у исходных родителей. Возможно, только некоторые из генов были активными. (Следует указать, что если бы были активными все дополнительные гены, вряд ли сохранился бы нормальный регуляторный ответ, поскольку большое число промоторов способно связать большинство регуляторных молекул.)

Чем вызваны вариации в экспрессии генов? Одно из объяснений, часто используемое для случаев трансфицированных генов и интегрированных геномов ретровирусов, связано с предполагаемой важной ролью сайта интеграции. Возможно, что ген выражается, если он интегрирован только в пределах активного домена. Другая возможность заключается в наличии эпигенетической модификации. Некоторые из трансфицированных генов мышей оказались метилированными. (Так как донорная ДНК не несла метильные группы, это наблюдение позволяет предположить, что существует фермент, способный осуществлять метилирование *de novo*.) Изменения в способе метилирования генов могут обуславливать

изменение их активности. С другой стороны, гены, которые являются активными у родителей, могут быть делетированными или амплифицированными у потомства.

На примере этих экспериментов мы видим возможность расширения круга объектов для исследования регуляции от культивируемых клеток до животных. Возможность введения ДНК в геном позвоночных позволяет нам производить в нем изменения, добавлять новые гены, способные нести определенные модификации, введенные в ДНК *in vitro*. С помощью этого метода может оказаться возможным выявить причины тканеспецифичной экспрессии генов. Вероятно, в конце концов станет реальностью и замена дефектных генов в генотипе.

Рекомендуемая литература

Природа ретровирусов подробно проанализирована в сб. "Mobile Genetic Elements" (ed. Shapiro, Academic Press, New York, 1983). Сведения о геноме *D. melanogaster* приводятся в обзоре Спрадлинга и Рубина (*Spradling, Rubin, Ann. Rev. Genet.*, **15**, 219–264, 1981); эффекты, связанные с мутацией *ocelliless*, описываются в работе Спрадлинга и др. (*Spradling et al., Cell*, **27**, 193–202 и 203–210, 1981). События, вовлекаемые в амплификацию гена *dhfr*, рассматривает Шимке (*Schimke et al., Cold Spring Harbor Symp. Quant. Biol.*, **45**, 785–797, 1981). Возможности трансфекции анализирует Пеллицер (*Pellicer et al., Science*, **209**, 1414–1422, 1980). О методике инъекции генов мышей в клетки сообщает Полмитер (*Palmüter et al., Cell*, **29**, 701–710, 1982).

Глава 39

КАК ФОРМИРУЕТСЯ МНОГООБРАЗИЕ АНТИТЕЛ

Иммунный ответ млекопитающих проявляется в безграничном разнообразии **антител**, каждое из которых синтезируется в ответ на появление нового **антигена**. Обычно антиген представляет собою белок (или связанную с белком субстанцию) чужеродного происхождения, который попадает в кровяное русло животного или человека. Например, таким антигеном может быть белковая оболочка инфицирующего вируса. Появление антигена индуцирует образование антитела, *специфически распознающего только данный антиген*. Антитела синтезируются в В-лимфоцитах, и они (совместно с другими белыми кровяными клетками, Т-лимфоцитами и макрофагами) участвуют в удалении из организма антигена. Если в организме однажды уже синтезировались антитела на определенный антиген, то способность к производству данного антитела сохраняется, чтобы отражать дальнейшие повторные атаки уже знакомого антигена. Одной из наиболее загадочных особенностей этого процесса является способность образовывать нужные антитела всякий раз при попадании в организм нового антигена. Каким образом организм может быть подготовлен к производству антител, каждое из которых предназначено специфически распознавать антиген с непредсказуемой структурой?

Принято считать, что организм способен образовывать более 10^6 различных антител. Каждое антитело представляет собой тетрамерную молекулу **иммуноглобулина**, состоящую из двух идентичных **легких** (L) и двух идентичных **тяжелых** (H) цепей. Согласно комбинаторной теории, любой тип легкой цепи может соединиться с любым типом тяжелой цепи. Следовательно, для образования 1 000 000 возможных антител требуется больше 1000 типов легких и больше 1000 типов тяжелых цепей.

Структура иммуноглобулинового тетрамера показана на **рис. 39.1**. Каждая белковая цепь состоит из двух принципиально различающихся областей: N-концевой **вариабельной** (V) области и С-концевой **константной** (C) области. Они были определены при сравнении аминокислотных последовательностей разных иммуноглобулиновых пептидов. Как следует из названия, вариабельные области значительно отличаются по последовательностям у разных белков, в то время как константные области, наоборот, обнаруживают существенную гомологию. Сочетание соответствующих областей легких и тяжелых цепей образует различные домены в иммуноглобулиновой молекуле.

Вариабельный домен формируется при ассоциации вариабельных областей легкой и тяжелой цепей. V-домен отвечает за распознавание антигена. Таким образом, эти белки обнаруживают максимальную пластичность в той области, которая взаимодействует с разными мишенями в каждой индивидуальной реакции антиген—антитело. Способность отвечать на различные антигены создается за счет V-доменов соответствующей специфичности.

В пределах самой вариабельной области некоторые участки более вариабельны, чем другие. Наиболее вариабельные участки обозначаются как **гипервариабельные (HV)**. Менее вариабельные получили название **каркасных (FR)** областей. Вариабельную область можно схематически изобразить как серию чередующихся субобластей: FR1—HV1—FR2—HV2—FR3—HV3—FR4. Гипервариабельные участки короче, чем каркасные. Общее число вариабельных областей насчитывает несколько сотен как для легких, так и для тяжелых белковых цепей.

Число различных константных областей невелико. В зависимости от него легкие цепи делятся на два типа: каппа (κ) и лямбда (λ). Каппа-тип характеризуется единственной константной областью; лямбда-тип обладает несколькими родственными типами константных областей.

Тяжелые цепи имеют большее число константных областей. Различные классы антител различаются константными районами тяжелых цепей, которых существует около восьми. Константные области тяжелых цепей образуют несколько индивидуальных доменов в молекуле иммуноглобулина. Первый домен формируется из константной области легкой цепи и CH1-участка тяжелой цепи. Структура остальной части тяжелой цепи у разных классов несколько различна. На рис. 39.1 изображен короткий шарнирный пептид, связывающий первую половину молекулы с двумя константными доменами, каждый из которых образован соответствующими областями (CH2 и CH3) тяжелой цепи. Константные домены выполняют эффекторные функции, необходимые для осуществления иммунного ответа. Эти области молекулы консервативны. Фактически разные классы иммуноглобулинов имеют родственные, но не обязательно идентичные эффекторные функции. В конечном счете характер эффекторной функции определяется типом константной области тяжелой цепи.

Сравнивая характерные особенности вариабельной и константной областей, мы сталкиваемся с центральной дилеммой структуры иммуноглобулиновых генов. Каким образом геном кодирует белки, у которых вариабельные области имеют более 1000 возможных последовательностей, тогда как константные области той же самой молекулы состоят менее чем из 10 последовательностей? Оказывается, что число кодирующих последовательностей для каждой области отражает степень ее вариабельности. Существует множество генов, кодирующих V-районы, и только несколько генов кодируют C-районы. В данном случае под словом «ген» понимают последовательность ДНК, кодирующую отдельный участок целого иммуноглобулинового пептида (тяжелой или легкой цепи). Таким образом, V-гены кодируют вариабельные, а C-гены—константные области молекулы, хотя ни один из этих генов не экспрессируется как независимая самостоятельная единица. Для образования генной единицы, которая могла бы функционировать, продуцируя полноценные легкие и тяжелые цепи, V-ген должен физически

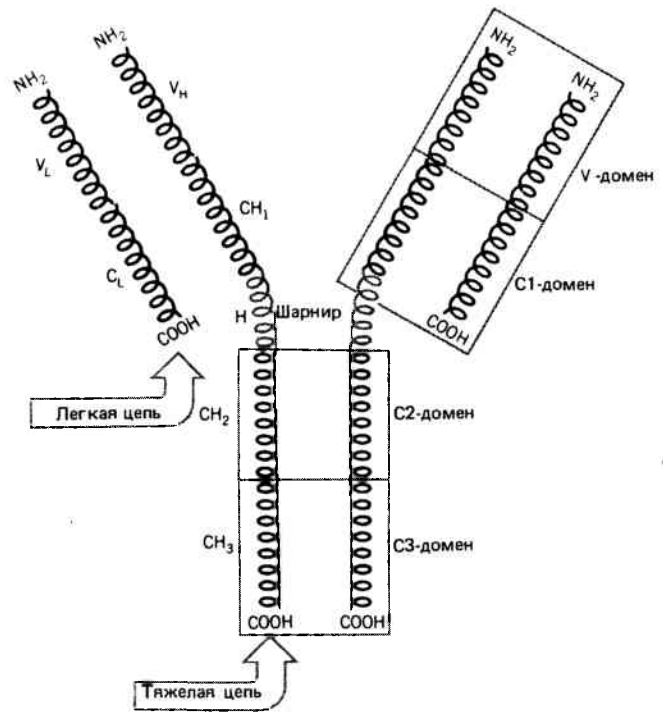


Рис. 39.1. Объединение тяжелых и легких цепей ведет к образованию молекулы иммуноглобулина, имеющей несколько обособленных доменов.

соединиться с C-геном. В этой системе два «гена» кодируют один полипептид.

Для каждого из трех семейств иммуноглобулиновых цепей—каппа, лямбда и тяжелых—только один из многих V-генов может присоединиться к одному из нескольких C-генов. Это возможно благодаря **соматической рекомбинации** соответствующих последовательностей ДНК, происходящей в *B-лимфоцитах, экспрессирующих антитела*. В большинстве случаев число V-генов очень велико и, видимо, составляет несколько сотен последовательностей, организованных в огромный генный кластер. (Исключением является цепь лямбда-типа мыши, состоящая всего из нескольких генов.) Исходное разнообразие V-генов частично определяет разнообразие иммуноглобулинов. Однако оно объясняется не только набором последовательностей, запрограммированных в геноме; определенный вклад вносят изменения, происходящие при создании функционального гена.

Кластеры V-генов, кодирующих иммуноглобулины,—не единственные примеры громадных (с молекулярной точки зрения) генных областей, предназначенных для синтеза однотипных белков. Главный комплекс гистосовместимости также связан с образованием белков, участвующих в иммунном ответе. Он охватывает очень протяженную область ДНК, в пределах которой расположено много генов, имеющих родственную структуру и функции. (Однако для экспрессии этих генов не требуется перестройка последовательностей ДНК.)

О важности иммунного ответа у млекопитающих свидетельствует огромное число генов, кодирующих иммуноглобулины и относящихся к локусу гистосовместимости, что заставляет нас более подробно остановиться на характеристике этих обширных генных кластеров.

Создание функционального иммуноглобулинового гена на первый взгляд напоминает ламарковский процесс

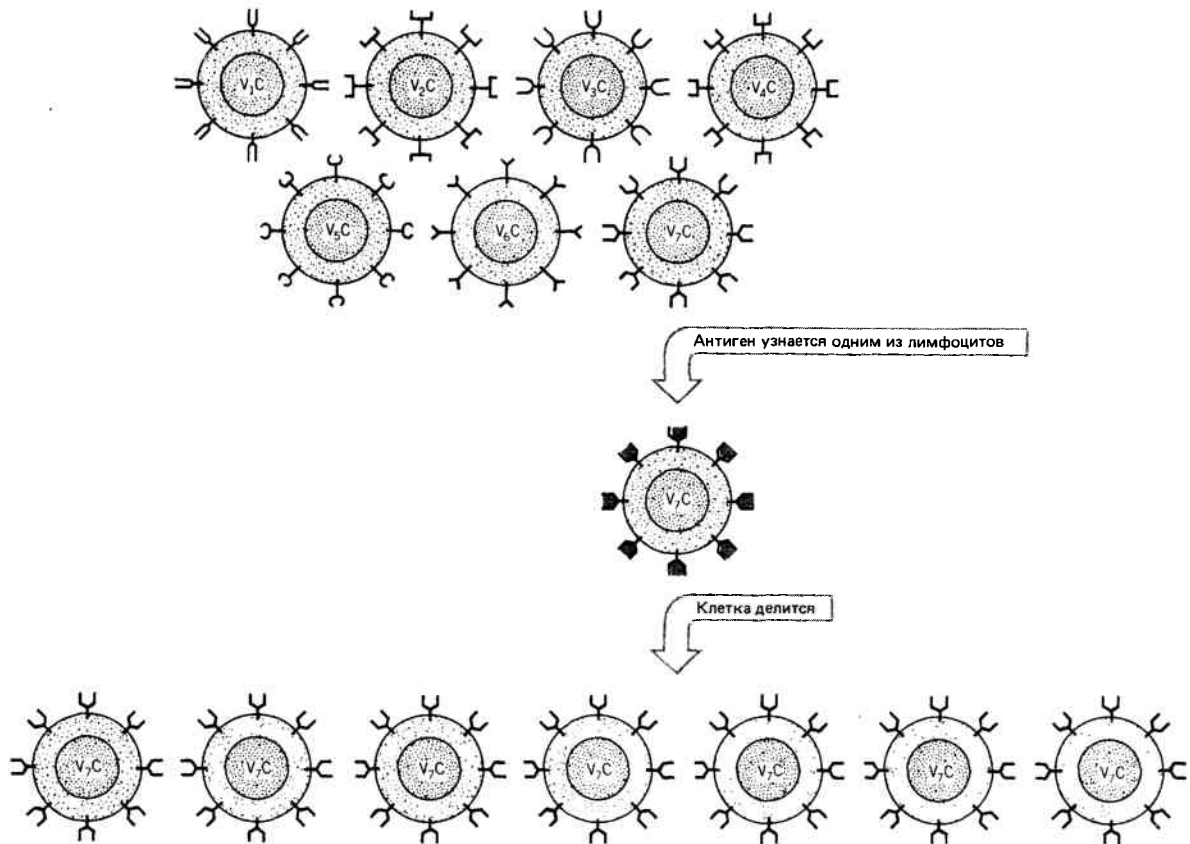


Рис. 39.2. Реакция с антигеном приводит к появлению клона лимфоцитов, образующих соответствующее антитело.

изменения в геноме в ответ на действие внешнего фактора (антигена). Рождаясь, организм не имеет функционирующего гена, производящего нужное антитело. Последующее образование этого гена дает возможность антителу синтезироваться таким образом, что оно способно взаимодействовать только с данным антигеном.

Путь от антигена к синтезу антитела — индуцируемый процесс. Однако этот феномен может быть объяснен и без ссылки на ламаркизм при помощи **клональной селекционной теории**, основные положения которой суммированы на рис. 39.2. Рекомбинация V- и C-генов, приводящая к образованию функционального гена, происходит в популяции незрелых В-лимфоцитов. Каждая клетка образует антитело только одной специфичности, что предполагает только одну перестройку в генах легких цепей и одну перестройку в генах тяжелых цепей. Разные клетки продуцируют разные антитела, поскольку при каждом новом акте реконструкции соединяются разные V- и C-гены. При появлении антигена клетка, вырабатывающая антитело, специфичное к данному антигену, стимулируется к делению, возможно, благодаря какому-то сигналу, поступающему с клеточной поверхности, где происходит взаимодействие антитела с антигеном. В результате последующих клеточных делений появляется большое число новых лимфоцитов, секретирующих данное антитело в таких огромных количествах, что оно может стать доминирующим в популяции антител. Следует заметить, что весь этот процесс происходит исключительно в соматических клетках и не затрагивает клетки зародышевой линии; таким образом, ответ организма на антиген по наследству не передается.

Имуноглобулиновые гены образуются путем соединения ранее независимых частей

Материал для исследования иммуноглобулиновых генов получают из миелом или плазматом — опухолевых клеточных линий, которые произошли из лимфоидных клеток, экспрессирующих иммуноглобулины. Каждая опухолевая клеточная линия обычно синтезирует антитело только одного вида. Этот белок является основным продуктом клетки, и кодирующая его мРНК может составлять до 10% от общего количества poly(A)⁺-мРНК клетки. мРНК выделяют обычным способом и клонированные образцы V- и C-районов используют в качестве зондов для идентификации соответствующих геномных последовательностей. Организация этих последовательностей в клетках, производящих иммуноглобулины (в миеломах, плазматомах или отдельных лимфоцитах), отличается от их организации в других соматических клетках или клетках зародышевой линии. Большая часть информации о генах иммуноглобулина получена при их изучении у мыши и человека.

В клетке, синтезирующей антитело, каждая цепь иммуноглобулина кодируется одним геном, имеющим прерывистое строение. Экзоны этого гена в точности соответствуют функциональным доменам белка. Первый экзон варибельной области кодирует сигнальную последовательность (необходимую для прикрепления к мембране), второй — основную часть V-области (размером менее 100 кодонов). Структура константной области зависит

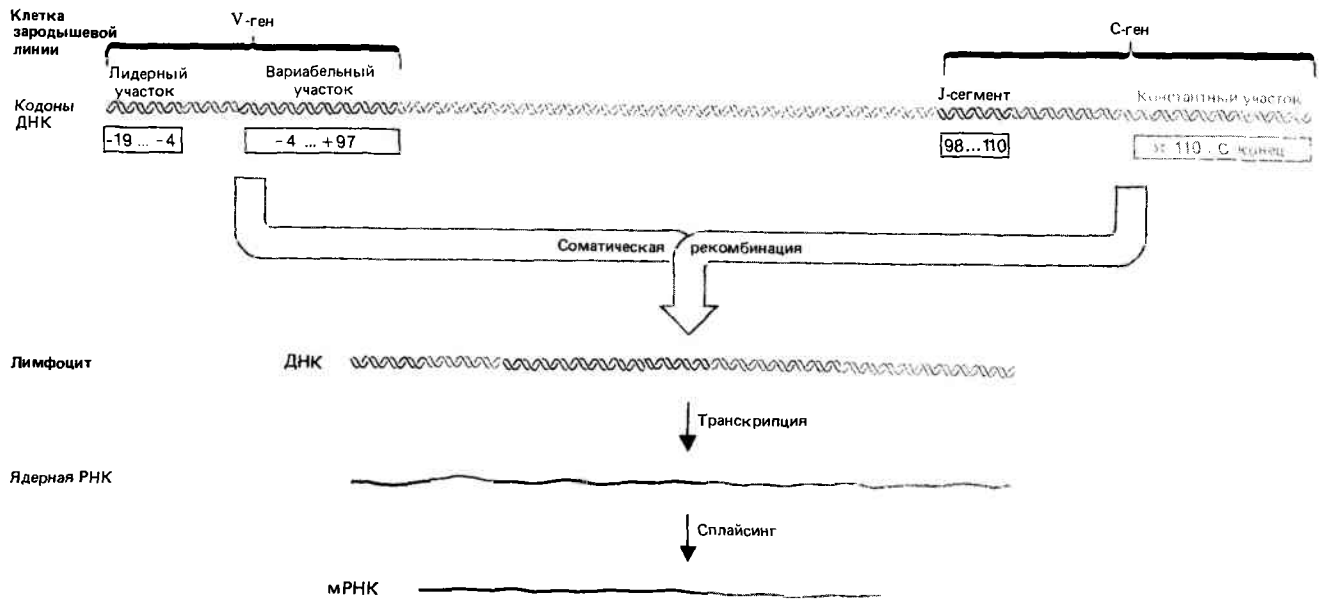


Рис. 39.3. В клетках зародышевой линии V- и C-гены семейства лямбда расположены на одной хромосоме; в результате их ре-

комбинации в лимфоците образуется функциональный ген легкой цепи лямбда-типа.

от типа цепей. В случае легких цепей она кодируется одним экзоном (третьим по счету в интактном гене). Константная область тяжелых цепей кодируется несколькими экзонами, соответствующими белковым доменам, изображенным на рис. 39.1. Для каждой из белковых областей, СН1, шарнирной, СН2 и СН3, имеется свой экзон. Длина каждого из этих экзонов составляет примерно 100 кодонов, шарнирный участок более короткий. Интроны обычно относительно невелики по размеру (~300 н.п.). Экспрессия иммуноглобулиновых генов происходит обычным образом. Ядерный транскрипт подвергается сплайсингу, приводящему к образованию непрерывной молекулы мРНК (содержащей кроме кодирующих последовательностей нетранслируемые лидерные и терминальные районы, входящие в состав концевых экзонов).

В клетках, не экспрессирующих иммуноглобулины, генные последовательности, детектируемые при помощи V- и C-зондов, фланкированы участками ДНК, которые отличаются от тех, что окружают функционирующий ген. Эти последовательности пространственно разобщают V- и C-гены. И хотя для каждого из семейств — лямбда-, каппа- и H-цепей — V- и C-гены расположены на одной хромосоме, расстояния, разделяющие их, неизвестны. Гены всех трех семейств иммуноглобулиновых цепей находятся на разных хромосомах. Такой тип организации, при котором V- и C-гены пространственно разобщены, характеризует клетки зародышевой линии (она также обнаружена во всех соматических клетках организма, не относящихся к иммунной системе).

Лямбда-цепь собирается из двух частей, как показано на рис. 39.3. V-ген состоит из лидерного экзона (L), отделенного одним интроном от переменного (V) сегмента. C-ген состоит из J-участка, отделенного одним интроном от константного (C) экзона. Название J-сегмента происходит от слова *joining* (соединение) и обозначает ту область, к которой присоединяется V-сегмент. J-сегмент представляет собой короткую последовательность, кодирующую несколько (13) концевых аминокислотных остатков переменной области. В функционирующем гене V—J сегмент представляет собой

единный экзон, кодирующий целиком всю переменную область. Интересно, что C-ген как бы отдает свой J-сегмент для создания законченной последовательности V-области. V_λ-ген может соединиться с любым из двух C_λ-генов, каждый из которых имеет свой J-сегмент, образуя двухкомпонентную J—C-структуру. Имея дело с такой многокомпонентной системой, чтобы избежать путаницы, мы будем называть V, J или C «сегментами», когда речь идет о последовательностях ДНК, и «областями», когда имеются в виду полипептиды.

Каппа-цепь также собирается из двух частей, однако в данном случае имеются отличия в структуре C-района. Группа из пяти J-сегментов занимает область ДНК в 500–700 п.н. и отделена от уникального C_κ-экзона интроном в 2–3 т.п.н. Показано, что у мыши центральный J-сегмент является нефункциональным (ψJ3). V_κ-сегмент может соединиться с любым из сегментов J. Этот процесс схематически изображен на рис. 39.4.

Любой из используемых J-сегментов становится концевым участком интактного переменного экзона. Все J-сегменты, расположенные слева от него, элиминируются (на рис. 39.4 это J1 и J2). Сегменты, расположенные справа, рассматриваются как часть интрона между переменным и константным экзонами (на рис. 39.4 J5 относится к той области, которая удаляется во время сплайсинга). Таким образом, все функциональные J-сегменты имеют сигнал для рекомбинации с V-сегментом на левой границе и сигнал для сплайсинга с C-экзоном на правой. Каждый из J-сегментов, который выбран для построения V-гена, использует свой сигнал сплайсинга.

В построении гена тяжелой цепи участвует дополнительный сегмент (рис. 39.5). Открытию этого D (от англ. *diversity*)-сегмента способствовало обнаружение в белках нескольких аминокислотных остатков (от 2 до 13), локализованных между теми участками полипептида, которые кодируются V- и J-сегментами. Между V_H- и J_H-сегментами на хромосоме расположено множество D-сегментов. В результате рекомбинационных актов сегмент V_H соединяется с одним из сегментов D, а он в свою очередь — с одним из четырех J_H-сегментов (ко-

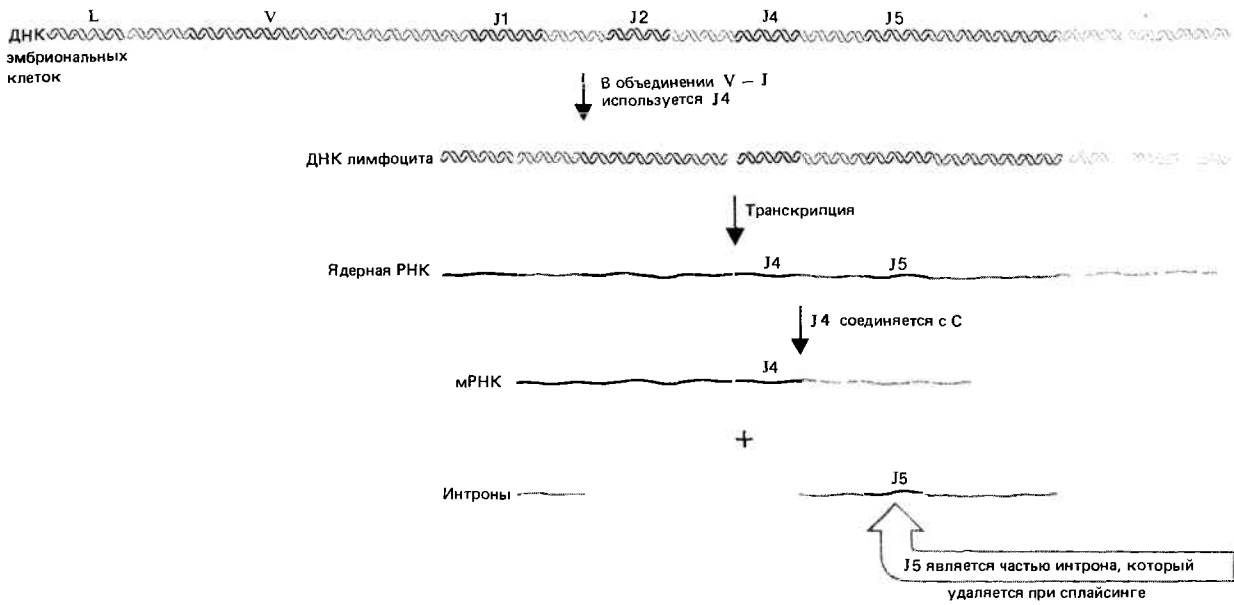


Рис. 39.4. В клетках зародышевой линии С-ген каппа-типа располагается за кластером из нескольких фрагментов; в лимфоци-

те для образования комплекса V—J выбирается любой из С-участков в результате процессинга РНК.

которые отличаются по длине на 4–6 кодонов). Эта генная перестройка ведет к экспрессии соседнего C_H -гена (состоящего из нескольких экзонов).

Организация D-сегментов еще полностью не раскрыта, но многие из них, по-видимому, расположены последовательно друг за другом, образуя кластер. У мыши локус генов тяжелых цепей содержит около десяти D-сегментов различной длины. Вероятно, существует неизвестный еще механизм, обеспечивающий участие одного и того же D-сегмента в реакциях соединения между V—D и D—J.

мы должны проанализировать различные типы V- и С-генов. Иммуноглобулиновые гены наиболее полно охарактеризованы у мыши; почти столько же известно о соответствующих локусах человека. Структура двух семейств легких цепей иммуноглобулинов мыши показана на рис. 39.6.

У мыши почти все легкие цепи представлены каппа-типом и только около 5% относится к типу лямбда. У человека около 60% легких цепей — каппа и около 40% — лямбда. Каждый лимфоцит образует легкие цепи только одного типа (либо каппа, либо лямбда), которые соединяются с тяжелыми цепями.

Природное разнообразие иммуноглобулиновых генов в клетках зародышевой линии

Для того чтобы оценить, какой вклад в вариабельность антител вносит исходное разнообразие кодирующих областей, имеющееся в клетках зародышевой линии,

Каппа-локусы мыши и человека имеют сходную структуру. Как мы уже описывали, несколько J-фрагментов (по 13 кодонов каждый) расположены перед уникальным C_H -геном. Множество различных вариабельных последовательностей может быть присоединено к этому J—С-локусу. Каков источник разнообразия V_H -генов? По этому поводу было выдвинуто несколько гипотез. Пред-

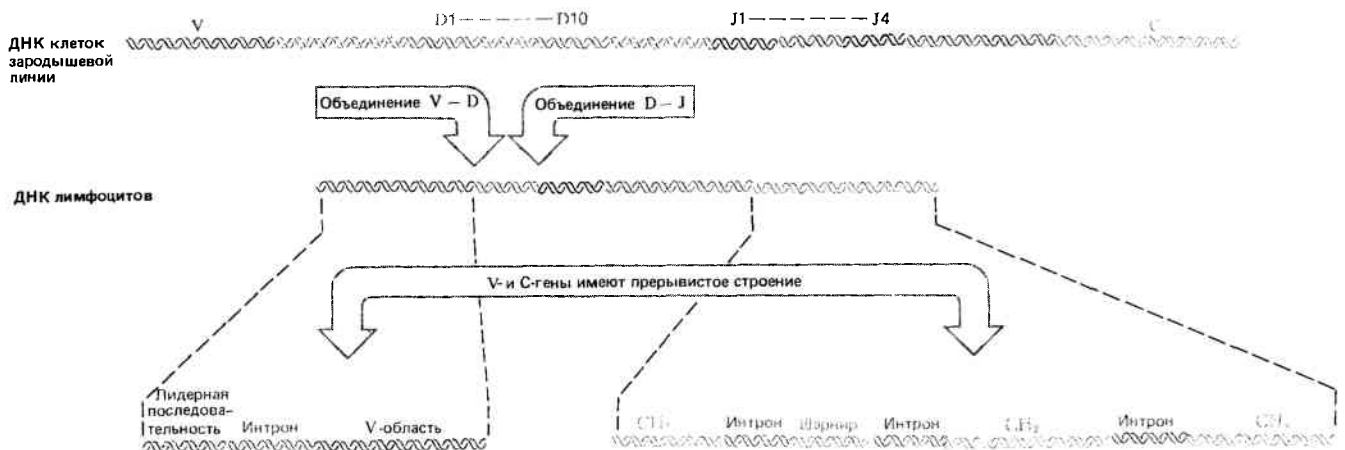


Рис. 39.5. Образование гена тяжелой цепи происходит путем соединения V-гена с D-сегментом и присоединения D-сегментов к одному из сегментов J, расположенных перед С-геном.

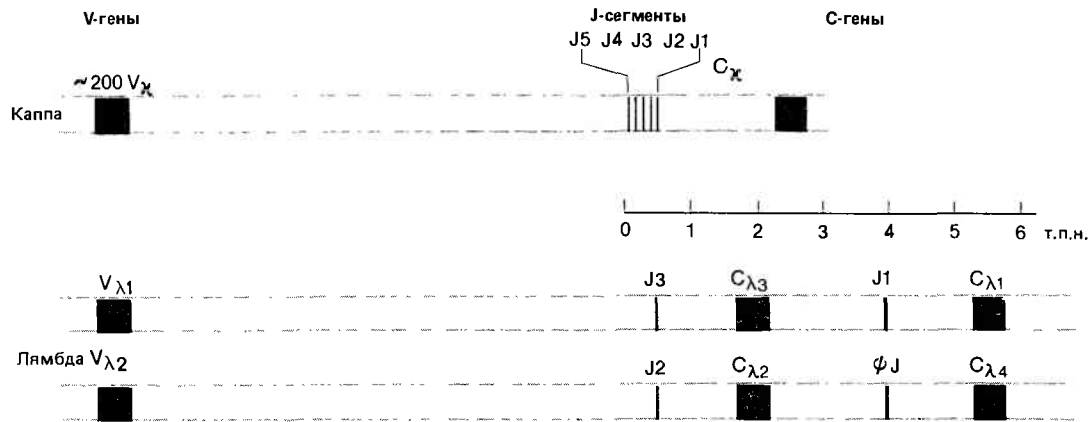


Рис. 39.6. Различное строение легких цепей каппа- и лямбда-типов у мыши.

полагалось, что все V_{κ} -последовательности могут содержаться в зародышевой линии, а новые V_{κ} -гены образуются путем соматических рекомбинаций между исходными эмбриональными. Оказалось, что эти предположения неправильны. По-видимому, разнообразие полноценных кодирующих районов создается в процессе соматической рекомбинации между случайно выбранными парами индивидуальных V- и J-сегментов, а соматические мутации могут приводить к появлению новых V_{κ} -сегментов за счет введения каких-то изменений в уже существующие.

При секвенировании белковых цепей выявляются группы варибельных близкородственных областей, происходящих, вероятно, одна из другой в результате небольших замен в аминокислотной последовательности. Каждая такая группа называется **изотипом**. Все известные V_{κ} -белковые последовательности классифицированы примерно на 26 изотипов, хотя мы предполагаем, что в конечном счете будет доказано существование 50 изотипов. Различия между изотипами слишком велики, чтобы объясняться только соматическими изменениями, скорее всего каждый изотип кодируется по крайней мере одним геном в зародышевой линии.

Генные пробы некоторых V_{κ} -изотипов использовались для определения количества гибридизующихся фрагментов ДНК зародышевой линии мыши. Количество генов для разных изотипов варьирует от одного до десяти. Таким образом, некоторые изотипы могут кодироваться единственным геном, другие же — несколькими генами. В этом смысле выделение изотипа является несколько условной классификацией, основанной только на сходстве аминокислотной последовательности. Взаимосвязь изотипа с генетической структурой зависит от генной варибельности. Если какой-либо V_{κ} -ген не имеет родственных генных последовательностей, то он представляет один изотип. Если существует несколько близкородственных V_{κ} -генов, то их белки также могут образовывать единый изотип.

Прямое определение числа V_{κ} -генов в клетках зародышевой линии затруднено из-за различной степени их дивергенции. Дело в том, что индивидуальной пробой можно тестировать несколько генов; если же один из этих генов используется в качестве новой пробы, то он может взаимодействовать с рядом других генов и т. д. В данном случае налицо классический пример генного семейства,

некоторые гены которого обладают высокой степенью родства, тогда как другие — только незначительной гомологией. Если мы предположим, что существует 50 изотипов, каждый из которых кодируется в среднем 6 генами, то в эмбриональных клетках будет исходно содержаться около 300 V_{κ} -генов.

За счет комбинации любого из 300 V_{κ} -генов с любым из четырех J-сегментов геном мыши может потенциально образовать примерно 1200 каппа-цепей. Однако в некоторых случаях количество белков в каком-либо изотипе, по-видимому, превышает число соответствующих V_{κ} -генов. Новые гены образуются за счет соматических изменений в индивидуальных генах зародышевой линии (но не за счет рекомбинации между уже существующими V_{κ} -генами). Механизм этого процесса будет обсуждаться в последующих разделах.

К сожалению, далеко не все известно об организации этих генов в клетках зародышевой линии. Считается, что V_{κ} -гены занимают огромный участок хромосомы, находящийся на неизвестном расстоянии от 5'-конца гена константной области. При этом не удивительно, если несколько близкородственных генов организованы в субкластеры, образовавшиеся в результате дупликации и дивергенции генов-предшественников.

Следует отметить, что в организации лямбда-генов у мыши и человека наблюдаются значительные различия. Число V_{λ} -генов у мыши необычайно мало по сравнению с количеством соответствующих областей у человека (число которых сравнимо с каппа). Дело в том, что в геноме мыши содержится только два V_{λ} -гена. Предполагается, что в прошлом у этого организма в области V_{λ} -генов произошла обширная делеция, и, возможно, поэтому большая часть легких цепей у мыши относится к каппа-типу.

Лямбда-гены мыши образуют две группы, вероятно возникшие в результате дупликации первоначальной генной единицы. Каждый из V_{λ} -генов связан с двумя константными областями, причем каждая из этих областей образует J—C-структуру. Ген V_{λ} может соединиться либо с J—C λ 3-, либо с J—C λ 1-областями. Следовательно, этот V-ген выбирает один из двух J-сегментов. (В этом его принципиальное отличие от каппа-гена, где несколько разных J-сегментов связаны с одним C κ -геном.) Ген V_{λ 2 обычно соединяется только с J—C λ 2. Область C λ 4 не ис-

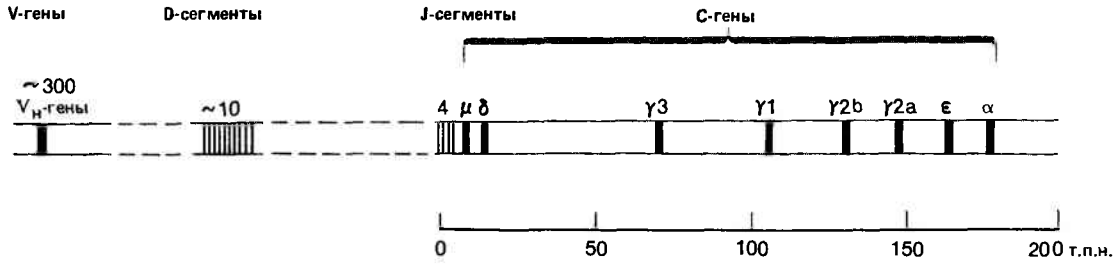


Рис. 39.7. Вся информация для сборки гена тяжелой цепи иммуноглобулина мыши заложена в ее единственном генном кластере.

пользуется, поскольку ее J-сегмент имеет несовершенную структуру. Иногда V_{λ2} объединяется с J—C_{λ1} или J—C_{λ3}. Относительное расположение обеих групп лямбда-генов неизвестно; возможно, они локализуются рядом. У человека имеется не только большее число V_λ-генов; установлено существование по крайней мере шести C_λ генов, которые расположены тандемно, занимая область ДНК в 25 т. п. н. (Дополнительное разнообразие возникает благодаря соматическим мутациям; например, обнаружено несколько вариантов V_{λ1}.)

Локус генов тяжелой цепи состоит из нескольких отдельных участков. Их структура показана на рис. 39.7. В геноме мыши содержится около 300 V_H-генов. На некотором расстоянии от них расположен кластер D-сегментов. Возможно, не очень далеко от него находится кластер сегментов J. Далее в пределах 170 т. п. н. ДНК расположены все C_H-гены. За счет комбинаций из 300 V_H-генов, десяти D-сегментов и четырех J-фрагментов геном мыши может потенциально обеспечить 12 000 переменных участков, каждый из которых присоединяется к любому из C_H-генов.

Хотя все восемь C_H-генов мыши функциональны, обычно только один из них экспрессируется в данной клетке. Этот кластер эволюционировал за счет обычных процессов дупликации, дивергенции и перераспределения экзонов при помощи неравного кроссинговера или геной конверсии. По-видимому, наиболее близки между собой четыре гамма-гена, что следует из анализа их белковой структуры.

У человека имеется большее число C_H-генов, среди которых гены мю и дельта представлены в единственном числе; имеются четыре копии гамма-генов, обозначенных 1–4, и один гамма-псевдоген; кроме того, в геноме человека присутствует три эпсилон-гена, из которых по крайней мере один — псевдоген, и два альфа-гена.

Реакция объединения генных сегментов — дополнительный источник разнообразия антител

Известно, что разнообразие антител может быть обусловлено комбинацией соответствующих компонентов легкой или тяжелой цепей, но и сама по себе реакция воссоединения этих участков служит источником соматических изменений. В случае легких цепей пограничный район между V- и J-сегментами относится к третьей гипервариабельной области. Первый кодон у J-сегментов различен, хотя в целом все они в значительной мере кон-

сервативны; гомологичны 10 позиций из 13 (допуская незначительные замены третьего нуклеотида в кодоне). Конец J-сегмента соответствует четвертой каркасной области вариабельного участка.

Точка соединения V—J сегментов каппа-цепей попадает на последовательность, кодирующую аминокислотные остатки 95 и 96. Если точка стыковки сегментов не фиксирована, а может сдвигаться на несколько пар нуклеотидов, то в районе каждой потенциальной V—J-рекомбинации могут образовываться разные аминокислоты. Эта ситуация показана на рис. 39.8. Использование пяти возможных рекомбинационных рамок приводит к появлению трех различных аминокислот в позиции 96, одна из которых (аргинин) *исходно не кодировалась геном*. Поскольку другие V_κ и J_κ имеют в этих позициях разные кодоны, точка соединения сегментов может быть мощным источником разнообразия. Интересно, что аминокислота 96 участвует в формировании той области молекулы антитела, которая связывает антиген, а также в образовании контактов между легкой и тяжелой цепями. То же самое относится и к местам стыковок с J-сегментами у легких лямбда и тяжелых цепей.

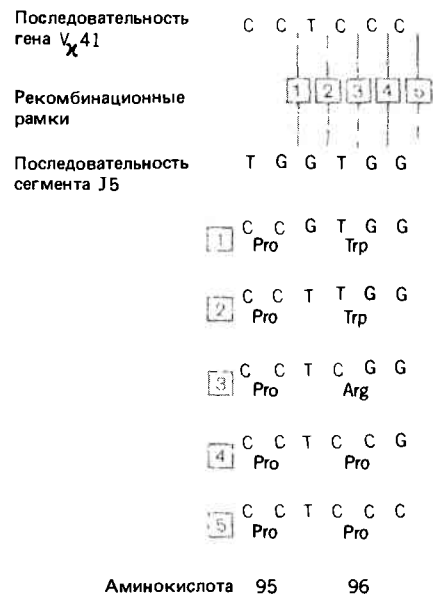


Рис. 39.8. Изменение рекомбинационных рамок между линейно расположенными сегментами V и J приводит к образованию новых триплетов, кодирующих аминокислоты 95 и 96 легкой цепи каппа-типа. Аналогичная ситуация имеет место и при объединении V—J и V—D—J других типов цепей.

Еще большее разнообразие может возникнуть в том случае, если рекомбинация не связана с положением первоначальных рамок считывания, а происходит между любой парой нуклеотидов в соответствующих V- и J-областях. Доказано, что такие события имеют место в действительности. Иногда это приводит к делеции кодона из соматического гена. В других случаях V- и J-сегменты соединяются таким образом, что трансляция J-последовательности происходит в неправильной рамке считывания. В результате образуется нефункциональный ген, который не экспрессируется из-за появления бессмысленного кодона.

Не известно, какая часть актов соединения приводит к образованию неполноценных генов. Возможно, это зависит от того, существуют ли преимущества для каких-то определенных видов реакций. Однако само по себе появление непригодных генов является дорогой ценой, которую клетка платит за увеличение разнообразия, получаемое ею в результате регулировки сайта рекомбинации.

Подобным образом, но в значительно большей степени вносят разнообразие реакции соединения с участием D-сегмента тяжелой цепи. Однако и в этом случае остается та же проблема появления нефункциональных генов за счет тех актов рекомбинации, которые нарушают фазу считывания последующих J- и C-сегментов.

Удивителен тот факт, что в некоторых тяжелых цепях имеются аминокислоты, находящиеся в соответствующих районах, но не относящиеся ни к одному из известных D-сегментов. Для объяснения этого факта были выдвинуты две гипотезы. Возможно, что новые D-сегменты образуются в соматических клетках в результате D—D-рекомбинации. Эта рекомбинация может контролироваться иным механизмом, чем тот, который обеспечивает соединение V—D—J. Другая гипотеза предполагает, что в процессе рекомбинации в места стыков V—D или D—J могут внедряться дополнительные нуклеотиды. Эта версия основана на том, что включение новых нуклеотидов происходит благодаря активности фермента дезоксирибонуклеотидтрансферазы (известно, что этот фермент активен в лимфоцитах), который использует образующиеся при этом свободные 3'-концы. Во всех случаях существование новых последовательностей указывает на наличие особого механизма образования разнообразия.

Рекомбинация между V- и С-генами вызывает делеции и перестройки последовательностей ДНК

По-видимому, сборка частей иммуноглобулиновых генов как легких, так и тяжелых цепей происходит посредством одного и того же механизма, хотя количество фрагментов различно в каждом случае. Использование того же самого механизма подтверждается присутствием одинаковых канонических последовательностей, находящихся на границах всех иммуноглобулиновых сегментов зародышевой линии, участвующих в сборке. Каждая такая последовательность состоит из семи-нуклеотидного палиндрома, отделенного от девятичленной нуклеотидной последовательности участками из 12 или 23 пар оснований.

На рис. 39.9 показана взаимосвязь канонических последовательностей в Ig-локусе мыши. В каппа-локусе за каждым V_{κ} -геном следует последовательность со спейсером из 12 пар нуклеотидов. Перед каждым J_{κ} -сегментом расположена каноническая последовательность, имеющая спейсер из 23 пар оснований. Последовательности, прилежащие к V- и J-сегментам, находятся в противоположных ориентациях. В лямбда-локусе обнаружена обратная организация: за каждым геном следует каноническая последовательность, со спейсером из 23 нуклеотидов, в то же время перед каждым J_{λ} -сегментом расположена последовательность, разделенная спейсером из 12 пар нуклеотидов.

Сравнение этих двух случаев указывает на существование правила, по которому происходит реакция соединения. Каноническая последовательность, имеющая спейсер одного типа, может соединиться только с последовательностью с другим типом спейсера. Поскольку канонические последовательности при V- и J-сегментах могут располагаться в любом порядке, то разные типы спейсеров не вносят какой-либо определенной информации. Однако они служат для того, чтобы препятствовать рекомбинации как между V-генами, так и между J-сегментами.

Структура компонентов генов тяжелых цепей подтверждает это положение. За каждым V_H -геном следует каноническая последовательность со спейсером из

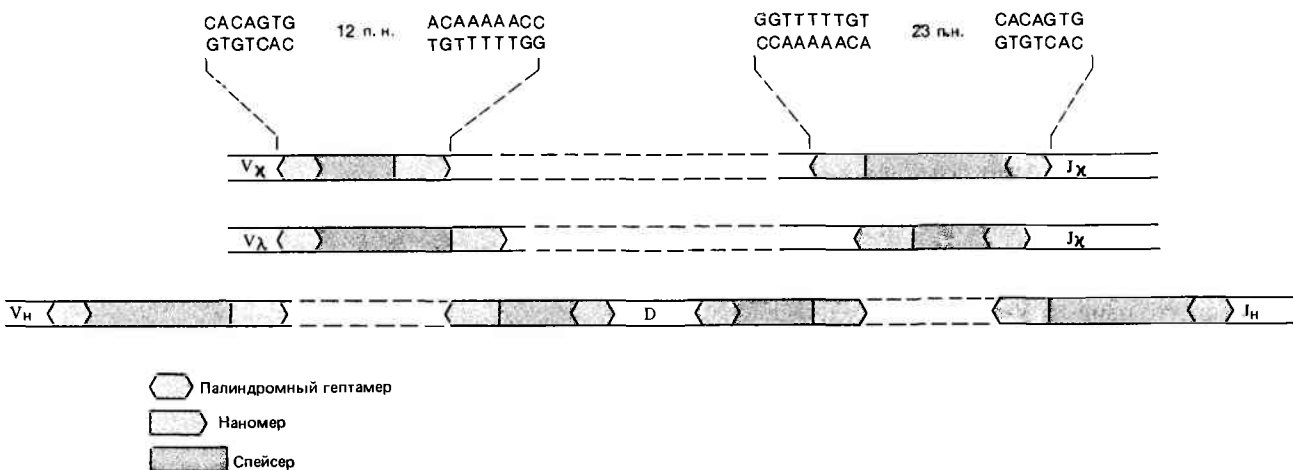


Рис. 39.9. Каждая пара рекомбинирующих сайтов содержит канонические последовательности, ориентированные в противоположных направлениях. У одного из членов пары эти последовательности

разделяются спейсером из 12 нуклеотидов, у другого — спейсером из 23 нуклеотидов.

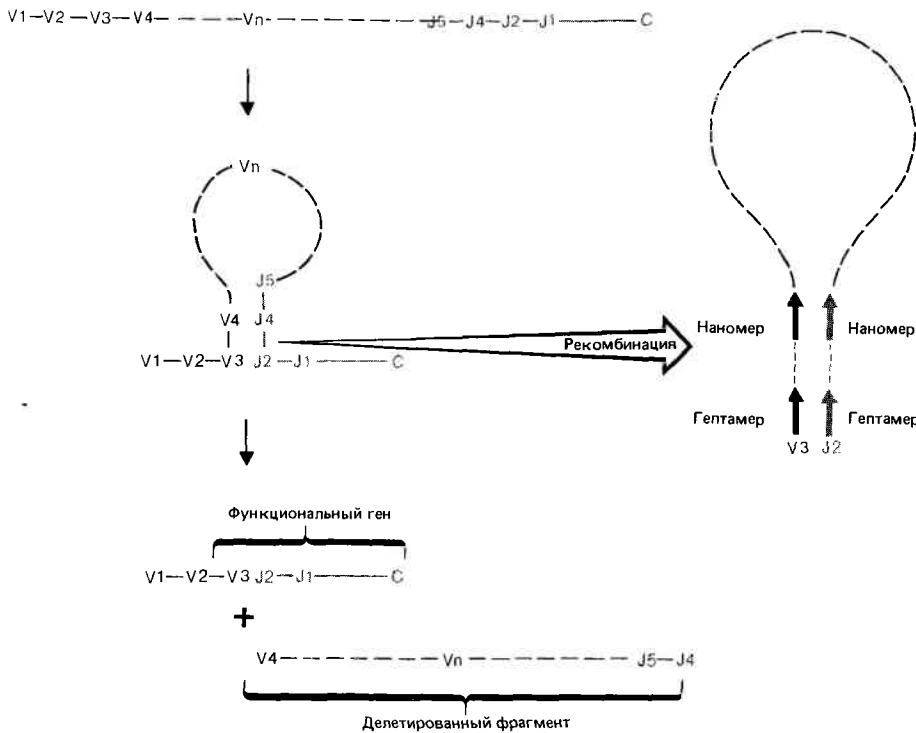


Рис. 39.10. Канонические последовательности могут использоваться для сближения V- и J-сегментов, образования разрыва с последующим воссоединением и удалением участка, расположенного между ними.

23 пар оснований. Канонические последовательности, фланкирующие D-сегменты, имеют 12-нуклеотидный спейсер, а те, что располагаются перед J-участками, несут спейсер из 23 п. н. Таким образом, V-ген должен быть соединен только с D-сегментом, а D в свою очередь только с J. V-ген не может непосредственно соединиться с J, поскольку они имеют канонические последовательности одного типа.

Расстояние между компонентами канонической последовательности почти точно соответствует одному или двум виткам двойной спирали. Эта особенность может отражать тот факт, что некоторые геометрические соотношения используются в рекомбинационных реакциях. Например, рекомбинационный белок (или белки) может присоединиться к одной стороне ДНК, что напоминает взаимодействие с другими сигнальными элементами генома, такими, как промоторы и операторы (гл. 11 и 14).

Употребляя термин «рекомбинация» для описания процесса сборки иммуноглобулиновых генов, мы не считаем, что эта реакция является реципрокной и происходит между гомологичными последовательностями. Хотя в данном случае налицо физическая перестройка последовательностей, включающая разрыв и воссоединение (а не явления, характерные для транспозиции), однако сам механизм этого молекулярного процесса пока еще неизвестен. На рис. 39.10–39.12 схематически изображены три возможных варианта.

Делеционная модель показана на рис. 39.10. Она предполагает, что один из V-генов непосредственно «сливается» с одним из J-сегментов. Все, что находится между двумя взаимодействующими последовательностями, удаляется и, возможно, теряется. Эти последовательности могут быть приближены одна к другой за счет спаривания между инвертированными каноническими последовательностями (разделенные цепи двухцепочечной молекулы при этом будут образовывать крестообразную структуру). Следует заметить, что в случае реципрокной рекомбинации, происшедшей в пределах канонических

последовательностей, область между V и J не была бы удалена, а продолжала бы существовать, но только в обратной ориентации. Следовательно, если спаривание канонических последовательностей приводит к делеции, то, вероятно, мы сталкиваемся с другим типом разрыва и воссоединения. Одна из возможностей заключается в том, что петля, образовавшаяся в результате спаривания нуклеотидов, будет отрезана у самого основания, а концы V- и J-областей последовательно соединятся друг с другом.

На рис. 39.11 представлена одна из возможных инверсионных моделей. До сих пор мы предполагали, что все V-гены находятся в той же ориентации, что и J-сегменты, хотя, как именно происходит соединение этих областей, точно не известно. Если какой-либо из V-генов находился в обратной ориентации, то при рекомбинации в результате инверсии промежуточного материала мог образоваться функциональный ген, что привело бы к соединению V с J. Если же часть V-генов находится в той же ориентации, что и J, а часть – в противоположной, то в некоторых

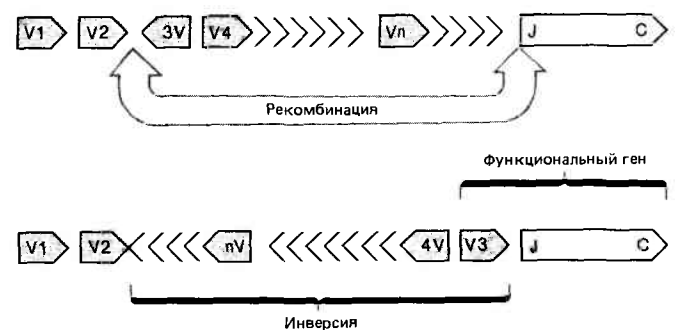


Рис. 39.11. В данной модели рассматривается реакция соединения, в которой участвует V-ген, ориентированный в противоположном направлении относительно J-сегмента. Образование функционального V—J—C-гена может произойти в результате инверсии области между V и J.

случаях соединение V—J сопровождалось бы делециями, тогда как в других случаях они сопровождалась бы инверсиями.

Модель неравного кроссинговера показана на рис. 39.12. Она предполагает неравную рекомбинацию между сестринскими хроматидами (реплицированными копиями гена). (В этой модели роль канонических последовательностей остается неясной.) В результате рекомбинации в одной из хроматид образуется функциональный ген; его образование сопровождается делецией участка между рекомбинирующими V- и J-сегментами. У другой хроматиды область между V и J дублирована. При этом V- и C-гены по обе стороны от необычной дублированной области способны к дальнейшей рекомбинации. Две хроматиды при митозе расходятся в дочерние клетки.

Обе модели, делеционная и неравного кроссинговера, предполагают одинаковую структуру активного гена, однако судьба потомства тех клеток, в которых произошли делеция и рекомбинация, различна. По делеционной модели все потомки такой клетки несут делецию. Модель неравного кроссинговера подразумевает, что половина потомков имеет делецию, а другая половина содержит потенциально активную дубликацию. Для выяснения истинной ситуации следовало бы провести генеалогический анализ клеток, в которых произошла рекомбинация. Инверсионная модель несколько отличается от двух предыдущих; она предполагает, что материал между V- и C-генами остается в инвертированной области.

Некоторые возможные случаи аллельного исключения

Каждая V—J- (или V—D—J) рекомбинация происходит только один раз. В данном лимфоците в результате перестройки соответствующего типа образуются один ген легкой и один ген тяжелой цепей. Поскольку в таком событии участвуют гены только одной из гомологичных хромосом, другой аллель в этой клетке не экспрессируется. Это явление получило название *аллельного исключения*.

Существование аллельного исключения вносит определенные трудности в анализ процесса соматической рекомбинации. Проба, которая гибридизуется с перестроенным участком одной из гомологичных хромосом,

будет также выявлять аллельную последовательность на другой хромосоме. Следовательно, мы вынуждены анализировать совместно разные события, происходящие с двумя гомологичными хромосомами. Однако обычно наблюдаемую картину перестройки активного гена можно интерпретировать как делецию материала между рекомбинирующими V- и C-локусами.

Иногда в функционирующей клетке обнаруживаются оба типа организации данного гена: эмбриональный и перестроенный. В таких случаях мы считаем, что перестройка последовательностей гена произошла на одной из хромосом, в то время как другая осталась неизменной. Реорганизация последовательностей ДНК могла происходить либо путем делеции, либо путем неравного кроссинговера.

В некоторых клеточных линиях обе хромосомы перестраиваются независимо. В ряде случаев материал хромосомы между рекомбинирующими V- и C-генами полностью утрачивается клеткой. Наиболее простое объяснение этому — независимые делеции, происходящие в каждой из гомологичных хромосом, но причиной может быть и неравный кроссинговер.

В тех случаях, когда эмбриональный тип организации отсутствует в обеих гомологичных хромосомах, можно считать, что только в одной из них произошла продуктивная перестройка, приведшая к образованию функционального гена. Перестройка в другой хромосоме не будет продуктивной, и, хотя здесь возможно несколько вариантов, ни в одном из них последовательность гена не является функциональной и не может экспрессироваться в иммуноглобулиновую цепь. Если при перестройке произошел сдвиг по тем нуклеотидам, которые участвуют в соединении генных компонентов, то структура образовавшегося гена на первый взгляд может не отличаться от нормальной. Заметны большие неправильные перестройки двух типов: либо реакция соединения прошла не до конца и ген только частично перестроился, например D соединился с J, но остался неприсоединенным к V, либо неправильные перестройки появляются в случае рекомбинации по участкам, отклоняющимся от нормальной канонической последовательности. Так, например, известен случай, когда V_{α} -ген присоединяется к участку интрона между кластером J_{α} - и C_{α} -геном; при этом образуется псевдоген, который не содержит J-сегмента.

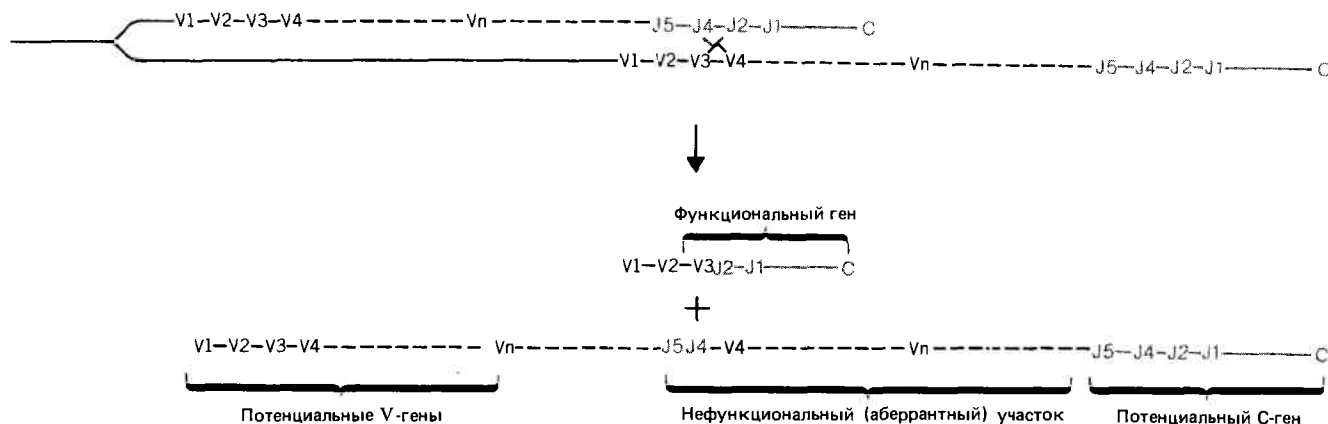


Рис. 39.12. Неравный кроссинговер между сестринскими хроматидами может привести к возникновению рекомбинантов двух типов; один из них должен содержать функциональный ген,

образовавшийся в результате делеции области между V и J, а у другого эта область будет дублирована.

Непродуктивные перестройки часто происходят в миеломах и в плазмацитомах. Согласно стохастической модели, они типичны не только для опухолевых клеток, а отражают общую картину перестройки последовательностей ДНК. В каждой клетке первоначально присутствуют два локуса, конфигурация которых соответствует эмбриональной и обозначается как Ig^0 . Любой из этих локусов может перестроиться и образовать продуктивный ген Ig^+ или непродуктивный Ig^- . В том случае, если перестройка оказалась удачной, синтез функциональной цепи не позволяет другому аллелю перестраиваться. Активные клетки имеют генотип Ig^0/Ig^+ . Если перестройка непродуктивна, то образуется клетка Ig^0/Ig^- . Однако это не служит препятствием для перестройки другого аллеля, характерного для клеток зародышевой линии. Если эта перестройка продуктивна, клетка, образующая иммуноглобулин, имеет генотип Ig^+/Ig^- . И в этом случае образование активной цепи подавляет возможные дальнейшие перестройки.

Непродуктивные перестройки обоих аллелей приводят к образованию клеток Ig^-/Ig^- . (Возможность существования таких клеток следует из стохастической модели.) Вероятно, в некоторых случаях клетка с конфигурацией аллелей Ig^-/Ig^- все же может предпринять попытку дальнейшей перестройки. Так, в некоторых случаях наблюдаемые образцы ДНК могли образоваться только в результате последовательных перестроек.

Если в одном из Ig^- -аллелей V- и J-сегменты стыковались неудачно, то возможна ситуация, когда другой V-ген совершит «скачок» и соединится с одним из оставшихся сегментов J, расположенных позади того, который перестроился ранее. Если такое соединение происходит путем неравного кроссинговера, Ig^- -локус, образованный в результате неправильной дупликации, все же способен обеспечивать соединение V- и C-генов, расположенных по обе стороны от этой дупликации. Эта модель объясняет природу необычных структур, обнаруживаемых в локусах с непродуктивной перестройкой. Они также могут быть объяснены сменяющимися друг друга сериями внутрихромосомных делеций и инверсий. В соответствии с данной моделью клетка осуществляет рекомбинацию V- и C-генов до тех пор, пока не будет достигнута продуктивная перестройка. Аллельное исключение обуславливается подавлением дальнейшей перестройки сразу же после образования активной цепи. Эта обратная связь осуществляется независимо для локусов тяжелых и легких цепей (гены тяжелых цепей обычно перестраиваются первыми), однако в случае легких цепей это правило должно выполняться в равной мере для обоих семейств (клетки могут иметь активную цепь либо каппа-, либо лямбда-типа). Вероятно, каппа-гены перестраиваются раньше, и перестройка генов лямбда происходит только в том случае, если обе попытки перестроить каппа-гены оказались неудачными.

Какая связь между соединением V- и C-генов и их активацией? Неперестроенные V_H -гены не продуцируют РНК. Только когда V_H -ген правильно соединяется с C_H -геном, образуемая генная единица транскрибируется. Однако последовательность, расположенная перед V-геном, не изменяется в процессе реакции соединения, и тот же самый промотор должен присутствовать в перестроенных, непродуктивно перестроенных и продуктивно перестроенных генах.

Означает ли это, что активация промоторов V-генов ка-

ким-то образом зависит от последовательностей, расположенных дальше по ходу транскрипции? Если неэкспрессируемый аллель C_H остается в эмбриональной конфигурации, он может транскрибироваться в непродуктивную РНК размером 8,4 т.п.н., которая содержит также последовательности, расположенные перед C_H -геном. Одно из возможных объяснений состоит в том, что активация промотора V_H происходит при транслокации этого гена к C_H -области. Большие непродуктивные перестройки большей частью не транскрибируются, и причина этого неизвестна. Возможно, в процессе таких перестроек теряются некоторые необходимые сайты. В редких случаях образуется нефункциональная РНК.

Дальнейшая рекомбинация ДНК обуславливает переключение классов иммуноглобулинов

Класс иммуноглобулина определяется типом C_H -области. Так, например, IgM соответствует C_H1 -последовательность, $IgD-C\delta$, $IgG-C\gamma$, $IgE-C\epsilon$ и $IgA-C\alpha$. В каждый данный момент времени лимфоцит продуцирует иммуноглобулины одного класса, но в процессе развития иммунокомпетентной клетки классы могут сменять друг друга. Это явление сопровождается изменением в экспрессии гена C_H -области; оно получило название **переключение класса**. Данный феномен затрагивает C_H -ген, в то время как V_H -ген остается неизменным. Таким образом, один и тот же V_H -ген может успешно экспрессироваться в сочетании с разными C_H -генами. Легкая цепь продолжает функционировать без изменений на протяжении всех стадий развития лимфоцита. Следовательно, переключение класса допускает изменение типа эффекторного ответа (обуславливаемого C_H -областью), но оставляет неизменным антигенное распознавание (контролируемое V-районами).

Экспрессия иммуноглобулинов незрелыми лимфоцитами всегда начинается с синтеза IgM . Клетки, образующие IgM , имеют эмбриональное строение C_H -генов, характерное для зародышевой линии (рис. 39.7). V—D—J-рекомбинации уже достаточно, чтобы запустить экспрессию C_H -гена. Изменения в экспрессии C_H -генов осуществляются двумя путями. Некоторые из них происходят на уровне процессинга РНК, однако большая часть этих изменений происходит путем дальнейших рекомбинаций ДНК. В этом процессе участвует система, отличная от той, которая функционирует при соединении сегментов V—D—J (она может вступать в действие только после этого соединения).

В клетках, экспрессирующих поздние C_H -гены, последовательности, соответствующие стадиям, которые предшествуют экспрессии этого C_H -гена (в частности, C_H1), обычно делетированы. Таким образом, переключение класса связано с рекомбинацией, в результате которой новый C_H -ген оказывается рядом с экспрессируемым комплексом V—D—J. Определение первичной последовательности перестроенных единиц V—D—J— C_H после переключения показывает, что сайты переключения расположены непосредственно перед C_H -генами. Эти сайты получили название **S-областей**. Рис. 39.13 иллюстрирует два (гипотетических) примера переключения.

При первом переключении экспрессия C_H1 -гена сменяется C_H1 . Активация C_H1 -гена достигается путем ре-

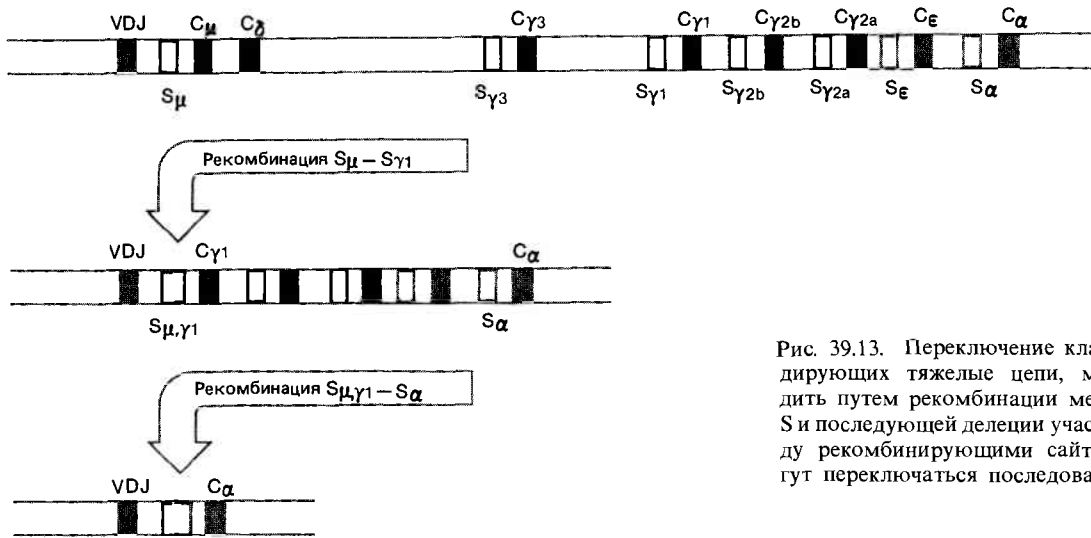


Рис. 39.13. Переключение классов генов, кодирующих тяжелые цепи, может происходить путем рекомбинации между областями S и последующей делеции участков ДНК между рекомбинирующими сайтами. Гены могут переключаться последовательно.

комбинации между сайтами S_{μ} и S_{γ_1} , а промежуточный материал делетируется. S_{μ} -сайт находится между областью V—D—J и C_{μ} -геном. Сайт S_{γ_1} расположен перед C_{γ_1} -геном. Вся область между V—D—J и C_{γ_1} удаляется как интрон при процессинге РНК.

Линейная делеционная модель накладывает определенные ограничения на локус C_H -генов. Если переключение класса уже произошло, то это означает, что любой C_H -ген, расположенный между C_{μ} - и новым C_H -геном, в данной клетке экспрессироваться не может. На рис. 39.13 показано, что клетки, в которых произошло переключение и функционирует ген C_{γ_1} , не могут давать начало клеткам, экспрессирующим C_{γ_3} -ген, так как он делетирован. Однако в принципе возможно другое переключение на любой другой ген, расположенный по ходу транскрипции за геном, экспрессируемым в данной клетке. На этом же рисунке показано второе переключение, приводящее к экспрессии C_{α} и сопровождающееся рекомбинацией между S_{α} и комплексной областью S_{μ, γ_1} , образованной при первом переключении.

Предполагается, что перед кодирующими последовательностями всех C_H -генов имеются S-области. Однако не известно, существуют ли какие-либо ограничения при их использовании. И, хотя переключения классов фактически происходят последовательно, мы не знаем, является ли это главным или второстепенным фактором, приводящим к экспрессии поздних C_H -генов, и возможно ли прямое переключение IgM на любой другой класс.

Кроме делеции материала перед экспрессируемым C_H -геном в клетке, в которой уже произошло одно переключение, возможны перестройки в области, расположенной по ходу транскрипции за функционирующим геном. (Это может влиять на способность к дальнейшим последовательным переключениям.) Согласно одной из моделей, первое переключение между S_{μ} - и другой S-областью активирует рекомбинационную систему, которая затем может осуществлять перестройки и дальнейшие переключения.

Известно, что границы сайтов переключения четко не определены, так как показано, что в разных клетках, экспрессирующих одинаковый C_H -ген, рекомбинация происходит в разных точках. Установить эти границы будет возможно только при определении первичной структуры большого числа S-областей, в пределах которых могут

происходить переключения. А пока мы можем лишь выявить в них какие-то общие особенности, которые, вероятно, используются в процессе переключения.

S-области располагаются перед C_H -генами на расстоянии примерно 2 т. п. н. Из трех разных S_{μ} -сайтов, которые были охарактеризованы, два обнаружили гомологию. Гомология существует также с некоторыми, но не всеми S-областями других C_H -генов. В некоторых S-областях присутствуют короткие тандемно повторяющиеся последовательности, однако определить их значение пока не удается.

Известно, что переключение C_H -генов, как и соединение V- и C-генов, сопряжено с делециями. Однако из этого не следует, что в основе механизма переключения лежат внутримолекулярные делеции; процесс может происходить и путем обмена между сестринскими хроматидами. Такой обмен, по-видимому, происходит в известном случае при переключении экспрессии C_{α} -гена на экспрессию C_{γ_1} . Этот пример нельзя объяснить с позиции линейной делеционной модели, но можно, если допустить незаконный переход материала с сестринской хроматиды. Трудность в интерпретации данных, получаемых на миеломных клетках, заключается в том, что не известно, в какой степени они отражают естественный процесс переключения (который для многих опухолевых клеток не характерен вовсе). Следовательно, проблема сводится к тому, чтобы изучать этот процесс на нормальных лимфоцитах.

Изменения в экспрессии ранних генов тяжелой цепи могут происходить за счет процессинга РНК

Период синтеза IgM, характерный для начальных стадий развития лимфоцита, делится на две части, в течение которых синтезируются разные варианты константной области μ -типа. По мере того как стволовая клетка дифференцируется в предшественник В-лимфоцита, включается синтез легкой цепи и на поверхности клетки появляется молекула IgM (μ_2L_2). Эта форма IgM содержит вариант μ_m -константной области (индекс m указывает на то, что IgM расположен на мембране). Далее, когда В-лимфоцит в процессе дифференцировки превращается

в плазматическую клетку, начинает экспрессироваться μ_s -вариант константной области. (На самом деле IgM секретируется в виде пентамера IgM₅, при этом J служит соединяющим пептидом, который образует дисульфидные связи с μ -цепями.) μ_m - и μ_s -варианты тяжелой μ -цепи различаются только по длине. μ_m длиннее, чем μ_s , что объясняется изменением аминокислотной последовательности на С-конце. У μ_m -цепи в этом районе находится гидрофобный участок из 41 аминокислоты, который, возможно, определяет удержание всего белка на мембране. В μ_s -цепи он заменен гидрофильной последовательностью из 20 остатков. Эта замена, возможно, обуславливает прохождение через мембрану тяжелой цепи μ -типа.

Цепи μ_m и μ_s кодируются разными мРНК. Эти молекулы идентичны почти до конца последнего константного домена. Далее они различаются тем, что μ_m имеет 41 дополнительный кодон, за которым располагается нетранслируемый участок, а μ_s содержит 20 кодонов, заканчивающихся другой конечной последовательностью.

Исследование соответствующей геномной области показало, что концевые участки μ_m и μ_s кодируются разными экзонами. Эта связь между структурой гена и мРНК продемонстрирована на рис. 39.14. При экспрессии мембранной формы константная область образуется за счет сплайсинга шести экзонов. Первые четыре из них кодируют четыре домена константной области. Два других,

M1 и M2, детерминируют гидрофобную С-концевую область, состоящую из 41 остатка, и нетранслируемый конец.

На стадии секреции константная область формируется только первыми четырьмя экзонами. К последнему из них присоединяется 20 дополнительных кодонов и нетранслируемая область, в результате чего последовательность экзона удлинняется по сравнению с той, которая экспрессировалась на предыдущей (мембранной) стадии.

Следовательно, различие между двумя мРНК зависит от того, будет ли экзон последнего константного домена сплайсироваться с M1- и M2-экзонами (тогда этот экзон будет короче и образуется последовательность для μ_m) или этого не произойдет (тогда экзон удлинится и образует последовательность для μ_s). Каким же образом контролируется выбор того или другого сайта сплайсинга?

Одна из возможных моделей представлена на рис. 39.14. На стадии связывания с мембраной участок полиаденилирования ядерной РНК располагается после M2 (либо за счет разрыва транскрипта в этом сайте, либо за счет терминации). Поскольку в ядерной РНК участок сплайсинга расположен перед началом экзона M1, она использует сайт сплайсинга, находящийся в последнем константном экзоне. На стадии секреции транскрипция ядерной РНК останавливается после последнего константного экзона. Поскольку в ядерной РНК на этой стадии уже

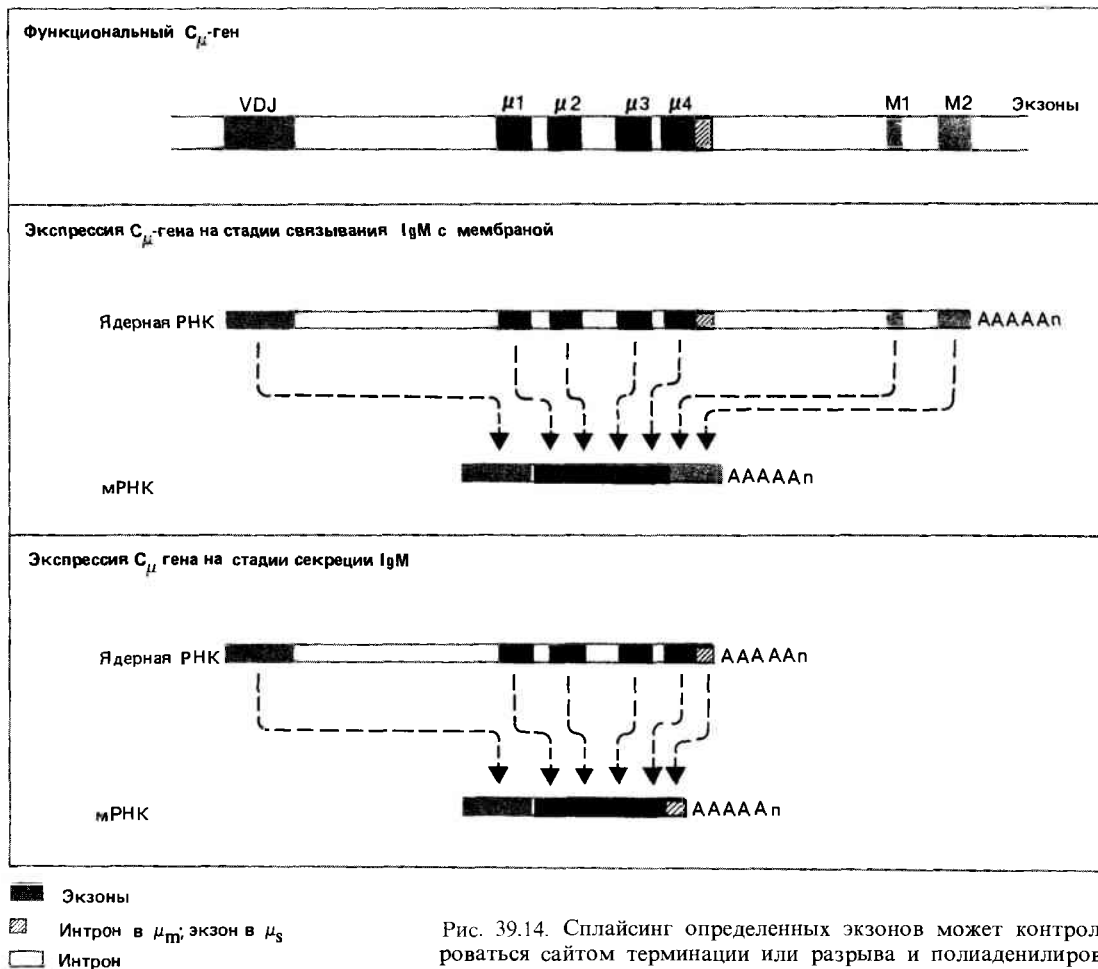


Рис. 39.14. Сплайсинг определенных экзонов может контролироваться сайтом терминации или разрыва и полиаденилирования таким образом, что будут экспрессироваться разные формы гена тяжелой цепи.

нет добавочных экзонов, сайт сплайсинга, присутствующий в последнем экзоне, не используется. Данная модель предполагает, что регуляция состоит в выборе сайта разрыва и полиаденилирования (или терминации), и исключает возможность разного типа сплайсинга для одной и той же РНК. Однако для подтверждения этой точки зрения необходимо дальнейшее изучение ядерной РНК.

Аналогичный переход от мембранной к секреторируемой форме обнаружен для других константных областей. Сохранение экзонной структуры подтверждает, что в данном случае функционирует сходный механизм. Экзоны M1 и M2 трех γ -цепей очень близки, они кодируют последовательность из 66 аминокислот, из которых первые аминокислотные остатки (в количестве 41) гомологичны мембранной последовательности μ -типа. Возможно, узнаваемым рецепторным белком. Этот белок и «заякоривает» в мембране *m*-форму иммуноглобулина. Можно предположить, что задержание белка в мембране необходимо для инициации клеточной пролиферации в ответ на первоначальное распознавание антигена.

Одновременный синтез IgM и IgD зрелыми В-лимфоцитами — это исключение из правила, согласно которому одна клетка синтезирует только один тип иммуноглобулина. μ - и δ -иммуноглобулины различаются константными областями тяжелой цепи, но идентичны в остальных участках белковой молекулы. По-видимому, совместная экспрессия двух C_H -генов является результатом альтернативного процессинга РНК. Константная область δ расположена близко от μ -гена; если транскрипция продолжается в эту область, VDJ-экзон может сплайсироваться с экзонами δ -гена. Это напоминает ситуацию с экспрессией поздних генов аденовируса (рис. 20.20). Образование разных полиаденилированных концов в ядерной РНК контролирует выбор кодирующих областей для сплайсинга с начальным экзоном. (Известен только один случай, когда выбор участков полиаденилирования недостаточен. Переключение от δ_m к δ_s может происходить только путем прямого выбора различных сайтов сплайсинга.)

Удивительный пример использования сплайсинга при экспрессии генов константных областей демонстрируют клетки лейкемии мышей, в которых переключение от μ к синтезу $\gamma 2b$ происходит при сохранении μ -гена. Одно из возможных объяснений этого явления состоит в том, что транскрипция продолжается через всю константную область до $\gamma 2b$ -гена, после чего происходит сплайсинг очень длинного (50 т.п.н.) транскрипта. (Возможно, однако, что ген $\gamma 2b$ может быть перенесен ближе к μ -гену.) Является ли такой путь нормальным альтернативным вариантом рекомбинации ДНК или же характерен только для тех особых клеток, в которых обнаруживается, — не известно.

Соматические мутации вносят определенный вклад в разнообразие антител

Сравнения последовательностей функционирующих иммуноглобулиновых генов и соответствующих V-генов в недифференцированной ткани показало, что в клеточной популяции, экспрессирующей антитела, появляются новые последовательности. Как мы уже упоминали, некоторое дополнительное разнообразие обусловлено сдвигом сайта рекомбинации при сборке экзонов V—J легкой

цепи и экзонов V—D—J тяжелой цепи. Однако некоторые изменения происходят в участках V-области, расположенных до точек рекомбинации, т.е. внутри самого варибельного домена. Они представляют собой **соматические мутации**, специфически индуцируемые в активном лимфоците. В случае тяжелых и легких цепей каппа-типа в создании варибельного домена могли бы участвовать многие потенциальные V-гены. Каждый образец, представляющий собой экспрессируемый V-ген, можно использовать для идентификации всех соответствующих сегментов зародышевой линии. Совокупность этих последовательностей должны составлять полный набор V-генов организма. Любой экспрессируемый ген с отличающейся первичной структурой, по-видимому, образуется путем соматических изменений. Однако основная экспериментальная трудность такого анализа состоит в том, что невозможно идентифицировать каждый потенциальный компонент V-генов зародышевой линии. Эта проблема просто разрешается на природной модели, представляющей систему мышинных генов лямбда-цепей. Исследования нескольких миелом, продуцирующих цепи λ_1 -типа, показали, что многие из них кодируются последовательностью единственного эмбрионального гена. Однако другие имеют отличающиеся последовательности, которые, вероятно, образовались в результате мутации исходного эмбрионального гена.

В ряде случаев для определения частоты соматического мутагенеза необходимо исследовать огромное число клеток, экспрессирующих один и тот же V-ген. Практический подход для идентификации такой группы состоит в характеристике иммуноглобулинов, получаемых от серии клеточных линий, осуществляющих иммунный ответ на одинаковый антиген. (Используемые для этой цели антигены представляют собой маленькие молекулы, имеющие дискретную структуру, которая, по-видимому, обуславливает стойкий иммунный ответ. Они отличаются от больших белков, отдельные части которых могут стимулировать образование разных антител. Эти маленькие молекулы получили название *гаптен*ов. Для того чтобы придать им свойства антигена, их соединяют с инертным белковым носителем. Иммунизируя таким антигеном мышь, получают реактивные лимфоциты, которые соединяют путем слияния с миеломными клетками для получения **гибридом**. Такие гибридные клетки продолжают независимый синтез желаемого антитела.)

Анализ 19 различных клеточных линий, синтезирующих антитела, связывающие гаптенфосфорилхолин, показал, что 10 из них имеют одну и ту же V_H -последовательность (экспрессируемую в сочетании с μ -, γ - или α -константными областями). В клетках зародышевой линии эта V_H -последовательность идентифицируется как T15 и является одной из четырех V_H -генов, гибридуемых с анализируемой V_H -пробой. Другие девять функционирующих генов отличаются друг от друга и от всех четырех членов семейства, тестируемых в эмбриональных клетках. Тем не менее они в большей степени родственны последовательности T15, чем каким-либо другим, и фланкированы теми же последовательностями, что и T15. Все это позволяет предполагать, что указанные девять генов имеют общего предка — члена семейства T15 и произошли от него путем соматических мутаций.

Последовательности этих экспрессируемых генов отличаются от эмбриональных в равной степени во всех областях варибельного домена: и в каркасных, и в ги-

первариабельных. Изменения последовательностей обнаруживаются как в кодирующих областях, так и в областях, расположенных после гена. Однако в области, предшествующей гену, они, по-видимому, не распространяются слишком далеко. Все изменения затрагивают отдельные нуклеотидные пары. Степень вариабельности в каждом случае различна. У одного из генов дивергенция достигает 3,8% по сравнению с соответствующим геном клеток зародышевой линии, у другого — всего лишь 1,4%. Это соответствует 8 и 3 аминокислотным заменам в белке. На аминокислотную последовательность влияет только незначительная часть таких соматических мутаций, поскольку многие из них затрагивают третье основание кодона или попадают в нетранслируемую область.

Существование большей части неэффективных мутаций указывает на то, что процесс соматического мутагенеза носит более или менее случайный характер и касается в основном V-гена или области, расположенной за ним. Для генов тяжелых цепей переключение классов может служить тем механизмом, который активирует мутагенез. V_H-гены, ассоциированные с константными областями μ -типа, редко мутируют, тогда как те же самые V_H-гены, соединенные с γ - или α -цепями, могут подвергаться мутациям значительно чаще.

Случайно распределяемые мутации непредсказуемым образом влияют на функцию белка. Некоторые из них могут вызывать его инактивацию, другие, наоборот, способствуют появлению высокой специфичности к определенному антигену. Таким образом, суть данного процесса сводится к тому, что среди популяции лимфоцитов селекционируются такие клетки-продуценты антител, в которых благодаря случайной мутации образовался V-домен, подходящий для связывания присутствующего антигена.

Однако это полностью не разрешает нашего вопроса о причине разнообразия антител. Можно считать, что за него ответственно огромное число V-генов, не говоря уже о дополнительном разнообразии, возникающем при соединении D- и J-сегментов. В случае фосфорилхолина показано, что при многочисленных независимых введениях антигена для образования антитела обычно используется один и тот же V_H-ген. Можно предполагать, что перестройки происходят в огромном числе лимфоцитов еще до появления антигена, поскольку всегда находится определенный V_H-ген либо в исходном, либо в мутированном виде, готовый вступить в действие. Означает ли это, что количество перестроек так велико, что каждый потенциальный V-ген фактически присутствует в экспрессируемой форме?

Строение главного локуса гистосовместимости

Главный локус гистосовместимости занимает небольшой участок на одной из хромосом мыши (локус H2) и человека (HLA). В пределах этого участка выявляется много генов, продукты которых выполняют функции, связанные с иммунным ответом. Для тех индивидуальных генных локусов, продукты которых идентифицированы, в популяции было обнаружено много аллелей. Локус считается высокополиморфным, и это означает, что индивидуальные геномы в популяции с большей вероятностью отличаются друг от друга по этому локусу. Некоторые типы функций, картируемые в H2 и прилежащем районе, суммированы на карте, представленной

на рис. 39.15. Классическая H2-область занимает 0,3 единицы карты и включает в себя три класса генов. Прилежащая область попадает в другую единицу карты; в ней располагаются гены, родственные генам H2-области. В молекулярных масштабах этот маленький участок хромосомы имеет весьма значительные размеры: в общей сложности 1,3 единицы карты составляют более 2500 т. п. н. ДНК.

Гены, относящиеся к классу I, кодируют **трансплантационные антигены**, специфические белки, присутствующие на поверхности всех клеток млекопитающих. Как следует из их названия, эти белки ответственны за отторжение чужеродной ткани, которая отличается от собственных тканей организма определенным набором трансплантационных антигенов. Эти белки занимают важное место в иммунной системе организма, и их присутствие на поверхности (цитотоксических) T-лимфоцитов играет определенную роль в процессе уничтожения клеток-мишеней. Все белки этого типа кодируются областями K, D, L и R и их функции выявлены серологическим методом (т. е. по антигенным свойствам). Каждая линия мышей имеет только один из нескольких возможных аллелей, отвечающих за любую из этих функций.

Гены, принадлежащие классу II, кодируют белки, обнаруживаемые на поверхностях B- и T-лимфоцитов, а также макрофагов. Они принимают участие во взаимодействии между клетками, что необходимо для осуществления иммунного ответа. Область I генов этого класса подразделяется на ряд подобластей, обозначаемых как A, B, J, E и C (на карте не показаны).

Гены класса III кодируют **белки комплемента**. Этот участок локуса также известен как S-область. Название области происходит от слова *serum* (*сыворотка*) и указывает на то, что эти белки представляют собой ее компоненты. Их функция заключается в том, чтобы взаимодействовать с комплексом антитело—антиген и вызывать лизис клеток.

Справа от H2-области расположены локусы Qa и Tla, кодирующие белки, обнаруживаемые на гематопозитических клетках. Эти белки описываются как дифференцировочные антигены, т. к. каждый из них присутствует только на поверхности определенной группы клеток крови, вероятно имеющих отношение к их функции. Поскольку по своей структуре они родственны белкам класса I области H2, можно считать, что кодирующие их гены являются продолжением этой области хромосомы, но выполняют более специфические функции, связанные с развитием лимфоцитов и макрофагов.

Трансплантационные антигены представляют собой трансмембранные белки, состоящие из двух цепей. Одна из них представлена $\beta 2$ -микроглобулином, белком с мол. массой 12000 (он кодируется единственным геном, расположенным на другой хромосоме). Этот компонент нужен для того, чтобы димерный белок расположился на клеточной поверхности. Трансмембранный компонент представлен полипептидной цепью в 45000 дальтон, кодируемой локусом гистосовместимости. У белка имеются три **наружных домена** (примерно по 90 аминокислот каждый; один из этих доменов взаимодействует с $\beta 2$ -микроглобулином), **трансмембранная область** из 40 аминокислотных остатков и короткий **цитоплазматический домен** из 30 остатков, который располагается внутри клетки.

Организация генов класса I соответствует структуре белка. Первый экзон кодирует сигнальную последовательность (которая отщепляется от белка во время про-

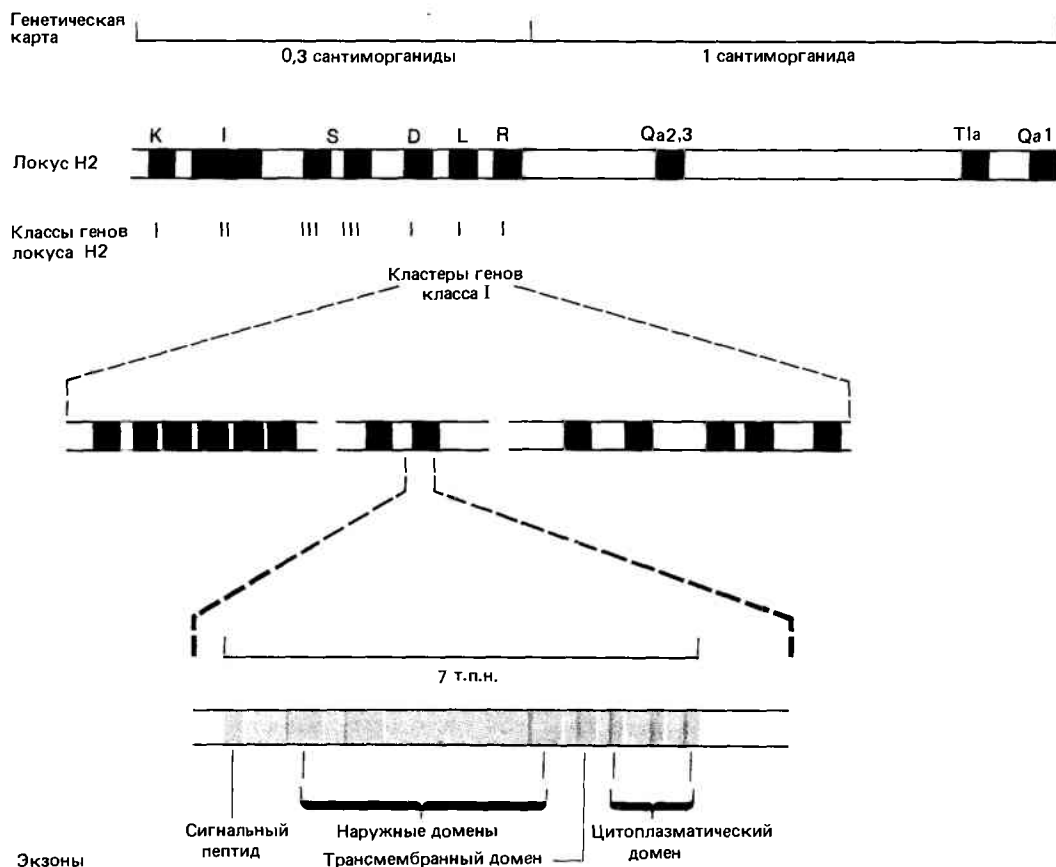


Рис. 39.15. Локус гистосовместимости содержит ряд генетических участков; гены класса I организованы в огромные кла-

стеры; существует определенное соответствие между структурой гена и структурой кодируемого им белка.

хождения через мембрану). Следующие три экзона кодируют по одному наружному домену. Последовательность трансмембранного домена определяется пятым экзонам. И наконец, последние три довольно маленьких экзона совместно кодируют цитоплазматический домен. Структура гена трансплантационного антигена человека отличается только тем, что цитоплазматический домен кодируется двумя экзонами.

У мыши гены класса I организованы в кластеры. Недавно в этой области было выявлено 36 генов, которые занимают более 850 т.п.н. ДНК (они образуют отдельные кластеры, однако порядок их взаимного расположения пока не установлен). Самая большая индивидуальная группа состоит из 7 генов и занимает более 190 т.п.н. В каждом из кластеров гены ориентированы в одном направлении. Прилежащие друг к другу гены обнаруживают более высокое сходство в первичной структуре, что предполагает их происхождение из общего предка путем тандемных дупликаций.

Для установления относительной локализации индивидуальных кластеров внутри генетического локуса могут быть использованы различия в рестрикционных картах ДНК разных линий мышей. (Заметим, что при этом рестрикционные различия используются как маркеры в генетическом картировании.) С помощью такого подхода было показано, что упомянутый выше кластер из 7 генов локализован в области Qa 2,3; другие гены обнаруживаются в области T1a; один аллель попадает в область H2—L. Родство между этими генами указывает на то,

что Qa и T1a совместно с H2-локусом образуют единую группу, относящуюся к классу I. Это означает, что гены класса I располагаются по обоим концам обширного кластера, а гены класса II и класса III либо с самого начала эволюционировали внутри кластера, либо были транслоцированы в то место, которое они занимают сейчас. Наряду с функциональными генами в исследуемой области присутствуют и псевдогены. В настоящее время существует способ предварительной оценки общего числа активных генов области.

У генов класса I экзон, кодирующий третий наружный домен, значительно более консервативен по сравнению с другими экзонами. Возможно, что консервативный домен представляет собой участок, взаимодействующий с $\beta 2$ -микроглобулином; и именно этим объясняется постоянство его структуры. Он также гомологичен доменам константной области иммуноглобулинов.

Была установлена структура гена $\beta 2$ -микроглобулина. Он состоит из четырех экзонов; первый из них кодирует сигнальную последовательность, второй — основную функциональную часть белка (от 3-й до 95-й аминокислоты), третий кодирует последние четыре аминокислоты полипептидной цепи и некоторую часть нетранслируемого концевой участка, а оставшаяся часть этого участка детерминируется последним четвертым экзонам. Длина $\beta 2$ -микроглобулина соизмерима с длиной иммуноглобулинового гена; имеется также определенное сходство в аминокислотном составе этих белков и некоторая (хотя и ограниченная) гомология нуклеотидной последователь-

ности между геном $\beta 2$ -микроглобулина и константными доменами Ig или геном типа I третьего наружного домена. Возможно, что все три упомянутые группы генов произошли от общего предшественника, который кодировал некий примитивный домен.

Структурная организация локусов иммуноглобулинов и гистосовместимости является яркой демонстрацией того, какое огромное расстояние пройдено от первоначальной концепции гена как обособленного неизменного участка ДНК, кодирующего один-единственный белок. Теперь нам известно, что гены могут образовывать огромные кластеры, состоящие из многих родственных последовательностей. Гены могут экспрессироваться, образуя альтернативные формы белка, и, наконец, они могут формироваться путем физических перестроек последовательностей ДНК в процессе развития соматических клеток. Таким образом, в течение последнего десятилетия представления об основном предмете генетики претерпели существенные изменения.

Рекомендуемая литература

Описание структуры генов Ig можно найти в книге Льюина (*Lewin, Gene Expression, 2, Eucaryotic Chromosomes, Wiley, New York, 1980, pp. 847–861*). Для углубленного изучения вопроса необходимо обратиться к оригинальным работам, резюмированным в ряде коротких обзоров, опубликованных в журнале "Cell". Деле-

ционная модель соединения V—J-сегментов была выдвинута Максом, Сейдманом и Ледером (*Max, Seidman, Leder, Proc. Nat. Acad. Sci. USA, 76, 3454, 1979*); модель, в основу которой положен обмен между сестринскими хроматидами, предложил Ван Несс (*Van Ness et al., Proc. Nat. Acad. Sci., USA, 79, 262–266, 1982*). О структуре генов тяжелых цепей сообщает Сакано (*Sakano et al., Nature, 286, 676–683, 1980*); в этой же работе анализируется роль канонических последовательностей при соединении V—(D)—J. В обзоре Эрли и Худа (*Early, Hood, Cell, 24, 1–3, 1981*) освещены явления аллельного исключения и перестройки иммуноглобулиновых генов; aberrантное соединение участков гена описали Макс с соавт. (*Max et al., Cell, 21, 793–799, 1980*); вопросы транскрипции исследовали Перри с соавт. (*Perry et al., Proc. Nat. Acad. Sci., USA, 77, 1937–1941, 1980*). В обзоре Дэвиса, Кима и Худа (*Davis, Kim, Hoode, Cell, 22, 1–2, 1980*) анализируется проблема переключения классов. Модель линейной делеции первоначально предложили Хоньо и Катаока (*Honjo, Kataoka, Proc. Nat. Acad. Sci., USA, 75, 2140–2144, 1978*). О результатах изучения процессинга РНК сообщают Эрли с соавт. (*Early et al., 20, 313–319, 1980*), Маки и др. (*Maki et al., Cell, 24, 353–356, 1981*) и Катаока с соавт. (*Kataoka et al., Proc. Nat. Acad. Sci., USA, 79, 2623–2627, 1982*). Соматические мутации рассмотрены в обзоре Балтимора (*Baltimore, Cell, 26, 295–296, 1981*). Организации генов гистосовместимости посвящен обзор Худа (*Hood et al., Cell, 28, 685–687, 1982*).

Словарь терминов

АКЦЕПТОРНАЯ ТОЧКА СПЛАЙСИНГА. Участок между правым концом интрона и левым концом экзона.

АЛЛЕЛЬ. Одна из двух (или нескольких) альтернативных форм гена.

АЛЛЕЛЬНОЕ ИСКЛЮЧЕНИЕ. Экспрессия в лимфоците *только одного* аллеля, кодирующего иммуноглобулин.

АЛЛОСТЕРИЧЕСКИЙ КОНТРОЛЬ. Способность эффекторов при взаимодействии с одним участком фермента оказывать влияние на активность другого.

АМБЕР-КОДОН. Триплет UAG, один из трех бессмысленных кодонов, обуславливающих терминацию белкового синтеза.

АМБЕР-МУТАЦИЯ. Любое изменение в ДНК, приводящее к появлению амбер-кодонов вместо триплета, кодирующего аминокислоту.

АМБЕР-СУПРЕССОРЫ. Мутантные гены, кодирующие тРНК с измененными антикодонами, способными узнавать кодон UAG. В некоторых случаях сохраняется способность узнавать исходный кодон.

АМИНОАЦИЛ-тРНК-СИНТЕТАЗЫ. Ферменты, осуществляющие ковалентное присоединение аминокислот к 2'- или 3'-ОН концам тРНК.

АМПЛИФИКАЦИЯ. Образование дополнительных копий хромосомных последовательностей, обнаруживаемых в хромосомной или внехромосомной ДНК.

АНТИГЕН. Чужеродная молекула, проникновение которой в организм вызывает синтез антитела (иммуноглобулина).

АНТИКОДОН. Триплет, занимающий определенное и постоянное положение в структуре молекулы тРНК; комплементарно взаимодействует с кодоном (или кодонами) мРНК.

АНТИТЕРМИНАТОРНЫЕ БЕЛКИ. Позволяют РНК-полимеразе продолжать транскрипцию, проходя определенные сайты терминации транскрипции.

АПОИНДУКТОР. Белок, взаимодействующий с ДНК для включения транскрипции РНК-полимеразой.

АССИМИЛЯЦИЯ ЦЕПИ. Способность ResA-белка обменивать одноцепочечную ДНК на гомологичную цепь из дуплекса; за счет этого одноцепочечная нить входит в состав дуплекса.

АТТЕНУАЦИЯ. Регуляция транскрипции на уровне терминации, осуществляемая при экспрессии некоторых бактериальных оперонов.

АТТЕНУАТОР. Терминаторная последовательность, на которой происходит аттенуация.

АУТОГЕННЫЙ КОНТРОЛЬ. Действие генного продукта, который либо подавляет (отрицательный аутогенный контроль), либо активирует (положительный аутогенный контроль) экспрессию своего собственного гена.

АУТОСОМЫ. Все хромосомы, кроме половых; в диплоидной клетке имеется по две копии каждой аутосомы.

АЦЕНТРИЧЕСКИЙ ФРАГМЕНТ. Фрагмент хромосомы, лишенный центромеры и поэтому теряющийся в процессе клеточного деления.

ALU-СЕМЕЙСТВО. Семейство гомологичных диспергированных последовательностей в геноме человека.

ALU-гомологичное семейство. Семейство последовательностей в геномах млекопитающих, родственное Alu-семейству человека.

АР-ЭНДОНУКЛЕАЗЫ. Ферменты, разрезающие ДНК в апуриновых или апириимидиновых участках с образованием 5'-концов.

att-САЙТЫ. Участки фаговой и бактериальной хромосом, комбинация между которыми приводит к интеграции или исключению фага.

α-АМАНИТИН. Бициклический октапептид, получаемый из гриба *Amanita phalloides*; ингибирует транскрипцию определенных эукариотических РНК-полимераз, в особенности РНК-полимеразы II.

БАКТЕРИОФАГИ (ФАГИ). Вирусы, инфицирующие бактерии. **БЕЛКИ, ОСУЩЕСТВЛЯЮЩИЕ ПОЗИТИВНУЮ РЕГУЛЯЦИЮ.** Необходимы для активации транскрипционной единицы.

БЕЛОК-РЕПРЕССОР. Способен связываться с оператором на ДНК или с РНК, предотвращая соответственно транскрипцию или трансляцию.

БИБЛИОТЕКА ГЕНОМА. Набор клонированных фрагментов ДНК, содержащий весь геном.

БЕССМЫСЛЕННАЯ МУТАЦИЯ. Изменение в ДНК, приводящее к замене смыслового кодона, соответствующего какой-либо аминокислоте, на бессмысленный (терминирующий).

БЕССМЫСЛЕННЫЙ КОДОН. Один из трех триплетов, UAG, UAA, UGA, вызывающих терминацию синтеза белка (UAG известен как *amber*-кодон, UAA – как *ochre*-кодон, UGA – как *opal*-кодон).

БИВАЛЕНТ. Структура, содержащая все четыре хроматиды (по две гомологичные пары); образуется в начальной стадии мейоза.

БЛОК ПРИБНОВА. Каноническая последовательность ТА-ТААТГ, находящаяся на расстоянии около 10 п. н. перед стартовой точкой бактериальных генов. Представляет собой часть промотора, отвечающую за связывание РНК-полимеразы.

БЛОТТИНГ ДНК ПО САУЗЕРНУ. Процедура переноса денатурированной ДНК из агарозного геля на нитроцеллюлозный

фильтр для гибридизации с комплементарным нуклеотидом.

БЛОТТИНГ РНК. Перенос РНК из агарозного геля на нитроцеллюлозный фильтр для последующей гибридизации с комплементарной ДНК.

БРЕШЬ (ПРОБЕЛ) В ДНК. Отсутствие одного или нескольких нуклеотидов.

БЫСТРО ЛИЗИРУЮЩИЕ (r) МУТАНТЫ. Мутанты Т-четных фагов, обуславливающие изменение зоны лизиса *E. coli* в конце инфекции.

БЫСТРО РЕАССОЦИИРУЮЩИЙ КОМПОНЕНТ. Последовательности ДНК, которые ренатурируют первыми; содержат высокоповторяющуюся ДНК.

ВВЕДЕНИЕ КОНЦЕВОЙ МЕТКИ. Методика введения радиоактивно меченной группы в один из концов (5' или 3') ДНК.

ВВЕДЕНИЕ МЕТКИ В ДНК ПРИ СМЕЩЕНИИ ОДНОЦЕПЧЕЧНОГО РАЗРЫВА (НИК-ТРАНСЛЯЦИЯ). Способность ДНК-полимеразы I *E. coli* использовать разрыв в одной из цепей ДНК для того, чтобы начиная с него последовательно осуществить деградацию этой цепи, одновременно заменяя ее новосинтеризированной; используется для введения в ДНК радиоактивно меченных нуклеотидов *in vitro*.

ВЕДУЩАЯ ЦЕПЬ. Цепь ДНК, синтезирующаяся непрерывно в 5'—3'-направлении.

ВЕКТОР ДЛЯ КЛОНИРОВАНИЯ. Любая плазида или фаг, в которые может быть встроена чужеродная ДНК с целью клонирования.

ВЕРЕТЕНО. Структура, образующаяся в процессе деления эукариотической клетки; после растворения ядерной оболочки к веретену с помощью микротрубочек прикрепляются хромосомы.

ВИРУЛЕНТНЫЕ ФАГОВЫЕ МУТАНТЫ. Мутанты, неспособные лизогенизировать бактерии, а вызывающие их лизис.

ВИРУС-ПОМОЩНИК. Обладает функциями, отсутствующими у дефектного вируса, компенсирует их при смешанной инфекции.

ВОЗВРАТНОЕ СКРЕЩИВАНИЕ. Скрещивание линии неизвестного генотипа с рецессивной гомозиготой. В таком скрещивании фенотип потомства определяется только хромосомами родителя с неизвестным генотипом.

ВСТАВКИ (ИНСЕРЦИИ). Обнаруживаются благодаря присутствию в ДНК дополнительных пар оснований.

ВЫБОРОЧНАЯ РЕПЛИКАЦИЯ (*saltatory replication*). Побочная амплификация, приводящая к появлению большого количества копий какой-либо последовательности.

ВЫРЕЗАНИЕ. Процесс вырезания фага, эписомы или какой-либо иной последовательности ДНК из хромосомы «хозяина»; связан с их переходом в состояние автономно существующих молекул.

ВИРОЖДЕННОСТЬ ГЕНЕТИЧЕСКОГО КОДА. Соответствие нескольких кодонов одной аминокислоте. Замена в третьем основании кодона не всегда приводит к замене аминокислоты.

ВЫСОКОПОВТОРЯЮЩАЯСЯ ДНК. Компонент тотальной ДНК, который реассоциирует первым; идентичен сателлитной ДНК.

ВЫТЕСНЕНИЕ МЕТКИ (*pulse-chase*). Клетки непродолжительное время инкубируют с радиоактивно меченым предшественником (участвующим в каких-либо реакциях или же являющимся предшественником ряда макромолекул); затем исследуют судьбу метки, продолжая инкубацию с немеченым предшественником.

ВЫТЕСНЕНИЕ ЦЕПИ. Характерный для ряда вирусов способ репликации ДНК, при котором синтезирующаяся цепь ДНК вытесняет прежнюю (гомологичную) цепь двухцепочечной молекулы.

V-ГЕН. Последовательность, кодирующая основную часть варибельной (N-концевой) области иммуноглобулиновой цепи.

ГАМЕТА. Половая клетка, содержащая гаплоидный набор хромосом.

ГАПЛОИДНЫЙ НАБОР ХРОМОСОМ. Содержит по одной

копии каждой аутосомы и одну половую хромосому; гаплоидное число хромосом (n) является характеристикой гамет.

ГАПТЕН. Небольшая молекула, присоединенная к инертному белковому носителю и выполняющая функцию антигена.

ГЕН (ЦИСТРОН). Фрагмент ДНК, участвующий в образовании полипептидной цепи; в его состав входят участки, расположенные перед кодирующей последовательностью и после нее (лидерная и концевая области), а также инсерционные последовательности (интроны).

ГЕННАЯ КОНВЕРСИЯ. Изменение одной из цепей гетеродуплексной ДНК, приводящее к комплементарности с другой цепью в тех положениях, где есть неспаренные основания.

ГЕНОМНЫЕ (ХРОМОСОМНЫЕ) КЛОНЫ ДНК. Последовательности геномной ДНК, клонированные в составе вектора.

ГЕНОТИП. Совокупность генов организма.

ГЕТЕРОГАМЕТНЫЙ ПОЛ. Характеризуется хромосомным набором $2A + XY$.

ГЕТЕРОГЕННАЯ ЯДЕРНАЯ (ГЯ) РНК. Представлена продуктами транскрипции ядерных генов, осуществляемой РНК-полимеразой II; эти транскрипты гетерогенны по размерам и обладают низкой стабильностью.

ГЕТЕРОДУПЛЕКСНАЯ (ГИБРИДНАЯ) ДНК. Образуется при спаривании одноцепочечных комплементарных ДНК из разных родительских дуплексов; возникает в процессе генетической рекомбинации.

ГЕТЕРОЗИГОТА. Особь с различными аллелями в каком-либо определенном локусе.

ГЕТЕРОМУЛЬТИМЕРНЫЕ БЕЛКИ. Состоят из различных субъединиц (кодируемых разными генами).

ГЕТЕРОХРОМАТИН. Генетически неактивные участки хромосом; постоянно находятся в конденсированном состоянии.

ГИБРИДИЗАЦИЯ. Процесс взаимодействия комплементарных цепей РНК и ДНК, образующих гибрид РНК—ДНК.

ГИБРИДИЗАЦИЯ В ЖИДКОЙ СРЕДЕ. Реакция между комплементарными цепями нуклеиновой кислоты, происходящая в растворе.

ГИБРИДИЗАЦИЯ В УСЛОВИЯХ НАСЫЩЕНИЯ. Гибридизация, проводимая в присутствии большого избытка одного из реагирующих компонентов. При этом все комплементарные последовательности другого компонента переходят в двухцепочечную форму.

ГИБРИДИЗАЦИЯ НА ФИЛЬТРАХ. Осуществляется путем инкубации препаратов денатурированной ДНК, иммобилизованной на нитроцеллюлозном фильтре, с радиоактивно меченной РНК или ДНК.

ГИБРИДИЗАЦИЯ С КОЛОНИЯМИ. Метод гибридизации *in situ*, используемый для выявления бактериальных колоний, содержащих рекомбинантные плазмиды, которые несут фрагменты ДНК, гомологичные определенным последовательностям.

ГИБРИДИЗАЦИЯ IN SITU. Реакция гибридизации между денатурированной ДНК клеток, раздавленных на предметном стекле микроскопа и меченной радиоактивными изотопами одноцепочечной РНК или ДНК; гибридизация обнаруживается благодаря последующей радиоавтографии.

ГИБРИДНЫЙ ДИСГЕНЕЗ. Неспособность определенных линий *D. melanogaster* скрещиваться между собой: образующиеся гибридные особи стерильны (хотя в остальном они могут быть фенотипически нормальными).

ГИБРИДОМА. Клеточная линия, образующаяся при слиянии клеток миеломы и лимфоцита; такая линия постоянно экспрессирует иммуноглобулины обоих родителей.

ГИПЕРХРОМИЗМ. Увеличение оптической плотности при денатурации ДНК.

ГИПЕРЧУВСТВИТЕЛЬНЫЕ К ДЕЙСТВИЮ ДНКазы I САЙТЫ. Небольшая область хроматина, определяемая по ее повышенной чувствительности к ДНКазе I и другим нуклеазам; возможно, не содержит нуклеосом.

ГИПОТЕЗА КАЧАНИЙ. Объясняет способность тРНК узнавать более чем один кодон благодаря неканоническому (отличному от G—C, A—T) спариванию с третьим основанием кодона.

ГИРАЗА. Топоизомераза II типа из *E. coli*. Фермент способен вносить отрицательные супервитки в ДНК.

ГИСТОНЫ. Эволюционно консервативные белки эукариот, связывающие ДНК; участвуют в формировании нуклеосомы, основной структурной субъединицы хроматина.

ГЛАВНЫЙ ЛОКУС ГИСТОСОВМЕСТИМОСТИ. Протяженная область хромосомы, содержащая большой кластер генов, кодирующих трансплантационные антигены и другие белки, обнаруживаемые на поверхности лимфоцитов.

ГОМОГАМЕТНЫЙ ПОЛ. Характеризуется хромосомным набором 2A + XX.

ГОМОЗИГОТА. Особь с одинаковыми аллелями в соответствующем локусе гомологичных хромосом.

ГОМОЛОГИ. Хромосомы, имеющие одинаковые генетические локусы; диплоидная клетка обладает двумя копиями каждого гомолога, по одной от каждого родителя.

ГОМОМУЛЬТИМЕРНЫЙ БЕЛОК. Состоит из идентичных субъединиц.

ГОРЯЧАЯ ТОЧКА. Участок, в котором частота возникновения мутаций (или рекомбинаций) очень велика.

ГРУППА КОМПЛЕМЕНТАЦИИ. Серия мутаций, относящихся к одной группе, не способных комплементировать друг друга в *транс*-конфигурации; они составляют генетическую единицу (цистрон), которую, может быть, лучше было бы назвать группой некомплементации.

ГРУППА СОВМЕСТИМОСТИ. Плазмиды, которые не могут существовать в одной и той же бактериальной клетке.

ГРУППА СЦЕПЛЕНИЯ. Включает все локусы, которые могут быть объединены (прямо или косвенно) на основании сцепленности; термин эквивалентен хромосоме.

G-ОКРАШИВАНИЕ ХРОМОСОМ. Метод специфического окрашивания метафазных хромосом; позволяет идентифицировать отдельные хромосомы гаплоидного набора.

G1. Период клеточного цикла между последним митозом и началом репликации ДНК.

G2. Период клеточного цикла после окончания репликации ДНК и до начала следующего митоза.

ДВУНАПРАВЛЕННАЯ РЕПЛИКАЦИЯ. Репликация, при которой две репликационные вилки движутся в противоположных направлениях от общего старта.

ДЕЛЕЦИИ. Удаление последовательности ДНК; области, фланкирующие удаленные участки, соединяются.

ДЕНАТУРАЦИЯ БЕЛКА. Переход из физиологической конформации в другую (неактивную) конформацию.

ДЕНАТУРАЦИЯ ДНК ИЛИ РНК. Переход этих молекул из двухцепочечной формы в одноцепочечную; разделение цепей наиболее часто достигается нагреванием.

ДЕРЕПРЕССИРОВАННОЕ СОСТОЯНИЕ. То же самое, что и индуцированное, если употребляется для описания нормального состояния гена. При описании эффекта мутаций имеет то же значение, что и конститутивное.

ДИВЕРГЕНЦИЯ. Процент различий в нуклеотидных последовательностях двух ДНК или в аминокислотных последовательностях двух белков.

ДИСКИ. Участки *политенных хромосом*, имеющие вид плотных образований. Содержат большое количество ДНК; представлены участками, задерживающими краситель при определенных химических обработках.

ДИПЛОИДНЫЙ НАБОР ХРОМОСОМ. Содержит по две копии каждой аутосомы и две половые хромосомы.

ДИЦЕНТРИЧЕСКАЯ ХРОМОСОМА. Образуется в результате объединения двух хромосомных фрагментов, каждый из которых содержит центромеру. Такая хромосома нестабильна и может быть разорвана в процессе митоза.

ДНК-ЗАВИСИМАЯ ГИБРИДИЗАЦИЯ. Гибридикация ДНК с РНК в условиях избытка ДНК.

ДОБРОВОЛЬНЫЕ ИНДУКТОРЫ. Напоминают истинные индукторы транскрипции, но не являются субстратами индуцируемых энзимов.

ДОМЕН В МОЛЕКУЛЕ БЕЛКА. Участок аминокислотной последовательности, связанный с определенной функцией.

ДОМИНАНТНЫЙ АЛЛЕЛЬ. Определяет фенотип гетерозиготы (в отличие от рецессивного аллеля).

ДОМЕН В ХРОМОСОМЕ. Может обозначать либо область, внутри которой суперспирализация происходит независимо от других участков, или область, включающую в себя экспрессируемый ген и обладающую повышенной чувствительностью к ДНКазе I.

D-ПЕТЛЯ. Область внутри митохондриальной ДНК, в которой небольшой участок РНК взаимодействует с одной из цепей ДНК, вытесняя исходную комплементарную цепь. Этот же термин используется при описании события, катализируемого RecA-белком, которое заключается в замене одной цепи ДНК другой одноцепочечной ДНК, захваченной извне.

dna-МУТАНТЫ БАКТЕРИЙ. Не способны синтезировать ДНК при 42°C, но осуществляют этот процесс при 37°C.

ЗАКРЫТАЯ РАМКА СЧИТЫВАНИЯ. Содержит кодоны преждевременной терминации, не позволяющие мРНК транслироваться в белок.

ЗАТРАВКА. Короткая последовательность (часто это РНК), комплементарно взаимодействующая с одной из цепей ДНК; образует свободный 3'-ОН-конец, используя который ДНК-полимераза начинает синтез дезоксирибонуклеотидной цепи.

ЗИГОТА. Оплодотворенная яйцеклетка; образуется в результате слияния двух гамет.

ИЗОАКЦЕПТОРНЫЕ tРНК. Молекулы tРНК, соответствующие одной и той же аминокислоте.

ИЗОТИП. Группа близкородственных иммуноглобулиновых цепей.

ИММУНИТЕТ ПРИ ЛИЗОГЕНИИ. Способность профага предотвращать развитие в бактерии другого фага того же типа.

ИММУНИТЕТ. Свойство профага препятствовать инфекции бактериальной клетки фагом того же типа; свойство плазмиды предотвращать размножение другой плазмиды того же типа. Эти свойства обусловлены различными механизмами. Кроме того, описана транспозиционная иммунность. Этот термин означает невозможность встраивания транспозона в ДНК, если в ней уже присутствует транспозон такого же типа.

ИНВЕРТИРОВАННЫЕ ПОВТОРЫ. Две копии одной и той же последовательности ДНК в составе одной молекулы, находящиеся в противоположной ориентации. Прилежащие друг к другу инвертированные повторы образуют палиндром.

ИНВЕРТИРОВАННЫЕ ТЕРМИНАЛЬНЫЕ ПОВТОРЫ. Близкие или идентичные короткие последовательности ДНК, находящиеся на концах некоторых транспозонов в противоположной ориентации.

ИНГИБИРОВАНИЕ КОНЕЧНЫМ ПРОДУКТОМ. Способность конечного продукта метаболического пути ингибировать активность фермента, катализирующего одну из ранних реакций этого процесса.

ИНДУКТОР. Небольшая молекула, включающая транскрипцию гена за счет связывания с регуляторным белком.

ИНДУКЦИЯ. Свойство клеток (бактериальных или дрожжевых) синтезировать определенные ферменты только при наличии соответствующих субстратов; применительно к экспрессии генов термин означает включение транскрипции в результате взаимодействия индуктора с регуляторным белком.

ИНДУКЦИЯ ПРОФАГА. Вырезание профага из генома клетки-хозяина, разрушение лизогенного репрессора и начало литического (инфекционного) цикла.

ИНДУЦИРОВАННЫЕ МУТАЦИИ. Возникают под действием мутагенного фактора.

ИНТЕГРАЦИЯ. Внедрение вирусной или иной последовательности ДНК в геном клетки-хозяина, приводящее к ковалентному соединению с хозяйской последовательностью.

ИНТЕРФАЗА. Фаза клеточного цикла между митотическими делениями клетки; подразделяется на G1, S и G2.

ИНТРОН. Транскрибируемый участок ДНК, который удаляется из состава транскрипта при сплайсинге; в результате последовательности, находящиеся по обе стороны от интрона (экзоны), объединяются.

IS. Это сокращение используется для обозначения инсерционной (вставочной) последовательности небольшого бакте-

риального транспозона, кодирующего только генетические функции, необходимые для транспозиции.

КАНОНИЧЕСКАЯ (УСРЕДНЕННАЯ) ПОСЛЕДОВАТЕЛЬНОСТЬ. Некая идеальная последовательность, каждая позиция в которой представлена нуклеотидом, наиболее часто встречающимся в реальных последовательностях.

КАПСИД. Внешняя белковая оболочка вирусной частицы.

КАТАБОЛИТНАЯ РЕПРЕССИЯ. Ослабление экспрессии многих бактериальных оперонов, происходящее при добавлении глюкозы; вызывается уменьшением уровня циклической АМР в клетке и инактивацией вследствие этого регуляторного САР-белка.

КАТЯЩЕЕСЯ КОЛЬЦО. Способ репликации, при котором репликационная вилка совершает множество оборотов на циркулярной матрице; синтезирующаяся в каждом цикле цепь ДНК вытесняет цепь, синтезированную в предыдущем цикле, образуя хвост, состоящий из линейного набора последовательностей, комплементарных одноцепочечному матричному кольцу.

кДНК. Одноцепочечная ДНК, синтезированная *in vitro* путем обратной транскрипции; комплементарна РНК.

кДНК-КЛОН. Двухцепочечная ДНК, находящаяся в составе клонирующего вектора и соответствующая последовательности РНК.

КИНЕТИЧЕСКАЯ СЛОЖНОСТЬ. Сложность последовательностей некоторой фракции ДНК, определяемая по кинетике реассоциации.

КИНЕТОХОР. Особая структура в хромосоме, к которой прикреплены микротрубочки или митотическое веретено.

КЛЕТОЧНЫЙ ЦИКЛ. Последовательность событий, происходящих в клетке между двумя делениями.

КЛОН. Большое число клеток или молекул, идентичных одной родоначальной клетке или молекуле.

КОДИРУЮЩАЯ ЦЕПЬ. Та цепь ДНК, последовательность которой идентична мРНК.

КОДОМИНАНТНЫЕ АЛЛЕЛИ. Аллели, совместно проявляющиеся в гетерозиготе. Ни один не доминирует над другим.

КОДОН. Триплет нуклеотидов, соответствующий определенной аминокислоте или терминирующему сигналу.

КОИНТЕГРАТ. Структура, образуемая при слиянии двух репликонов, только один из которых исходно содержит транспозон; в коинтеграте две копии транспозона расположены у границ соединения репликонов и ориентированы в одном направлении.

КОКОНВЕРСИЯ. Одновременная коррекция двух сайтов в процессе конверсии гена.

КОЛЬЦО БАЛЬБИНИ. Гигантский пуф на политенной хромосоме.

КОМПЛЕМЕНТАРНАЯ ЦЕПЬ. Одна из цепей ДНК, используемая в качестве матрицы для синтеза РНК и комплементарная ей.

КОМПЛЕМЕНТАЦИЯ. Способность независимых (неаллельных) генов, находящихся в *транс*-конфигурации, детерминировать диффундирующие продукты, которые обуславливают появление дикого фенотипа.

КОНКАТЕМЕР ДНК. Образован из нескольких тандемно повторяющихся единиц генома.

КОНКАТЕНАТЫ. Кольцевые молекулы ДНК, соединенные между собой подобно кольцам в цепи.

КОНСЕРВАТИВНАЯ РЕКОМБИНАЦИЯ. Рекомбинация путем разрыва и последующего воссоединения существующих цепей ДНК без дополнительного синтеза новых участков.

КОНСТИТУТИВНЫЕ ГЕНЫ. Гены, экспрессия которых зависит только от взаимодействия РНК-полимеразы с промотором и не регулируется дополнительными факторами; иногда так называют гены, слабо экспрессирующиеся во всех типах клеток.

КОНСТИТУТИВНЫЕ МУТАЦИИ. Приводят к тому, что гены, экспрессия которых регулируется, начинают экспрессироваться без какого-либо контроля.

КОНСТИТУТИВНЫЙ ГЕТЕРОХРОМАТИН. Конденсированное (биологически неактивное) состояние хроматина, содержащего неэкспрессирующиеся последовательности, обычно сателлитную ДНК.

КОНТРОЛИРУЮЩИЕ ЭЛЕМЕНТЫ. Подвижные элементы генома кукурузы, первоначально идентифицированные по их генетическим свойствам.

3'-КОНЦЕВАЯ ПОСЛЕДОВАТЕЛЬНОСТЬ. Нетранслируемая последовательность на 3'-конце мРНК, следующая за терминирующим кодоном.

КООРДИНИРОВАННАЯ РЕГУЛЯЦИЯ. Означает общий контроль экспрессии группы генов.

КОРДИЦЕПИН. 3'-Дезоксиаденозин, ингибитор полиаденилирования РНК.

КОРЕПРЕССОР. Маленькая молекула, которая включает механизм репрессии транскрипции, связываясь с регуляторным белком.

КОР-ПОСЛЕДОВАТЕЛЬНОСТЬ ДНК. 146 нуклеотидных пар ДНК, находящихся в составе кор-частицы.

КОРРЕКЦИЯ. Любой механизм, исправляющий ошибки в процессе синтеза белка или нуклеиновой кислоты. Характерное свойство этого процесса – проверка правильности каждого компонента после его включения в полимерную цепь.

КОСМИДЫ. Плазмиды, содержащие встроенный *cos*-участок фага лямбда, благодаря чему плазмидная ДНК может быть упакована *in vitro* в оболочку фага.

КОР-ЧАСТИЦА. Продукт частичной нуклеазной деградации нуклеосомы, содержащий гистоновый октамер и 146 нуклеотидных пар ДНК; его структура сходна со структурой нуклеосомы.

КРЕСТООБРАЗНАЯ СТРУКТУРА. Структура ДНК, образуемая инвертированными повторами, при условии что комплементарное спаривание происходит между повторяющимися последовательностями, расположенными в одной цепи ДНК (в отличие от обычного комплементарного взаимодействия со своим партнером в другой цепи дуплекса).

КРИПТИЧЕСКИЙ САТЕЛЛИТ. Последовательность сателлитной ДНК, которая не обнаруживается в виде обособленного пика в градиенте плотности, т.е. этот сателлит не отделяется от основной полосы ДНК.

КРОССИНГОВЕР. Обмен материалом между гомологичными хромосомами, происходящий в процессе мейоза и лежащий в основе генетической рекомбинации.

КЭП. Структура на 5'-конце эукариотических мРНК; образуется после транскрипции за счет присоединения 5'-конца гуанинового нуклеотида к концевому основанию мРНК. Эта структура может быть метилирована по крайней мере по той молекуле гуанина, которая присоединилась. Кэп имеет следующее строение ${}^7\text{MeG}^5\text{ppp}^5\text{Np}..$

САР-БЕЛОК (БАК). Активируется циклической АМР и осуществляет позитивную регуляцию; необходим для инициации транскрипции РНК-полимеразой некоторых катаболит-чувствительных оперонов *E. coli*.

СААТ. Участок консервативной последовательности, расположенной примерно на расстоянии 75 пар оснований перед стартовой точкой в транскрипционных единицах эукариот; возможно, принимает участие в связывании РНК-полимеразы II.

C_0t . Основной параметр реакции реассоциации, вычисляемый как произведение концентрации ДНК на время инкубации в реакционной смеси.

$C_0t_{1/2}$. Значение C_0t , характеризующее завершение реакции наполовину. Это значение прямо пропорционально общей длине уникальных последовательностей реассоциирующей ДНК.

ЛЕВАЯ ГРАНИЦА СПЛАЙСИНГА. Граница между правым концом экзона и левым концом интрона.

ЛЕТАЛЬНЫЙ ЛОКУС. Любой ген, в котором может быть получена летальная мутация.

ЛИГИРОВАНИЕ. Образование фосфодиэфирной связи между двумя основаниями одной цепи ДНК, разделенными разрывом. (Этот термин употребляют также в случае соединения тупых концов и при образовании связи в РНК.)

ЛИГИРОВАНИЕ ПО ТУПЫМ КОНЦАМ. Реакция соединения двух молекул двухцепочечной ДНК по концам, ни один из которых не выступает над другим.

ЛИДЕР. Нетранслируемая последовательность, находящаяся на 5'-конце мРНК и предшествующая иницирующему кодону.

ЛИЗИС. Разрушение бактериальных клеток в конце фагового инфекционного цикла, выход зрелых фаговых частиц.

ЛИЗОГЕНИЯ. Способность фага существовать в бактерии в виде профага, являющегося компонентом бактериального генома.

ЛИЗОГЕННАЯ БАКТЕРИЯ. Содержит в своем геноме профаг в репрессированном состоянии.

ЛИЗОГЕННЫЙ РЕПРЕССОР. Белок, блокирующий индукцию профага.

ЛИНКЕР. Синтетический, короткий двухцепочный олигонуклеотид, содержащий сайты узнавания для ряда рестриционных эндонуклеаз; может быть присоединен к концам фрагмента ДНК, полученного с помощью какой-либо другой рестрицирующей эндонуклеазы, в процессе реконструирования рекомбинантной ДНК.

ЛИНКЕРНАЯ ДНК. ДНК нуклеосомы, выходящая за пределы кор-частицы (минимальной нуклеосомы) длиной 146 п.н.

ЛИПКИЕ КОНЦЫ. Комплементарные одноцепочечные участки ДНК, выступающие на противоположных концах двухцепочечной молекулы; возникают в результате ступенчатых разрывов в двухцепочечных молекулах ДНК.

ЛИТИЧЕСКАЯ ИНФЕКЦИЯ. Фаговая инфекция, заканчивающаяся разрушением бактериальных клеток и выходом фагового потомства.

ЛОКУС. Место в хромосоме, в котором картируется ген, отвечающий за определенный признак; locus может быть представлен любым аллелем данного гена.

LEAKY-МУТАЦИИ. Мутации, обуславливающие сохранение остаточного уровня экспрессии.

LTR (*long terminal repeat*). Длинные концевые повторы, находящиеся в прямой ориентации на обоих концах ДНК ретровируса.

МАЖОРНЫЕ мРНК. мРНК, представленные в клетке большим числом копий.

МАРКЕР (ДНК). Фрагмент ДНК известного размера, используемый для калибровки фрагментов в электрофоретическом геле.

МАРКЕР (ГЕНЕТИЧЕСКИЙ). Любой аллель, используемый в эксперименте.

МАТЕРИНСКОЕ НАСЛЕДОВАНИЕ. Предпочтительное проявление в потомстве генетических маркеров матери.

МЕДЛЕННЫЙ КОМПОНЕНТ В РЕАКЦИИ РЕАССОЦИАЦИИ. Компонент, который реассоциирует последним; обычно состоит из неповторяющейся ДНК.

МЕЖАЛЛЕЛЬНАЯ КОМПЛЕМЕНТАЦИЯ. Изменение свойств гетеромультимерного белка в результате взаимодействия субъединиц, кодируемых двумя различными мутантными аллелями; гибридный белок может быть более или менее активен по сравнению с белком, состоящим из субъединиц только одного типа.

МЕЖДИСКИ. Менее плотные области политенных хромосом между дисками.

МЕЖЦИСТРОННАЯ ОБЛАСТЬ. Расстояние между терминирующим кодоном одного гена и иницирующим кодоном следующего гена в полицистронной транскрипционной единице.

МЕЙОЗ. Два последовательных деления клетки (I и II мейотические деления), в результате которых исходное число хромосом $4n$ уменьшается до $1n$ в каждой из четырех образовавшихся клеток. Эти клетки созревают и превращаются в гаметы (сперматозоиды или яйцеклетки).

МЕТОД ОТПЕЧАТКОВ. Используется для обнаружения участка на молекуле ДНК, который связывается с некоторым белком. Этот прием основывается на том, что белки защищают последовательность ДНК в области связывания от воздействия нуклеаз.

МЕТОД ПОЛУЧЕНИЯ УНИКАЛЬНОЙ ФРАКЦИИ ДНК. Стоит в получении неповторяющейся фракции ДНК путем гибридизации ДНК с РНК и последующей проверки гибридизованной ДНК в реакции реассоциации.

МЕТОД СЛУЧАЙНОГО РАЗРЕЗАНИЯ ГЕНОМА РЕСТРИКТА-

ЗАМИ (*shotgun experiment*). Клонирование всего генома в виде случайного образованных фрагментов.

МИГРАЦИЯ ВЕТВИ. Способность одноцепочечной ДНК, частично спаренной с комплементарной цепью в двухцепочечной молекуле ДНК, продолжать комплементарное взаимодействие и вытеснять гомологичную цепь из дуплекса.

МИЕЛОМА. Линия опухолевых клеток, происшедшая из лимфоцита, обычно продуцирует один тип иммуноглобулинов.

МИКРОКОККОВАЯ НУКЛЕАЗА. Эндонуклеаза, расщепляющая ДНК; при обработке этим ферментом хроматина гидролиз ДНК происходит преимущественно между нуклеосомами.

МИНИ-ХРОМОСОМА. Кольцевая ДНК (например, вирусов SV40 или полиомы) в нуклеосомной форме.

МИТОЗ. Деление эукариотической соматической клетки.

МНОЖЕСТВЕННОСТЬ мРНК. Среднее число молекул мРНК какого-либо одного вида на клетку.

МОДИФИКАЦИЯ ДНК ИЛИ РНК. Изменения нуклеотидов после их первоначального включения в полинуклеотидную цепь.

МОДИФИЦИРОВАННЫЕ ОСНОВАНИЯ. Все основания, за исключением четырех канонических, из которых синтезируется ДНК (Т, С, А, G) или РНК (U, С, А, G); модифицированные основания появляются в результате изменений, возникающих в нуклеиновой кислоте после завершения ее синтеза.

МОЗАИЧНЫЙ ФЕНОТИП. Образуется в результате мутации в соматической клетке в ходе развития.

МОЛЧАЩИЕ ГЕННЫЕ САЙТЫ. Такие точки в гене, мутации в которых не вызывают изменения продукта.

МОЛЧАЩИЕ МУТАЦИИ. Не изменяют продукт, кодируемого геном.

МОНОЦИСТРОННЫЕ мРНК. Кодируют один белок.

МУЛЬТИКОПИЙНЫЕ ПЛАЗМИДЫ. Присутствуют в бактериальных клетках в количестве более одной на хромосому.

МУЛЬТИМЕРНЫЕ БЕЛКИ. Состоят более чем из одной субъединицы.

МУТАГЕНЫ. Факторы, увеличивающие частоту возникновения мутаций, вызывая изменения в ДНК.

МУТАНТЫ С БЫСТРОЙ ОСТАНОВКОЙ СИНТЕЗА ДНК. *dna*-Мутанты *E. coli*, прекращающие репликацию сразу же после повышения температуры до 42°C.

МУТАНТЫ С ЗАМЕДЛЕННОЙ ОСТАНОВКОЙ СИНТЕЗА ДНК. Мутанты *dna E. coli*, способные завершить начавшийся цикл бактериальной репликации, но неспособные при 42°C иницировать следующий раунд.

МУТАНТЫ С ОСЛАБЛЕННЫМ КОНТРОЛЕМ. Штаммы *E. coli*, неспособные к строгому ответу при голодании по аминокислотам (или другим питательным веществам).

МУТАЦИЯ. Любое изменение в последовательности ДНК.

МУТАЦИЯ СДВИГА РАМКИ. Деления или вставки, размеры которых не кратны трем основаниям, приводят к изменению рамки считывания при трансляции триплетов в белок.

МУТАЦИИ, УСИЛИВАЮЩИЕ ПРОМОТОР. Увеличивают частоты инициации транскрипции.

НЕИНДУЦИРУЕМЫЕ МУТАЦИИ. Мутации, которые не могут быть индуцированы.

НЕЙТРАЛЬНЫЕ ЗАМЕНЫ В БЕЛКЕ. Замены аминокислот, не сказывающиеся на активности белка.

НЕПЕРМИССИВНЫЕ УСЛОВИЯ. Условия, не позволяющие выживать условно летальным мутантам.

НЕПОВТОРЯЮЩАЯСЯ ДНК. Характеризуется кинетикой реассоциации, свойственной уникальным последовательностям.

НЕПОЛНОЕ ДОМИНИРОВАНИЕ. (См. кодоминантность.)

НЕПРЯМОЕ КОНЦЕВОЕ МЕТЧЕНИЕ (*indirect end labeling*). Метод изучения ДНК, при котором производится разрез в определенном участке молекулы и выявляются все фрагменты, содержащие последовательность, располагающуюся только с одной стороны от данного разреза; с помощью такого подхода определяют расстояния от данного разрыва в молекуле ДНК до следующего.

НЕРАВНЫЙ КРОССИНГОВЕР. Рекомбинационное событие, при котором точки рекомбинации находятся не в идентичных локусах двух родительских молекул ДНК.

НЕРАСХОЖДЕНИЕ. Неспособность хроматид (дублированных хромосом) расходиться к противоположным полюсам во время митоза или мейоза.

НЕСОВМЕСТИМОСТЬ. Невозможность совместного существования в одной клетке определенных бактериальных плазмид.

НЕТРАНСКРИБИРУЕМЫЙ СПЕЙСЕР. Область между транскрибируемыми единицами в геномном кластере.

НУКЛЕОТИД. Компактное образование у бактерий, содержащее ДНК.

НУКЛЕОСОМА. Основная структурная единица хроматина, состоящая из ~ 200 нуклеотидных пар ДНК и октамера гистоновых белков.

НУЛЕВАЯ ТОЧКА РЕАССОЦИАЦИИ ДНК. Дуплексы ДНК, существующие в самом начале реакции реассоциации.

ОБМЕН ЦЕПИ. Реакция, в которой одна из цепей двухцепочечной ДНК освобождается из дуплекса, образуя пару с комплементарной цепью другой молекулы, заменяя своего гомолога во втором дуплексе.

ОБРАТНАЯ МУТАЦИЯ. Мутация, которая восстанавливает дикий фенотип, снимая действие мутации, инактивировавшей ген.

ОБРАТНАЯ ТРАНСКРИПЦИЯ. Синтез ДНК на матрице РНК; осуществляется ферментом обратной транскриптазой.

ОБЯЗАТЕЛЬНЫЕ (ГЕНЫ «ДОМАШНЕГО ХОЗЯЙСТВА»). Гены, кодирующие жизненно важные функции; экспрессируются в клетках всех типов.

ОДНОКОПИЙНЫЕ ПЛАЗМИДЫ. Плазмиды, присутствующие в бактериальной клетке в виде единственной копии.

ОДНОНАПРАВЛЕННАЯ РЕПЛИКАЦИЯ. Единственная репликационная вилка движется от определенной точки, называемой местом начала репликации.

ОПЕРАТОР. Участок ДНК, связываясь с которым белок-репрессор предотвращает инициацию транскрипции на прилежащем промоторе.

ОПЕРОН. Единица транскрипции и регуляции у бактерий, состоящая из структурных генов, регуляторного гена (генов) и контролирующих элементов, узнаваемых продуктами регуляторного гена.

ОСНОВНАЯ ПОЛОСА. Широкая полоса, образующаяся в градиенте плотности и состоящая из геномной ДНК, за исключением некоторых видимых сателлитных ДНК, формирующих отдельные полосы.

ОСТАНОВКА ТРАНСЛЯЦИИ В РЕЗУЛЬТАТЕ ОБРАЗОВАНИЯ ГИБРИДА (Hybrid-arrested translation). Метод, позволяющий идентифицировать кДНК, соответствующую определенной мРНК. Основывается на том, что комплементарное взаимодействие кДНК с РНК ингибирует трансляцию этой РНК.

ОТКРЫТАЯ РАМКА СЧИТЫВАНИЯ. Состоит из ряда триплетов, кодирующих аминокислоты; не содержит каких-либо терминирующих кодонов; эта последовательность потенциально может транслироваться в белок.

ОТРИЦАТЕЛЬНАЯ КОМПЛЕМЕНТАЦИЯ. Подавление активности субъединицы дикого типа в мультимерном белке мутантной субъединицей в результате межаллельной комплементации.

ОТРИЦАТЕЛЬНАЯ СУПЕРСПИРАЛИЗАЦИЯ. Введение в двухцепочечную ковалентно замкнутую ДНК супервитков, направление которых противоположно направлению витков цепей молекулы.

ОТСТАЮЩАЯ ЦЕПЬ. Должна удлиняться в 3'-5'-направлении, поэтому синтезируется прерывисто в виде коротких фрагментов (5'-3'), которые затем ковалентно соединяются.

ОХРА-КОДОН. Триплет UAA, один из трех бессмысленных кодонов, вызывающих терминацию синтеза белка.

ОХРА-МУТАЦИЯ. Изменение в ДНК, приводящее к появлению UAA-кодона в сайте, первоначально занятом другим кодоном.

ОХРА-СУПРЕССОР. Ген, кодирующий мутантную тРНК, способную узнавать UAA-кодон, благодаря чему синтез белка может быть продолжен.

ПАЛИНДРОМ. Последовательность ДНК, которая остается неизменной, если на одной из цепей ДНК ее читать слева направо, а на другой справа налево; состоит из прилежащих друг к другу инвертированных поворотов.

ПАПОВАВИРУСЫ. Вирусы, инфицирующие животных, характеризующиеся небольшим геномом; к ним относятся вирус SV40 и вирус полиомы.

ПАРА ОСНОВАНИЙ (А-Т ИЛИ С-G). Являются партнерами в двойной цепи ДНК. Другие типы пар встречаются в определенных условиях в РНК.

ПЕРВИЧНЫЙ ТРАНСКРИПТ. Первоначально синтезированная немодифицированная молекула РНК, соответствующая транскрипционной единице.

ПЕРЕКЛЮЧЕНИЕ КЛАССА. Изменение экспрессии генов С-области иммуноглобулиновой цепи в процессе дифференцировки лимфоцита.

ПЕРИОДИЧНОСТЬ ДНК. Число пар оснований, приходящихся на один виток спиральной молекулы ДНК.

ПЕРМИССИВНЫЕ УСЛОВИЯ. Позволяющие выживать условно летальным мутантам.

ПЕТИТ-ШТАММЫ. Мутанты дрожжей с нарушенной функцией митохондрий.

ПЕТЛЯ. Одноцепочечная область на конце шпильки, образующейся в РНК (или одноцепочечной ДНК); соответствует последовательности между инвертированными повторами в двухцепочечной ДНК.

ПЛАВЛЕНИЕ ДНК. Денатурация ДНК.

ПЛАЗМИДА. Кольцевая внехромосомная ДНК, способная к автономной репликации.

ПЛЕЙОТРОПИЯ. Способность гена оказывать влияние более чем на один признак (которые непосредственно не связаны между собой).

ПОБОЧНЫЙ СИНТЕЗ ДНК. Любой синтез ДНК в эукариотической клетке, происходящий вне S-фазы.

ПОВТОРЯЮЩАЯСЯ ДНК. Ведет себя в реакции реассоциации так, как если бы большое количество копий последовательностей (похожих или идентичных) данной ДНК присутствовало в образце, что позволяет любой паре комплементарных последовательностей реассоциировать.

ПОВТОРЯЮЩАЯСЯ ЕДИНИЦА. Длина повторяющейся последовательности тандемного кластера. Ее рестрикционная карта выглядит аналогично рестрикционной карте кольцевой молекулы ДНК.

ПОЗДНЯЯ СТАДИЯ. Период инфекционного цикла фага после начала репликации ДНК.

ПОЛИАДЕНИЛИРОВАНИЕ. Присоединение последовательности полиаденилиновой кислоты к 3'-концу эукариотической РНК после завершения ее синтеза.

ПОЛИМОРФИЗМ. Одновременное существование в популяции нескольких аллельных вариантов какого-либо гена; обнаруживается либо по различию в фенотипах, соответствующих разным аллелям, либо по характеру рестрикции ДНК, несущей разные аллели.

ПОЛИПЛОИДНАЯ КЛЕТКА. Содержит более чем два гаплоидных набора хромосом.

ПОЛИПРОТЕИН. Продукт трансляции; расщепляется на несколько отдельных белков.

ПОЛИСОМА (ПОЛИРИБОСОМА). мРНК, ассоциированная с несколькими рибосомами, участвующими в трансляции.

ПОЛИТЕННЫЕ ХРОМОСОМЫ. Образуются в результате последовательных репликаций хромосом без их последующего расхождения.

ПОЛОВЫЕ ХРОМОСОМЫ. Хромосомы, различающиеся у двух полов. Обычно такие хромосомы обозначаются как X и Y (или W и Z), один пол имеет XX- или WW-набор, другой - XY- или WZ-набор.

ПОЛУКОНСЕРВАТИВНАЯ РЕПЛИКАЦИЯ. Осуществляется за счет разделения цепей исходной двухцепочечной молекулы и последующего использования каждой из них в качестве матрицы для синтеза комплементарной цепи.

ПОЛУПРЕРЫВИСТАЯ РЕПЛИКАЦИЯ. Одна из цепей синтезируется непрерывной, тогда как другая - в виде фрагментов.

ПОЛЯРНОСТЬ. Влияние мутации, возникающей в каком-либо гене, на экспрессию (на транскрипционном или трансляционном уровне) генов, расположенных в той же транскрипционной единице.

ПОСЛЕДОВАТЕЛЬНОСТЬ ШАЙНА-ДАЛЬГАРНО. Вся или только часть полипуриновой последовательности AGGAGG, находящейся на мРНК непосредственно перед иницирующим AUG-кодонам, комплементарна последовательности на 3'-конце 16S-рРНК; принимает участие в связывании рибосомы с мРНК.

ПРАВИЛО GT-AG. Данные динуклеотиды всегда присутствуют в первых двух и последних двух положениях интронов, входящих в состав ядерных генов.

ПРАЙМОСОМА. Комплекс белков, принимающих участие в иницировании синтеза фрагментов Оказаки в процессе прерывистой репликации ДНК; праймосома может перемещаться вдоль ДНК, участвуя в последующих актах инициации.

ПРЕРЫВИСТАЯ РЕПЛИКАЦИЯ. Осуществление синтеза ДНК путем образования коротких фрагментов (Оказаки), которые затем соединяются в единую нить.

ПРОВИРУС. Двухцепочечная последовательность ДНК, встроенная в хромосому эукариот и соответствующая геномной РНК ретровирусов.

ПРОГУЛКА ПО ХРОМОСОМЕ (chromosome walking). Термин означает последовательное выделение клонов, содержащих перекрывающиеся фрагменты ДНК. С помощью такого подхода можно картировать ДНК.

ПРОКАРИОТЫ. Организмы (бактерии), не имеющие ядра.

ПРОМЕЖУТОЧНЫЕ (ВСТАВОЧНЫЕ) ПОСЛЕДОВАТЕЛЬНОСТИ (intervening sequence). Такие последовательности называются *интронами*.

ПРОМОТОР. Участок ДНК, ответственный за связывание РНК-полимеразы, иницирующей транскрипцию.

ПРОСТАЯ ПОСЛЕДОВАТЕЛЬНОСТЬ ДНК. Понятие, эквивалентное сателлитной ДНК.

ПРОФАГ. Фаговый геном, интегрированный в бактериальную хромосому.

ПРЯМЫЕ МУТАЦИИ. Инактивируют ген дикого типа.

ПРЯМЫЕ ПОВТОРЫ. Идентичные (или похожие) последовательности, присутствующие в двух или более копиях, одинаково ориентированные, находящиеся на одной молекуле ДНК; примыкающие друг к другу необязательно.

ПСЕВДОГЕНЫ. Неактивные, но стабильные элементы генома, возникшие в результате мутаций в ранее работающем гене.

ПУФ. Необычайное увеличение диска на политенной хромосоме, связанное с активным синтезом РНК в каком-то локусе этого диска.

pBR322. Один из стандартных плазмидных векторов, используемых при клонировании.

РАДИОАВТОГРАФИЯ. Обнаружение радиоактивно меченных молекул по их способности оставлять след на фотографической пленке.

РАДИОАКТИВНЫЙ СВИДЕТЕЛЬ. Радиоактивно меченный компонент нуклеиновой кислоты, включаемый в реакцию реассоциации в очень малых количествах, не влияющих на дальнейшее протекание реакции.

РАЗРЫВ В ДУПЛЕКСЕ ДНК. Отсутствие фосфодиэфирной связи между двумя соседними нуклеотидами в одной из цепей двухцепочечной ДНК.

РАЗРЫВ И ВОССОЕДИНЕНИЕ. Тип генетической рекомбинации, при котором две молекулы двухцепочечной ДНК разрываются в соответствующих точках и затем соединяются крестнакрест.

РАМКА СЧИТЫВАНИЯ. Один из трех возможных способов считывания нуклеотидной последовательности в виде последовательного ряда триплетов.

РАННЯЯ СТАДИЯ. Период фаговой инфекции до начала репликации ДНК.

РАСКРУЧИВАНИЕ ДНК. Процесс, происходящий в результате отрицательной суперспирализации (так как сама двойная цепь закручивается в направлении, противоположном спирализации нитей ДНК).

РАСПОЛОЖЕНИЕ ПРОТИВ ХОДА ТРАНСКРИПЦИИ (downstream). Условное обозначение расположения последовательностей в направлении, противоположном направлению экспрессии; например, бактериальный промотор находится против хода транскрипции от транскрипционной единицы, а иницирующий кодон находится против хода транскрипции от кодирующей области.

РАССТОЯНИЕ ПО КАРТЕ. Измеряется в сМ (сантиморганадах), соответствует проценту рекомбинаций.

РАСХОДЯЩАЯСЯ ТРАНСКРИПЦИЯ. Транскрипция с двух промоторов в противоположных направлениях.

РАСХОЖДЕНИЕ ХРОМОСОМ. Перемещение членов гомологичной пары хромосом к противоположным полюсам во время клеточного деления. При митозе и втором мейотическом делении этот термин употребляется по отношению к сестринским хроматидам; при первом мейотическом делении он применим к парам сестринских хроматид.

РЕАКЦИЯ, ПРОИСХОДЯЩАЯ ВХОЛОСТУЮ (idling reaction). Образование рибосомами pppGpp и ppGpp, когда незаряженная тРНК находится в участке А; определяет возможность строгого ответа.

РЕАССОЦИАЦИЯ ДНК. Взаимодействие комплементарных цепей с образованием двухцепочечной спирали.

РЕВЕРСИЯ МУТАЦИИ. Замена в ДНК, которая или исправляет первоначальное повреждение (истинная реверсия) или компенсирует его (в результате вторичной мутации в данном гене).

РЕВЕРТАНТЫ. Клетки или организмы, у которых произошла реверсия мутации.

РЕГУЛЯТОРНЫЙ ГЕН. Кодирован РНК или белок, чья функция состоит в контроле экспрессии других генов.

РЕДКО ВСТРЕЧАЮЩАЯСЯ мРНК. Фракция мРНК, состоящая из большого числа индивидуальных видов мРНК, каждая из которых присутствует только в нескольких копиях на клетку.

РЕЗОЛВАЗА. Фермент, обеспечивающий сайтспецифическую рекомбинацию между двумя транспозонами, находящимися в виде прямых повторов в коинтеграционных молекулах.

РЕКОМБИНАЦИОННАЯ РЕПАРАЦИЯ. Способ залечивания бреши в одной из цепей двухцепочечной ДНК за счет замещения участком гомологичной цепи из другой молекулы.

РЕКОМБИНАЦИОННЫЕ УТОЛЩЕНИЯ (УЗЕЛКИ) (recombination nodules). Плотные образования, присутствующие в составе синаптического комплекса; возможно имеют отношение к кроссинговеру.

РЕНАТУРАЦИЯ. Реассоциация денатурированных комплементарных цепей ДНК с образованием двухцепочечной молекулы.

РЕПЛИКАЦИОННАЯ ВИЛКА. Точка, в которой цепи родительской двухцепочечной ДНК расходятся для того, чтобы могла произойти репликация.

РЕПЛИКАЦИОННЫЙ ГЛАЗОК. Область реплицирующейся ДНК внутри протяженного нереплицирующегося района.

РЕПЛИКОН. Единица генома, способная к автономной репликации ДНК; содержит точку инициации репликации.

РЕПЛИСОМА. Мультиферментный комплекс, формирующийся в бактериальной репликационнойвилке для осуществления синтеза ДНК. Содержит ДНК-полимеразу и ряд других белков.

РЕПРЕССИЯ. Ингибирование транскрипции (или трансляции) за счет связывания белка репрессора со специфическим сайтом на ДНК (или мРНК).

РЕСТРИКЦИОННАЯ КАРТА. Линейная последовательность сайтов, в которых ДНК расщепляется различными рестриктазами.

РЕТРОВИРУС. РНК-содержащий вирус, цикл размножения которого проходит через стадию двухцепочечной ДНК.

РЕТРОРЕГУЛЯЦИЯ. Способность дистально расположенной последовательности мРНК регулировать ее трансляцию.

РЕЦЕССИВНО ЛЕТАЛЬНЫЙ АЛЛЕЛЬ. Аллель, летальный в гомозиготном состоянии.

РЕЦЕССИВНЫЙ АЛЛЕЛЬ. Не проявляющийся в гетерозиготе. Отсутствие проявления обусловлено инактивацией или отсутствием продукта рецессивного аллеля.

РЕЦИПРОКНАЯ РЕКОМБИНАЦИЯ. Образование новых генотипов с противоположным расположением аллелей по сравнению с их организацией у матери или отца.

РИФАМИЦИНЫ. Группа антибиотиков, ингибирующих транскрипцию у бактерий.

РНК-ЗАВИСИМАЯ ГИБРИДИЗАЦИЯ. Реакция гибридизации, проводимая в условиях избытка РНК. Этот тип гибридизации используется в тех случаях, когда необходимо, чтобы все комплементарные последовательности одноцепочечной ДНК участвовали в реакции.

м-РНК (МАЛАЯ ЯДЕРНАЯ РНК). Одна из многих маленьких РНК, содержащихся в ядре; ее функция неясна, возможно, принимает участие в сплайсинге.

РОДСТВЕННЫЕ тРНК. Молекулы тРНК, которые узнаются определенной аминокислотой-синтетазой.

Р-ПЕТЛЯ. Структура, образующаяся при гибридизации РНК с комплементарной цепью двухцепочечной ДНК; при этом происходит вытеснение исходной цепи ДНК в виде петли, расположенной в области гибридизации.

РесА-БЕЛОК. Кодируется гесА-локусом *E. coli*; обладает двумя активностями: функционирует как протеаза и участвует в обмене цепей между различными молекулами ДНК. Протеазная активность контролирует SOS-функцию; нуклеиновые кислоты благодаря этому белку участвуют в процессах рекомбинации и репарации.

ρ-НЕЗАВИСИМЫЕ ТЕРМИНАТОРЫ. Последовательности ДНК, на которых РНК-полимераза прекращает транскрипцию в отсутствие ρ-фактора.

ρ-ФАКТОР. Белок, помогающий РНК-полимеразе прекращать транскрипцию в определенных (ρ-зависимых) сайтах.

САЙТСПЕЦИФИЧЕСКАЯ РЕКОМБИНАЦИЯ. Происходит между двумя определенными (не обязательно гомологичными) последовательностями, например наблюдается при интеграции и исключении фага или при разрешении коинтегральных структур в процессе транспозиции.

САЙТЫ ЗАМЕЩЕНИЯ. Сайты в гене, мутации в которых приводят к замене кодируемой аминокислоты.

САМОКОМПЛЕМЕНТАРНАЯ ДНК. Состоит из обращенных повторов, которые ренатурируют за счет реассоциации комплементарных последовательностей, расположенных на одной цепи в денатурированной ДНК.

САТЕЛЛИТНАЯ ДНК. Состоит из большого числа тандемных повторов (идентичных или похожих) основной короткой последовательности.

СУПЕРВИТКИ В ДНК. Образуются в результате суперспирализации ДНК; способствуют дальнейшему закручиванию двух цепей дуплекса относительно друг друга.

СЕЛЕКЦИЯ. Создание условий, при которых выживает только определенный тип клеток.

СЕСТРИНСКИЕ ХРОМАТИДЫ. Копии хромосомы, образующиеся при ее репликации.

СИГМА-ФАКТОР. Субъединица бактериальной РНК-полимеразы, необходимая для инициации; влияет преимущественно на выбор сайтов связывания (промоторов).

СИГНАЛЬНАЯ ГИПОТЕЗА. Объясняет необходимость N-концевой последовательности секретируемых белков для прикрепления новообразованного полипептида к мембране; согласно гипотезе, мРНК и рибосома прикрепляются к мембране в процессе синтеза белка благодаря его N-концевой части.

СИНАПСИС. Конъюгация двух пар сестринских хроматид гомологичных хромосом, происходящая во время мейоза; образующаяся структура называется бивалентом.

СИНАПТОМАЛЬНЫЙ КОМПЛЕКС. Структура, возникающая при конъюгации хромосом.

СИНТЕННЫЕ ЛОКУСЫ. Генетические локусы, относящиеся к одной и той же хромосоме.

СЛОЖНОСТЬ ПОСЛЕДОВАТЕЛЬНОСТЕЙ ДНК. Суммарная длина уникальных последовательностей ДНК данного препарата.

СЛОЖНЫЕ ТРАНСПОЗОНЫ. Центральная часть таких транспозонов фланкирована инсерционными последовательностями, которые и обуславливают способность к транспозиции.

СЛОЖНЫЙ ЛОКУС. Локус (например, у *D. melanogaster*), генетические свойства которого не согласуются с принципом «один ген — один белок».

СОГЛАСОВАННАЯ ЭВОЛЮЦИЯ. Означает, что два родственных гена эволюционируют совместно, как если бы они относились к одному локусу.

СОМАТИЧЕСКИЕ КЛЕТКИ. Все клетки организма, за исключением половых клеток.

СПЕЦИАЛИЗИРОВАННЫЕ ГЕНЫ (ГЕНЫ «РОСКОШИ»). Гены, кодирующие белки, синтезирующиеся (обычно) в больших количествах в клетках определенных типов.

СПЛАЙСИНГ. Процесс удаления интронов и объединения экзонов в мРНК.

СПОНТАННЫЕ МУТАЦИИ. Возникают в отсутствие каких-либо факторов, увеличивающих частоту мутирования.

СПОРУЛЯЦИЯ. Образование бактериальных (в результате изменения морфологии) или дрожжевых (в результате мейоза) спор.

СТАРТОВАЯ ТОЧКА (ИНИЦИИРУЮЩИЙ САЙТ). Обозначает участок ДНК, соответствующий первому основанию, включаемому в РНК.

СТЕБЕЛЬ. Двухцепочечный участок шпильки, образованный спаренными основаниями.

СТЕПЕНЬ СКРУЧЕННОСТИ. Число раз, которое ось двойной спирали пересекает саму себя в пространстве.

СТРЕПТОЛИДИГИНЫ. Группа антибиотиков, ингибирующих элонгацию транскрипции, осуществляемую бактериальной РНК-полимеразой.

СТРОГИЙ КОНТРОЛЬ РЕПЛИКАЦИИ. Ограничение репликации однокопийных плазмид, согласованное с репликацией бактериальной хромосомы.

СТРОГИЙ ОТВЕТ. Способность бактерии прекращать синтез тРНК и образование рибосом в голодной среде.

СТРУКТУРА ТИПА ВОСЬМЕРКИ. Образуется из двух замкнутых молекул ДНК, сцепленных вместе в результате рекомбинационных событий, которые полностью не завершились.

СТРУКТУРНЫЙ ГЕН. Кодирует РНК или белок.

СТУПЕНЧАТЫЕ РАЗРЕЗЫ. Возникают в двухцепочечной ДНК, в том случае, когда ее цепи расщепляются в точках, находящихся недалеко друг от друга.

СУПЕРСПИРАЛИЗАЦИЯ. Означает закручивание кольцевой двухцепочечной ДНК в пространстве таким образом, что при этом она пересекает собственную ось дуплекса.

СУПРЕССИЯ. Изменения, которые устраняют проявление мутации, не исправляя при этом первоначального нарушения в ДНК.

СУПРЕССОР (ВНЕГЕННЫЙ). Обычно ген, кодирующий мутантную тРНК, которая прочитывает мутантный кодон согласно значению первоначального кодона или включает похожую аминокислоту.

СУПРЕССОР (ВНУТРИГЕННЫЙ). Компенсирующая мутация, восстанавливающая первоначальную рамку считывания.

СУПРЕССОР БЕССМЫСЛЕННОГО КОДОНА. Ген, кодирующий мутантную тРНК, способную узнавать бессмысленный кодон.

СЦЕПЛЕННОСТЬ. Свойство генов одной хромосомы наследоваться совместно; измеряется в процентах рекомбинации между локусами.

СЦЕПЛЕНИЕ С ПОЛОМ. Способ наследования, характерный для генов, находящихся в половых хромосомах (обычно в X-хромосоме).

СФЕРОПЛАСТ. Бактериальная или дрожжевая клетка, лишенная или почти лишенная клеточной стенки.

S-ФАЗА. Период клеточного цикла, в течение которого происходит синтез ДНК в эукариотической клетке.

SI-НУКЛЕАЗА. Фермент, который специфично деградирует неспаренные (одноцепочечные) последовательности ДНК.

SOS-БЛОК. Последовательность ДНК (оператор) длиной около 20 п. н., узнаваемая белком-репрессором LexA.

SOS-ОТВЕТ. Координированная индукция синтеза многих ферментов, в том числе и репарационных, в ответ на радиационное или иное повреждение ДНК у *E. coli*; является следствием активации протеазной активности RecA-белка, расщепляющей LexA-репрессор.

SSB-БЕЛОК. Белок *E. coli*, связывающийся с одноцепочечной ДНК.

ТАЛАССЕМИЯ – заболевание человека, вызванное отсутствием α - или β -глобина в его эритроцитах.

ТАНДЕМНЫЕ ПОВТОРЫ. Множественные копии одинаковых последовательностей, расположенных одна за другой, ориентированных в одном направлении.

ТЕЛОМЕРА. Естественный конец хромосомы.

ТЕМПЕРАТУРА ПЛАВЛЕНИЯ ($T_{пл}$). Соответствует среднему значению температурного интервала, в пределах которого происходит плавление ДНК.

ТЕМПЕРАТУРОЧУВСТВИТЕЛЬНЫЙ МУТАНТ. Не способен развиваться при повышенной температуре.

ТЕРМИНАЛЬНАЯ ИЗБЫТОЧНОСТЬ. Повторение одних и тех же последовательностей на обоих концах, например, хромосомы фага.

ТЕРМИНАТОР. Последовательность ДНК, находящаяся на конце транскрипта и ответственная за прекращение транскрипции.

ТЕРМИНИРУЮЩИЕ (ОСВОБОЖДАЮЩИЕ) ФАКТОРЫ. Узнают бессмысленные кодоны, вызывая освобождение синтезированной полипептидной цепи и рибосом от мРНК.

ТЕРМИНИРУЮЩИЙ КОДОН. Один из трех триплетов UAG (*amber*), UAA (*ochre*) или UGA, вызывающих терминацию синтеза белка; их также называют бессмысленными кодонами.

ТЕСТ НА КОМПЛЕМЕНТАЦИЮ IN VITRO. Наличие в клетках дикого фенотипа такого компонента, который восстанавливает исходную активность в экстракте, приготовленном из мутантных клеток.

ТОПОИЗОМЕРАЗА. Фермент, изменяющий число зацеплений в ДНК (топоизомеразы типа I изменяют число зацеплений последовательно на единицу; топоизомеразы типа II изменяют число зацеплений последовательно на два).

ТОПОЛОГИЧЕСКИЕ ИЗОМЕРЫ. Одинаковые молекулы ДНК, различающиеся только числом зацеплений.

ТОЧКА НАЧАЛА РЕПЛИКАЦИИ (*ori*). Последовательность ДНК, в которой происходит инициация репликации.

ТОЧКОВЫЕ МУТАЦИИ. Изменение одной пары оснований.

ТРАНЗИЦИЯ. Мутация, в результате которой происходит замена пиримидина на другой пиримидин или же пурина на другой пурин.

ТРАНСВЕКЦИЯ. Свойство локуса влиять на активность аллельного локуса, расположенного в другом гомологе. Эффект наблюдается только при синапсисе двух хромосом.

ТРАНСВЕРСИЯ. Мутация, в результате которой пурин замещается пиримидином или же наоборот.

ТРАНСДУКЦИЯ. Перенос бактериального гена от одной бактерии к другой при помощи фага. Фаг, осуществляющий этот перенос, носит название трансдуцирующего. Его геном содержит как собственные гены, так и гены клетки-хозяина.

ТРАНС-КОНФИГУРАЦИЯ. Расположение двух сайтов на двух разных молекулах ДНК (хромосомах).

ТРАНСКРИБИРУЕМЫЙ СПЕЙСЕР. Часть транскрипционной единицы рРНК, входящая в состав транскрипта; отсутствует в зрелой молекуле рРНК.

ТРАНСКРИПЦИОННАЯ ЕДИНИЦА. Участок ДНК между сайтами инициации и терминации РНК-полимеразы; может содержать несколько генов.

ТРАНСКРИПЦИЯ. Процесс синтеза РНК на ДНК-матрице.

ТРАНСЛОКАЦИЯ РИБОСОМЫ. Перемещение кодона вдоль мРНК после включения каждой новой аминокислоты в полипептидную цепь.

ТРАНСЛОКАЦИЯ ХРОМОСОМНАЯ. Перестройка, при которой часть хромосомы отрывается, а затем присоединяется к другой хромосоме.

ТРАНСЛЯЦИЯ. Процесс синтеза белка на матричной мРНК.

ТРАНСПЛАНТАЦИОННЫЙ АНТИГЕН. Белок, кодируемый главным локусом гистосовместимости, присутствует на поверхности всех клеток млекопитающих, участвует во взаимодействии между лимфоцитами.

ТРАНСПОЗАЗА. Фермент, участвующий в интеграции транспозона в новый сайт.

ТРАНСПОЗИЦИОННЫЙ ИММУНИТЕТ. Свойство определенных транспозонов блокировать встраивание родственных транспозонов в ту же молекулу ДНК.

ТРАНСПОЗОН. Последовательность ДНК, способная реплицироваться и внедрять одну из копий в новое место генома.

ТРАНСФЕКЦИЯ ЭУКАРИОТИЧЕСКИХ КЛЕТОК. Приобретение нового генетического маркера в результате включения экзогенной ДНК.

ТРАНСФОРМАЦИЯ БАКТЕРИАЛЬНЫХ КЛЕТОК. Приобретение нового генетического маркера в результате включения экзогенной ДНК.

ТРАНСФОРМАЦИЯ ЭУКАРИОТИЧЕСКИХ КЛЕТОК. Переход в состояние неконтролируемого роста; имеет много общего и ли совпадает с опухолевым перерождением клеток.

ТАТА-ПОСЛЕДОВАТЕЛЬНОСТЬ (БЛОК ХОГНЕССА). А—Т-богатая семичленная последовательность, находящаяся на расстоянии около 25 п. н. перед стартовой точкой каждой транскрипционной единицы, транскрибируемой РНК-полимеразой II; вероятно, необходима для такого расположения фермента, при котором он может осуществлять правильную инициацию.

УМЕРЕННО РЕАССОЦИИРУЮЩИЙ КОМПОНЕНТ. Обладает скоростью реассоциации, промежуточной между быстро (сателлитная ДНК) и медленно (уникальная ДНК) ренатурирующими компонентами; содержит умеренно повторяющуюся ДНК.

УПАКОВОЧНЫЙ КОЭФФИЦИЕНТ. Отношение длины ДНК к длине структуры, которая ее содержит.

УСИЛИТЕЛИ ТРАНСКРИПЦИИ (*enhancer*). Участки ДНК, усиливающие транскрипцию с ряда эукариотических промоторов, находящихся по отношению к ним в *цис*-положении. Эти элементы оказывают свое действие независимо от того, с какой стороны промотора они располагаются.

УСЛОВНО ЛЕТАЛЬНЫЕ МУТАЦИИ. Вызывают гибель клетки или вируса только в определенных (непермиссивных) условиях, но не проявляют своего летального действия в других (пермиссивных) условиях.

УЧАСТКИ СПЛАЙСИНГА. Последовательности, непосредственно окружающие границы между экзонами и интронами.

УЧАСТОК РЕКОМБИНАЦИОННОГО ПЕРЕКРЕСТА (*recombinant joint*). Место соединения двух рекомбинирующих молекул двухцепочечных ДНК (граница гетеродуплексной области).

ФАГ (БАКТЕРИОФАГ). Бактериальный вирус.

ФАЗОВЫЕ ВАРИАЦИИ. Изменение типа бактериальных жгутиков.

ФАКТОРЫ ИНИЦИАЦИИ (IF). Белки, которые специфически связываются с малой субчастицей рибосомы на стадии инициации белкового синтеза.

ФАКТОРЫ ЭЛОНГАЦИИ. Белки, циклично ассоциирующие с рибосомой в соответствии с включением каждой новой аминокислоты в полипептидную цепь.

ФАКУЛЬТАТИВНЫЙ ГЕТЕРОХРОМАТИН. Неактивное состояние последовательностей хроматина, который также представлен и активной формой – например, одна из X-хромосом у самок млекопитающих.

ФЕНОТИП. Совокупность внешних признаков организма; формируется в результате взаимодействия генотипа и среды.

ФЕРМЕНТЫ РЕСТРИКЦИИ. Узнают определенные короткие последовательности, как правило, в неметилированной ДНК и расщепляют ее иногда в месте связывания, а иногда в каком-либо другом месте (это зависит от типа фермента).

ФЕРМЕНТЫ С ПРОЦЕССИВНЫМ МЕХАНИЗМОМ ДЕЙСТВИЯ (*processive enzymes*). Ферменты, которые действуют на

определенный субстрат, не отделяясь от него между повторяющимися каталитическими событиями.

ФИКСАЦИЯ КРОССОВЕРА. Возможное последствие неравного кроссинговера, в результате которого мутация в одном члене tandemного кластера может распространиться на весь кластер (или элиминироваться).

ФРАГМЕНТЫ ОКАЗАКИ. Короткие фрагменты ДНК длиной 1000–2000 оснований; образуются в результате прерывистой репликации; впоследствии ковалентно соединяются в непрерывную цепь.

ХИАЗМА. Участок, в котором две гомологичные хромосомы обмениваются своим наследственным материалом при мейозе.

ХИМИЧЕСКАЯ СЛОЖНОСТЬ. Количество ДНК, измеренное химическим методом.

ХОЛОДОЧУВСТВИТЕЛЬНЫЙ МУТАНТ. Не способен развиваться при низкой температуре.

ХРОМАТИДЫ. Хромосомные копии, образующиеся при репликации.

ХРОМАТИН. Комплекс между ДНК и белками в ядрах интерфазных клеток; выявляется по способности окрашиваться красителями, специфическими для ДНК.

ХРОМОМЕРА. Интенсивно окрашиваемая гранула; ее можно различить как составную часть хромосомы при определенных условиях (особенно на ранних стадиях мейоза).

ХРОМОСОМА. Дискретная единица генома, состоящая из молекулы ДНК и белков; содержит большое количество генов.

ХРОМОСОМА С МНОЖЕСТВОМ РЕПЛИКАЦИОННЫХ ВИЛОК (В БАКТЕРИИ). Существование нескольких репликационных вилок, обеспечивает возможность инициации нового цикла репликации до того, как закончится предыдущий цикл.

ХРОМОСОМНЫЙ ОСТОВ. Белковая структура, имеющая форму пары сестринских хроматид; возникает при удалении гистонов из хромосомы.

ХРОМОЦЕНТР. Агрегат гетерохроматина из различных хромосом.

ЦЕНТРОМЕРА. Область хромосомы, в которую входит участок прикрепления к митотическому или мейотическому веретену.

ЦИКЛИЧЕСКАЯ АМР. Молекула АМР, в которой фосфатная группа соединена как с 3'-, так и с 5'-положениями рибозы; связывание с циклическим АМР активирует САР-белок, являющийся положительным регулятором транскрипции у прокариот.

цис-АКТИВНЫЙ ЛОКУС. Влияет на функционирование последовательностей ДНК, расположенных с ним в одной молекуле; предполагается, что этот локус не кодирует белок.

цис-КОНФИГУРАЦИЯ. Термин применяется по отношению к двум сайтам, расположенным в одной молекуле ДНК.

ЦИСТРОН. Генетическая единица, выявляемая путем комплементационного теста; эквивалентна гену и означает единицу ДНК, кодирующую белок.

ЦИТОПЛАЗМАТИЧЕСКИЙ СИНТЕЗ БЕЛКА. Заключается в трансляции молекул мРНК, соответствующих ядерным генам; процесс осуществляется рибосомами, прикрепленными к цитоскелету.

ЦИТОПЛАЗМАТИЧЕСКОЕ НАСЛЕДОВАНИЕ. Характерно для признаков, определяемых митохондриальными генами, и генами, локализованными в хлоропластах (или в любых других внеядерных оргanelлах).

ЦИТОСКЕЛЕТ. Сеть фибрилл в цитоплазме эукариотических клеток.

С-ГЕНЫ. Кодируют константные области иммуноглобулиновых белковых цепей.

С-ЗНАЧЕНИЕ. Общее количество ДНК в гаплоидном геноме.

С-СЕКМЕНТАЦИЯ. Характерная картина, возникающая при специфическом окрашивании околоцентромерной области.

ЧАСТОТА ПОВТОРЯЕМОСТИ. Число копий данной последовательности в гаплоидном геноме; для неповторяющейся ДНК равно 1, для повторяющейся ДНК ≥ 2 .

ЧЕРЕДОВАНИЕ ПОСЛЕДОВАТЕЛЬНОСТЕЙ С ДЛИННЫМ ИНТЕРВАЛОМ (*long-period interspersion*). Тип организации генома, при котором чередуются длинные участки умеренно повторяющейся ДНК с уникальной ДНК.

ЧЕРЕДОВАНИЕ ПОСЛЕДОВАТЕЛЬНОСТЕЙ С КОРОТКИМ ИНТЕРВАЛОМ (*short-period interspersion*). Тип организации генома, для которого характерно чередование умеренно повторяющихся последовательностей ДНК длиной около 300 п. н. с неповторяющимися последовательностями длиной около 1000 п. н.

ЧИСЛО ЗАЦЕПЛЕНИЙ. Количество раз, которое две цепи ковалентно замкнутой двухцепочечной ДНК пересекают друг друга.

ЧИСТЫЕ ЛИНИИ. Организмы, гомозиготные по изучаемым признакам.

ШПИЛЬКА. Представляет собой двухцепочечную область, образующуюся за счет спаривания оснований между соседними (инвертированными) комплементарными последовательностями в одноцепочечной РНК или ДНК.

ЭВОЛЮЦИОННОЕ ВРЕМЯ. Определяется скоростью накопления мутаций в данном гене.

ЭКЗОН. Любой отдельный фрагмент прерывистого гена, который сохраняется в зрелой РНК.

ЭКЗОНУКЛЕАЗЫ. Ферменты, последовательно отщепляющие нуклеотиды с концов полинуклеотидной цепи; могут быть специфичными в отношении 5'- или 3'-концов ДНК или РНК.

ЭКСЦИЗИОННО-РЕПАРАЦИОННЫЕ СИСТЕМЫ. Удаляют одноцепочечный участок двухцепочечной ДНК, содержащий поврежденные или неправильно спаренные основания, и заменяют его, синтезируя последовательность, комплементарную той, которая присутствует в оставшейся цепи.

ЭНДОНУКЛЕАЗЫ. Ферменты, расщепляющие связи полинуклеотидной цепи нуклеиновых кислот; могут быть специфичны в отношении РНК одноцепочечных или двухцепочечных ДНК.

ЭПИГЕНЕТИЧЕСКИЕ ИЗМЕНЕНИЯ. Изменения фенотипа особи без изменения ее генотипа.

ЭПИСОМА. Плазмида, способная интегрировать в бактериальную ДНК.

ЭУХРОМАТИН. Представляет собой весь генетический материал интерфазного ядра, за исключением гетерохроматина.

ЭФФЕКТ ПОЛОЖЕНИЯ. Изменение экспрессии гена в результате его транслокации в новый участок генома; например, ранее активный ген при перемещении в гетерохроматиновый район может быть инактивирован.

ЯДЕРНЫЙ МАТРИКС. Сплетение фибрилл, окружающих и пронизывающих ядро.

ЯДРЫШКО. Обособленная область ядра, образуемая при транскрипции генов рРНК.

ЯДРЫШКОВЫЙ ОРГАНИЗАТОР. Область хромосомы, содержащая гены, кодирующие рРНК.

Предметный указатель

- Автокатализ 322
Аденилатциклаза 201
Аденин 25, 96
Аденовирусы 430
Аденозинтрифосфат *см.* АТР
Аденозинтрифосфатаза *см.* АТРаза
5-Азацитидин 388
Азотистые основания, комплементарные 27
-- строение 25
Акридины 38
- мутагенный эффект 57, 58
АКТГ 342
Активатор 339
Аланин 56
Алкаптонурия 17
Аллели 9
- множественные 19
- расщепление 10
Алелизм, тест 19
Аллельное исключение 511–512
Аллостерический контроль 178
Амбер (*amber*)-мутации 61
Амилаза, синтез, генетический контроль 256–257
Аминоацил-тРНК 59
- супрессорная 61
Аминоацил-тРНК—синтетазы 67–68, 92, 283
- узнавание тРНК 94
Аминокислоты 55–56
- кодирование 59
- строение 56
Амплификация 495–497. *См. также* Кольца Бальбиани; Хромосомные пуфы
- рДНК 294
Амплифицированные последовательности, селекция 497–500
Анализирующее скрещивание *см.* Возвратное скрещивание
Анафаза мейоза 13
- митоза 12
Антигены 502
- трансплантационные 516
- Т и t (большие и малые) 257–258
Антикодоны 59, 68
- петля, стебель 90
- спаривание с интроном 318
Антипараллельные цепи 27
Антирепрессор 218
Антисмысловая последовательность 217
Антитела 502–518. *См. также* Гистосовместимость
Антитерминация 169–171
Апоиндукторы 189
Апопротеин 284
Аргинин 56
- в гистонах 359
- кодоны 60, 61
Аристотель (теория информации о развитии эмбриона) 55
Аспарагин 56
Аспарагиновая кислота 56
Аспартат-карбамоилтрансфераза 497
Аттенуатор 191
Аттенуация 191–196
Аутогенный контроль 196, 201–204
- цикл 210–213
-- и лизогения 213, 219
Аутосомы 13
Ацентрический фрагмент 352
Ацетилаза 179
АТР (аденозинтрифосфат) в лигазной реакции 319
АТРаза (аденозинтрифосфатаза) 284
- олигомицинчувствительная 284, 285
- репликационного аппарата фага 428

Бактериальные рРНК-гены 296–297
- тРНК-гены 296–297
Бактериальный геном 32
-- репликация 398–399
-- упаковка 347–349
Бактерии, гибридизация колоний 244
- клонирование в них ДНК 236–238
- репликационный аппарат 420–421
- экспрессия эукариотических генов 244–245
Бактериофаги *см.* Фаги
Батарей (генов) 340
Белки, богато представленные в клетке, кодирование 279–280
- высококонсервативные 275
- дивергенция 275
- капсида 345
- комплемента 516
- мультимерные 18, 55
- репликации 422–425. *См. также* Белок, связывающийся с одноцепочечной ДНК
- рибосомные 201–203, 283
- связанные с 5'-концевым основанием 429
- синтез 59, 74
-- в митохондриях 283
-- в рибосомах 64–65
-- инициация 74, 75
-- направление 59
-- цитоплазматический 281
- хроматина 358–361. *См. также* Гистоны; Негистоновые белки
- HMG 360
Белковый комплекс, воспроизведение 393
Белок БАК (БРЦ) 201, 215
- Cro 170, 171, 215, 220
- IHF 456
- Int 456
- Lex 441, 442
- pN 170, 171
- RecA 448–452
- Rep 424–425, 427
- SSB (связывающийся с одноцепочечной ДНК) 421–424, 428, 450, 451
- α 378
Белок-репрессор 178. *См. также* Репрессор
Берка-Шарпа метод 249
Библиотеки генов 243–244
Биваленты 11, 14, 16
Биосинтез, генетический контроль 18
Блоттинг по Саузерну 242, 243. *См. также* Нозерн-блоттинг
Бляшки (стерильные пятна) 211
Бромистый этидий 348
Бромурацил 39
Бэкросс (возвратное скрещивание) 15

Валин 56
Величина С 222–223
Виозин 89
Вирулентная (+) цепь 421. *См. также* Кольцевая ДНК; Комплементарная (-) цепь
Вирулентные мутации 211
Вирус желтой мозаики турнепса 345

- полиомы, безнуклеосомные пробелы 391
- табачной мозаики (ВТМ) 344, 345
- SV40 как вектор 326, 328
- Вирус-помощник 492
- Вирусы, дефектные по репликации 492.
См. также Ретровирусы
- Вирусные геномы кольцевые 32
- упаковка 344–347
- Внегенные супрессоры *см.* Супрессорные мутации
- Внехромосомные элементы 206. *См. также* Микрохромосомы; Плазмиды; Подвижные (мобильные) элементы; Транспозоны
- Внехромосомный палиндром 293
- Внеядерная (цитоплазматическая) наследственность 281. *См. также* Митохондриальный геном; Геном хлоропластов
- Водородные связи в ДНК 27
- Водянка плода 272
- Возвратное (анализирующее) скрещивание (бэкресс) 10, 15
- Вставки (включения, инсерции) 38, 40
- Вставочные последовательности *см.* Интроны
- ВТМ *см.* Вирус табачной мозаики
- Вторичная структура *см.* Полипептидная цепь, структура
- Вырожденность кода 60
- третьего основания 60, 95
- Высокоповторяющиеся последовательности (ДНК) 227, 231, 300–301. *См. также* Сателлитная ДНК
- Высшие растения 282
- Вытеснение метки 333–334
- β-Галактозидаза 179
- Гаметофит 9
- Гаметы (половые клетки) 8, 11
- Гаплоидное число хромосом 11, 13
- Гаплоидные ткани 9
- Гаплоидный геном 223
- Геликаза АТР-зависимая 423
- Гель-электрофорез 44
- rPHK 318, 319
- Гемоглобин анти-Lepore 273, 274
- плода 269
- строение 18
- Lepore 272
- Гемоглобинопатия Н 272, 273
- Гемоглобины, типы 268–269, 272–273
- Ген *см.* Гены
- алкогольдегидрогеназы (АДГ) 264
- альтернативные формы 9
- гипотеза один ген–один белок 17–18
- один ген–один фермент 18
- один ген–одна полипептидная цепь 18
- миоглобина 265
- молекулярная структура 42–54
- природа, генетическая точка зрения 8–21. *См. также* Цистрон
- история вопроса 8, 21
- размер 223
- структура, прямые исследования 43–44.
См. также Рестрикционная карта
- что это такое? 54
- Генетика, основная концепция 8
- Генетическая инженерия 236
- информация, природа 7
- карта 17, 42, 46, 51. *См. также* Кортирование генов; Рестрикционная карта;
- Цитологическая карта хромосом
- линейность расположения генов 16–18
- Генетическая рекомбинация 16. *См. также* Кроссинговер; Рекомбинация
- Генетический код 55, 60, 95. *См. также* Вырожденность кода
- дешифровка 57
- знаки пунктуации 62
- история вопроса 57
- триплетная природа 57–58
- маркер, термин 20
- материал у прокариот и эукариот 22, 24
- Ген-интегратор 339, 340
- Генная конверсия 227
- экспрессия *см.* Экспрессия генов
- Генные локусы 15. *См. также* Локус; Локусы
- сегменты 508–509
- Геном(ы) 344–358. *См. также* Бактериальные геномы; Вирусные геномы
- вирусов, бактерий, эукариот, особенности 32
- динамичность 457–518
- млекопитающих 36
- размеры 242
- независимые домены 348–349
- органелл 280–288. *См. также* Митохондриальный геном
- тканеспецифические вариации (у дрозофилы) 495–497
- эукариот, размеры 225–226. *См. также* Величина С
- строение 221–226
- хлоропластов 282
- Генотип 9, 11, 20
- Гены *см.* Библиотека генов; Гистоновые гены; Глобиновые гены; Гомеозисные гены; Кластеры генов; Митохондриальные гены; Множественные гены; Экспрессия генов
- активация 338
- альтернативные и перекрывающиеся 53–54
- внехромосомные *см.* Микрохромосомы
- выделение 242–243
- «домашнего хозяйства» (конститутивные) 235
- задержанно ранние (средние) 207, 212, 213
- инсулина 265
- картирование *см.* Картирование генов
- кодирование белков 18, 279–280
- кодирующие rPHK 285, 379
- контроль метаболических реакций 18
- независимость комбинации 9, 11
- овальбумина 336
- органелл 281. *См. также* Митохондриальные гены
- организация в геноме 8
- расположение в хромосоме 12, 14–17
- рибосомных белков 202
- сцепленные, измерение 15
- строение и размеры 42–54, 252–254, 279
- участвующие в репарации, номенклатура 439
- эукариот 52–53. *См. также* Прерывистые гены
- экспрессия в бактериях 244–245
- lac-Гены *см.* Лактозные гены
- 5S'-Гены 294–295
- Гептамер 510
- Гетерогаметный пол 13
- Гетерогенная ядерная РНК (гяРНК, hnRNA) 331–338
- размеры 332–333
- сложность строения 336
- Гетеродуплексы ДНК (гибридная ДНК) 445–446
- Гетерозигота 9
- Гетерополимер 60
- Гетеротипичный тетрамер 369, 370
- Гетерохроматин 301, 350
- связь с сателлитной ДНК 301–302
- Гибридизация ДНК векторной и чужеродной 239
- *in situ* (цитологическая) 301, 357
- колоний бактерий 244
- нуклеиновых кислот в растворе 224
- методы 36
- РНК–ДНК, исследование гибридов в электронный микроскоп 246–248
- кинетика, оценка числа генов 232–234
- насыщающая, метод 232
- poly(A) с затравкой 241
- Гибридная ДНК *см.* Гетеродуплексы ДНК
- Гибридный дисгенез 480
- и транспозиция 480–482
- Гипохромный эффект 35
- Гираза 412–413
- Гистидин 56
- Гистон *см.* Гистоны
- H1 359–361
- роль в митотической конденсации хромосом 385, 386
- H2A 348, 359, 360
- связывание с убиквитином (UH2A) 385–386
- H2B 348, 359, 360
- H3 359, 360
- H4 359, 360
- H5 360
- Гистоновые гены 289–291, 390
- Гистоновые октамеры 363–365, 370–371, 392
- Гистоны 289, 359–361. *См. также* Гистон; Гистоновые гены
- модификации 384–385
- организация в нуклеосоме 368–369
- типы 359, 361
- Гистосовместимость, главный комплекс 503
- Главный локус гистосовместимости 516–518
- Глицин 56
- Глобиновые гены 250, 253, 255, 265, 390
- кластеры 269–270
- происхождение 274
- транскрипция 336
- чувствительность к ДНКазе 382
- Глобины 268–272. *См. также* Глобиновые гены
- Глутамин, глутаминовая кислота 56
- Гомео-бокс 264
- Гомеозисные гены (мутации) 262, 263
- *Antennapedia (Antp)* 263
- *proboscipedia (pb)* 263
- *Sex combs reduced (Sxr)* 263
- Гомогаметный пол 13
- Гомозигота 9

- Гомологичные хромосомы 443, 444
 Гомомультимер 18
 Гомотипические тетрамеры (нуклеосом) 368
 Горячие точки 38–40, 452
 -- транспозиции 460
 Градиент плотности CsCl 300
 Группы комплементации 19, 53
 -- сцепления 16
 Гуанин 25, 96
 гяРНК *см.* Гетерогенная ядерная РНК
 гяРНК 333
 ГТФаза 428
- Двойная спираль правосторонняя 27
 -- свободная энергия 35
 -- структура 28, 29
 -- суперспирализация *см.* Суперспирализация двойной спирали
 Двухцепочечная ДНК, разрыв, воссоединение и рекомбинация 445–448, 451
 Дезаминирование оснований 40
 3'-Дезоксиаденозин (кордицепин) 335
 Дезоксирибоза 25
 Дейтеростомы, размер генома 223
 Делеции 38, 40, 41
 -- при неравном кроссинговере 271–273
 Делеционная модель рекомбинации между V- и C-генами 510
 Денатурация белка и ДНК 57. *См. также* ДНК, денатурация
 Детерминация 393
 Дигидроуридин 88
 Дигидрофоллятредуктаза (DHFR) 497–500
 Диккий тип, термин и обозначения 12, 20
 Динамичность генома 457–518
 Диплоид частичный 179
 Диплоидная клетка 12
 Диплоидные ткани 9
 Диплоидный набор хромосом 10
 -- организм 8
 Диски хромосом *см.* Хромосомные диски и междиски
 Дифференциация 393
 ДНК 236–245. *См. также* Вирулентная ДНК; Двухцепочечная ДНК; кДНК; Клонирование ДНК; Комплементарная цепь ДНК; мДНК; Метилирование ДНК; Одноцепочечная ДНК; Н-Цепь ДНК; L-Цепь ДНК безнуклеосомные пробелы и сверхчувствительность к ДНКазе 390–392
 вируса SV40 364
 восстановление и рекомбинация 443–456. *См. также* Рекомбинация вторичная структура 164
 высокоповторяющаяся *см.* Сателлитная ДНК
 генома, постоянное изменение *см.* Амплификация; Модификация ДНК; Мутации; Подвижные (мобильные) элементы; Рекомбинация; Ретровирусы; Трансдукция; Транспозиция; Транспозоны; Трансфекция; Фазовая вариация
 гетеродуплексные (гибридные) молекулы 445–446
 двойная спираль 26–29. *См. также* ДНК, В-форма
 двуспиральная структура, модели 29–32
 -- деметилирование и экспрессия гена 386–387
 -- денатурация 35–36, 57, 229
 -- и ренатурация 35–36
 -- димерные кольца 447
 -- и РНК, различия 24
 -- как генетический материал, история вопроса 21–24,
 -- как трансформирующий фактор 22
 -- кодирующая и антикодирующая цепи 59
 -- количество кодирующих генов 280
 -- в ядре *см.* Величина С
 -- кольцевая молекула 32
 -- компоненты 24–26
 -- линкерная 363
 -- митохондрий *см.* Митохондриальная ДНК
 -- недокрученная и перекрученная 32
 -- «нулевая» 235
 -- нуклеосом 362–371
 -- ограниченный гидролиз 363
 -- одноцепочечные разрывы (ники) 366
 -- определение нуклеотидной последовательности 49–51
 -- основания, модифицирование *см.* Метилирование ДНК; Модификация ДНК
 -- перестройки 8
 -- полуметилированный дуплет 386
 -- право- и левосторонние спирали 33
 -- простые последовательности 298–308
 -- разделение цепей 409
 -- различные формы спирали 29, 30
 -- разрыв и воссоединение 445–447. *См. также* Рекомбинация ДНК
 -- реассоциация, кинетика в растворе 224–225, 228, 229
 -- определяемая ее концентрацией 231
 -- ренатурация 35–36. *См. также* ДНК, денатурация; ДНК, реассоциация
 -- обращенных повторов 299–300
 -- репликативная форма (РФ) 421. *См. также* Релаксированное кольцо (РФП); Суперспирализованное кольцо (РФИ)
 -- сверхчувствительность к ДНКазе 390–392
 -- синтез *см.* Репликация ДНК; Фрагменты Оказаки
 -- инициация 421–422
 -- полуконсервативный 396–398
 -- полупрерывистый 416–418
 -- репаративный 38
 -- система защиты 431–442. *См. также* Модификация ДНК; Репарация
 -- структура *см.* ДНК, двойная спираль; ДНК, двуспиральная структура; ДНК, разные формы
 -- креста 32, 163, 164. *См. также* Шпильки
 -- ступенчатые разрывы 238
 -- супервитки (суперспирали) отрицательные 32, 33, 412–413
 -- положительные 32, 33, 412–413
 -- сцепленная молекула 445
 -- топология и перестройки 410–412
 -- упаковка 344–349. *См. также* Хромосомы
 -- участки, связывающие белки 141, 142, 145
 -- фага лямбда 212
 -- В-форма 27
 -- фрагменты 250–252. *См. также* Фрагменты Оказаки
 -- функция 55–57
 -- химерная 238–241
 -- хлоропластов 282
 -- центрифугирование в градиенте плотности 300–301
 -- частота повторяемости нуклеотидных последовательностей 226
 ДНКаза I 365–367, 388, 390, 392
 ДНКаза II 365, 366, 390
 ДНК-зонд 231
 ДНК-копии *см.* кДНК
 ДНК-лигаза 418, 419
 ДНК-полимераза фага Т4 428
 ДНК-полимеразы прокариот 414–416
 -- эукариот 413–414
 -- I, II, III 414, 415, 423
 -- α , β , γ 414
 ДНК-топоизомеразы 410
 Доминантность 18
 -- аллеля 9
 Доминантные признаки 8
 Доминирование неполное (частичное) 9
 Дрейф генов 275
 Дрожжи гетероталлические и гомоталлические 486
 -- жизненный цикл 486
 -- локусы типа спаривания 485–489
 -- структура теломеры 354
 -- точки начала репликации 403–404
 -- трансформация 403–404
 -- элементы Ту 472–474
 Дупликация 274
- Единица генетической карты 42
 -- мутации 19
 -- рекомбинации 15, 19
 -- *Сведберга* (S) 65
 -- транскрипции (транскрипционная единица) 139, 289, 292
 -- у фага лямбда 169
 -- функции 19
 -- эволюционного времени (ЕЭВ) 275
 Естественный отбор и накопление вредных мутаций 295–296
- Закон независимого расщепления 9
 Зародышевые линии (клеток) и трансфекция 501–502
 -- иммуноглобулиновые гены 506–508
 Зигота 8
- Изолейцин 56
 N⁶-Изопентениладенозин (i⁶A) 88
 Изопептидная связь 385
 Изопропилтиогаляктозид (ИПТГ) 176, 181, 185
 Изосхизомеры 386
 Икосаэдральные фаги 206
 Икосаэдрическая симметрия 345
 Иммуниет 213
 -- транспозиции 469
 Иммуности область 213
 Иммуный ответ 502
 Иммуноглобулин 502, 503
 Иммуноглобулиновые гены 504–508, 512
 -- переключение классов 512–513
 Инверсии, влияние на амплификацию 495–497
 -- двойные 465
 -- и фазовая вариация (у сальмонелл) 470–472

- Инверсионная модель рекомбинации между V- и C-генами 510
 «Инверсия знака», модель 413
 Инвертированные (обращенные) повторы (палиндромы) 163–166. *См. также* Шпильки
 -- короткие, концевые 460, 465
 -- ренатурация 299–300
 Инвертируемый сегмент G у фага Mu 470
 Ингибирование конечным продуктом 190, 191
 Индуктор(ы) 176, 179
 – «добровольные» 177, 181
 – связанные с ДНК 185, 186
 Индукция, контроль 176–177, 188–189
 – парадокс 187–188
 – профага 206
 Инициации факторы 74–78
 Инициация и амплификация 496
 – кодон 63
 – контроль 168, 173
 – синтеза ДНК 421–422
 – репликации 425–427
 – транскрипции 338, 426
 Инозин 88, 96
 Инсерции 38
 Инсерционные последовательности (IS) 460–461
 Интеграция профага 206, 453
 – ретровирусов 492
 Интерфаза 10
 – мейоза 13
 Интерфазный хроматин 349–352, 373
 Интерферон 279. *См. также* Кластеры генов интерферона
 Интрон *см.* Экзон–интрон, граница
 – регуляторного белка 258–262
 – спаривание с антикодоном 318
 Интроны 52, 53, 246–262. *См. также* Прерывистые гены
 – генов, кодирующих рРНК и тРНК 254–255
 – порядок удаления при сплайсинге 323–324
 – тРНК, спаривание с антикодоном 318
 – эволюция 255, 265
 Инфекция 206
 Информационная РНК (матричная РНК) *см.* мРНК
 Информация генетическая, история вопроса 55
 ИПТГ *см.* Изопропилтиогаляктозид
 Исключение профага 453, 454
 – транспозона 465–466
- Капсид 345
 – аденовируса и фага T4 344
 Карбоксифенилаланин-1-дезоксирибулозофосфат (КДРФ) 190
 Каркасные (FR) области 503
 Карта безнуклеосомного пробела 391
 – генома митохондрий (дрожжей и человека) 285, 286
 -- фага лямбда 211, 212
 ---- T4 209
 – рекомбинации 42
 -- гена триптофан-синтазы 51
 – сцепления 16, 17
 Картирование генов 15, 42, 246–252
- R-петель 246–248
 Кассетная модель 486–488
 Катаболитная репрессия 201
 – экпрессия 201
 Катализ биохимический, понятие 321–322
 Катенан 412
 Катящееся колесо 404–406
 Квеозин 89
 кДНК (ДНК-копии) 241–242
 – подавление трансляции 242
 – рестрикционная карта 250
 КДРФ *см.* Карбоксифенилаланин-1-дезоксирibuлозофосфат
 Кинетохор 352
 Кистозный фиброз 48
 Кластеры генов 177
 -- внезапная коррекция, модель 296
 -- глобиновых 269–275
 -- и неравный кроссинговер 270–271
 -- интерферона 279
 -- координированная регуляция 177
 -- тандемных 289–298
 --- тРНК 297–298
 -- фиксация при кроссинговере 296
 -- эволюционное древо 274–275
 -- эволюция 295–296
 Кластеры мутаций синтеза цитохрома b 258–259
 – рДНК 293
 – сходных последовательностей ДНК 267–308
 – тандемных генов 289–298
 Клеверный лист, структура 67
 Клетка «типичная» 10
 Клетки соматические 10, 55
 Клетки HeLa 282, 344
 L-Клетки 282
 Клеточная масса, влияние на репликацию 401
 Клеточная теория наследственности 10
 Клеточное деление *см.* Клеточный цикл
 -- согласованность с репликацией ДНК 399–402
 Клеточные органеллы 280–288
 Клеточный цикл 10
 – эукариотической клетки 402
 Клон (ы) 236–244
 – геномной (хромосомной) ДНК 243
 – кДНК 242
 Клональная селекционная теория 504
 Клональный анализ 483–484
 Клонирование ДНК 236, 243–244
 -- в бактериях 236–238
 -- индивидуальных генов 242–243
 Клонирование векторы 236. *См. также* Плазмиды; Фаговые векторы
 Код генетический *см.* Генетический код
 Кодовое отношение 57
 Кодоминантность 9
 Кодон (ы) 57–60
 Кодоны-синонимы 60
 Коинтеграция 466
 – разрешение 466–469
 Коинтегрированная структура *см.* Коинтеграция; Репликоны, слияние
 Колинеарность генов и белков 51–52
 Коллаген 264
 Кольца Бальбиани 357
 Кольцевая хромосома 380
 Кольцевые геномы (кольцевые ДНК) ди-
 --мерные и мультимерные 447
- образование θ -структуры 398
 -- одноцепочечные, превращение в двух-
 цепочечные 421
 -- органелл 282
 -- разрезание «восьмерки» 448
 Комплементарная (–) цепь 421–422
 Комплементационный анализ 421
 Комплементация генетическая 19, 41. *См. также* Межаллельная комплементация
 Конверсия гена 452–453
 -- совместная 453
 Конкатемерные молекулы 345, 346
 Консервативная рекомбинация 443
 Константная цепь 502–503
 Конститутивные гены (гены «домашнего хозяйства») 235
 – мутации 178–181
 Конститутивный способ выражения гена 179
 – хроматин 350
 Контролирующие элементы 481–482
 Контртранскрипт 407
 Конформация белка 55
 Концевая метка (концевое мечение) 46, 365, 366, 388, 389
 Концевое присоединение, метод 240
 Концевые повторы длинные (LTR) 461, 491, 492
 -- короткие инвертированные 460
 Конъюгация хромосом 11
 Кооперативное связывание 214–216
 Координированная регуляция 177
 -- модификация 191
 Копийность (представленность) 234
 Кордицепин *см.* 3'-Дезоксиаденозин
 Корепрессоры 176, 189
 Короткие последовательности 229
 Кор-последовательности (коровые последовательности), определение 454
 – ступенчатый разрыв и воссоединение 454–456
 Кор-частицы *см.* Нуклеосомный кор
 Космиды 238
 Козюльды 277
 Кроссинговер 14, 16, 271. *См. также* Неравный кроссинговер
 – в кор-последовательности 454
 – фиксация (закрепление) 296, 308
 Крисс-кросс 14
 Кукуруза, контролируемые подвижные элементы 481–482
- Лактозные (*lac*) гены 176–188
 Левоспиральная и правоспиральные формы ДНК 30
 Лейцин 56
 Лигирование («сшивание») РНК 318
 – турых концов 240
 Лидерная последовательность 194
 Лидерные пептиды 192, 194, 197
 Лизин 56
 – в гистонах 359, 360
 --- ацетилирование 384
 Лизис 205
 -- быстрый 17
 Лизогения, аутогенный цикл 210–213
 – против лизиса 219–220
 Лизогенная репрессия 214
 Лики (*leaky*)-мутации (мутанты) 57–58
 В-Лимфоциты 503, 504
 Липкие концы 238, 241
 -- комплементация 238

- β-Липотропин 342
 Литическая инфекция 205
 Литический каскад 205–220
 -- системы 211
 -- фага T4 210
 – цикл 206, 213, 453
 Лocus *Antennapedia* (*Antp*) 263, 264
 – *bithorax* 262–263, 476
 – *fushi tarazu* (*ftz*) 264
 – *MAT* 488–489
 – *Notch* (*N*) 263
 – *rudimentary* 476
 – *white* (*w*) 476–480
 Локусы гомогенно окрашенные (HSP) 497–499
 – летальные 358
 – молчание и активные (у дрожжей) 485–488
 – сложные 262–264, 476–480
 -- рестрикционное картирование 476–477
 – старение 17
 Магические пятна 204
 Малые цитоплазматические РНК (мцРНК, scRNA) 328
 – ядерные РНК (мяРНК, snRNA) и сплайсинг 328–331
 Маркеры генетические 46
 Материнское наследование 281
 Матрица при синтезе ДНК 29
 Матричная РНК *см.* мРНК
 мДНК 235
 Межаллельная комплементация 19, 184
 – рекомбинация 452–453
 Межплоскостные взаимодействия (стэкинг) 27, 29
 Мейоз 11, 14
 – поведение хромосом 354
 – схема 13
 – участие каждой цепи ДНК 452. *См. также* Постмейотическая сегрегация
 Мейотический бивалент 355. *См. также* Синаптонемальный комплекс
 Меланотропины (МСГ), α, β, γ 342
 Менделя работы 8, 10
 – законы 8, 9, 14
 Метаболические реакции, генетический контроль 18
 Металлотионеин 497
 Метафаза мейоза 13
 – митоза 12
 N⁶Метиладенозин 88
 Метилаза 40, 433
 7-Метилгуанозин 88
 Метилирование ДНК 387–388
 -- в системе защиты 431–433
 5-Метилцитозин 40
 Метионин 56
 Метод насыщающей гибридизации 232
 Метотрексат (*mtx*) 497–499
 Миграция ветви 445, 446
 Миеломы 504
 Микрогетерогенность 290
 Микронуклеус тетрахимен 320
 Микрохромосомы двойные 497–500
 Минимальные хромосомы (мини-хромосомы) 362–366, 386
 -- вируса SV40 380, 391
 -- переход в кольцевую структуру 365
 Миоглобин 274
 Миссенс (*missens*)-мутации 61–62
 – глобиновых генов 272–274
 – горячие точки *см.* Горячие точки
 – единица *см.* Единица мутации
 – индукция 38, 39
 -- фагом Mu 469
 – индуцированные и спонтанные 37
 ---- молекулярная основа 37–38
 – лики *см.* Лики (*leaky*)-мутации
 – локуса *white* 476–480
 – миссенс *см.* Миссенс (*missens*)-мутации
 – митохондриальных генов 284, 285, 287
 – молчание 41
 – неиндуцибельные 181, 182
 – нейтральные 275
 – нонсенс *см.* Нонсенс (*nonsens*)-мутации
 – опал *см.* Опал (*opal*)-мутации
 – остановки репликации 425, 427
 – охра *см.* Охра (*ochre*)-мутации
 – петит (*petite*) у дрожжей 285, 288
 – рецессивные 18
 – сдвига рамки считывания 38
 – соматические 515–516
 – супрессорные *см.* Супрессорные мутации
 – типы 41
 – точковые *см.* Точковые мутации
 – частота 41–42
 – ядерных генов, влияние на митохондриальные гены 284
 Мутация
 – термин 13
 – *anterobithorax* (*abx*) 262
 – *bithoraxoid* (*bdx*) 262, 263
 – *contrabithorax* (*cbx*) 262, 263
 – *dhfr* 497–500
 – *Hyperabdominal* (*Hab*) 262
 – *infraabdominal* (*iab2*) 262
 – *Miscadestrial pigmentation* (*Mcp*) 262
 – *Notch* (*N*) 263
 – *ocellules* 495–497
 – *postbithorax* (*pbx*) 262, 263
 – *Ultrabithorax* (*Ubx*) 262–264
 Мхи, гаплоидное состояние 9
 мцРНК (scRNA) *см.* Малые цитоплазматические РНК
 мцРНК (scRNP) 328
 Мышь, эволюция сателлитной ДНК 304
 мяРНК (snRNA) *см.* Малые ядерные РНК
 мяРНК (snRNP) 328, 331
 Нанометр 510
 Наследование, сцепленное с полом 14
 Наследственность, элементарный фактор 8
 Неаллельные копии 269
 Негистоновые белки (негистоны) 360–372
 -- и чувствительность хроматина к ДНКазе 383–384
 Негистоновый α-белок 378
 Нейтральные замены 41
 Неоднозначное соответствие, гипотеза 94–96
 Непермиссивные штаммы 61
 Неповторяющиеся последовательности (неповторяющаяся ДНК, уникальные последовательности) 225–226
 -- гибридизация с мРНК 232
 -- синтез мРНК 231
 -- структурных генов 231–232
 Неравный кроссинговер 270–272, 306–308

- и образование иммуноглобулиновых генов 511
 Нерцепрокные рекомбинантные хромосомы 271
 Нетранскрибирующиеся спейсеры 289, 293, 294
 Ники 366
 Нитевидные фаги 206
 Нитроцеллюлоза, нитроцеллюлозный гель (фильтр) 242, 243
 Нозерн-блоттинг 243, 324
 Нонсенс (*nonsens*)-мутации (мутанты) 61–62, 99
 Нонсенс-супрессоры 98
 Нуклеаза микрококков 360–361, 363, 366, 390
 -- специфичность 378–379
 Нуклеазы 365–367
 Нуклеация, центр 345
 Нуклеиновые кислоты, гибриды *см.* Гибридизация нуклеиновых кислот; Гибридизация РНК—ДНК; Отжиг
 -- длина 344
 Нуклеозид 25, 26
 Нуклеоид (бактерий) 92, 344, 347–349
 Нуклеоплазмин 372
 Нуклеосомные мономеры 363
 Нуклеосомный кор (минимальная нуклеосома), консерватизм структуры 362–363
 Нуклеосомы (нуклеосомные частицы) 358–375. *См. также* Фазирование нуклеосом
 - в активном хроматине 376–393
 - в сверхчувствительных участках 390
 - модель 368
 - определение 360
 - репродукция 369–372
 - сборка 372
 Нуклеотидные последовательности, делегирование и транспозиция 465
 -- дивергенция 275–278
 -- идентичные 308, 350
 -- канонические и сплайсинг 325–328
 -- контролируемых элементов 481
 -- локусов спаривания 486–488
 -- митохондриальной ДНК 286
 -- сателлитной ДНК 302–304
 -- функционально активные 277
 -- центромеры 353
 -- Alu 298–299
 Нуклеотиды 25, 27
 Область иммунности 213
 Обратные мутации 41
 Обращенные повторы *см.* Инвертированные повторы
 Овомукоид 324, 333
 Однопочечное перемещение 411
 Однопочечные обмены 438
 Онкогенез 492
 Онкогены 492–493
 Опал (*opa*)-мутации (мутанты) 61
 Оператор 178, 182–184
 - арабинозный (*ara*) 197–200
 - связывание репрессора 182–184, 214
lac-Оператор 183
 Оперон(ы) 176–192. *См. также* Лактозный оперон; Триптофановый оперон
 - дерепрессированный 189
 - классическая модель 178
 - регуляция 188–205
 - репрессибельный 189
 - *rrn* 296–297
 - *trp см.* Триптофановый оперон
 - *tyrU* 297, 298
 Ослабленный контроль 204
 Основания азотистые модифицированные 88
Bam-Островки 294
 Отжиг 36
 «Отпечатков», метод 142
 Охра (*ochre*)-мутации (мутанты) 61
 Палиндромы *см.* Инвертированные (обращенные) повторы
 Парадокс величины С 222
 Паранемное соединение 450
 Пентозы в ДНК и РНК 24
 Пептидная связь 55
 -- образование 57, 82–83
 Первичная структура *см.* Полипептидная цепь, структура
 Первичный транскрипт 246
 Перекрывающиеся гены 53–54
 Пермеаза 179
 Пермиссивные и непермиссивные условия 42
 Пермиссивные и непермиссивные штаммы 61
 R-Петли, картирование 246–248
 D-Петля 90, 286, 404–406, 449
 TFC-Петля 90
 Пиримидиновые основания (пиримидины) 25
 Плавление ДНК *см.* ДНК, денатурация
 Плазмацитомы 504
 Плазмидные векторы 237, 239, 240
 Плазмиды 179, 206, 458
 - гибридные (химерные) 236, 244
 - мультикопийные 237, 396
 - несовместимость 406–408
 - однокопийные 237, 396
 - ослабленный контроль репликации 237
 - «челночные» 238
 Плектонемное соединение 450
 Пневмококки, вирулентность 22
 Поверхностное исключение 406
 Повторяющаяся единица 289, 303–306
 Повторяющиеся последовательности 225–229. *См. также* Тандемные повторы
 Подвижные (мобильные) элементы *см.* Транспозоны
 -- автономные и неавтономные 481, 482
 -- способные к перемещению в пределах генома и вне его 490–518. *См. также* Ретровирусы; Тканеспецифичные вариации; Трансфекция
 -- у кукурузы 481–482
 -- у зукариот 472–489
 -- Alu 494
 -- *coria* 474, 475, 477, 491
 -- Ds, способность к транспозиции или разрывам хромосом 482–485
 -- FB 475
 -- IS 460–462
 -- номенклатура 460
 -- IS-подобные 461
 -- IS10R 463
 -- IS50L 464
 -- P 475, 480–481
 -- Tu 473–474
 Поздние гены 207, 211, 212
 -- аденовируса 257
 Полиаденилирование 335–336
 Полиморфизм генетический 46, 275
 - рестрикционный 48
 Полинуклеотидная цепь 25, 26
 Полинуклеотид-фосфоорилаза 60
 Полипептидная цепь, конформация 55
 Полипротеины 341–342
 Полирибосомы (полисомы) 69
 Политенные хромосомы 355–357
 -- и амплификация 495–496. *См. также* Кольца Бальбиани; Хромосомные пуфы
 Полифункциональные ферменты 434
 Половые клетки *см.* Гаметы
 Положительная интерференция 16
 Полуконсервативный синтез ДНК 29, 396–398
 Полунепрерывная репликация 418
 Полухроматиды 11
 ПОМК *см.* Проопиомеланокортин
 Постмейотическая сегрегация 452
 Пострепликационная репарация 438
 Праймаза 421–422
 Праймер 241
 Праймосома 422–425
 Предранние гены 207, 211
 Предпраймерная РНК 423
 Прерывистая репликация 417
 Прерывистые гены 52–53, 246–254, 264–265. *См. также* Интроны; Экзоны
 -- рестрикционное картирование 248–254
 -- электронная микроскопия 246–248
Прибнова блок (последовательность) 200
 Признаки доминантные и рецессивные *см.* Доминантные признаки; Рецессивные признаки
 Прокариоты 21
 Пролин 56
 Промотор, определение границ 151
 - точки контакта с РНК-полимеразой 146–147
 Промоторные мутации 144–146
 Промоторы для РНК-полимеразы 153
 Проопиомеланокортин (ПОМК) 341, 342
 Простые последовательности 227
 Протостомы, размер генома 223
 Протуберанец 104
 Профаг 206
 Профаза мейоза 13
 - митоза 12
 Процессинг РНК 309–342
 -- его регуляция 331–343
 -- и экспрессия генов 513–515
 - транскрипта как этап экспрессии гена 338
 Процессуемые псевдогены 278
 Прямые мутации 41
 - повторы 460, 465
 Псевдогены 278–279, 294–295, 494
 Псевдоурдин 88
 Птицы, эволюция глобиновых генов 274
 Пуриновые основания (пурины) 25
 Пуфы *см.* Хромосомные пуфы
 Разрешение 468
 Разрыв – воссоединение 445–446
 Разрыв – слияние – мост 484
 Рамка считывания 58. *См. также* Мутации рамки считывания
 -- автономных элементов 483
 -- открытая 63

- при сплайсинге 255, 261
- Ранние гены 207, 212
- вирусов SV40 и полиомы 257
- рДНК 292–295, 354, 476
- макронуклеуса 392
- Реверсии 13, 17
- в локусе *w* 479
- вторичные 41
- истинные (истинные обратные мутации) 41
- Регулон 197
- Регуляторные белки 215, 392
- гены 178, 392
- Регуляторный фактор 156
- Резолваза TrpR 467
- Рекомбинантная РНК 236
- Рекомбинантный фенотип 9–10
- Рекомбинационная единица *см.* Единица рекомбинации
- инверсия 472
- карта 51
- репарация 438, 440
- Рекомбинационное расстояние 42, 48
- Рекомбинационные утолщения 444
- Рекомбинация 443–456. *См. также* Нерцепицидная рекомбинация; Реципрокная рекомбинация
- ассимиляция (поглощение) одной цепи 449, 450
- генетическая 16
- частота 42–43
- гомологичная 493
- ДНК и переключение классов иммуноглобулинов 512–513
- и транспозиция 467. *См. также* Репликативная рекомбинация
- инициация 446–447
- консервативная 443
- негомологичная 493, 499
- незаконная 454
- органелл 287
- промежуточные продукты 447–448
- соматическая 503
- между двумя двухцепочечными ДНК 444–445
- профагом и бактериальной ДНК 454
- V- и С-генами 509–511
- общая 443
- Релаксация 413
- Ренатурация ДНК *см.* ДНК, денатурация и ренатурация
- Релаксированное кольцо ДНК (РФ II) 421
- Репарационные системы 442
- Репарация (повреждений ДНК) 431. *См. также* Пострепликационная репарация
- короткими и длинными последовательностями 439
- механизм 436–438
- эксцизионная 437–440
- SOS-Репарация 441
- Репарирующая система 431
- Репликативная рекомбинация 443
- и транспозиция 464–466
- Репликационная вилка 397–398, 402, 404, 416–417
- при амплификации 496
- Репликационный аппарат 421–425
- фагов 427–429
- Репликационный глазок 397–399
- Репликация ДНК *см.* Амплификация; Полунепрерывная репликация; Прерывистая репликация; Репликационная вилка; Репликационный глазок; Репликон; Точки начала репликации; Точки окончания репликации
- белок-инициатор 401
- и ацетилирование гистонов 384, 385
- и клеточное деление 399–402
- и структура хроматина 370–371, 376
- механизм 30
- нуклеосом 370–371
- однонаправленная и двунаправленная 397–399
- плазмиды 407
- реинициация 498, 499. *См. также* Амплификация
- скачкообразная 304
- топология 409–419
- ферментативный аппарат 420–431. *См. также* Репликационный аппарат
- Ds-элементов 484–485
- Репликоны 396–408
- линейные 429–430
- слияние 466
- Реплисомы 420
- Репрессия, контроль 188–189
- мультивалентная 196
- ферментов 176–177
- Репрессор 178–187. *См. также* Белок-репрессор
- взаимодействие субъединиц 184
- димеры 213–214, 217
- отделение от ДНК 185–186
- связывание 181–185, 214–216
- на фильтрах 181
- синтез 216–218
- узнавание операторов 182
- фага лямбда 211–218
- С- и N-концевые фрагменты 184–185, 187
- cI-Репрессор 211
- trp-Репрессор 198
- Рестриктаза Bam 294
- EcoRI 239
- HpaII 386, 387
- MspI 386, 387
- Рестриктазы 43–46, 238, 431–436
- Рестрикционная карта tandemных генов 290. *См. также* Карта рестрикции
- Рестрикционное картирование 43–49
- методом «прогулки по хромосоме» 477, 478
- прерывистых генов 248–252
- сложных локусов 476–477
- Рестрикционный полиморфизм *см.* Полиморфизм рестрикционный
- Рестрикция ДНК 44–49. *См. также* Сайты рестрикции
- и модификация оснований 432–433
- системы 431
- частичная 243
- Ретикулоциты, факторы инициации 78–79
- Ретровирусы, жизненный цикл 490–492
- способность к трансдукции 492–494
- Рецепторный сайт 339, 340
- Редессивные аллели 9
- признаки 8
- Реципрокная рекомбинация 10, 271, 447, 458. *См. также* Нерцепицидная рекомбинация
- в коровых последовательностях 455
- и однонаправленная транспозиция 488
- исключение профага 453
- между инвертированными повторами 465
- Рибоза 25
- Рибонуклеазы 310. *См. также* РНКаза II; РНКаза III РНКаза D; РНКаза Р
- Рибонуклеиновые частицы *см.* мцРНП; мяРНП
- Рибосома, субчастицы, форма 104, 105
- 30S-субчастица 104, 109
- Рибосомные белки (р-белки) 108, 202
- контроль трансляции 201–203
- Рибосомные РНК *см.* рРНК
- Рибосомы 59
- бактерий и эукариот 65
- Рибосомы, белки *см.* Рибосомные белки
- и терминация аттенуации 192
- метод их связывания 60
- митохондрий 65, 283–284
- регуляция конформации trp-мРНК 195
- самосборка 110–111
- синтез белков 64–65, 284–285
- строение 64–65
- участки связывания с тРНК 73
- хлоропластов 65
- Риботимидин 88
- Рибулдизобисфосфат-карбоксилаза кукурузы 284
- Рифампицин 348
- РНК 65. *См. также* Малые цитоплазматические РНК; Малые ядерные РНК; РНК-зонд; РНК-полимераза; Сплайсинг РНК
- адапторная (транспортная) *см.* тРНК
- вируса табачной мозаики 345
- вторичная структура 33–35, 164
- двойная спираль, свободная энергия 34, 35
- гибридизация с ДНК 231–234. *См. также* Гибридизация нуклеиновых кислот; Гибридизация РНК–ДНК
- как генетический материал 24
- катализатор 321–322
- матричная *см.* мРНК
- ретровирусов 490–492
- рибосом *см.* рРНК
- синтез, образование пuffed 358
- созревание *см.* Процессинг РНК
- спаривание оснований 34
- стабильная 310–317
- транспортная *см.* тРНК
- упаковка в вирусах 345
- хроматина 358
- ядерная 300
- РНКаза II 317
- РНКаза III 311–315
- РНКаза D 317
- РНКаза H 408
- РНКаза Р 316–317
- РНК-активатор 339
- РНК-затравка 407–408, 428
- РНК-зонд 231
- РНК-матураза 259–262
- РНК-лигаза 318
- РНК-полимераза, взаимодействие с димерами 217
- гены 155, 202
- и нуклеосомы 379
- как негистоновый белок 360
- передвижение 154, 193
- промотеры 153–155
- связывание с ДНК 141, 145, 170, 171, 183

- синтез в цитоплазме 283
- субъединицы дополнительные 171–172
- функция 173
- РНК-посредник 51, 65. *См. также* мРНК
- рРНК 65, 291–296. *См. также* Интроны
- генов, кодирующих рРНК
- необычный сплайсинг 320–322
- процессинг 313–315
- сборка рибосом 203
- зукариот, созревание 314–315
- РФ *см.* ДНК, репликативная форма
- РФ1 *см.* Суперспирализованное кольцо ДНК
- РФII *см.* Релаксированное кольцо ДНК

- Сайт внедрения 477
- res-Сайт 468
- Сайтспецифическая рекомбинация 443, 453–454
- Сайты *см.* Рецепторный сайт; Сенсорный сайт
- замещения 275
- дивергенция 277
- метилированные 388
- молчание 275, 276
- присоединения, вторичные 455
- разрезания 433
- расщепления 314–315
- рестрикции 389
- как генетические маркеры 46–49
- сверхчувствительные к ДНКазе 388–392
- связывания ДНК в нуклеосоме 365
- РНК-полимеразы 140–143, 145
- скручивания 298
- узнавания 433
- att-Сайты 454–456
- cos-Сайты 345, 346
- Сальмонеллы 470–472
- Сантиморганда 15
- Сателлитная ДНК 300–306
- локализация 301
- млекопитающих 302–304
- околоцентромерных областей 352
- членистоногих 302
- эволюция 304–306
- α-Сателлитный хроматин 378
- Сборка макромолекул 203–204
- Сведберга единица *см.* Единица Сведберга
- Сенсорный сайт 339, 340
- Серин 56
- фосфорилирование 384
- Серповидноклеточная анемия 18
- Сестринские хроматиды 350
- Сигма-фактор 160
- Сигналы терминации 60–62
- Сигнальные последовательности 298
- Синапсис гомологических молекул ДНК и рекомбинация 443–445
- Синаптономемальный комплекс 443, 444
- Синтез ДНК, механизм 27, 29. *См. также* ДНК, синтез; Репликация ДНК
- Сканирование линкером, метод 152
- Смещение разрыва 415, 416
- Соматическая сегрегация 281
- Спаривание оснований 27
- ошибочное 38
- Сперматозоид 9
- Сплайсинг РНК 52, 53, 256, 317–331
- две стадии 318–320
- нарушения 328
- последовательность реакций 322–324
- Спорофит 9
- Споруляция 158
- Споры 9
- «Стебель» 34
- Строгий ответ 204–205
- Структурные гены 178, 230–236, 339–340
- внутренняя организация 245–266. *См. также* Интроны; Прерывистые гены
- организация родственных генов 268–280
- Стэкинг-взаимодействия *см.* Межплоскостные взаимодействия
- Суперрепрессирующие условия 189
- Суперспирализация двойной спирали 32–33
- Суперспирализованное кольцо (РФ1) 421
- Супрессия 8–10, 288
- Супрессор 58, 61
- Супрессорные мутации (внегенные супрессоры) 41
- Сцепление 14
- «Сшивание» (лигирование) тупых концов 240

- Талассемия 272–273
- Тандемные гены 289–298
- рРНК 291–296
- повторы 354
- Таутомерные формы азотистых оснований 27
- Теломеры 353–354
- Телофаза митоза 12
- Терминатор, зависимость от фактора ро 165
- Терминация преждевременная 61
- сигналы 60–61
- у зукариот 172–173
- Терминология 20–21
- Тест на комплементацию (*цис-транс-тест*) 19, 20
- Тетрамеры (нуклеосом) 368–370
- Тимин 25
- Тиминовые димеры 436
- 4-Тиоуридин 88
- Тип спаривания 485–489
- кассетная модель 486–488
- Тирозин 56
- Тканеспецифичные вариации 495–497
- Топоизомеразы 410, 450, 451
- бактерий 411–412
- Топологические изомеры 410
- Точка плавления ДНК 35, 36
- Точки начала репликации 298, 396, 400
- двухпочечной ДНК 425–427
- у дрожжей 403–404
- окончания репликации 396, 400
- Точковые мутации 37, 38, 40, 41, 46, 58, 145
- Транзиция и трансверсия 37
- Транс-активная функция 180–182
- Трансвекция 479
- Трансверсия 37
- Транс-доминантность 181, 182
- Трансдукция 454
- ретровирусами 492–494
- Транскрибирующийся спейсер 292, 293
- Транскрипт *см.* Первичный транскрипт
- процессинг и транспорт 338
- Транскрипционный матрикс 380
- Транскрипция 59. *См. также* Инициация транскрипции
- блокирование 181–182, 205
- генов фага SpO1 160
- единица *см.* Единица транскрипции
- и структура нуклеосом 376, 379–381
- обратная 494
- промоторы 139–156, 298
- регуляция специфичности 168
- стартовая точка 139, 140
- Транслокации 14
- Трансляция 59
- аутогенный контроль 201–204
- оперонов р-белка 203
- подавляемая гибридом 242
- при перекрывающихся рамках считывания 62–64
- у бактерий 65
- Транспозаза 466
- Транспозирующиеся элементы *см.* Транспозоны
- Транспозиция 274. *См. также* Подвижные (мобильные) элементы; Транспозоны
- и репликативная рекомбинация 464–466
- образование новой копии 459
- однонаправленная инициация локусом MAT 488–489
- определение 458, 464
- путем разрешения коинтеграта 466–469
- РНК-зависимая 494–495
- у бактерий 458–460
- у дрожжей 485–489
- Транспозон (Th), TnA 467
- Tn3 466–469
- Tn5 462–464
- Tn9 461, 462
- Tn10 462–463
- Tn903 462
- Транспозоны (транспозирующиеся элементы) 458–470
- в геноме кукурузы 481–485
- в геноме дрозофилы 474–476. *См. также* Гибридный дисгенез
- номенклатура 462
- сложные 461–463
- у зукариот 472–489
- Транспортная РНК *см.* тРНК
- Трансфекция 23, 25, 497, 500–502
- схема 26, 501
- Трансфераза терминальная 239
- Трансформация 22, 403, 404
- Трансформирующий фактор 22
- Треонин 56
- Третичная структура *см.* Полипептидная цепь, структура
- Трипсин 184
- Триптофан 18, 56
- Триптофановый (*trp*) оперон 190–197
- Триптофан-синтаза 51
- тРНК 59. *См. также* Интроны генов, кодирующих рРНК и тРНК
- дрожжей, сплайсинг 318–320
- инициаторная 74
- как адаптор 67–68
- мутантные 98–100
- нонсенс-супрессорные 98
- процессинг 310, 313, 315–317
- структура 67
- клеверного листа 87
- третичная 90, 91
- супрессорные 100–101
- «сшивание» (лигирование) 318, 319

- узнавание синтетазой 94
- условные наименования 68
- тРНК-нуклеотидилтрансфераза 316, 317

- Убиквитин (U) 385–386
- Умеренно повторяющиеся нуклеотидные последовательности 227–228
 - синтез мРНК 231
 - чередование с неповторяющимися последовательностями 229
 - Alu 298–299
- Уникальные гены 232
- Уникальные последовательности 226. *См. также* Неповторяющиеся последовательности
 - на границах экзон-интрон 255–256
- Урацил 25
 - вырезание из ДНК 418
- Урацил-ДНК-гликозидаза 40, 417
- Уридин 96
- Условно-летальные мутации 42
- УФ-индуцированный мутагенез 440

- Фаг(и) лямбда (λ) 169, 170, 208–210
 - внедрение в ДНК клетки 453
 - литический цикл 453
 - процессинг мРНК 312
 - созревание 451
 - упаковка ДНК 345, 346
 - λ 344, 450
 - Mu транспозирующий 469–470
 - M13 450
 - T2 23
 - генетический материал 24
 - T4, литический каскад 208–210
 - репликационный аппарат 427–429
 - упаковка ДНК 345, 346
 - T7, геном 208
 - процессинг мРНК 311–313
 - репликационный аппарат 429
 - SPO1, гены транскрипции 160
 - ϕ X174 63, 450
- Фаги, кодирование ДНК-полимераз 416
 - литическое развитие 206–207
 - морфогенез 345–347
- Фаговые векторы 238
- Фазирование нуклеосом 376–378, 392
- Фазовая вариация 470–472
- Фактор диссоциации 76
 - строгого контроля 204, 205
 - элонгации 83
 - ρ (ρ) и терминатор 165–167
 - σ F2 79, 88
 - IF3 76
- Факультативный хроматин 350
- Фенилаланин 56
- Фенотип дикий и мутантный 12
 - обозначения 20
- Ферментный круговорот 413
- Ферменты рестрикции *см.* Рестриктазы
- Флагеллин, флагеллиновый ген 470–472
- Фланкирующие последовательности 223, 492
- Формилметионин 74
- Фосфодиэфирная связь 311, 318
 - перенос при сплайсинге 323
- Фотореактивация 437
- Фрагменты Оказаки 417–419

- Хиазма 14, 16, 444
 - образование 16
 - хромосом типа ламповых щеток 355
- Химерные плазмиды 237, 244
- Хлорамфеникол 284
- Хлоропласты *см.* Геном хлоропластов; Рибосомы хлоропластов
- Холостое взаимодействие 204
- Хорея Гентингтона 48
- Хроматиды сестринские 10
- Хроматин 10, 344, 349–352. *См. также* Конститутивный хроматин; Факультативный хроматин
 - дискретные частицы 360–361
 - в нуклеосомах, репродукция (репликация) 369–372
 - домены, чувствительные к ДНКазе 381–382
 - нити размером 10 нм 372–373
 - 30 нм 374
 - петли, домены и остов 373
 - роль в активации гена 392–393
 - синаптомембранного комплекса 443, 444
 - структура 358–375
- Хромеры 354
- Хромосома, лишённая гистонов 374, 375
- X-Хромосома, амплификация генов 495
 - внедрение нуклеотидных последовательностей 476
- Хромосомная теория наследственности 10
- Хромосомные диски и междиски 32, 356
 - соответствие генам 358
 - пuffy 357–358, 495. *См. также* Кольца Бальбиани
- Хромосомный набор человека 351
- Хромосомы 8, 10, 42, 344–358. *См. также* Интерфазные хромосомы; Микрохромосомы; Митотические хромосомы
 - в соматических клетках 55
 - гигантские *см.* Политенные хромосомы
 - гомологичные 10
 - дифференциальное окрашивание 350–352
 - независимая сегрегация 12
 - половые 13, 14. *См. также* X-хромосома
 - синапсис 443
 - скручивание 298
 - типа ламповых щеток 354–357
 - число 10–11
 - репликаонов в них 402–403
- Хромосомный центр 350, 356

- Центромера 352–353
 - локализация сателлитной ДНК 301
 - политенных хромосом 356
- N-Цепь ДНК 286
- L-Цепь ДНК 287
- Циклогексимидаз 284
- Цис-доминантность 180–181
- Цистеин 56
- Цис-транс-тест *см.* Тест на комплементарность
- Цистрон 18–20
- Цитозин 25, 96
- Цитокинез 12
- Цитоплазматическая (внеядерная) наследственность 281. *См. также* Митохондриальный геном; Геном хлоропластов

- M-Цитотип 480–481
- Цитохром, комплекс bc_1^M 284
 - b 284
 - мутации, кодирование 258–261, 285
- Цитохромоксидаза 258–259, 285
- Цитохром-c-оксидаза 284

- Человек, митохондриальный геном 286
 - сплайсинг РНК 329
 - ядерная РНК 300
- Четвертичная структура *см.* Полипептидная цепь, структура
- Членистоногие 302

- «Шотган»-эксперименты 243
- Шпильки 34, 163–165, 173, 300
 - терминаторные 192–195

- Эволюционное древо 274–277
- Эволюция и накопление вредных мутаций 295–296
 - и дивергенция нуклеотидных последовательностей 275–278
 - и псевдогены 278–279
 - нулевая 276
 - сателлитной ДНК 304–306
 - согласованная (сопряженная) 277
 - тупики 278–279
- Экзон-интрон, граница 255–256, 324–325
- Экзоны 52, 53, 246, 251–254
- Экспрессия генов, альтернативные пути 54
 - и деметилирование ДНК 386–387
 - основной (репрессированный) уровень 191
 - предположения о природе 392
 - регуляция 175–220
 - модели 337–341
 - уровни 234
 - эукариотических в бактериях 244–245
- Экспозиционная репарация 437–440
- ЭЛГ 342
- Электронная микроскопия генов 246–249
 - ДНК и хроматина 360, 373, 374, 380
 - индивидуальных молекул ядерных РНК 323
 - хромосомы 375
- Эндонуклеазы 433, 489
- Эндорибонуклеаза 312
- Энергетические изменения 281
- Энкефалин 341
- Эритромицин 284
- Эритроциты эмбриональные 268
- Эукариотические геномы *см.* Геномы эукариот
 - клетки 22
- Эухроматин 301, 349–350
- Эффект положения 497

- Ядерные гены 253
 - сравнение с генами органелл 284
- Ядерный матрикс 333
 - гидролизат, электрофореграмма 366
- Ядро, окраска по Фельгену 351
- Ядрышко 293
- Ядрышковые гены 253
 - организаторы 293
- Яйцеклетка 9

Указатель латинских названий

- Aplysia* 342
Ascomyces 452
Aspergillus nidulans 284
- Bacillus subtilis** 416
- Caenorhabditis elegans* 222
Canis domesticus 292
Chironomus tentans 357
Chlamidomonas reinhardii 254, 281, 282, 287
- Dictyostelium** 354
– *discoideum* 222, 291–294
– *dodiscoideum* 254
Drosophila melanogaster 12, 14, 17, 18, 32, 46, 223, 244, 254, 262–264, 282, 289, 291, 296, 356–358, 374, 380, 386, 388, 389, 402, 403, 425, 474–482, 491, 495–497
– *virilis* 302
- Escherichia coli* 17, 23, 38, 40, 46, 59, 61, 62, 75, 103, 176, 190, 222, 244, 274, 310, 311, 316, 317, 321, 344, 348, 349, 398–400, 403, 411–416, 418, 421, 425–435, 437–441, 452, 469, 470
- Euglena gracilis* 282
- Haemophilis parainfluenzae** 46
- Lei chmania* 354
- Mus musculus** 292, 301, 303
- Neotellia** 444
Neurospora 18
– *crassa* 254, 284
Nicotiana tabacum 292
Notophthalmus viridescens 291, 292, 355
- Oxytricha** 392
- Paramecium** 429
Pisum sativum 282
Physarum 354
– *polycephalum* 254
- Rhynchosciara** 495
- Saccharomyces** 354
– *cerevisiae* 222, 238, 254, 282, 284, 287, 288, 291, 292, 402–404, 472–474
Salmonella typhimurium 190, 399, 400, 458
- Strongylocentrotus purpuratus* 227, 228, 232, 284, 290, 291
- Tetrahymena** 331, 354
– *pigmentosa* 254
– *pyriformis* 392
– *thermophila* 320–322
Trypanosoma 354
- Ustilago maydis** 451
- Xenopus** 319, 331, 372, 403
– *borealis* 295
– *laevis* 124, 228, 264, 274, 282, 289, 291–296, 336, 378, 386

Оглавление

Предисловие редактора перевода	5	Определение нуклеотидной последовательности ДНК	49
Предисловие	6	Коллинеарны ли гены и белки?	51
ЧАСТЬ I. ПРИРОДА ГЕНЕТИЧЕСКОЙ ИНФОРМАЦИИ	7	Гены эукариот могут быть прерывистыми	52
Глава 1. Что такое ген? Генетическая точка зрения	8	Перекрывающиеся и альтернативные гены	53
Элементарный фактор наследственности	8	Что такое ген?	54
Независимость различных генов	9	Рекомендуемая литература	54
Роль хромосом в наследственности	10	Глава 4. Расшифровка генетического кода	55
Гены располагаются в хромосомах	12	ДНК нужна только для того, чтобы кодировать последовательность аминокислот	55
Гены линейно выстроены вдоль хромосом	14	Генетический код считывается триплетами	57
Генетические карты непрерывны	16	Аппарат для последовательного белкового синтеза	59
Один ген – один белок	17	Кодоны, соответствующие аминокислотам	59
Новое определение: цистрон	18	Природа сигналов терминации	60
Замечания, касающиеся терминологии	20	Универсален ли код?	62
Рекомендуемая литература	21	Трансляция при перекрывающихся рамках считывания	62
Глава 2. Что такое ген? Биохимическая точка зрения	21	Рекомендуемая литература	64
Генетический материал – это ДНК	21	Глава 5. От гена к белку	64
Дальнейшие доказательства роли ДНК	23	Синтез белка происходит в рибосомах	64
Компоненты ДНК	24	Поиски посредника	65
ДНК – двойная спираль	26	Транспортная РНК – адаптор	67
Об альтернативных двуспиральных структурах	29	Рибосомы передвигаются, как конвоиры	69
Двойная спираль может подвергаться суперспирализации	32	Рекомендуемая литература	70
Суперспирализация влияет на структуру двойной спирали	33	ЧАСТЬ II. СИНТЕЗ БЕЛКОВ	71
РНК тоже имеет вторичную структуру	33	Глава 6. Конвейер для сборки полипептидных цепей	72
ДНК можно денатурировать и ренатурировать	35	Функциональные участки рибосомы	72
Нуклеиновые кислоты гибридизуются путем спаривания оснований	36	Инициация: специальная иницирующая тРНК	73
Молекулярная основа мутаций	37	В инициации принимают участие 30S-субчастицы и вспомогательные факторы	75
Мутации концентрируются в горячих точках	38	Недолговечная «свобода» 30S-субчастиц	76
Частота мутирования	41	Освобождение инициаторной тРНК	76
Рекомендуемая литература	42	В инициации у эукариот участвует много факторов	77
Глава 3. Что такое ген? Молекулярная структура	42	Важная роль фактора eIF2 в синтезе белка	78
Прямые исследования структуры гена	43	Последовательность событий в прокариотах и эукариотах	79
Рестриктирующие ферменты расщепляют ДНК на специфичные фрагменты	44	Элонгация: поступление аминоацил-тРНК в А-участок	79
Построение рестрикционной карты	45	Гидролиз GTP происходит после присоединения аминоацил-тРНК	81
Некоторые тонкости рестрикционного картирования	45	За образование пептидной связи ответственна рибосома	81
Сайты рестрикции можно использовать в качестве генетических маркеров	46		

Стадия транслокации	82	ЧАСТЬ III. СИНТЕЗ РНК	131
Извлечение энергии, необходимой для работы рибосомы	84	Глава 10. РНК-полимеразы – транскрипционный аппарат клетки	132
Терминация: завершение синтеза белка	85	Что представляет собой РНК-полимераза?	132
Рекомендуемая литература	85	Субъединичная структура бактериальной РНК-полимеразы	133
Глава 7. Транспортная РНК: трансляционный посредник	85	Сигма-фактор контролирует связывание РНК-полимеразы с ДНК	133
Универсальная структура клеверного листа тРНК содержит много модифицированных оснований	86	Рабочий цикл сигма-фактора	135
L-образная пространственная структура тРНК	87	Минимальный фермент синтезирует РНК	135
Синтетазы ответственны за подбор соответствующих друг другу аминокислот и тРНК	89	Функции субъединиц минимального фермента РНК-полимеразы фагов, возможно, являются «минимальными» ферментами	137
Стадия активации	91	Сложные эукариотические РНК-полимеразы	137
Кодон-антикодонное узнавание и гипотеза неоднозначного соответствия	92	Рекомендуемая литература	138
Модификация оснований может контролировать узнавание кодона	94	Глава 11. Промоторы: сайты инициации транскрипции	139
Митохондрии содержат минимальный набор тРНК	96	Определение стартовой точки <i>in vivo</i> и <i>in vitro</i>	139
Мутантные тРНК способны прочитывать различные кодоны	97	Сайт связывания РНК-полимеразы	140
Конкуренция между супрессорными и обычными тРНК	98	Консервативная последовательность в промоторах <i>E. coli</i>	143
Транспортная РНК может изменить рамку считывания	100	Промоторные мутации, усиливающие и ослабляющие экспрессию генов	144
Рекомендуемая литература	101	Основные точки контакта в промоторе	146
	102	Узнавание промоторов и расплетание двойной спирали ДНК	147
		Позитивная регуляция работы промоторов	148
		Возможные консервативные последовательности для РНК-полимеразы II	149
		Системы транскрипции <i>in vitro</i> и <i>in vivo</i>	149
		В системе <i>in vitro</i> РНК-полимераза II функционирует правильно	150
		Промоторы РНК-полимеразы II многокомпонентны	151
		Промотор РНК-полимеразы III расположен в самой транскрипционной единице	154
		Рекомендуемая литература	156
Глава 8. Рибосомы как фабрики белкового синтеза	102	Глава 12. Системные переключения иницирования транскрипции	156
Рибосомы – компактные рибонуклеопротеиновые частицы	103	Спорообразование	157
Структура рибосомной РНК	105	Сигма-факторы, специфические для различных стадий фаговой инфекции	159
Каждый рибосомный белок характеризуется специфической локализацией	107	Для каждого сигма-фактора может существовать своя собственная консервативная последовательность – 35 и – 10	159
Взаимодействие рибосомных белков и рРНК	107	Новая фагоспецифическая РНК-полимераза	161
Диссоциация и реконструкция рибосомных субчастиц	109	Рекомендуемая литература	161
Мутации, влияющие на самосборку рибосомы	109	Глава 13. Терминация и антитерминация	162
Порядок самосборки определяется пространственной организацией субчастиц	110	Обнаружение терминаторов в системе <i>in vitro</i>	162
Мутационные изменения могут затрагивать все компоненты рибосомы	111	Существуют р-зависимые и р-независимые терминаторы	162
Рибосомы содержат несколько активных центров	112	Немного об инвертированных повторах	163
Связывание 30S-субчастиц с мРНК	113	Достигнув полиндрома, минимальный фермент приостанавливается	164
Точность трансляции	114	Как работает фактор ρ?	166
Рекомендуемая литература	115	Мутации по гену фактора ρ	167
		Механизм антитерминации, контролируемый фаговым геномом	168
		Антитерминация зависит от определенных сайтов в ДНК	169
		Существуют ли дополнительные субъединицы у РНК-полимеразы?	171
		Трудности в изучении терминации у эукариот	172
		Рекомендуемая литература	173
Глава 9. Информационная РНК в качестве матрицы для синтеза белка	115	ЧАСТЬ IV. КОНТРОЛЬ ГЕННОЙ ЭКСПРЕССИИ У ПРОКАРИОТ	175
Недолговечность бактериальных мРНК	116	Глава 14. Оперон на примере организации лактозных генов	176
Строение бактериальной мРНК	116		
Трансляция полицистронной мРНК	116		
Функциональное определение эукариотической мРНК	119		
3'-конец эукариотических мРНК может быть полиаденилирован	120		
Выделение мРНК с использованием р _{оу} (A)-конца	121		
Эукариотические мРНК имеют метилированный «кэп» на 5'-конце	122		
Возможности трансляционных систем <i>in vitro</i>	123		
Для инициации, по-видимому, необходимо комплементарное взаимодействие между мРНК и рРНК	124		
Малые субъединицы могут перемещаться в сайты инициации эукариотических мРНК	126		
Связь белкового синтеза с внутриклеточной локализацией	127		
Рекомендуемая литература	130		

Индукция и репрессия контролируется малыми молекулами	176		
Кластеры генов регулируются координированно	177		
Регуляторный ген контролирует структурные гены	178		
Контролирующая система оперона	178		
Конститутивные мутации определяют действия репрессора	178		
Функция оператора <i>cis</i> -доминантна	180		
В промоторе или гене репрессора могут встречаться неиндуцибельные мутации	181		
Каким путем репрессор блокирует транскрипцию?	181		
Контакты в операторе	182		
Взаимодействие субъединиц репрессора	184		
Репрессор – белок, связывающийся с ДНК	184		
Отделение репрессора от ДНК	185		
Накопление излишков репрессора	186		
Парадокс индукции	187		
Рекомендуемая литература	188		
Глава 15. Системы контроля: средства регуляции оперонов	188		
Различия между позитивным и негативным контролем	188		
Триптофановый оперон является репрессибельным	190		
Модификация координированной регуляции Триптофановый оперон контролируется с помощью аттенуации	191		
Аттенуация контролируется с помощью альтернативных вторичных структур	192		
Широкое распространение явления аттенуации Репрессия может иметь место для множества локусов	194		
Арабинозный оперон находится под двойным контролем	196		
Сложная организация регуляторной области <i>ara</i> -оперона	197		
Двойной промотор галактозного оперона	200		
Катаболитная репрессия способствует преимущественному использованию глюкозы	201		
Аутогенный контроль трансляции рибосомных белков	201		
Аутогенный контроль и сборка макромолекул Неблагоприятные условия определяют строгий ответ	203		
Рекомендуемая литература	204		
Глава 16. Литический каскад и лизогенная репрессия	205		
Литический цикл состоит из отдельных стадий	206		
Литическое развитие подвержено каскадной регуляции	207		
Образование кластеров генов с родственными функциями у фагов T7 и T4	208		
О том, как фаг лямбда осуществляет свой литической каскад	208		
Лизогения поддерживается благодаря аутогенному циклу	210		
Репрессор – димер с различными доменами Репрессор связывается кооперативно в каждом операторе	213		
Как запускается синтез репрессора?	214		
Для литической инфекции необходим антирепрессор	216		
Чувствительный баланс: лизогения против лизиса	218		
Рекомендуемая литература	219		
	220		
		ЧАСТЬ V. СТРОЕНИЕ ГЕНОМА ЭУКАРИОТ	221
		Глава 17. Геномы эукариот: множество последовательностей	222
		Парадокс величины <i>S</i> характеризует различия в размерах геномов	222
		Кинетика реассоциации зависит от генетической сложности последовательностей ДНК	224
		Эукариотические геномы состоят из последовательностей нескольких типов	224
		Размер генома можно оценивать по сложности неповторяющейся ДНК	225
		Геномы эукариот содержат повторяющиеся последовательности	226
		Умеренно повторяющаяся ДНК состоит из множества различных последовательностей	227
		Члены семейств повторяющихся последовательностей сходны, но не идентичны	228
		Участки умеренно повторяющейся ДНК чередуются с участками неповторяющейся ДНК	229
		Рекомендуемая литература	229
		Глава 18. Структурные гены: как они представлены в мРНК	230
		Являются ли структурные гены уникальными или повторяющимися?	230
		Большая часть структурных генов относится к неповторяющейся ДНК	231
		Сколько уникальных генов экспрессируется? Оценка числа генов по кинетике реакции, определяемой концентрацией РНК	232
		Уровни экспрессии генов сильно различаются	234
		Перекрывание популяций мРНК	234
		Рекомендуемая литература	236
		Глава 19. Исследование ДНК	236
		Любая последовательность ДНК может быть клонирована в бактериях	236
		Получение химерной ДНК	238
		Получение ДНК-копий на матрице мРНК	241
		Выделение из генома индивидуальных генов	242
		Клонирование всей ДНК генома («шотган») с образованием библиотек генов	243
		Эукариотические гены могут экспрессироваться в бактериях с образованием белка	244
		Рекомендуемая литература	245
		Глава 20. Структурные гены: внутренняя организация	245
		Обнаружение прерывистых генов с помощью электронной микроскопии	246
		Рестрикционное картирование прерывистых генов	248
		Характеристика фрагментов геномной ДНК	250
		Гены имеют самое разнообразное строение и размеры	252
		Интроны генов, кодирующих рРНК и тРНК	254
		Интроны – неповторяющиеся и быстро эволюционирующие компоненты генома	255
		На границах экзон-интрон имеется каноническая последовательность	255
		Интрон одного гена может быть экзоном другого гена	256
		Интрон, который может кодировать регуляторный белок	258
		Сложные локусы имеют очень большие размеры и участвуют в регуляции	262
		Как появились прерывистые гены?	264
		Рекомендуемая литература	266

ЧАСТЬ VI. КЛАСТЕРЫ СХОДНЫХ ПОСЛЕДОВАТЕЛЬНОСТЕЙ ДНК	267	Высокоповторяющаяся ДНК образует сателлитную ДНК	300
Глава 21. Структурные гены: организация родственных генов	268	Сателлитная ДНК часто располагается в области гетерохроматина	301
Множество типов глобинов	268	Сателлитная ДНК членистоногих состоит из очень коротких идентичных повторов	302
Гены глобинов организованы в виде кластеров	269	Сателлитная ДНК млекопитающих состоит из иерархически организованных повторов	302
Неравный кроссинговер приводит к перестройке кластеров генов	270	Реконструкция этапов эволюции сателлитной ДНК мыши	304
Многие формы α -талассемии – результат неравного кроссинговера	272	Различия в существующей в настоящее время повторяющейся единице сателлитной ДНК	305
Новые гены, образующиеся при β -формах талассемии	272	Роль неравного кроссинговера	306
Кластеры генов подвержены постоянным перестройкам	273	Фиксация при кроссинговере может обеспечивать существование идентичных повторов	308
Эволюционное древо глобиновых генов	274	Рекомендуемая литература	308
Дивергенция нуклеотидных последовательностей указывает на различие путей эволюции организмов	275	ЧАСТЬ VII. СОЗРЕВАНИЕ РНК: ПРОЦЕССИНГ	309
Два типа дивергенций последовательностей ДНК	275	Глава 25. Образование стабильной РНК путем разрезания и подравнивания предшественника	310
Использование часов для изучения эволюции генов глобина	276	Фосфоэфирные связи могут расщепляться с обеих сторон	311
Механизмы, обеспечивающие сохранение в геноме функционально активных последовательностей	277	РНКаза III «вырезает» ранние мРНК фага T7 из полицистронного продукта транскрипции	311
Псевдогены – тупики эволюции	278	Рибосомные РНК образуются из своих предшественников под действием РНКазы III	313
Семейства генов обычно кодируют белки, богато представленные в клетке	279	Сайты расщепления при созревании эукариотической рРНК	314
Рекомендуемая литература	280	тРНК разрезаются и подравниваются несколькими ферментами	315
Глава 22. Геномы клеточных органелл	280	Рекомендуемая литература	317
Гены органелл не подчиняются законам Менделя	281	Глава 26. Механизмы сплайсинга РНК	317
Геномы органелл представляют собой кольцевые молекулы ДНК	282	Сплайсинг дрожевой тРНК включает разрезание и сшивание	318
В органеллах экспрессируются их собственные гены	282	Необычный сплайсинг рРНК <i>Tetrahymena</i>	320
Митохондриальный геном дрожжей имеет большие размеры	284	РНК как катализатор: расширение понятия биохимического катализа	321
Компактная организация генома митохондрий млекопитающих	285	Реакция сплайсинга РНК осуществляется в определенной предпочтительной последовательности	322
ДНК некоторых органелл участвует в процессе рекомбинации	287	Границы сплайсинга могут быть взаимозаменяемыми	324
Перестройки митохондриальной ДНК дрожжей	288	Мутации в канонических последовательностях могут влиять на сплайсинг	325
Рекомендуемая литература	288	Участвуют ли в сплайсинге малые ядерные РНК?	328
Глава 23. Сходство и различия кластеров tandemных генов	289	Рекомендуемая литература	331
Гены гистонов образуют повторы	289	Глава 27. Регуляция процессинга РНК	331
Разнообразие кластеров tandemных генов гистона	290	гяРНК имеет большие размеры и нестабильна	332
рРНК и тРНК кодируются повторяющимися генами	291	мРНК образуется из гяРНК	333
Тандемно повторяющаяся единица включает оба рРНК гена	292	Значение полиаденилирования	335
Некоторые рРНК-гены располагаются не в хромосомах	293	гяРНК устроены более сложно, чем мРНК	336
О нетранскрибирующихся спейсерах и промоторах	294	Осуществляется ли регуляция на посттранскрипционном уровне?	336
5S-гены и псевдогены перемежаются	294	Модели регуляции экспрессии генов	337
Эволюционная дилемма	295	Роль клеточных полипротеинов	341
Бактериальные рРНК-гены и тРНК-гены входят в состав одних и тех же оперонов	296	Рекомендуемая литература	342
тРНК-гены могут быть организованы в виде кластеров	297	ЧАСТЬ VIII. УПАКОВКА ДНК	343
Рекомендуемая литература	298	Глава 28. О геномах и хромосомах	344
Глава 24. Организация простых последовательностей ДНК	298	Упаковка вирусных геномов в оболочку	344
Семейство Alu	298	Бактериальный геном свернут в нуклеоид	347
Обращенные повторы мгновенно ренатурируют	299	Нуклеоид содержит много суперспирализованных петель	348
		Различия между интерфазным хроматином и митотическими хромосомами	349

Эукариотическая хромосома как единица сегрегации	352	Глава 32. Топология репликации ДНК	409
Деспирализованное состояние хромосом «ламповых щеток»	354	Описание топологии ДНК	410
Гигантские хромосомы образуются в результате политемии	355	Топологические перестройки ДНК	410
Нарушение хромосомной структуры при транскрипции	357	Гириза вводит отрицательные суперспирали в ДНК	412
Рекомендуемая литература	358	ДНК-полимеразы эукариот	413
Глава 29. Нуклеосомные частицы и структура хроматина	358	ДНК-полимеразы прокариот проявляют несколько ферментативных активностей	414
Белковые компоненты хроматина	359	Синтез ДНК является полунепрерывным	416
Хроматин содержит дискретные частицы	360	Синтез фрагментов Оказаки инициируется РНК	418
Нуклеосома – основная субъединица всего хроматина	361	Рекомендуемая литература	419
Частицы минимальной нуклеосомы высококонсервативны	362	Глава 33. Ферментативный аппарат репликации ДНК	420
ДНК закручена вокруг гистонового октамера ДНК, симметрично обработанная нуклеазами	363	Сложность репликационного аппарата бактерий	420
Нерешенный вопрос о периодичности ДНК	367	Инициация синтеза одиночной цепи ДНК	421
Организация гистонов и ДНК	368	Движение праймосомы	422
Сборка нуклеосом и репродукция хроматина	369	Инициация репликации в точках начала репликации двухцепочечной ДНК	425
Для сборки нуклеосом нужны негистоновые белки	372	Репликационный аппарат фага T4	427
Нуклеосомы в нитях хроматина	372	Репликационный аппарат фага T7	429
Петли, домены и остовы	373	Проблема линейных репликонов	429
Рекомендуемая литература	375	Рекомендуемая литература	431
Глава 30. Нуклеосомы в активном хроматине	376	Глава 34. Система защиты ДНК	431
Существует ли упорядоченность в расположении нуклеосом?	376	Процессы рестрикции и модификации	432
Специфичность нуклеазы микробов	378	Альтернативные активности ферментов типа I	433
Сохраняют ли транскрибируемые гены нуклеосомную структуру?	379	Два вида активности ферментов типа III	435
Домены, чувствительные к ДНКазе в транскрибируемом хроматине	381	Механизмы репарации повреждения ДНК	436
Чувствительность к ДНКазе обусловлена негистоновыми белками	383	Системы эксцизионной репарации у <i>E. coli</i>	438
Гистоны подвергаются кратковременной модификации	384	Системы репарации у <i>E. coli</i> , включающие рекомбинацию	440
В некоторых нуклеосомах гистон H2A связывается с убиквитином	385	SOS-репарация	441
Экспрессия гена связана с деметилированием	386	Репарационные системы млекопитающих	442
Некоторые модели контроля метилирования	387	Рекомендуемая литература	442
Сайты, сверхчувствительные к ДНКазе, расположены перед активными промоторами	388	Глава 35. Восстановление и рекомбинация ДНК	443
В сверхчувствительных участках нет нуклеосом	390	Для осуществления рекомбинации необходим синтапс гомологичных молекул ДНК	443
Предположения о природе активации гена	392	Разрыв и воссоединение осуществляется через гетеродуплексные ДНК	445
Рекомендуемая литература	393	Действительно ли двухцепочечные разрывы инициируют рекомбинацию?	446
ЧАСТЬ IX. СОХРАНЕНИЕ ДНК В РЯДУ ПОКОЛЕНИЙ	395	Выделение промежуточных продуктов рекомбинации	447
Глава 31. Репликон: единица репликации	396	Обмен между цепями, осуществляемый при участии белка Rec A	448
Синтез ДНК является последовательным и полуконсервативным	396	Белок Rec A и условия рекомбинации	450
Бактериальный геном представлен одним репликоном	398	Конверсия гена ответственна за межлалельную рекомбинацию	452
Согласованность процессов репликации ДНК и клеточного деления	399	Специализированная рекомбинация узнает специфические сайты	453
Каждая хромосома эукариот содержит много репликонов	402	Ступенчатый разрез и воссоединение в коре	454
Выделение точек начала дрожжевых репликонов	403	Рекомендуемая литература	456
Репликация может происходить по типу «глазков», «катящихся колец» или D-петель	404	ЧАСТЬ X. ДИНАМИЧНОСТЬ ГЕНОМА: ПОСТОЯННОЕ ИЗМЕНЕНИЕ ДНК	457
Несовместимость плазмид связана с числом их копий	406	Глава 36. Транспозирующиеся элементы бактерий	458
Рекомендуемая литература	408	Открытие транспозиции у бактерий	459
		Инсерционные последовательности – это простейшие транспозоны	460
		Сложные транспозоны содержат IS-модули	461
		Только один модуль транспозона Tn10 функциональный	462
		Модули транспозона Tn5 почти идентичны, однако функционально очень различаются	463
		Транспозиция включает репликативную рекомбинацию	464

Транспозиция Tn3 происходит путем разрешения коинтеграта	466	В клетке могли происходить РНК-зависимые транспозиции	494
Некоторые необычные свойства транспозирующего фага Mu	469	Тканеспецифичные вариации в геноме дрожофилы	495
Фазовая вариация у сальмонелл определяется инверсией	470	Селекция амплифицированных последовательностей генома	497
Рекомендуемая литература	472	Введение экзогенных последовательностей посредством трансфекции	500
Глава 37. Мобильные элементы эукариот	472	Трансфицируемая ДНК способна включаться в геном клеток зародышевой линии	501
Дрожжевые элементы Tu напоминают бактериальные транспозоны	473	Рекомендуемая литература	502
В геноме <i>D. melanogaster</i> содержится несколько типов мобильных элементов	474	Глава 39. Как формируется многообразие антител	502
Сложные локусы и «прогулка по хромосоме»	476	Иммуноглобулиновые гены образуются путем соединения ранее независимых частей	504
Внедрения в локус w определяют сложность мишени	477	Природное разнообразие иммуноглобулиновых генов в клетках зародышевой линии	506
Роль мобильных элементов в гибридном дисгенезе	480	Реакция объединения генных сегментов – дополнительный источник разнообразия антител	508
Контролирующие элементы кукурузы способны транспозироваться	481	Рекомбинация между V- и С-генами вызывает делеции и перестройки последовательностей ДНК	509
Элемент <i>Ds</i> способен транспозироваться или вызывать хромосомные разрывы	482	Некоторые возможные случаи аллельного исключения	511
Транспозиция <i>Ds</i> -элемента связана с репликацией	484	Дальнейшая рекомбинация ДНК обуславливает переключение классов иммуноглобулинов	512
Молчащие и активные локусы дрожжей, контролирующие тип спаривания	485	Изменения в экспрессии ранних генов тяжелой цепи могут происходить за счет процессинга РНК	513
Молчащие и активные кассеты имеют одинаковые последовательности	486	Соматические мутации вносят определенный вклад в разнообразие антител	515
Однонаправленная транспозиция инициируется реплицируемым локусом <i>MAT</i>	488	Строение главного локуса гистосовместимости	516
Рекомендуемая литература	489	Рекомендуемая литература	518
Глава 38. Элементы, способные к перемещению в пределах генома и вне его	490	Словарь терминов	519
Жизненный цикл ретровирусов связан с событиями, напоминающими транспозицию	490	Предметный указатель	529
Ретровирусы способны трансдуцировать клеточные последовательности	492	Указатель латинских названий	538

Монография

Бенджамин Льюин

ГЕНЫ

Спец. редактор Г. Н. Ениколопов
 Ст. научн. редактор Л. Г. Тер-Саркисян
 Научный редактор М. Р. Погосбекова
 Мл. научн. редактор З. В. Соллертинская
 Художник Ю. С. Урманчеев
 Художественный редактор А. Я. Мусин
 Технический редактор А. Л. Гулина
 Корректор Н. А. Мистрюкова

ИБ № 5535

Сдано в набор 19.05.86. Подписано к печати 26.08.87. Формат 60 × 90¹/₈. Бумага офсетная № 1. Печать офсетная. Гарнитура таймс. Объем 34,00 бум. л. Усл. печ. л. 68,00. Усл. кр.-отт. 136,25. Уч. изд. л. 87,72. Изд. № 4/4114. Тираж 14 000 экз. Зак. 1102. Цена 6 р. 80 к.

ИЗДАТЕЛЬСТВО «МИР»

129820, ГСП, Москва, И-110, 1-й Рижский пер., 2

Можайский полиграфкомбинат Союзполиграфпрома при Государственном комитете СССР по делам издательств, полиграфии и книжной торговли. 143200, г. Можайск, ул. Мира, 93