
Глава 7. Структура гена	2
7.1. Развитие представлений о гене	2
7.2. Перекрывающиеся гены у вирусов и прокариот	4
7.3. Оперонный принцип организации генов у прокариот	6
7.4. Химический синтез генов	8
7.5. Клонирование и анализ ДНК	11
7.5.1. Ферменты рестрикции	11
7.5.2. Векторы для молекулярного клонирования	12
7.5.2.1. Плазмидные векторы	13
7.5.2.2. Фаговые векторы	17
7.5.2.3. Космидные векторы	18
7.5.2.4. Челночные векторы	18
7.5.2.5. Искусственные хромосомы дрожжей (YAC)	20
7.5.3. Создание геномных библиотек	21
7.5.4. Построение рестрикционных карт	23
7.5.5. Саузерн-блот анализ	25
7.5.6. “Хромосомная ходьба”	27
7.5.7. Нозерн-блот анализ	29
7.5.8. Полимеразная цепная реакция	29
7.5.9. Определение последовательности нуклеотидов (секвенирование)	32
7.5.10. Трансформация у дрозофилы	34
7.6. Расположение генов в хромосомах эукариот	38
7.7. Структура транскрипта: структурная и регуляторная части гена	42
7.8. Регуляторная часть гена	44
7.8.1. Промоторы и регуляторы	44
7.8.2. Метод репортерных генов для изучения промоторных участков генов	49
7.8.3. Энкхасерные участки гена	52
7.9. Структурная часть гена	58
7.9.1. Интроны и экзоны	58
7.9.2. Альтернативный сплайсинг	60
7.9.3. Локализация генов в интронах	63
7.9.4. Использование промоторов генов теплового шока	64
7.9.5. Участки терминирующие транскрипцию	66
7.9.6. Гомология генов	67
7.9.7. Псевдогены	67

7. Структура гена

7.1. Развитие представлений о гене

К концу 20-х годов сложилось представление о гене как материальной частице, лежащей в хромосоме, способной к саморепродукции и являющейся минимальной единицей рекомбинации, мутирования и генетической функции. Цепочка сцепленных генов представлялась как нитка корпускул или бусинок, механически связанных друг с другом. Ген, согласно этим представлениям, считался неделимым с помощью кроссинговера. Нормальному аллелю противопоставлялся мутантный. Для объяснения существования пар аллелей уже в 1902 г. К. Корренс и У. Бэтсон предложили теорию, согласно которой доминантный признак обуславливается наличием определенного гена, а рецессивный - выпадением, отсутствием этого гена. Эта теория получила название теории "присутствия - отсутствия". Она очень просто объясняла существование пар аллелей. Вначале признавал ее и Т. Морган. Однако, затем, когда были открыты большие серии промежуточных мутаций, в разной степени изменяющих признаки, т.е. серии аллельных генов, эта теория подверглась критике. Морган отказался от нее. Как писал А.С. Серебровский, "эпитафией к этой теории явились слова Моргана: "Не могут одному присутствию отвечать несколько отсутствий" (Из: Хесин, 1972, стр. 17).

В 1929 г. А.С. Серебровский и Н.П. Дубинин описали серию индуцированных мутаций гена *scute* (*sc*), нарушающих формирование разных щетинок у дрозофилы (Рис. 7.1.).

Но самое поразительное выяснилось, когда получили самок, у которой в одной X-хромосоме находился один аллель, а в другой хромосоме - другой аллель этого гена. В таких компаундах редуцированными были только те щетинки, которые нарушались обоими аллелями, а щетинки, развитые у мутантов по одному из них и редуцированные у мутантов по другому, в компаунде оказывались нормальными. Более того, если два аллеля нарушали развитие совсем разных щетинок, то в компаунде они давали возврат к норме: развивались все щетинки, например у особей sc^5/sc^6 , что было показано в работе Серебровского в 1930-м году, после того как он получил мутацию sc^6 . Обсуждая эти факты, авторы написали:

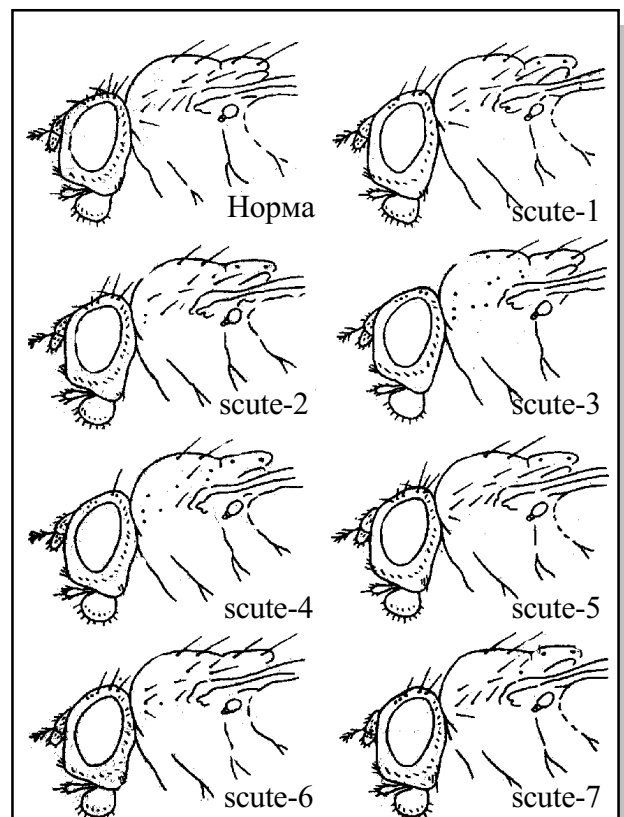


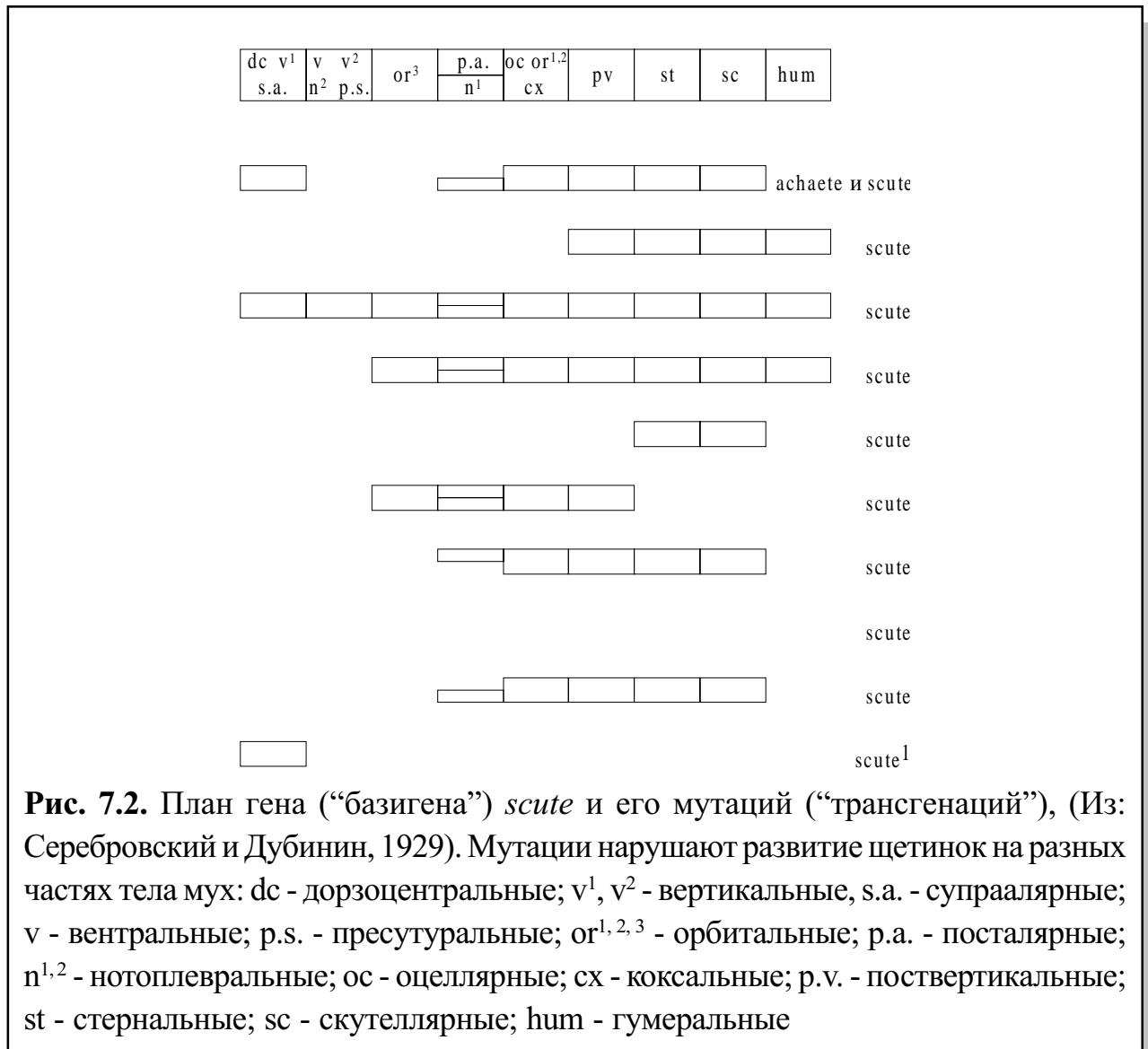
Рис. 7.1. Влияние разных мутаций гена *scute* на развитие щетинок у дрозофилы (Из: Серебровский, Дубинин, 1929)

“...явление частичного возврата к дикому типу может быть истолковано как обусловленное не полным аллелизмом двух аллелей. С этой точки зрения... общие части проявляются в силу того, что в обеих хромосомах имеются изменения одинаковых участков, другими словами, по этим участкам муха гомозиготна; не проявление же несовпадающих участков зависит оттого, что измененному участку одной хромосомы соответствует участок второй хромосомы, который не был затронут трансгенацией...”. И далее особенно важный вывод: “Это весьма ответственное воззрение, утверждающее

ДЕЛИМОСТЬ гена (трансгенацию по частям)...” Была построена карта нарушений щетинок (Рис. 7.2.).

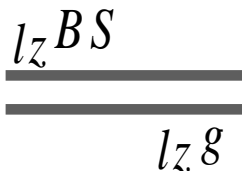
Авторы назвали ген *scute* базигеном, т.е. участком хромосомы, занимаемым всеми изменениями - трансгенами *scute*. Самостоятельные при трансгенациях (мутациях) элементарные участки внутри базигена были названы центрами. Мутации могут захватывать как отдельные центры, так и целые их группы. При этом аллели расположены ступенчато.

Эта теория сложного (центрального) строения, или делимости гена, имеющего одномерную протяженность, не была



поддержана крупнейшими генетиками того времени (А. Стертевантом, Дж. Шульцем, Р. Гольдшмидтом, Г. Меллером) (см. Хесин, 1972).

Много позже М. Грин и К. Грин (1949) показали делимость гена *lz* у дрозофилы посредством кроссинговера. Они получили гетерозигот,



0,1% потомков от которых обнаруживали либо нормальные глаза, либо более сильное мутантное проявление, чем любой из аллелей. Это могло произойти только в результате кроссинговера внутри этого гена. Чтобы спасти представление о неделимости гена, такие аллели стали называть псевдоаллелями. (См. также: Алиханян и др. 1985, стр. 219-220). (Функциональный тест на аллелизм; Цис-транс-тест, Инге-Вечтомов, 1989, стр. 374).

Огромный вклад в понимание структуры и функции гена внесли Дж. Бидл и Е. Татум, которые в начале 1940-х годов впервые исследовали биохимические мутации у *Neurospora crassa*. Эти исследования показали, что мутации ауксотрофности у нейроспоры прерывают цепи метаболизма на конкретных этапах. При этом аллельные мутации всегда затрагивали один и тот же этап биосинтеза. На основе своих результатов Дж. Бидл и Э. Татум сформировали принцип “один ген - один фермент”, означавший, что каждый ген контролирует синтез какого-либо определённого фермента.

В 1961 году С. Бензер изучал область гII фага Т4. Он сопоставил молекулярные размеры этой области -

Дополнение 7.1.

За открытие того, что ген действует, регулируя определенные химические события, Дж. Бидл и Э. Татум (G.W. Beadle, E.L. Tatum) получили Нобелевскую премию в 1958 году.

2700 п.н. и рекомбинационную длину - 10%. В экспериментах были получены минимальные частоты рекомбинации около 0,02%, т.е. 1/500 часть от всего рекомбинационного расстояния (10%). Если вся длина района 2700 п.н., значит рекомбинация происходит между каждыми 5-6 нуклеотидами (2700:500).

Позже У. Яновский показал, что кроссинговер происходит между любой парой нуклеотидов. Эта пара была названа реконом. (Детали этих опытов можно найти в любом учебнике).

Таким образом, не только ген, но и любая его часть могут быть разделены с помощью кроссинговера.

Литература к разделу 7.1.

Инге-Вечтомов С.Г. Генетика с основами селекции. Москва, Высшая школа, 1-592, 1989.

Серебровский А.С., Дубинин Н.П. Искусственное получение мутаций и проблема гена. Успехи эксп. биол., вып. 4, 235-247, 1929.

Хесин Р.Б. Теория гена в работах А.С. Серебровского. Природа N 8, 16-27, 1972.

7.2. Перекрывающиеся гены у вирусов и прокариот

(Из Алиханян и др. 1985, стр. 231-233).

Генетический материал мелкого бактериофага фХ174 представлен

одноцепочечной ДНК и состоит всего из 9 генов, продукты которых хорошо изучены. ДНК, необходимая для кодирования этих продуктов, должна бы состоять минимум из 6078 нуклеотидов. На самом же деле хромосома фага φX174 состоит из 5374 нуклеотидов. Этот парадокс удалось разрешить лишь после того, как в 1978 г. группой Ф. Сэнгера было проведено полное секвенирование ДНК этого фага. Оказалось, что кодирующие последовательности двух генов (**В** и **Е**) локализованы внутри кодирующих последовательностей двух других генов (**А** и **Д**). При этом рамка считывания (т.е. триплет, прочитываемый при трансляции) в каждом случае оказывалась сдвинутой на одну пару нуклеотидов. Например, в определённом участке внутри гена **Д** находится последовательность,

GTTTATGGTACG

которая в полипептиде **Д** кодирует последовательность валин-тирозин-глицин-треонин. Рамка считывания гена **Е** смещена вправо на один нуклеотид от рамки считывания гена **Д**. Поэтому триплет ATG распознается РНК-полимеразой как стартовый и в полипептиде **Е** появляется формилметионин, за которым последует валин, кодируемый триплетом GTA, и т.д.

Сходным образом кодирующая последовательность гена **В** оказывается внутри кодирующей последовательности гена **А**. В результате сдвига рамки считывания кодируемые перекрывающимися генами полипептиды полностью отличаются друг от друга по последовательностям аминокислот.

Вместе с тем в случае замены или делеции одного нуклеотида инактивируются сразу два гена.

Подобная ситуация “ген внутри гена” обнаружена и в ряде других случаев. Частично перекрывающиеся кодирующие последовательности обнаружены в ДНК вируса млекопитающих SV40. У РНКового фага MS2 один из генов перекрывает два других и, следовательно, не перекрывается лишь один из фаговых генов.

$$\begin{array}{c} 3 \\ \hline 1 \quad 2 \quad 4 \\ \hline \end{array}$$

При анализе одного из мобильных генетических элементов у бактерий - элемента IS5 - был обнаружен еще более яркий пример сверхкомпактной организации генетической информации. В этом случае одна из цепей дуплекса ДНК содержит два перекрывающихся гена, а комплементарный им участок второй цепи образует третий ген. Следовательно, обе цепи значимы и несут информацию, соответствующую трем генам.

Дополнение 7.2.

В 1980 году за выдающийся вклад в разработку методов экспериментальных манипуляций с ДНК Нобелевская премия по химии была присуждена П. Бергу, У. Гилберту и Ф. Сэнгеру (P. Berg, W. Gilbert, F. Sanger).

Литература к разделу 7.2.

Алиханян С.И., Акифьев А.П., Чернин Л.С. Общая генетика. Москва, Высшая школа, 1-446, 1985.

7.3. Оперонный принцип организации генов у прокариот

Бактериям необходимо уметь быстро отвечать на изменения в окружающей среде. Их выживаемость зависит от способности переключать метаболизм с одного субстрата на другой, поскольку поступление питательных веществ может постоянно меняться.

Бактерия не синтезирует ферментов того или иного метаболического пути в отсутствие соответствующего субстрата, но способна в любое время начать их синтез при появлении такового. Для осуществления подобного реагирования гены бактерий объединены в кластеры таким образом, что ферменты, необходимые для осуществления определенного пути биосинтеза, кодируются генами, сцепленными друг с другом.

Клетки *E. coli* в отсутствие бета-галактозида содержат очень незначительное число молекул фермента, разлагающего его - буквально несколько штук. Функция фермента сводится к разложению молекулы бета-галактозида на составляющие сахара. Например, лактоза разлагается на глюкозу и галактозу, которые в свою очередь используются в дальнейшем метаболизме. При добавлении галактозида в бактериальных клетках очень быстро усиливается активность фермента в результате синтеза новых молекул, которые появляются уже через 2-3 мин. Вскоре их число в клетке

превышает 5000. При удалении субстрата из среды синтез фермента прекращается так же быстро, как был начат.

Быстрое увеличение активности фермента за счет его синтеза под действием субстрата называется индукцией, а уменьшение активности, также под действием субстрата - репрессией генов.

Гены можно поделить на две группы по принципу действия их продукта. Гены, которые кодируют белки, необходимые для ферментативных или структурных функций, называются *структурными*. Большинство генов бактерий попадают в эту категорию, которая, таким образом, представлена огромным разнообразием функций и структур белков.

Гены, кодирующие белки, которые регулируют экспрессию других генов, называются *регуляторными*. Поскольку продукты регуляторных генов свободно диффундируют и находят подходящие мишени для активирования, такой тип взаимодействий генов называют *транс-регуляцией*. Транс-действующий белок может регулировать ген-мишень либо позитивно, если в результате взаимодействия регулируемый ген включается, или негативно, если он выключается. Белок-регулятор связывается с цис-регулирующей последовательностью регулируемого гена, которая располагается обычно (но не исключительно) выше гена-мишени (Из: Lewin, 1994, p. 413).

Структурные гены бактерий имеют тенденцию быть организованными в кластеры генов кодирующих белки, чьи функции связаны.

Примером кластерной организации у *E. coli* являются лактозные гены, которые индуцируются и

репрессируются под действием субстрата (см. Льюин, 1987, стр. 176-177).

Весь кластер генов занимает около 6000 п.н. (Рис. 7.3.). Ген *lacI* имеет свой промотор и терминатор. Конец района *lacI* непосредственно примыкает к промотору *lac(P)*, оператор *lac(O)* занимает первые 26 п.н. гена *lacZ*. После гена *lacZ*, который имеет необычно большую длину, расположены гены *lacY* и *lacA*, а также общий терминатор транскрипции.

Гены имеют следующие функции: продукт *lacZ* расщепляет бета-галактозид на составляющие его сахара, например, лактозу на глюкозу и галактозу, продукт гена *lacY* является бета-галактозид пермеазой, он транспортирует бета-галактозид в клетку. Ген *lacA* кодирует белок трансацетилазу, энзим, который переносит ацетильную группу с ацетил-СоА на бета-галактозид.

Мутации генов *lacZ* и *lacY* дают *lac*-фенотип, когда клетки не могут

использовать лактозу. У мутантов *lacZ* - отсутствует ферментативная активность бета-галактозидазы. Мутанты по гену *lacY* не могут усваивать лактозу из среды. Мутанты *lacA* - не проявляют мутантного фенотипа, что довольно неожиданно.

Вся эта система, включающая структурные гены и элементы, контролирующие их экспрессию, формирует общую единицу регуляции, называемую опероном. Модель оперона была предложена в 1961 году Ф. Жакобом и Ж. Моно. В короткое время эта модель стала в центре внимания не только генетиков, но и огромного числа биологов во всём мире. Она позволила увидеть реальный механизм регуляции активности генов (см. Brenner, 1997).

Транскрипция генов *lacZYA* контролируется белком-репрессором, кодируемым геном *lacI*. Гены *lacZYA* контролируются негативно: они транскрибируются до тех пор, пока не выключаются белком-регулятором.

	P	<i>lacI</i>	PO	<i>lacZ</i>	<i>lacY</i>	<i>lacA</i>
ДНК (длина в п.н.)		1040	82	3510	780	825
полипептид: аминокислоты дальтоны		360 38000		1021 125000	260 30000	275 30000
белок: структура дальтоны		тетрамер 152000		тетрамер 500000	мембранный белок 30000	димер 60000
функция белка		репрессор		β -галактозидаза	пермеаза	транс- ацетилаза
		мРНК:		Один транскрипт		

Рис. 7.3. Размеры генов и белков лактозного оперона у *E. coli* (Из: Lewin, 1994, р.416). P - промотор, O - оперон

Мутации гена-регулятора обуславливают постоянную транскрипционную активность генов *lacZYA*. Белок *lac* - репрессор связывается с оператором *O_{lac}* на старте транскрипции кластера *lacZYA*. Когда репрессор связывается с оператором, он предотвращает движение РНК полимеразы и инициацию транскрипции (Рис. 7.4.).

Белок-репрессор имеет два сайта связывания: один для индуктора, другой для оператора. Когда индуктор связывается со своим сайтом, это так изменяет конформацию молекулы белка, что изменяется связывающая способность другого сайта. Индуктор лактоза поступает в клетки, связывается с репрессором, в результате чего последний становится неспособным сдерживать транскрипцию и она начинается. При удалении индуктора репрессор вновь занимает положение на операторе, останавливая транскрипцию.

Основным преимуществом оперонной организации генов у микроорганизмов является координация регуляции активности: все гены экспрессируются или не экспрессируются в унисон. Матричная РНК транслируется всегда с 5' конца. Это объясняет почему индукция всегда обуславливает появление бета-галактозидазы, бета-галактозид пермеазы и только затем бета-галактозид трансацетилазы. А факт трансляции единой мРНК объясняет почему относительные количества всех трех ферментов остаются одинаковыми независимо от условий индукции (весь текст по: Lewin, 1994, pp. 413-421). [Другие примеры оперонов см Алиханян и др. 1985, стр. 239-252].

Дополнение 7.3.

За открытие механизмов генетического контроля синтеза ферментов Нобелевскую премию в 1965 году получили Ф. Жакоб, А. Львов и Ж. Моно (Francois Jacob, Andre Lwoff, Jaques Monod).

Литература к разделу 7.3.

- Алиханян С.И., Акифьев А.П., Чернин Л.С. Общая генетика. Москва, Высшая школа, 1-446, 1985 г.
- Льюин Б. Гены. Москва. Мир, 1-544, 1987.
- Brenner S. A night at the operon. Nature 386, 235, 1997.
- Curtis H., Barnes N.S. Biology. Fifth Edition. Worth Publishers, Inc. New York, 1989, p. 320-326.
- Lewin B. Genes V. Oxford, New York, Tokyo, Oxford University Press, 1-1272, 1994.

7.4. Химический синтез генов

Известно несколько способов получения генов:

1. Непосредственное выделение из природных источников (клонирование).
2. Химический синтез.
3. Копирование соответствующей гену мРНК для получения комплементарной ДНК-реплики (кДНК).

Искусственный синтез гена был впервые осуществлен химическим путем в 1969 году Г. Хораной и его сотрудниками. Группа Хораны синтезировала ген аланиновой тРНК дрожжей, структура которого к тому времени была полностью расшифрована. Синтезированный ген длиной 77 п.н. не содержал регуляторных последовательностей и поэтому не обладал функциональной активностью.

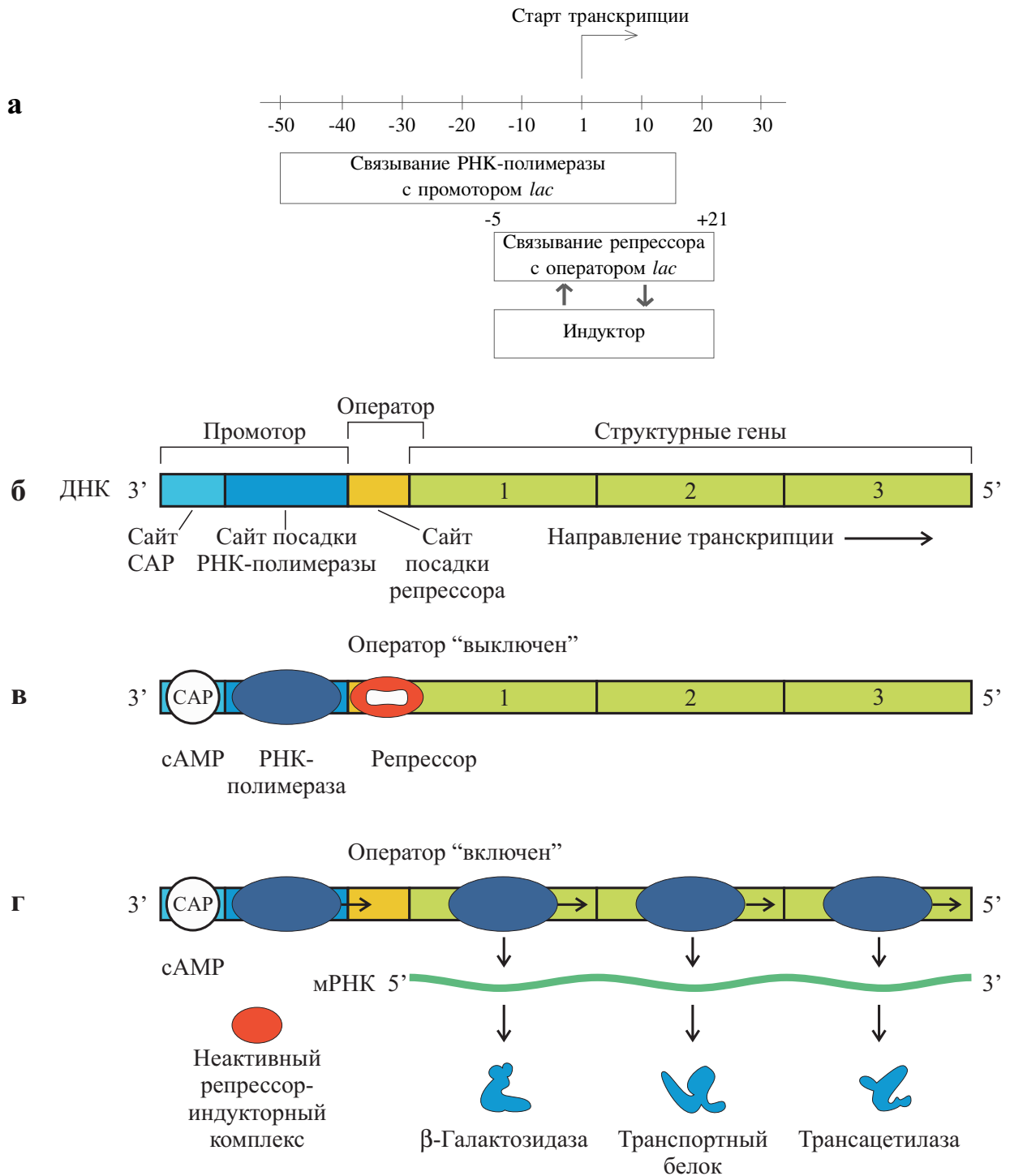


Рис. 7.4. Функционирование лактозного оперона у *E. coli*. **а.** локализация сайтов связывания молекул РНК-полимеразы и репрессора в регуляторной области гена *lacZ* (Из: Lewin, 1994, p.417). **б.** структура лактозного оперона. Как и у всех генов, регулируемых CAP и cAMP, промотор содержит два района: участок связывания с РНК-полимеразой и участок связывания с комплексом CAP-cAMP. **в, г.** негативная и позитивная регуляция *lac*-оперона. В отсутствие лактозы репрессор (продукт гена *I*) связывается с оператором. Хотя РНК-полимераза может связываться с промотором, она не может перемещаться далее репрессора. Оператор выключен, гены не работают. В присутствии индуцирующей молекулы репрессор инактивируется и более не связывается с оператором. Молекулы РНК-полимеразы перемещаются, и начинается транскрипция (б-г - из: Curtis, Barnes, 1989, p.325)

Позднее та же группа авторов синтезировала первый функционально активный ген - ген супрессорной тирозиновой тРНК *E. coli* длиной около 200 п.н.

Искусственно синтезированные гены могут экспрессироваться в составе гибридных молекул микроорганизмов. Первой удачной попыткой такого рода стала работа К. Итакуры и Г. Бойера с соавторами (1977) по экспрессии в *E. coli* гена, кодирующего гормон млекопитающих - соматостатин. Ген соматостатина был синтезирован на основе сведений о первичном строении этого пептидного гормона, состоящего всего из 14 аминокислот. В различных лабораториях были созданы штаммы *E. coli*, синтезирующие гормон роста человека - соматотропин, пептидные гормоны - брадикинин и ангиотензин, нейропептид лей-энкефалин, ген интерферона размером 200 п.н. и др.

Ген гормона роста человека длиной 584 п.н. - наиболее длинный из искусственно синтезированных в настоящее время. Он был встроен в плазмиду, реплицирующуюся в *E. coli* под контролем промотора триптофанового оперона. Трансформированные полученной химерной плазмидой клетки *E. coli* продуцировали при индукции промотора около 3 млн. молекул гормона роста человека в расчете на клетку. Этот полипептид, как было установлено в экспериментах на крысах с удаленным гипофизом, по функциям оказался полностью идентичен гормону роста человека.

Ген инсулина синтезировали в виде более чем 40 шестичленных олигонуклеотидов, которые затем объединяли в единую структуру с

помощью ДНК-лигазы. Полученные двуцепочечные полинуклеотиды длиной 271 и 286 п.н. были встроены в плазмиды. Туда же были встроены регуляторные участки ДНК, обеспечивающие экспрессию гибридных молекул. Клонированные гены кодировали синтез проинсулина, который путем несложной химической обработки можно превратить в активный инсулин, включающий две полипептидные цепи А и В из 21 и 30 аминокислотных остатков, соединенных между собой сульфгидрильными связями.

Используется также метод искусственного получения генов, основанный на их ферментативном синтезе с помощью механизма обратной транскрипции. Этот механизм связан с активностью РНК-зависимой ДНК-полимеразы или обратной транскриптазы (ревертазы) - фермента впервые обнаруженного при исследовании репликации РНК онкогенных вирусов. Фермент способен строить копии ДНК на разных РНК, включая синтетические полирибонуклеотиды. С помощью ревертазы можно синтезировать практически любой ген в присутствии соответствующих мРНК, методы выделения которых достаточно хорошо разработаны. Данный метод удобно применять для выделения генов, исключительно интенсивно транскрибируемых в определенной ткани.

Таким способом получены и клонированы гены, кодирующие глобины человека, других млекопитающих и птиц, белок хрусталика глаза быка, яичный белок, фиброин шелка и т.д. (Алиханян и др. 1985, стр. 415-418).

Литература к разделу 7.4.

Алиханян С.И., Акифьев А.П., Чернин Л.С. Общая генетика. Москва, Высшая школа, 1-446, 1985.

7.5. Клонирование и анализ ДНК

В процесс молекулярного клонирования (избирательного накопления) молекул ДНК входит несколько этапов: фрагментация ДНК путем обработки эндонуклеазой рестрикции, объединения этих фрагментов *in vitro* с векторной молекулой ДНК (способной к автономной репликации), введение вектора в реципиентный организм, в котором и происходит накопление рекомбинантных ДНК.

Формальной датой рождения генной инженерии считают 1972 год, когда группа П. Берга в США создала первую рекомбинантную ДНК *in vitro*, объединившую в своем составе генетический материал из трех источников: полный геном онкогенного вируса обезьян SV40, часть генома умеренного бактериофага лямбда и гены лактозного оперона *E. coli* (См. Дополнение 7.2). Сконструированная рекомбинантная молекула не была исследована на функциональную активность, поскольку у авторов возникли опасения, что методы генной инженерии могут привести к появлению микроорганизмов, опасных для здоровья человека, например, бактерий *E. coli*, способных перенести онкогенные вирусы животных в кишечник человека.

Функционально активные молекулы гибридной ДНК были впервые получены С. Козном, Д. Хелинским, Г. Бойером и их сотрудниками в 1972-1973 гг. (По: Алиханян и др., 1985, стр. 411-413).

7.5.1. Ферменты рестрикции

В. Арбер (когда?) доказал, что в бактериях действуют специальные ферменты, способные специфично отличать свою (бактериальную) ДНК от чужой (фаговой). Эти ферменты рестрицируют (т.е. ограничивают) возможность размножения фаговой ДНК в бактериях путем ее более или менее специфичной деградации. Такие ферменты были названы эндонуклеазами рестрикции или рестриктазами (Алиханян и др., 413-415). Естественной функцией рестриктаз является защита бактерии от инфекции вирусами. ДНК в сайтах рестрикции у самой бактерии модифицирована метилированием, так что фермент рестрикции не может порезать свою собственную ДНК. Однако вирусная ДНК не защищена, и ферменты ее расщепляют (Russell, 1998, p. 448).

Первая рестриктаза, специфично расщепляющая двухцепочечную ДНК в строго определенных сайтах, была выделена Г. Смитом и его сотрудниками в 1970 году из штамма *Haemophilus influenzae*. Она была обозначена *HindII*.

Известны три класса рестриктаз. В генно-инженерных работах используют ферменты второго класса, разрывающие двухцепочечную ДНК в зоне участка узнавания. Обычно фермент распознает специфичную последовательность из 4-6 п.н., являющуюся палиндромом и

Дополнение 7.4.

В 1978 году В. Арбер, Д. Натанс и Г. Смит (W. Arber, D. Nathans and H. Smith) получили Нобелевскую премию за открытие ферментов рестрикции и их применение в молекулярной генетике.

разрезает ее в середине или несколько в стороне. В последнем случае образуются выступающие одноцепочечные концы, получившие название “липких”. Такие концы, сформировавшиеся под действием одной и той же рестриктазы могут гибридизоваться между собой в силу комплементарности оснований. Эта способность обеспечивает возможность объединения различных молекул ДНК. Примером рестриктаз, образующих липкие концы, являются широко используемые в генно-инженерных работах рестриктазы *EcoRI* и *BamHI*. Первая из них кодируется R плазмидой, присутствующей в штамме *E. coli*, вторая обнаружена в клетках *Bacillus amyloliquefaciens*. Примером рестриктаз, образующих полностью двунитевые (“тупые”) концы, служит *HaeIII*, выделенная из штамма *Haemophilus aegyptium*. Последовательности, распознаваемые этими рестриктазами, показаны на Рис. 7.5. (Алиханян и др., 1985, стр. 413-415). Ферменты рестрикции обозначают по имени организма, из которого они изолированы. Используется трехбуквенная система (пишется курсивом), представляющая собой название вида бактерии, например *BgII*, т.е. из *Bacillus globigi*, *EcoRI* - из *E. coli*, *HindIII* - из *Haemophilus influenzae*. После

трех букв курсив следует какой-то буквенный символ с последующей римской цифрой. Буквенный символ обозначает генетическую линию или расу (Russell, 1998, p. 448).

В настоящее время известно более 400 рестриктаз, способных расщеплять ДНК в общей сложности по почти 120 различающимся последовательностям.

Специфические эндонуклеазы обнаружены у дрожжей родов *Saccharomyces* и *Pichia* (Текст: Алиханян и др., 1985, стр. 413-415; см. также: Russell, 1998, p. 448-452).

7.5.2. Векторы для молекулярного клонирования

Вектор - это молекула ДНК, переносщая клонируемый фрагмент ДНК. Любой вектор должен обеспечивать стабильное наследование рекомбинантных ДНК в автономном состоянии. Кроме того, вектор может иметь дополнительные характеристики, облегчающие генно-инженерные характеристики, например, содержать маркер, инактивирующийся при встройке в него чужеродной ДНК.

Векторные системы для молекулярного клонирования обычно создаются на основе репликонов плазмид и бактериофагов (Алиханян и др., 1985, с. 420-425).

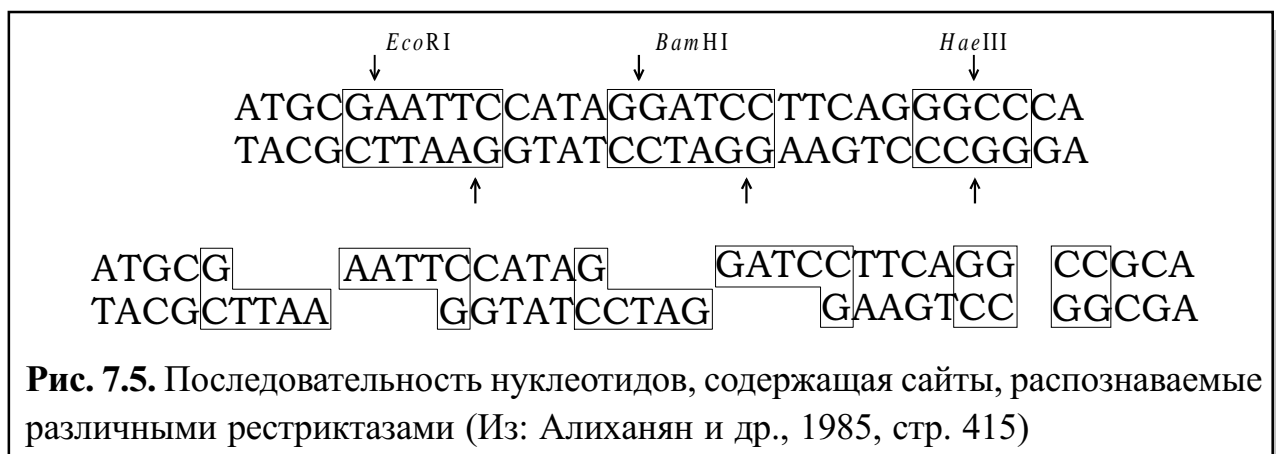


Рис. 7.5. Последовательность нуклеотидов, содержащая сайты, распознаваемые различными рестриктазами (Из: Алиханян и др., 1985, стр. 415)

7.5.2.1. Плазмидные векторы

Плазмида, используемая в качестве вектора при клонировании *E. coli*, должна иметь три характеристики:

1. Иметь сайт *ori* (участок инициации репликации) для нормальной репликации ДНК и размножения плазмиды в клетке *E. coli*.
2. Доминантный селективный маркер, который позволяет отобрать клетку, несущую плазмиду с встроенным фрагментом чужеродной ДНК.
3. Уникальный сайт рестрикции, т.е. встречающийся в векторе единственный раз (Из: Russell, 1998, p. 453).

Весьма распространенные плазмидные векторы (pBR322 и pUC9) ведут свое начало от небольшой плазмиды ColE1, реплицирующейся в клетках *E. coli* независимо от хромосомы и существующей в числе 10-20 копий на клетку. В определенных условиях выращивания бактерий можно добиться избирательной амплификации этих плазмид, в результате чего они образуют до 1000 копий на клетку. Плазмида pBR322 имеет гены устойчивости к антибиотикам ампициллину (Ap^r) и тетрациклину (Tc^r). В гене Tc^r имеется уникальный сайт, разрезаемый рестриктазой *Bam*HI (Рис. 7.6.)

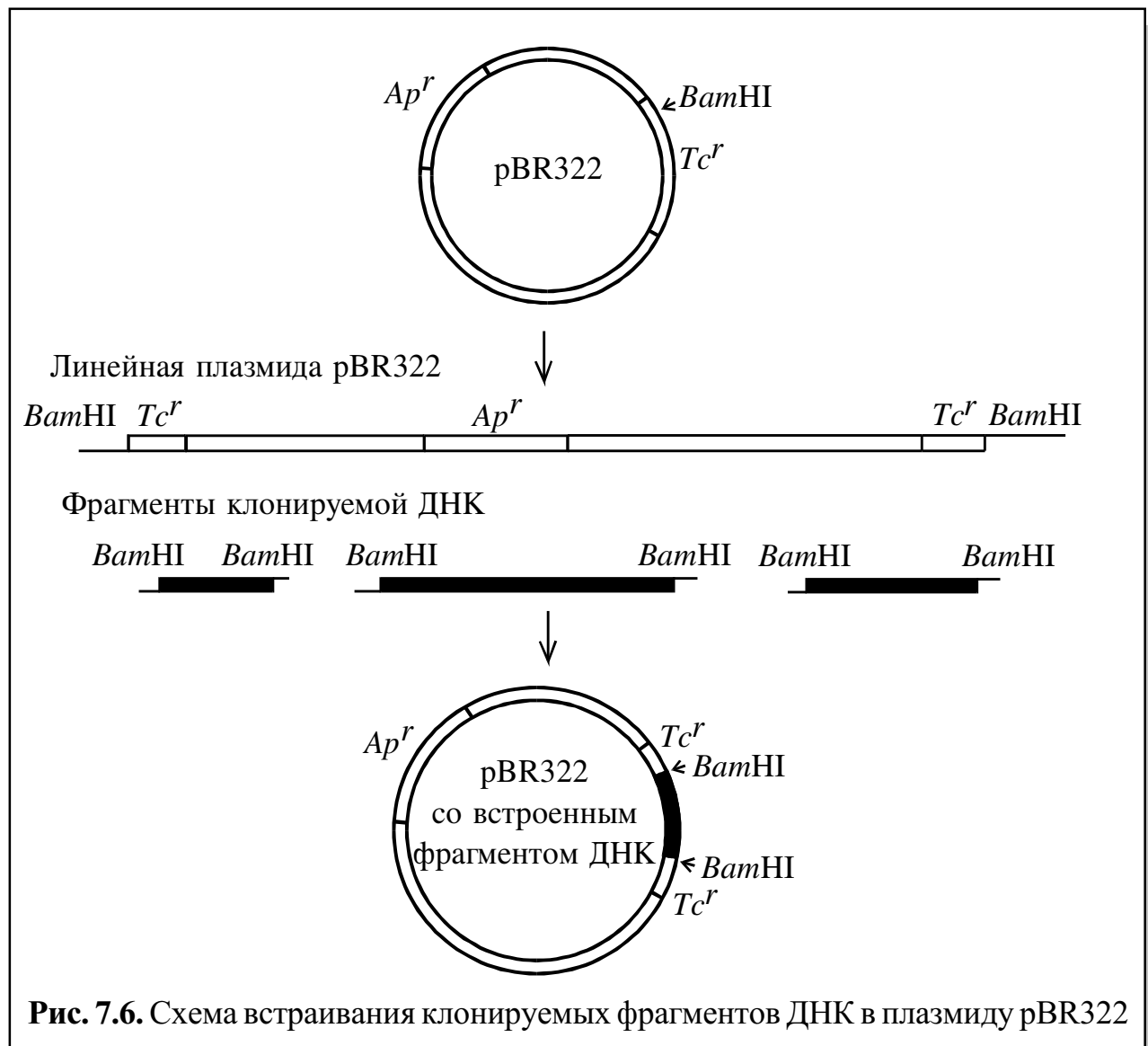


Рис. 7.6. Схема встраивания клонируемых фрагментов ДНК в плазмиду pBR322

После обработки ДНК плазмиды рестриктазой *Vam*HI получается линейная ДНК, имеющая специфичные “липкие” концы. Ее смешивают с фрагментами клонируемой ДНК, обработанной также *Vam*HI и имеющими те же “липкие” концы. Отжиг и обработка ДНК-лигазой приводит к образованию плазмиды, содержащей вставку чужеродной ДНК. Размер встроеного фрагмента - 3-5 т.п.н.

Как и сам процесс встраивания фрагмента в плазмиду, так и

трансформация плазмиды в *E. coli* являются процессами статистическими, в результате которых могут быть три типа клеток: не содержащие плазмиды, содержащие плазмиду без встройки, содержащие плазмиду со встройкой. Чтобы выбрать клетки, несущие плазмиду со встроеными последовательностями ДНК, используют следующую схему эксперимента (Рис. 7.7.). Максимальный размер встройки ДНК составляет 10 т.п.н.

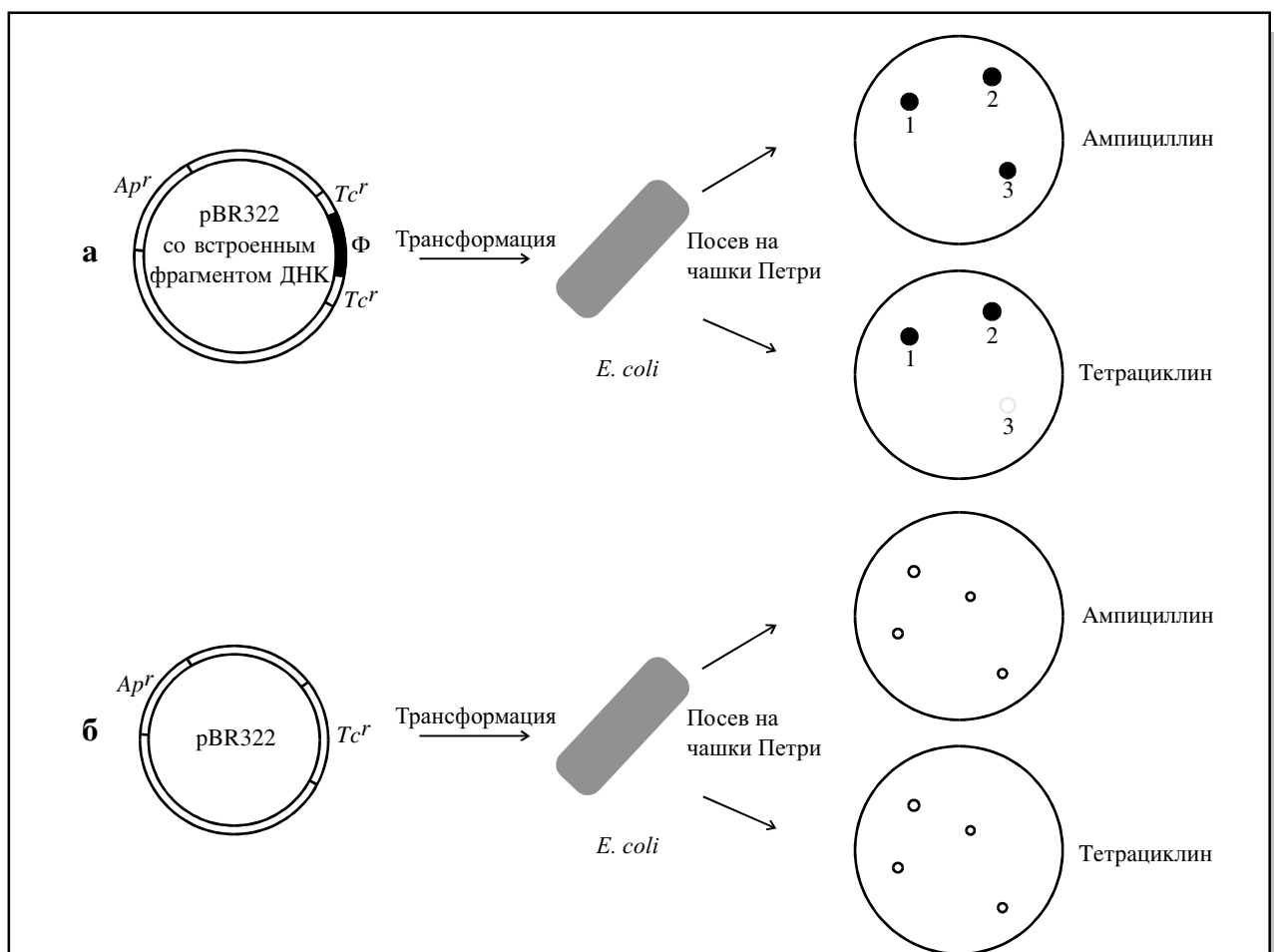


Рис. 7.7. Схема выявления плазмид, содержащих вставку клонируемых фрагментов ДНК. **а.** вставка фрагмента Φ в ген Tc^r нарушает его активность, и бактерия теряет устойчивость к этому антибиотику. После трансформации клеток *E. coli* плазмидами и последовательного посева этих бактерий на газон с ампициллином и тетрациклином, колонии образуются только у трансформантов, имеющих плазмиду. Если клонируемый фрагмент встроился в ген Tc^r , эти колонии (N3 на рисунке) будут отсутствовать. Бактерий из колонии 3 с ампициллинового газона собирают для опытов. **б.** Одинаковое расположение колоний на обоих газонах свидетельствует об отсутствии встройки чужеродной ДНК в плазмиду

Широко используется в качестве вектора плазида рBluescriptII KS(+/-). Это кольцевая молекула, имеющая длину 2961 п.н., создана на основе другой плазмиды - рUC19. Плазида рBluescript характеризуется следующими особенностями (Рис. 7.8.):

1. Число копий достигает 100 на клетку;
2. Селекционным маркером является ген резистентности к ампициллину - *amp^r*;
3. Она содержит целый ряд уникальных сайтов рестрикции, собранных в одном районе. Такие участки называют полилинкерными или участками множественного клонирования - multiple cloning site (MCS);
4. В плазмиде рBluescript участок MCS длиной около 200 п.н. содержит около 20 сайтов рестрикции;
5. Полилинкерный участок расположен в районе 5 конца гена *lacZ⁺*. Плазмиды

культивируют в клетках с мутантным проявлением гена *lacZ*. Поэтому нормальная активность гена *lacZ* достигается за счет работы *lacZ* в плазмиде. В результате инсерции чужеродной ДНК в полилинкер нарушается работа гена *lacZ* и в плазмиде. Колонии бактерий, содержащих *lacZ* и *lacZ⁺*, легко различаются, если поместить клетки на субстрат X-gal, который разрушается β-галактозидазой (продукт гена *lacZ*) с выпадением окрашенного в синий цвет нерастворимого в воде осадка.

Клетки, в которых произошла трансформация плазмидой, отбираются по их способности жить на среде с ампициллином. Те клетки, которые содержат плазмиды со встроенной ДНК, образуют неокрашенные колонии, если колонии

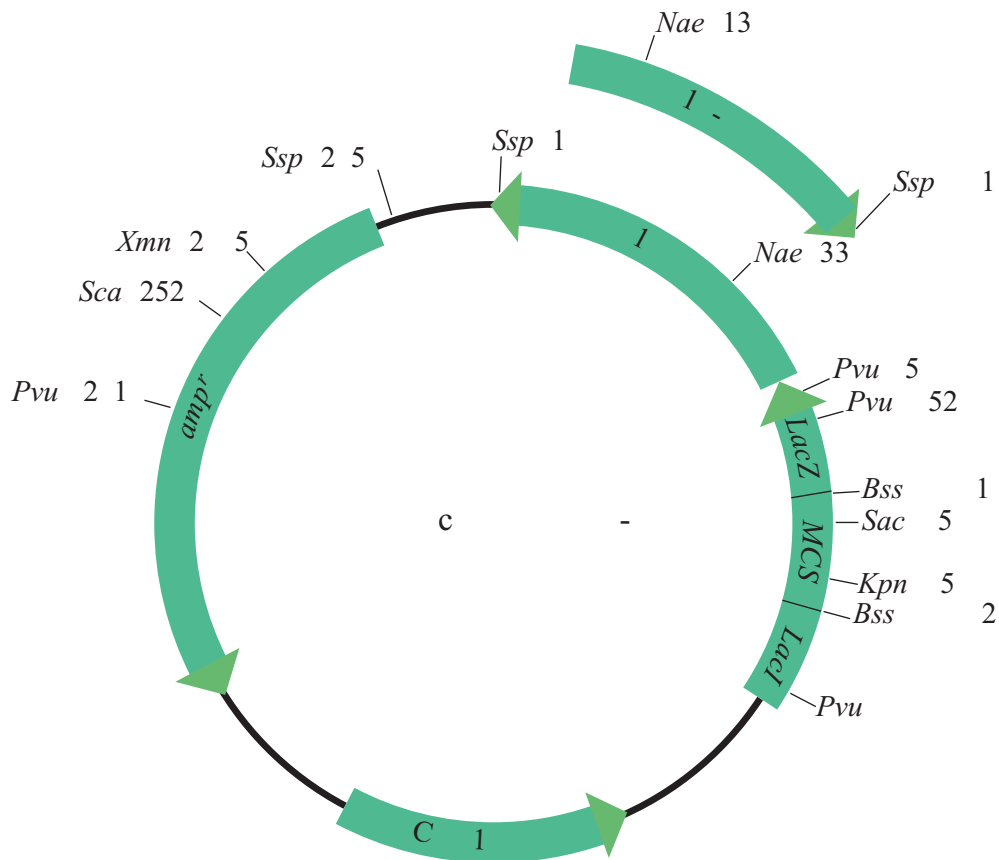


Рис. 7.8. Плазида рBluescript II KS (+/-)

окрашены в синий цвет, значит в этих клетках плазмиды не содержат встройки. Бактериальные искусственные хромосомы (BAC-bacterial artificial chromosome). Эти векторы приготовлены на основе плазмиды F-фактора у *E. coli* (Shizuya et al., 1992).

Эта плазида (Рис. 7.9.) имеет гены, которые не только контролируют репликацию ее ДНК, но и число копий в клетке. Регуляторные гены включают *oriS*, *rep E* (контроль однонаправленности репликации F-фактора), *parA* и *parB* (поддержание числа копий на уровне 1-2 на клетку *E.coli*). Кроме того в плазмиде

есть гены резистентности к хлорамфениколу (CMR) и сегмент клонирования. Последний включает сайты бактериофагов λ cosN и P1loxP (эти два сайта дают возможность получения концов, используемых для построения рестрикционных карт), два сайта инсерции (*HindIII* и *BamHI*), а также несколько C+G - богатых сайта рестрикции (*NotI*, *EagI*, *XmaI*, *SmaI*, *BglI* и *SfiI*) (Рис. 7.9.).

Размер клонируемого фрагмента составляет около 300 т.п.н. Наличие колоний клеток, содержащих инсерцию определялся по гибридизации на фильтрах с клонируемой ДНК (Из: Shizuya et al., 1992).

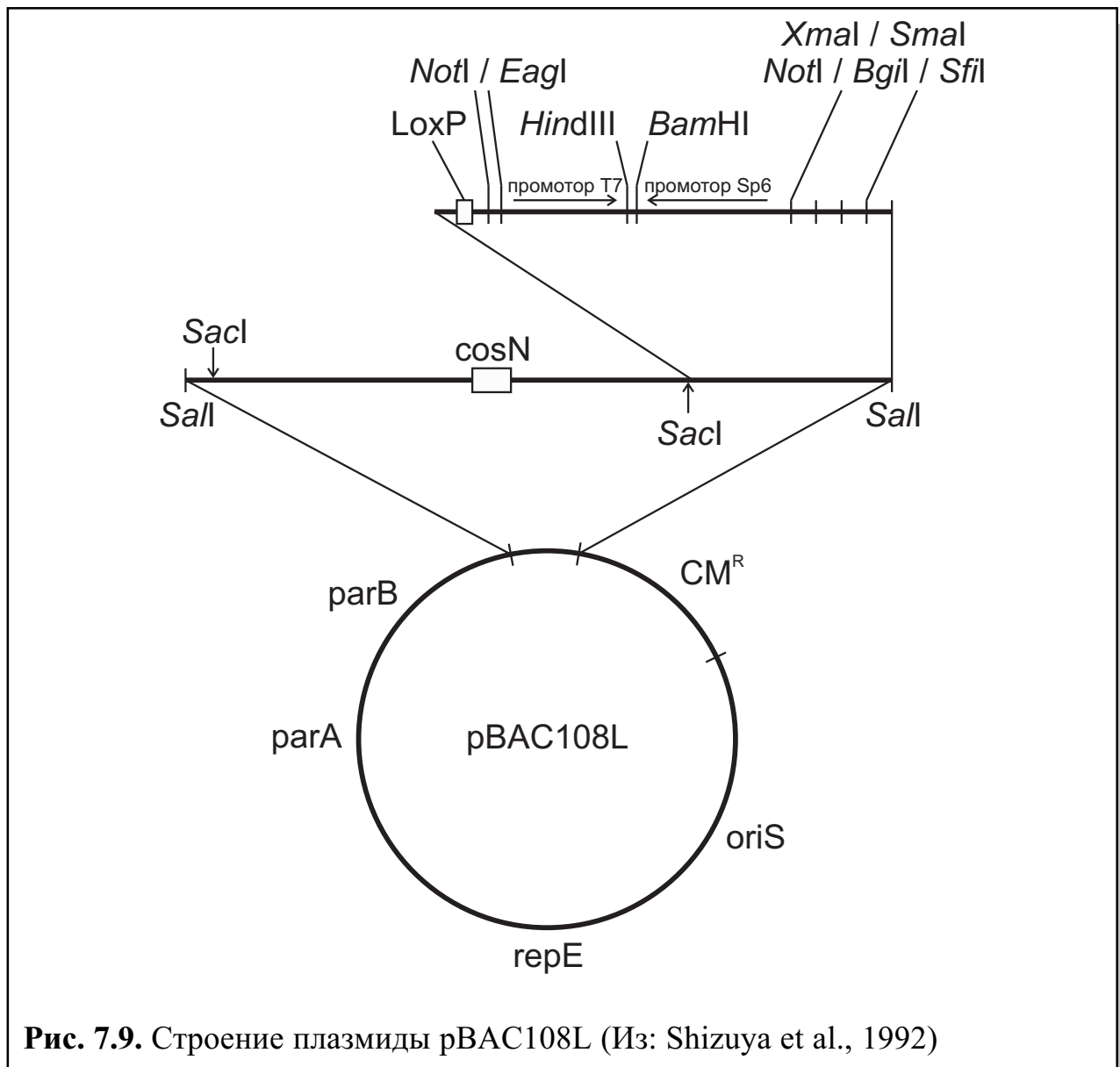


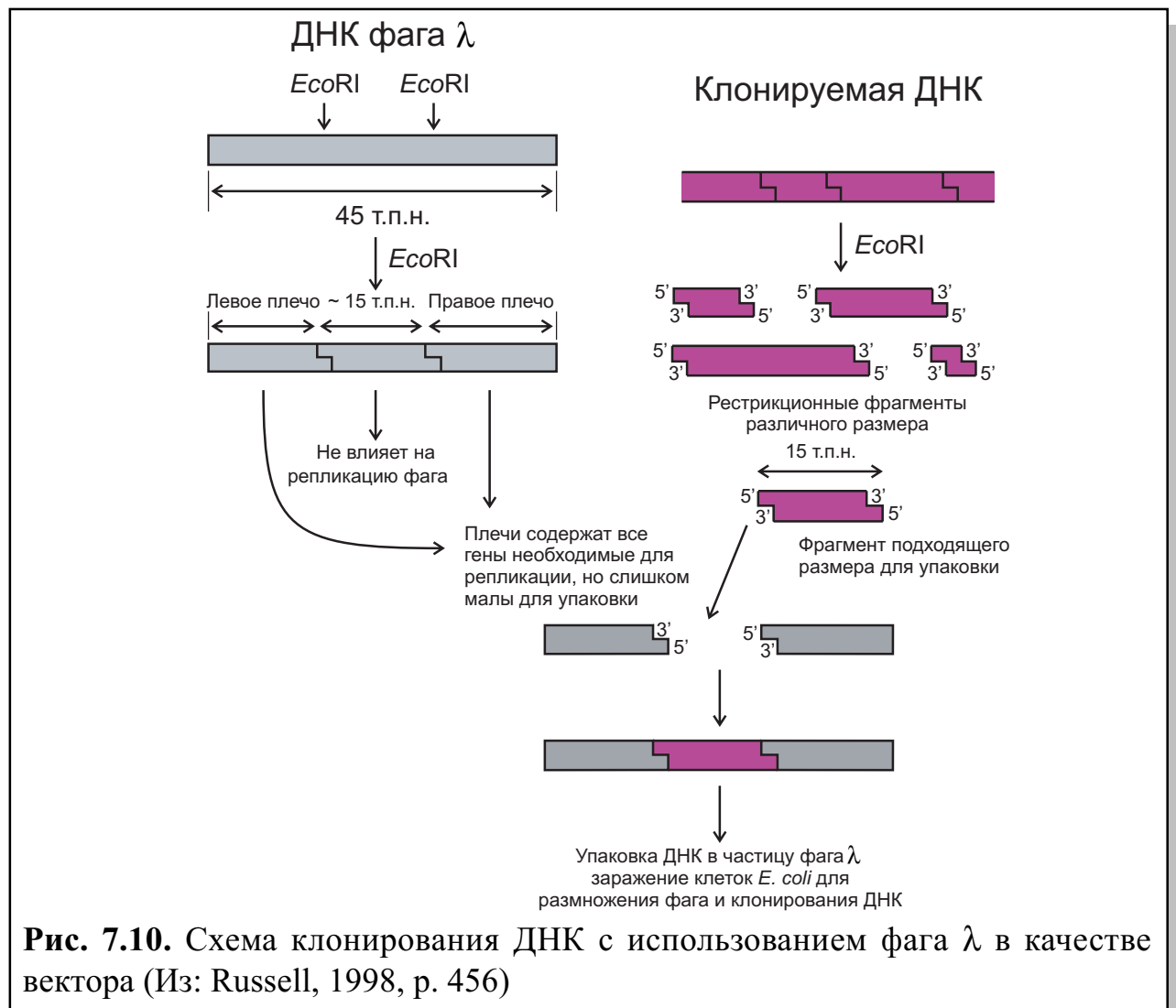
Рис. 7.9. Структура плазмиды pBAC108L (Из: Shizuya et al., 1992)

7.5.2.2. Фаговые векторы

Наиболее часто используемые фаговые векторы являются производными фага λ . Левое и правое плечи фага совместно имеют все гены, необходимые для литического цикла; напротив, гены, контролирующие лизогению, удалены. Поэтому что фаги проходят литический цикл, но лизогения не происходит. Между двумя плечами помещается заменяемый фрагмент ДНК, который не нужен для развития фага (Рис. 7.10.). На стыке заменяемого центрального фрагмента и каждого из плеч находится один и тот же сайт рестрикции, например, *EcoRI*. Другого такого сайта в плечах данного

фага нет. В ходе клонирования, ДНК фага режут рестриктазой *EcoRI*, клонируемый фрагмент ДНК режут этим же ферментом, затем плечи фага и клонируемый фрагмент смешивают, отжигают и затем лигируют стыки цепей.

В частицу λ фага может упаковаться ДНК длиной 37-52 т.п.н., на плечи приходится около 30 т.п.н., т.е. размер вставки может быть около 15 т.п.н. Фаги без вставок или, напротив со слишком большими вставками не развиваются и не поражают бактерии, на газонах *E. coli* выедаются т.н. бляшки (plaques) фагами, имеющими вставки нужного размера (По: Russell, 1998, p. 454).



Фаг P1

Некоторые бактериофаги имеют относительно большие геномы, что позволяет внедрить в них крупные фрагменты чужеродной ДНК. У одного из таких фагов, P1, геном размером 110-115 т.п.н. упаковывается в белковую оболочку. В генно-инженерных опытах компоненты P1 включены в кольцевую плазмиду (Рис. 7.11., см. след. стр.). Затем плазмидный вектор P1 разрезают на 2 плеча, в которых лигируют до 100 т.п.н. чужеродной ДНК. Затем рекомбинантный фаг поглощается подходящей клеткой-хозяином.

7.5.2.3. Космидные векторы

Космиды в природе не встречаются, они созданы искусственно в результате комбинации плазмид и фага λ . Космида имеет последовательность *ori*, позволяющую репликацию в *E. coli*, доминантный селекционный маркер *amp^R* и уникальные сайты рестрикции для инсерций и клонирования фрагментов ДНК (Рис. 7.12.). Кроме этого, космиды имеют *cos*-сайт, происходящий из фага λ . *Cos*-сайты - это участки, в которых многочисленные копии геномов λ -фага, объединенные в длинную цепь, называемую конкатамером, разрезаются на фрагменты длиной 48 т.п.н., пакуемые в фаговые головки. *Cos*-сайты в космидах позволяют упаковываться в λ фагов, что облегчает введение больших молекул ДНК в бактериальную клетку. *Cos*-сайты λ введены в плазмиду и получают небольшие космиды размером 5 т.п.н. Когда фрагменты ДНК размером 32-47 т.п.н. встраиваются в такую космиду, общая молекула достигает нужной длины, чтобы

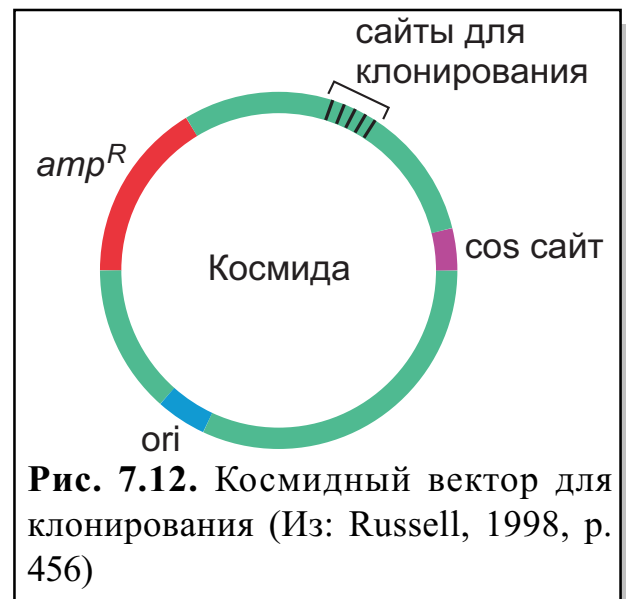


Рис. 7.12. Космидный вектор для клонирования (Из: Russell, 1998, p. 456)

упаковаться в фаговую частицу. Этот фаг затем вводится в клетку *E. coli*.

Размер клонируемого фрагмента ДНК в космидном векторе составляет 35-45 т.п.н. (Из: Russell, 1998, p.455).

4.5.2.4. Челночные векторы

Это векторы, которые способны реплицироваться в двух и более организмах-хозяевах. На рис. 7.13. показан вектор, способный размножаться в клетках дрожжей и *E. coli*.

YEp24 имеет последовательность *ori*, позволяющую реплицироваться в *E. coli* и селекционные маркеры *amp^R* и *tet^R* (*E. coli*), а также URA3 (контроль биосинтеза урацила для клеток дрожжей), последовательность, специфичную для дрожжей, - 2m circle, позволяющую реплицироваться автономно в дрожжевой клетке.

Таким образом, YEp24 способна реплицироваться как в клетках дрожжей, так и *E. coli*.

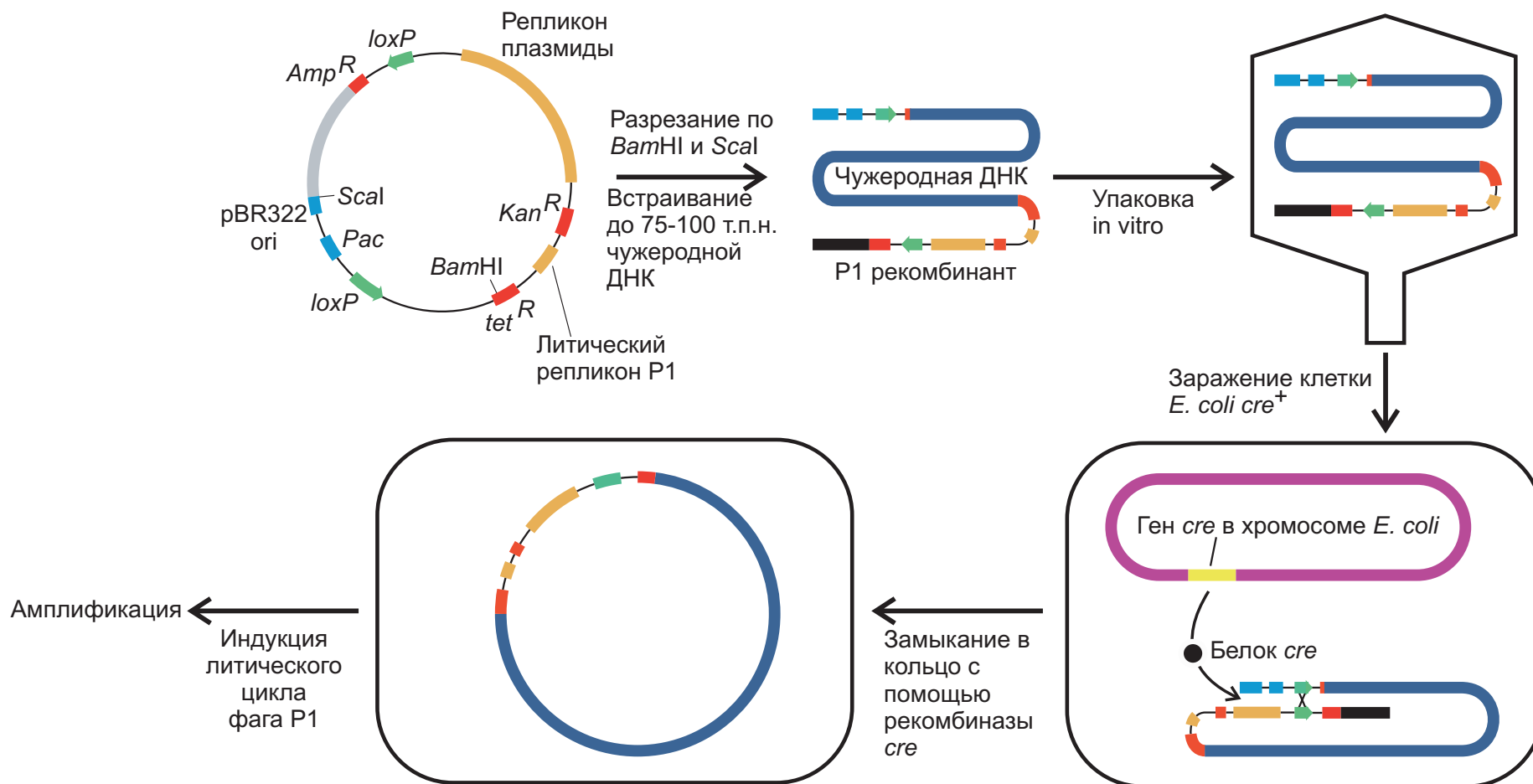


Рис. 7.11. Схема клонирования в векторе фага P1. Плазмида P1 содержит различные элементы генома P1; *pac*-сайт упаковки, два сайта *loxP* (опознаются фаговой рекомбиназой, кодируемой геном *cre*). Вектор режется так, что получают 2 плеча: короткое и длинное, в которое лигируются 85-100 т.п.н. ДНК. Упаковка рекомбинантной ДНК происходит in vitro с использованием упаковочных экстрактов P1: *pacase* расщепляет рекомбинантную ДНК в сайтах *pac* (Из: Strachan, Read, 1996, p.102)

4.5.2.5. Искусственные хромосомы дрожжей (YAC)

Векторы этого типа имеют следующие характеристики: 1. Теломеры на каждом конце; 2. Центромерные последовательности (CEN); 3. Селекционные маркеры на каждом плече (*TRP1* и *URA3* - независимость от наличия триптофана

и урацила, соответственно); 4. Автономно реплицирующиеся последовательности - ARS, позволяющие реплицироваться в дрожжевой клетке; 5. Рестрикционные сайты, уникальные для участков YAC, и пригодные для инсерций ДНК (Рис. 7.14).

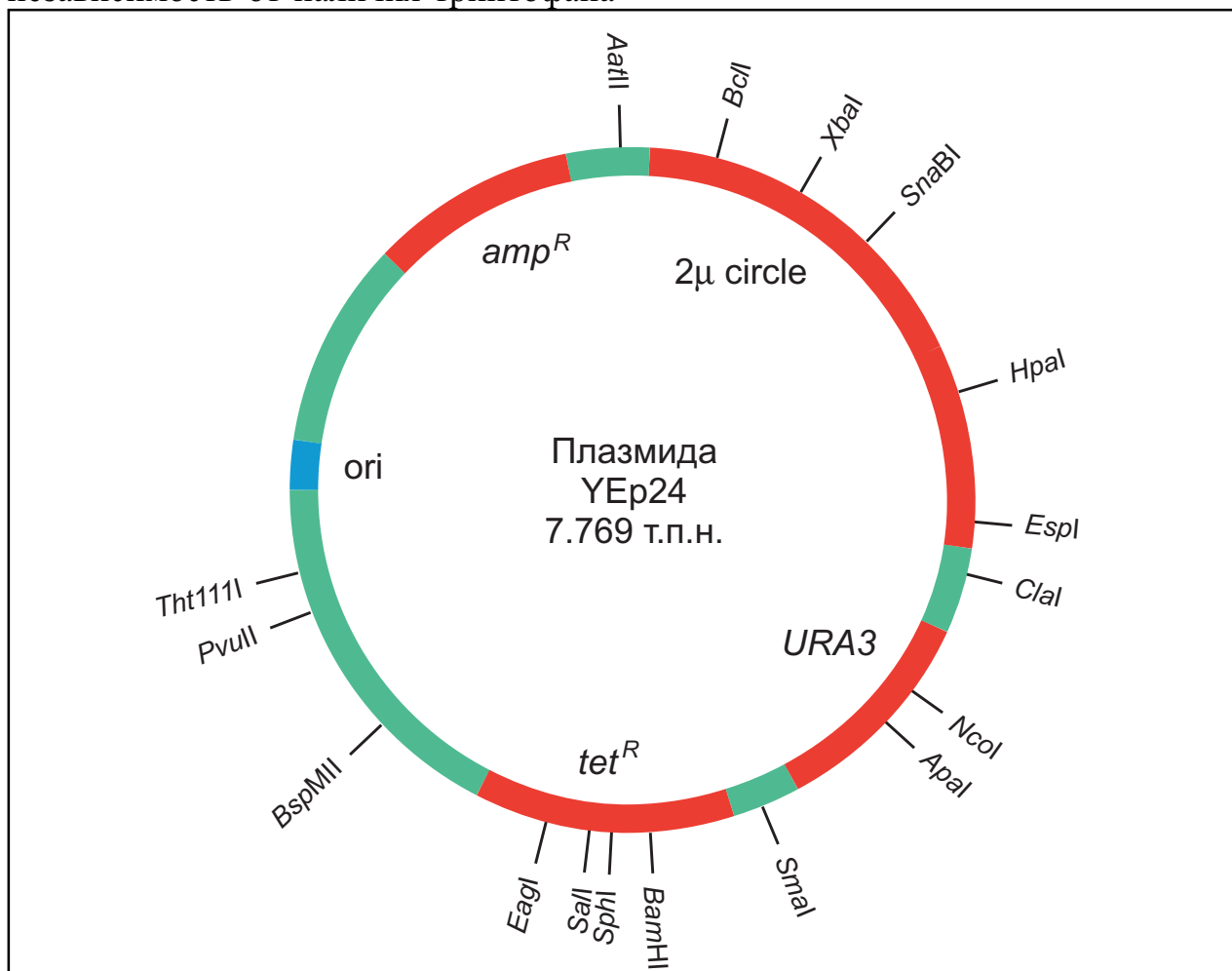


Рис. 7.13. Челночный вектор YEp24 между клетками *E. coli* и дрожжей. Показаны уникальные сайты рестрикции (Из: Russell, 1998, p.457)

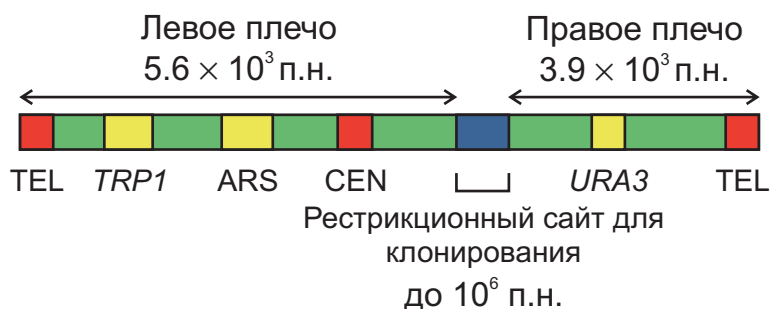


Рис. 7.14. Структура YAC-вектора (Из: Russell, 1998, p.458; Alberts et al., 1994, p.339, Burke et al., 1987)

Клоны вводятся в клетки дрожжей трансформацией. Отбирая по маркерам *TRP1* и *URA3*, можно быть уверенным, что инсерция ДНК имеет и левое, и правое плечо. Размеры клонируемых последовательностей ДНК велики: обычно несколько сот т.п.н. и даже больше.

(Текст: Алиханян и др., 1985, стр. 420-425. Другие векторы см. также стр. 421-423; Russell, 1998, p. 454-455).

7.5.3. Создание геномных библиотек

Коллекция клонов, содержащая хотя бы по одному экземпляру каждого из фрагментов ДНК, входящих в состав генома данного вида, называется геномной библиотекой.

Аналогичные коллекции, полученные из индивидуальных хромосом или их частей, называют хромосомными библиотеками.

Получение геномных библиотек включает выделение тотальной ДНК, фрагментацию ее с помощью рестриктаз, присоединение полученных фрагментов к векторным молекулам (плазмидного или фагового происхождения) и введение рекомбинантных ДНК в реципиентные бактерии (Алиханян и др., 1985, стр. 426). Известно три подхода, применяемых при создании геномных библиотек:

1. геномная ДНК полностью переваривается рестриктазой и затем заключается в клонирующий вектор. Недостатком этого метода является то, что размеры клонируемых инсерций довольно малы, около 4 т.п.н.
2. ДНК фрагментируется механическими воздействиями (пропускание через иглу шприца, обработка ультразвуком), что



Рис. 7.15. Частичное переваривание ДНК с помощью рестриктаз приводит к появлению серии перекрывающихся фрагментов, имеющих одинаковые липкие концы. Вертикальными стрелками помечены рестрикционные сайты, горизонтальными - фрагменты ДНК (Из: Russell, 1998, p.459)

позволяет получить более крупные фрагменты - 20-150 т.п.н., однако из-за отсутствия липких концов у фрагментов ДНК, необходимы дополнительные манипуляции.

3. Применяют частичное переваривание рестриктазами. Это означает, что только часть имеющихся сайтов рестрикции используется ферментом для разрезания (Рис. 7.15.) (Текст: Russell, 1998, pp.458-459).

Эта техника позволяет получить набор клонов бактерий или рекомбинантных фагов, различающихся по включенным фрагментам ДНК. Первую геномную библиотеку создали Д. Хогнесс с сотрудниками в 1974 году. Они использовали ДНК из генома *D. melanogaster*, которую расклонировали в клетках *E. coli* (Текст: Алиханян и др., 1985, стр. 426).

Копии ДНК, комплементарные РНК (комплементарная ДНК, кДНК или cDNA) могут составлять библиотеку кДНК. Анализ клонированных молекул кДНК дает информацию о гене, кодирующем эту кДНК. Поскольку

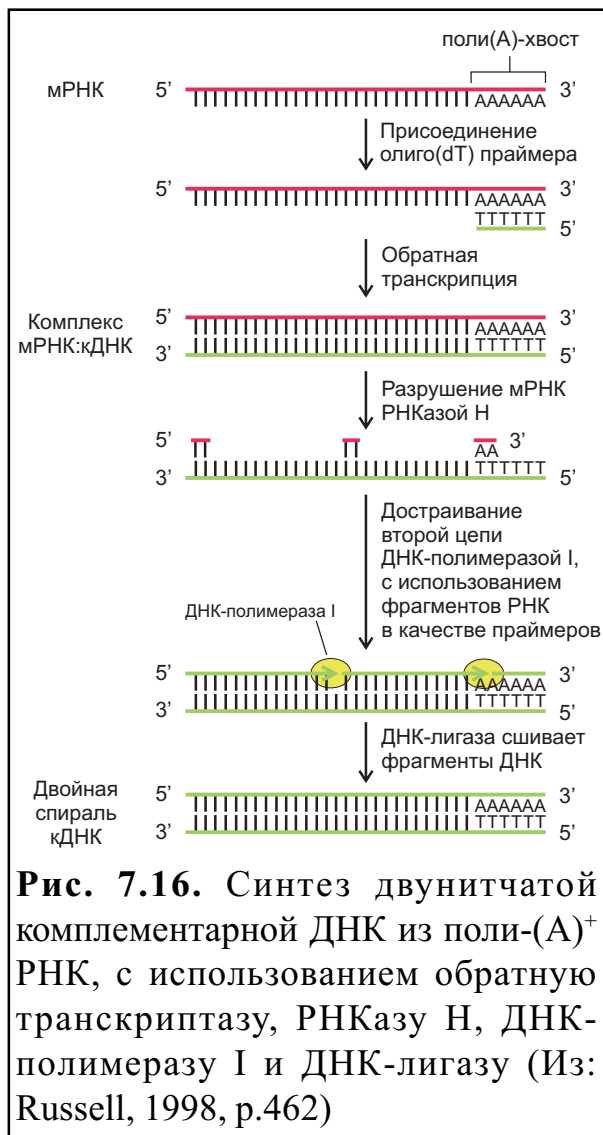
библиотека кДНК отражает спектр генной активности в данном типе клеток, из которых была выделена РНК, создание и анализ библиотек из кДНК особенно полезно для сравнений генной активности в клетках разных типов. Следует также учитывать, что клоны в библиотеке кДНК представляют зрелые, т.е. после процессинга, мРНК. Они не содержат интронов.

У эукариот только мРНК содержат поли(А)⁺-хвосты. Используя это качество, мРНК выделяют. Затем в системе *in vitro* к этой мРНК добавляют короткую цепь олиго(дТ) (Рис. 7.16.), которая после отжига служит в

качестве праймера для действия обратной транскриптазы (РНК-зависимой ДНК-полимеразы), синтезирующей комплементарную цепь ДНК на молекуле мРНК (происходит как бы обратная транскрипция: РНК → ДНК). Далее, с помощью РНКазы Н, ДНК-полимеразы I и ДНК-лигазы удаляется цепь РНК, на ее месте синтезируется новая ДНК-цепь, фрагменты которой лигируются (Рис. 7.16.). (Текст: Russell, 1998, p.461).

После того как молекулы кДНК синтезированы, их готовят для клонирования. К синтезированным молекулам кДНК с помощью T4 ДНК-лигазы добавляют линкерную ДНК - относительно короткий двуничатый фрагмент ДНК длиной 8-12 нуклеотидов (Рис. 7.17.). Линкер содержит сайт рестрикции одной из рестриктаз. По этому сайту кДНК-овые клоны лигируют с векторной ДНК (Из: Russell, 1998, pp.462-463).

Для того, чтобы выявить клоны, находящиеся в библиотеках, и выделить их, существуют различные способы. Рассмотрим как скринируют библиотеку, т.е. анализируют и выделяют клоны, заключенные в λ фагах (Рис. 7.18.). Сначала выращивают газон бактерий *E. coli*, на который помещают суспензию фагов, содержащих геномные клоны. В тех участках, где фаги инфицировали бактерий и размножились, образуется бляшка (plaque), т.е. как бы дырка на поверхности газона *E. coli*. Каждая бляшка состоит из большого числа потомков одиночной фаговой частицы, размножившихся в инфицированных и лизированных бактериях. Всякий раз, когда бактерия лизируется,



освобождаются не только упакованные фаговые частицы, но и неупакованная фаговая ДНК.

Мембранный фильтр кладут на этот газон на несколько минут, в результате чего ДНК из бляшки связывается с фильтром. Последующая обработка фильтра щелочным раствором приводит к денатурации ДНК на одиночные нити. После этого фильтры инкубируют с меченым зондом. Гибридизация меченого зонда

с ДНК из бляшки выявляется в виде темного пятна засветки на фотопленке. По положению засвеченного пятна находят положение бляшки на газоне. Фаги с этой бляшки можно собрать и, инфицировав новые бактерии, размножить фагов, а следовательно и клон до количеств, необходимых для молекулярного анализа.

7.5.4. Построение рестрикционных карт

Поскольку клонированные последовательности ДНК представляют собой гомогенную популяцию молекул ДНК, расщепление их с помощью рестриктаз приводит к образованию более мелких дискретных фрагментов. С помощью специального анализа можно определить расположение рестрикционных фрагментов в молекуле ДНК, т.е. построить рестрикционную карту.

Существует несколько способов построения рестрикционных карт. Для того, чтобы получить точную, детальную и полезную для работы карту, чаще всего нельзя ограничиться только одним из них. Наиболее употребительными являются следующие способы:

- 1) одновременное расщепление ДНК несколькими рестриктазами.
- 2) последовательное расщепление выделенного рестрикционного фрагмента второй рестриктазой.
- 3) частичное расщепление немеченой ДНК или ДНК, меченой по одному концу.
- 4) частичное расщепление ДНК экзонуклеазами с последующим расщеплением рестриктазой.

Какой из этих методов использовать в каждом конкретном случае, зависит от многих причин. По размерам

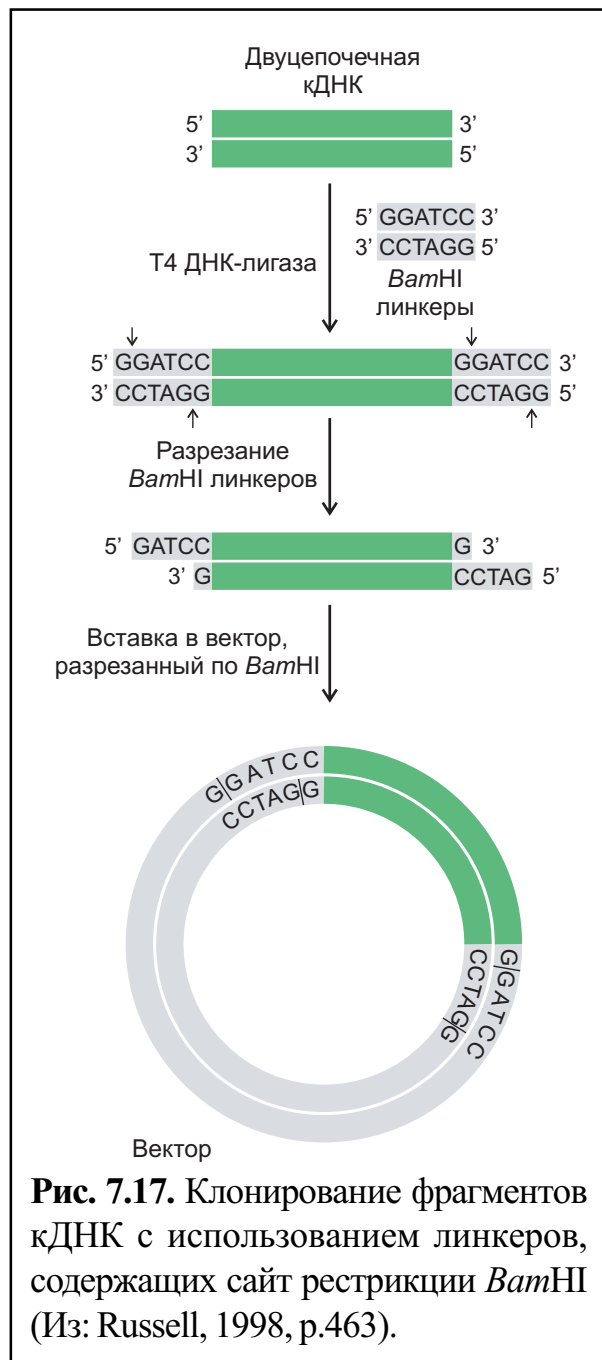


Рис. 7.17. Клонирование фрагментов кДНК с использованием линкеров, содержащих сайт рестрикции *Bam*HI (Из: Russell, 1998, p.463).

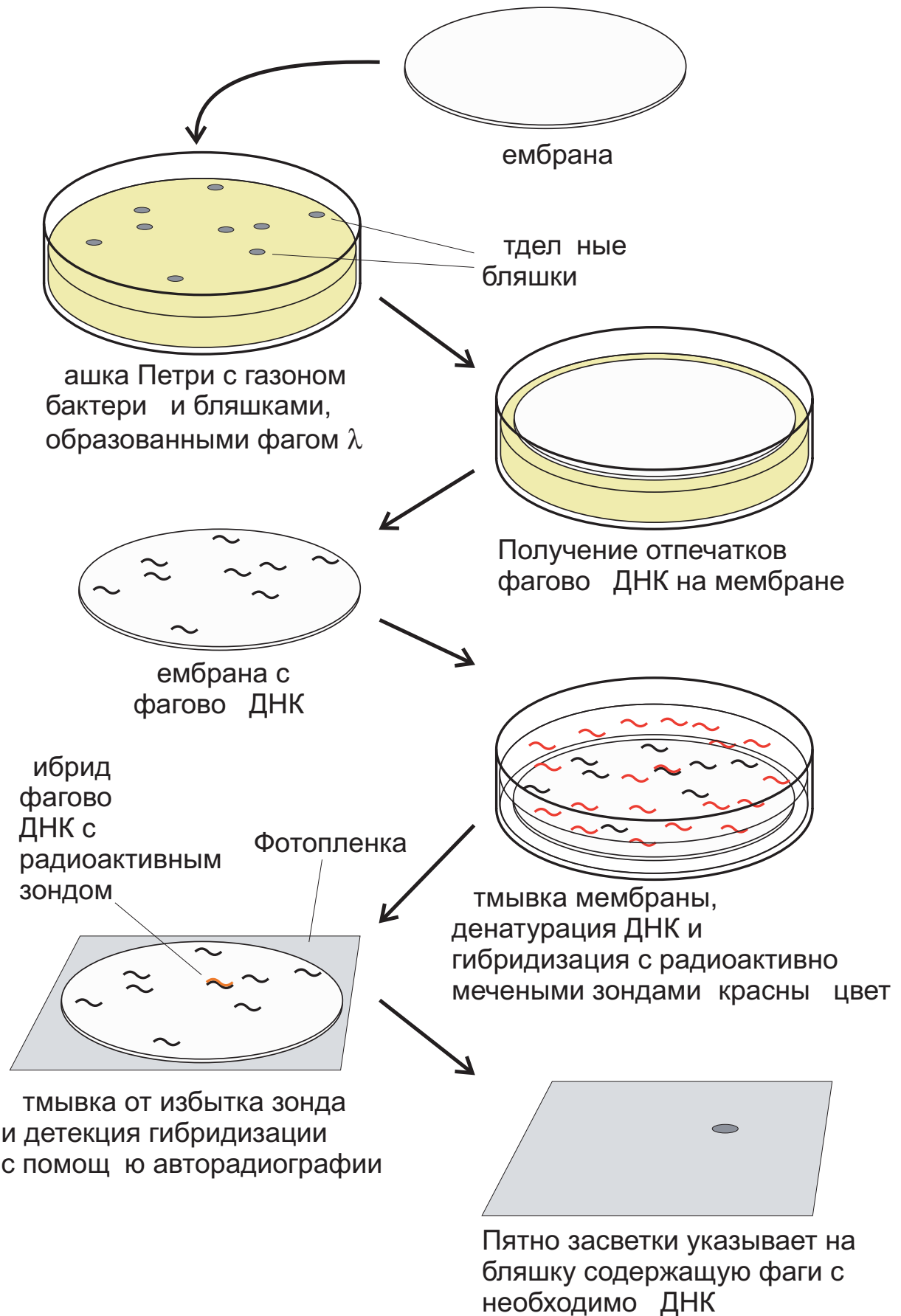


Рис. 7.18. Скрининг библиотеки λ фагов и выделение индивидуального фага, содержащего клон ДНК (Из: Russell, 1998, p.468)

образующихся фрагментов ДНК удается определить относительное местоположение по крайней мере некоторых участков расщепления. С увеличением числа парных комбинаций ферментов увеличивается число определенных участков вплоть до получения окончательной картины... В результате получается карта рестрикционных сайтов (Рис. 7.8.) (Из: Маниатис и др., 1984, стр. 337-338).

На Рис. 7.19. представлен простейший пример рестрикционного картирования. Допустим, что был клонирован фрагмент ДНК длиной 5 т.п.н. Теперь одну аликвоту ДНК режут рестриктазой *EcoRI*, другую - *BamHI*, третью - смесью обеих рестриктаз. После электрофоретического разделения фрагментов можно определить их размеры сопоставляя с размерами известных фрагментов ДНК (маркеры размеров ДНК). После электрофореза гель красят этидиум бромидом и фотографируют в ультрафиолетовом свете. На фотографиях измеряют расстояния, пройденные тем или иным фрагментом от старта. Для маркеров молекулярный размер каждого фрагмента известен, равно как и расстояние пройденное со старта. Поэтому легко составить калибровочную кривую (Рис. 7.19.). Затем на этой кривой находят положение каждого фрагмента по длине пройденного пути и определяют размер. В конкретном случае, изображенном на рис. 7.19., очевидно, что после обработки *EcoRI* появляются два фрагмента 4.5 и 0.5 т.п.н., это свидетельствует о наличии только одного сайта рестрикции и этот сайт находится на 0.5 т.п.н. от одного из концов исходного фрагмента. По этой же

логике есть только один рестрикционный сайт *BamHI*, локализованный на 2 т.п.н. от одного из концов. При совместном переваривании двумя рестриктазами получаются три фрагмента: 2.5, 2.0 и 0.5 т.п.н. Расположить эти фрагменты непротиворечиво с данными, полученными при действии каждой из рестриктаз в отдельности, можно лишь так, как нарисовано окончательной рестрикционной карте на Рис. 7.19.

В реальных экспериментах построение рестрикционных карт - значительно более сложный процесс (Текст: Russell, 1998, p. 469-471).

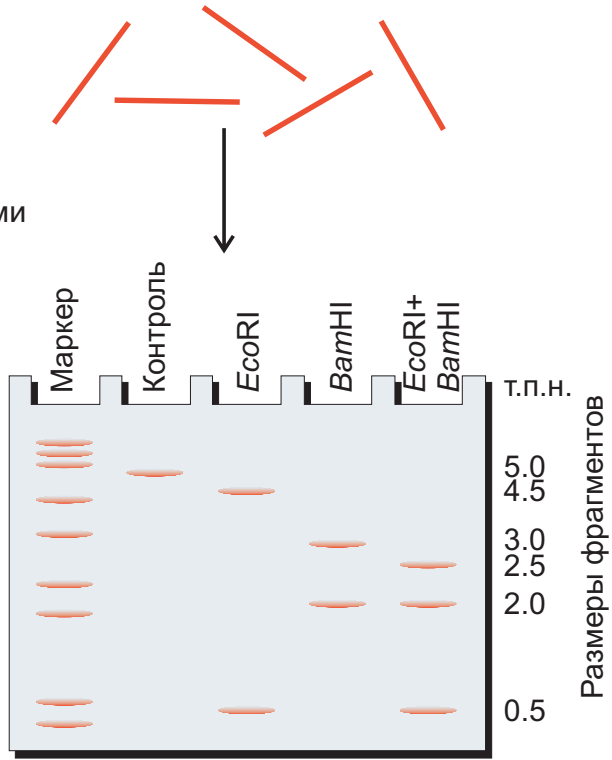
7.5.5. Саузерн-блот анализ

Для картирования хромосомных перестроек на физической карте, определения положения гена, выявления повторов, хромосомной ходьбы и т.д. используется Саузерн-блот анализ или геномный блот. Для разделения фрагментов ДНК обычно используют электрофорез в агарозном геле. Однако, дальнейшие процедуры с молекулами, находящимися в геле, невозможны, и поэтому применяют метод Саузерн-блот (назван по фамилии изобретателя - Эдварда Саузерна). ДНК в геле денатурируют, и к гелю прикладывают нитроцеллюлозный фильтр. ДНК переносится на фильтр, где фиксируется обработкой высокой температурой или ультрафиолетом. Эта процедура напоминает прикладывание промокательной бумаги к тексту с еще не высохшими чернилами (по-английски blotting - промокание). Таким образом, все фракции ДНК, разделившиеся в геле в результате электрофореза, оказались перенесенными на фильтр (Рис. 7.20.). Фильтр затем помещают в

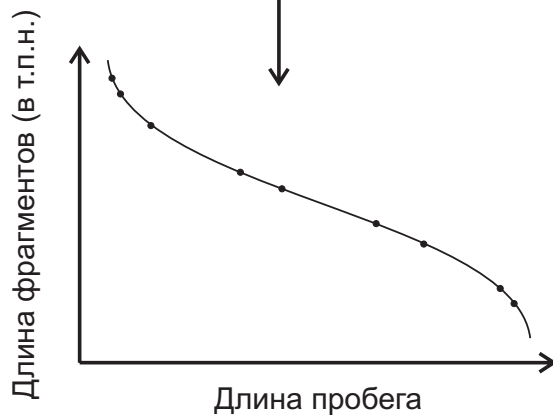
Клонированные линейные фрагменты ДНК длиной 5 т.п.н.

Разрезание рестриктазами и электрофорез в агарозном геле

Длинные фрагменты (-)
↓
Короткие фрагменты (+)



Построение калибровочной кривой по длине маркерных фрагментов и определение длины анализируемых фрагментов



Построение рестрикционной карты

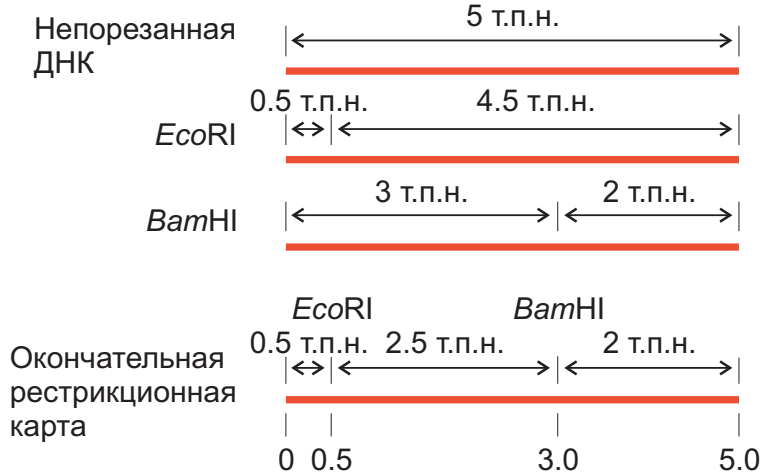


Рис. 7.19. Построение рестрикционной карты по сайтам рестриктаз *EcoRI* и *BamHI* (Из: Russell, 1998, p.471)

гибридизационный буфер, в котором находится зонд, меченый радиоактивным изотопом или химическим путем. Этот зонд свяжется с комплементарным участком ДНК. После промывки фильтра можно приступить к выявлению сигнала. Если зонд был радиоактивным, к фильтру прикладывают рентгеновскую пленку и выявляют участок засветки. Если зонд мечен с помощью химических модификаторов, то реакцию проводят с фильтром и на нем же выявляют сигнал.

Пример использования Саузерн-блотов представлены на Рис. 7.21.

7.5.6. “Хромосомная ходьба”

Процесс клонирования с помощью т.н. “хромосомной ходьбы” (walking) позволяет получать длинные тракты ДНК. Для начала “ходьбы” нужен специфический зонд, т.е. фрагмент ДНК из данного района хромосомы; это может быть ДНК гена, или ДНК Р-элемента при условии, что вследствие его инсерции в клонируемом гене индуцирована мутация. Таким образом “ходьба” - это процесс идентификации прилежащих клонов в геномной библиотеке.

Процесс “ходьбы” включает следующие этапы:

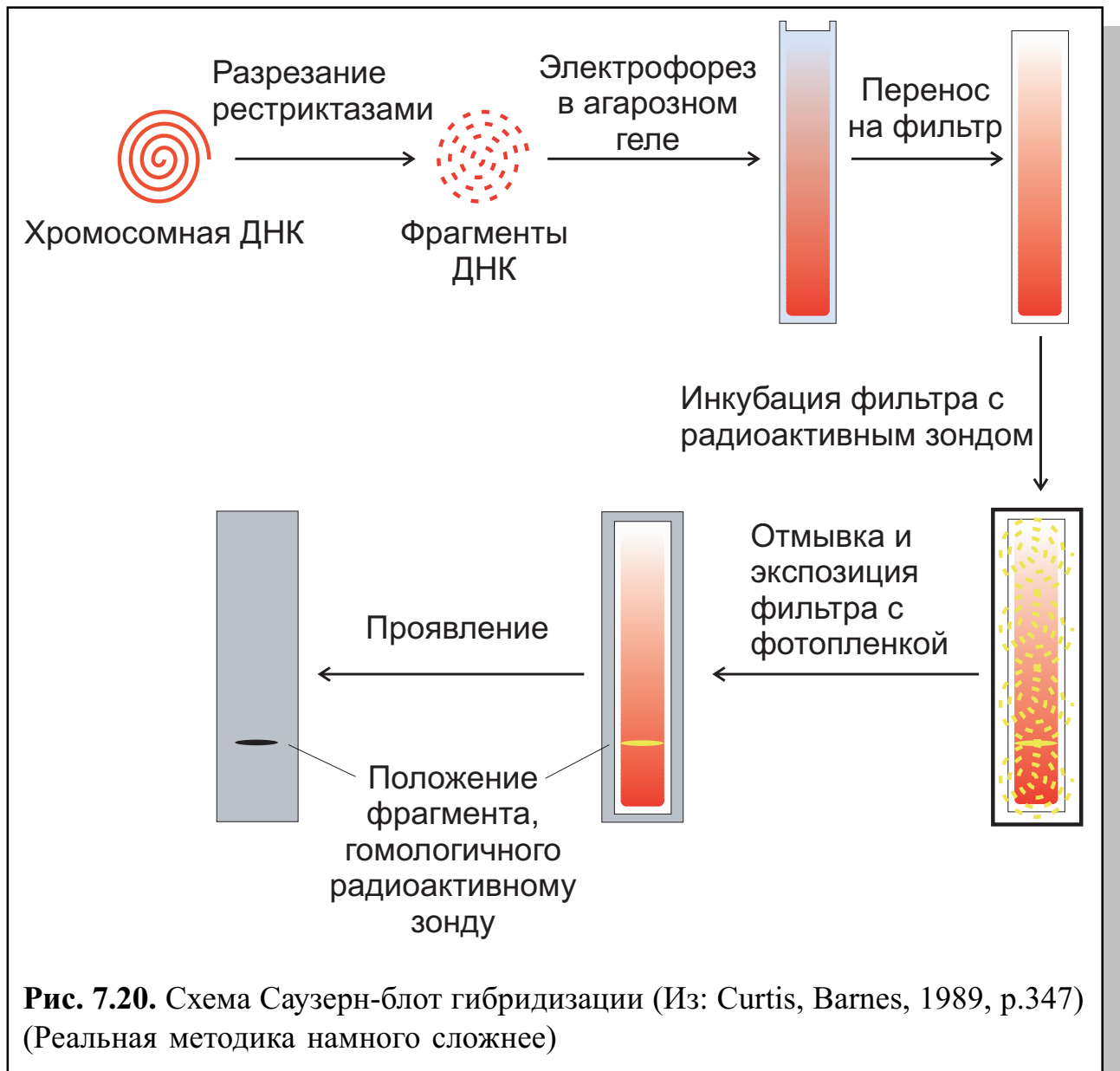
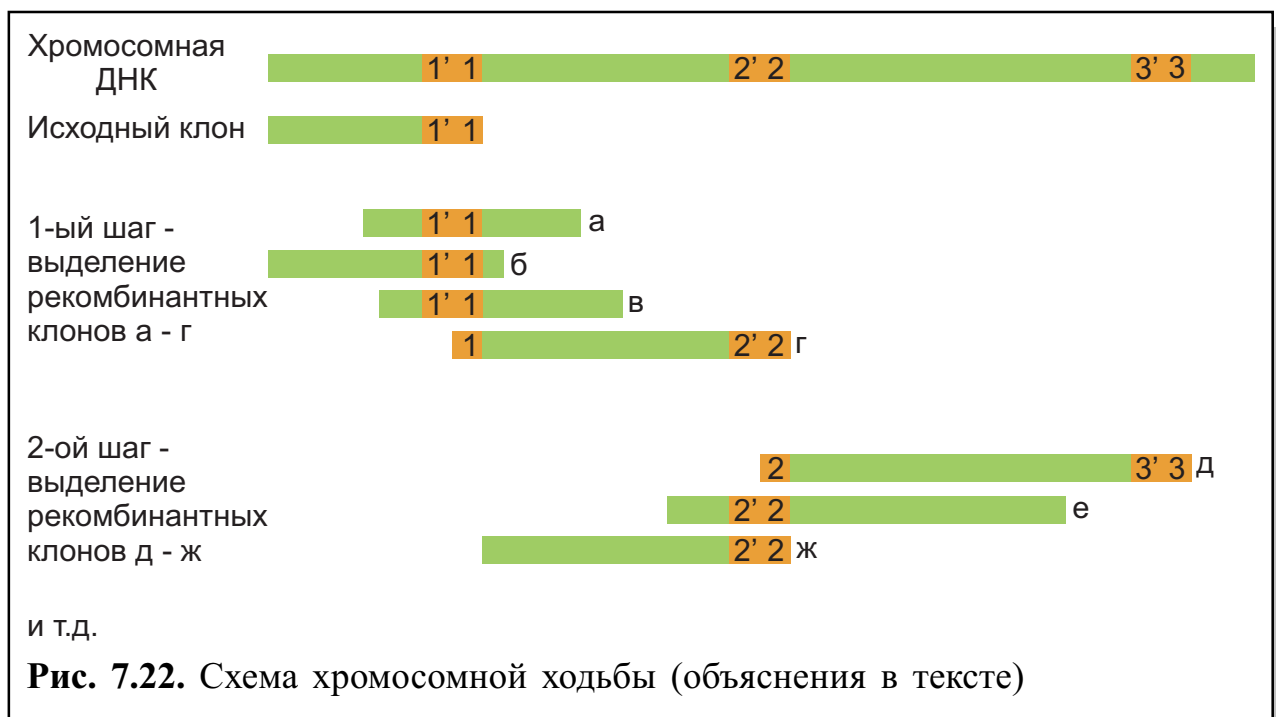


Рис. 7.20. Схема Саузерн-блот гибридации (Из: Curtis, Barnes, 1989, p.347) (Реальная методика намного сложнее)



ограничением является небольшая длина каждого шага.

7.5.7. Нозерн-блот анализ

Метод, похожий на Саузерн-блот, но разработанный для анализа РНК, называется Нозерн-блот (Northern). В данном случае РНК, выделенная из клетки, разделяется по размерам с помощью гель-электрофореза, затем молекулы переносятся на фильтр. После гибридизации с меченым одноцепочечным зондом выявляются сигналы в определенной области электрофореза, в которой обнаружена гомология РНК и ДНК зонда. Используя маркеры молекулярного веса, можно определить размер транскрипта какого-то гена.

Применяют Нозерн-блот в следующих типах экспериментов:

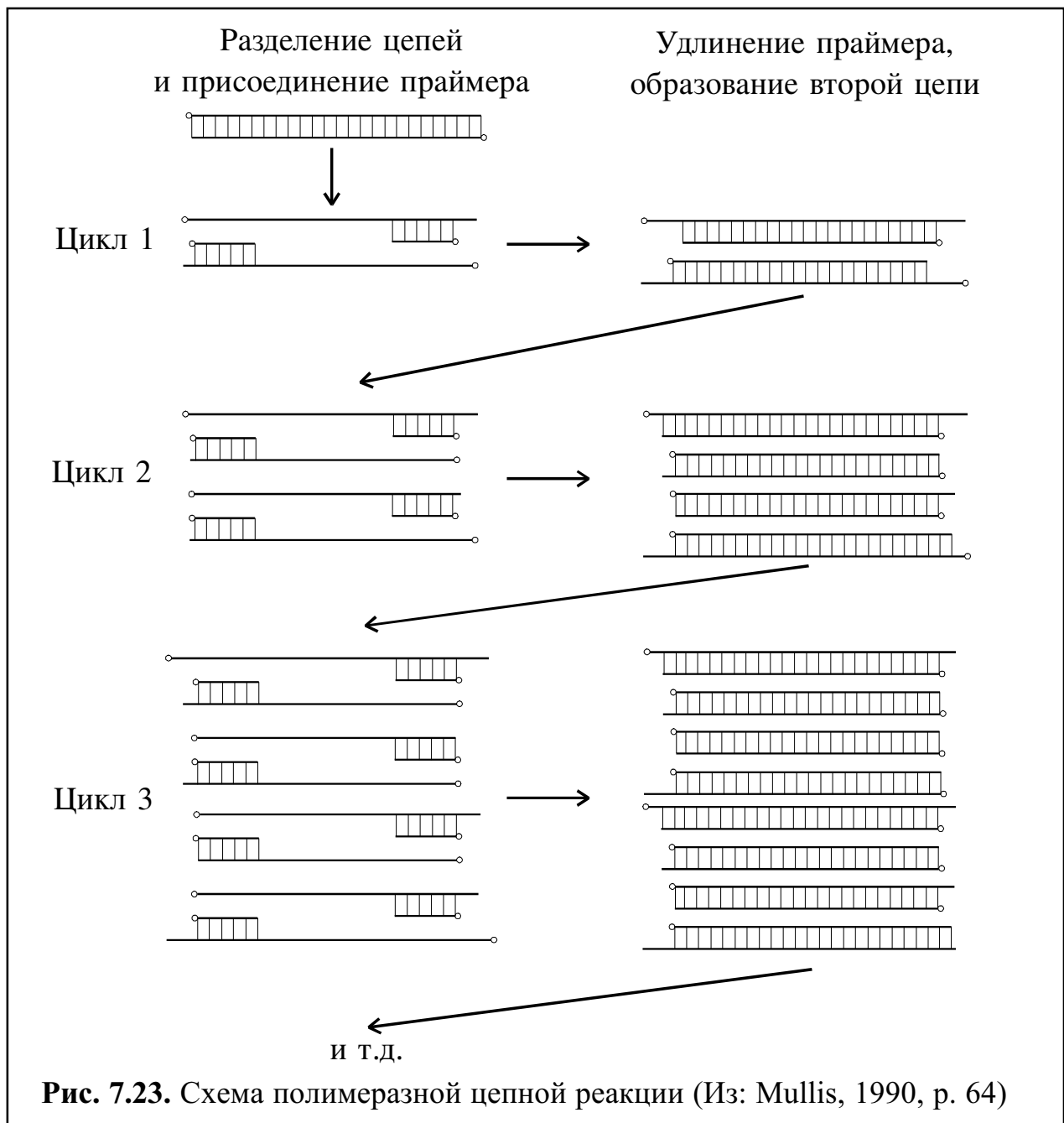
1. для определения размера или размеров специфических мРНК, кодируемых данным геном.
2. для выяснения того, присутствуют или нет в данном типе клеток мРНК,

считанные с данного гена, т.е. экспрессируется ген или нет.

3. как много присутствует этой РНК, как изменяется ее количество в развитии данного типа клеток (Из: Russell, 1998, p.473).

7.5.8. Полимеразная цепная реакция

Размножить определенный небольшой участок генома можно с помощью метода полимеразной цепной реакции (ПЦР), который был изобретён в 1983 году (Mullis, 1990) (Рис. 7.23.): исходную молекулу ДНК прогревают, чтобы разрушить водородные связи, соединяющие комплементарные нити. Затем реакционную смесь охлаждают в присутствии двух коротких (12-20 пар нуклеотидов) фрагментов ДНК, один из которых комплементарен участку ДНК слева от изучаемого локуса, а второй комплементарен участку другой нити справа от изучаемого локуса. Эти фрагменты, называемые праймерами, связываются с соответствующими



участками ДНК. Связавшиеся с нитью ДНК праймеры задают точку начала синтеза новой комплементарной нити на матрице ДНК. Осуществляет этот синтез фермент ДНК-полимераза. В следующем цикле раствор с полученными нитями ДНК снова прогревают и вновь синтезированные нити ДНК используют в качестве матрицы. Новая порция праймеров связывается с соответствующими участками, и происходит новый цикл

синтеза. В живой клетке ДНК-полимераза осуществляет редупликацию ДНК при делении клетки, то есть полностью воспроизводит геномную ДНК. При проведении ПЦР синтезируется только небольшой фрагмент, расположенный между двумя праймерами. За 30 циклов количество синтезированных фрагментов составит около 1 миллиарда.

Прогревание смеси производят при 98°C , а охлаждение - при 60°C ; эти

температуры весьма высоки для того чтобы обычная ДНК-полимераза нормально работала. Поэтому используют особый фермент (Тақ-ДНК-полимераза), выделенный из бактерий *Thermus aquaticus*, которые живут в горячих источниках (Mullis, 1990).

Метод ПЦР лежит в основе идентификации личности, установления родства людей, в судебной медицине. Так, при расследовании убийства семьи последнего российского царя, были использованы небольшие количества ДНК из костных останков 9 человек (в количествах, измеряемых миллиардными долями грамма - примерно 20 пикограмм, что составляет количество ДНК, содержащейся примерно в 10 клетках).

Определение половой принадлежности по ДНК основано на том, что половые *X* и *Y* хромосомы человека несут гомологичный ген, находящийся в них в двух аллельных состояниях. Этот ген кодирует компонент зубной эмали - амелогенин, в *Y* хромосоме он на 6 пар нуклеотидов длиннее, чем в *X* хромосоме. Для того, чтобы выявить, содержится ли в данном образце ДНК из

Y хромосомы, из всей геномной ДНК анализируют лишь область амелогенинового гена, содержащую участок, различный в *X* и *Y* хромосомах. Размер этой области составляет одну тридцатимиллионную долю от всего генома.

Для определения размера полученных ПЦР-фрагментов их разделяют с помощью электрофореза в агарозном геле. На Рис. 7.24. показаны результаты такого разделения фрагментов гена амелогенина после проведения ПЦР на ДНК из костных останков.

На девяти дорожках справа - фрагменты, полученные на ДНК из костных останков. Слева - две дорожки с контрольными образцами: фрагментами, синтезированными на геномной ДНК, выделенной из крови мужчины и женщины. У мужчин образуется два фрагмента: длиной 112 п.н., соответствующий *Y* хромосоме, и 106 п.н., соответствующий *X* хромосоме. У женщин имеется две *X* хромосомы, каждая из которых дает фрагмент 106 п.н. В исследованных девяти образцах в 4 случаях видны две

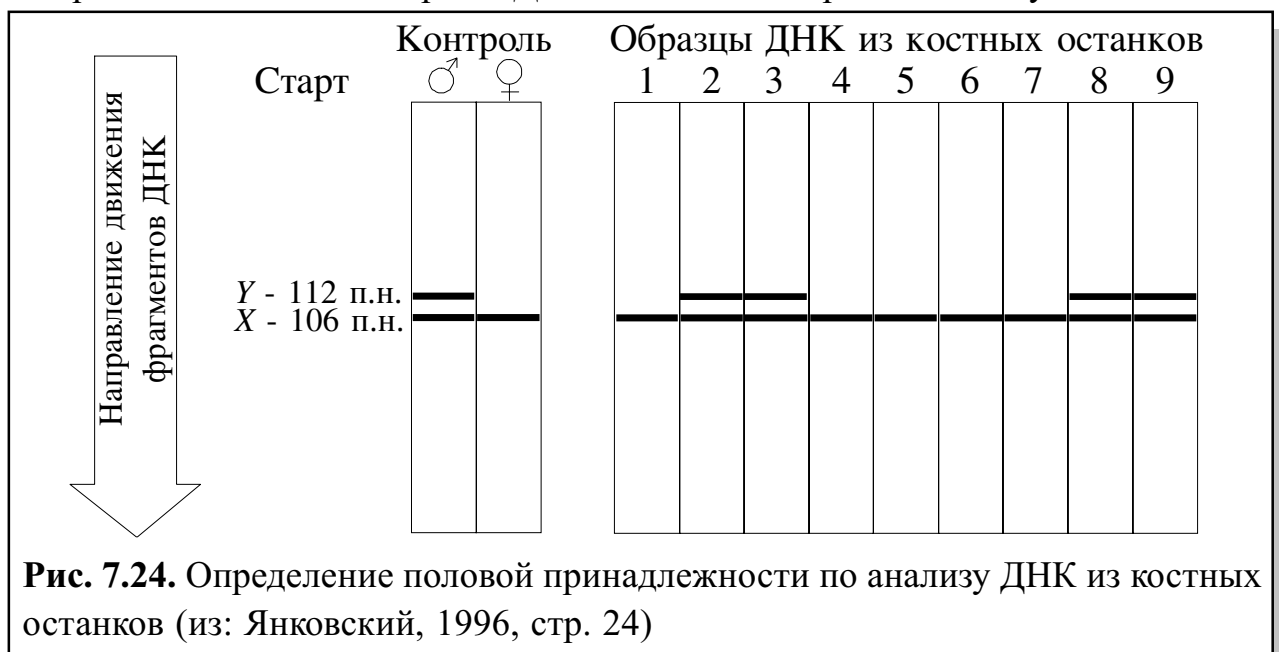


Рис. 7.24. Определение половой принадлежности по анализу ДНК из костных останков (из: Янковский, 1996, стр. 24)

полосы, а в 5 - одна полоса (Рис. 7.24.). Следовательно, среди убитых было 4 мужчины и 5 женщин. Именно столько женщин было в царской семье, расстрелянной в Екатеринбурге в июле 1918 года (Из: Янковский, 1996).

7.5.9. Определение последовательности нуклеотидов (секвенирование)

(Текста нет) Рис. 7.25.

Литература к разделам 7.5.1.-7.5.9.

- Алиханян С.И., Акифьев А.П., Чернин Л.С. Общая генетика. Москва, Высшая школа, 1-446, 1985.
- Alberts B., Bray D., Lewis J., Raff M., Roberts K., Watson J.D. Molecular biology of the cell (Third edition) Garland Publishing, Inc. New York, London, p. 287-301, 1994.
- Иванов П.Л. Молекулярно-генетическая индивидуализация человека. Автореферат докт. диссер., Москва, 1995.
- Лещинская И.Б. Генетическая инженерия. Соросовский образовательный журнал N1, 32-39.
- Маниатис Т., Фрич Э., Сэмбрук Дж. Молекулярное клонирование. Москва, Мир, 1-479, 1984.
- Свердлов Е.Д. Очерки современной молекулярной генетики. Молекулярная генетика, микробиология и вирусология, N2, N3, N4, 1995.
- Фаворова О.О. Лечение генами-фантастика или реальность? Сорос. образ. журн. N2, 21-27, 1997.
- Янковский Н.К. Молекулярно-генетические методы в руках детектива, или опыт исследования останков семьи последнего российского императора. Соросовский образовательный журнал N2, 21-27, 1996.
- Antonarakis S.E., Scott H.S. The human genome project and its impact in medicine. European Review 4, N4, 415-426, 1996.
- Birnstiel M.L. The genetic revolution 1950-1996: an introduction. European Review 4, N4, 333-334, 1996.
- Birnstiel M.L. Gene therapy. European Review 4, N4, 335-356, 1996.
- Curtis H., Barnes N.S. Biology. Fifth Edition. Worth Publishers, Inc. New York, 1989.
- Lewin B. Genes. Oxford, New York, Tokyo; Oxford University Press, pp. 1-1272, 1994.
- Mullis K.B. The unusual origin of the polymerase chain reaction. Scientific American, April, 1990, pp. 56-65.
- Murrey N.E. Gene cloning. European Review 4, N4, 357-370, 1996.
- Russell, P.J. Genetics Fifth edition. Addison Wesley Longman, Ins, Menlo Park, California, 1805, 447-453, 1998.
- Sibilia M., Wagner E.F. Transgenic animals. European Review 4, N4, 371-391, 1996.
- Shizuyer H., Birren B., Kim U.-J., Mancino V., Slepak T., Tachiiri Y., Simon M. Cloning and stable maintenance of 300-kilobase-pair fragments of human DNA in *Escherichia coli* using an F-factor-based vector. Proc. Natl Acad. Sci. USA V. 89, 8794-8797, 1992.
- Strachan T., Read A.P. Human Molecular Genetics. Bios Scientific Publishers limited. Oxford, p. 96-106, 1996.
- Walden R., Maas C., Martini N., Schell J. Transgenic plants. European Review 4, N4, 393-414, 1996.

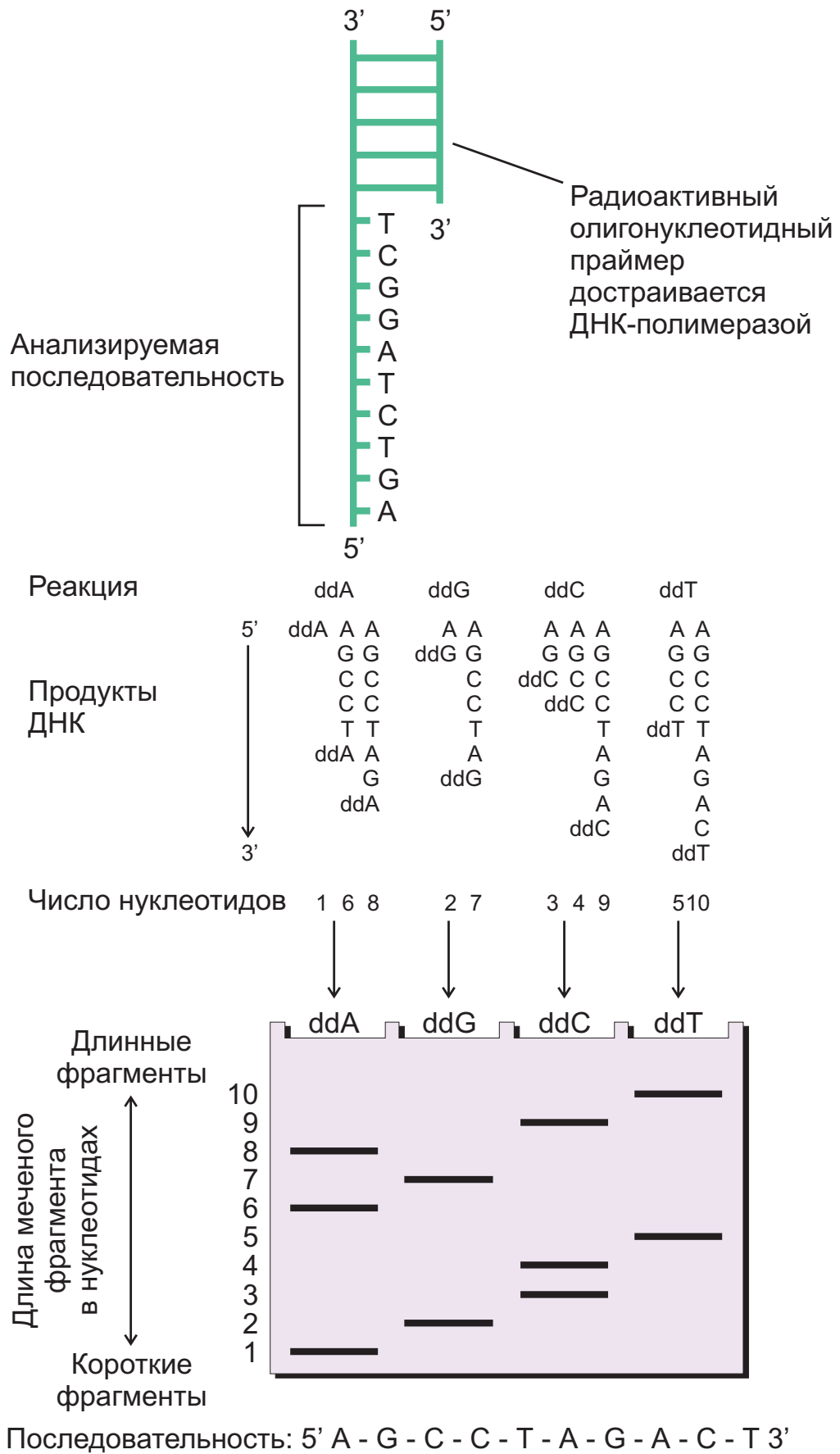


Рис. 7.25. Схема опыта по определению последовательности нуклеотидов

7.5.10. Трансформация у дрозофилы

Трансформацией в генетике называют перенос генов из одного организма в другой. Трансформации у прокариот посвящен раздел 5.10.

У эукариот трансформация в природе до сих пор не обнаружена. Что же касается экспериментальной передачи ДНК из одной клетки в другую, то на протяжении многих лет она у исследователей не получалась. Многие группы ученых азартно конкурировали друг с другом в конце 1970-х - начале 1980-х годов, пытаясь первыми осуществить перенос ДНК. Успехи были весьма редкими: молекулы ДНК, инъецированные микрохирургически в ранние эмбрионы некоторых видов млекопитающих, амфибий, морских ежей, иногда встраиваются в хромосомы клетки-хозяина. У дрозофилы искусственно введенная ДНК почти не включается в хромосомы эмбриона, хотя в культуре клеток типичная трансформация (или как говорят, трансфекция) происходит довольно часто. Однако, очевидно, что из такой клетки невозможно получить организм, и трансформированная ДНК не может передаваться потомству.

В конце концов повезло двум американским исследователям Дж. Рубину и А. Спрадлингу (Рис. 7.26.) в 1982 году. Для осуществления переноса ДНК вектора они использовали мобильный Р-элемент. Для осуществления транспозиции у Р-элемента существуют специально устроенные структуры - концевые повторы - по 150 п.н. определенной последовательности на каждом конце элемента совершенно необходимы для



Рис. 7.26. Аллан Спрадлинг

перемещений. Именно с этими участками связывается транспозаза и катализирует процесс встраивания-вырезания Р-элемента. Таким образом, для нормального перемещения Р-элемента нужно, чтобы были выполнены два условия: наличие у него неповрежденных концевых повторов и нормальная активность транспозазы. Если будет нарушен какой-либо из экзонов гена транспозазы или во время транскрипции и созревания матричной РНК транспозазы не будет удален один из интронов, полноценного фермента транспозазы не получится, и Р-элемент не будет перемещаться. Однако, недостаток транспозазы из одного Р-элемента может быть легко компенсирован, если в геноме присутствует еще один Р-элемент, нормальный, который и обеспечит наличие транспозазы.

Учитывая все перечисленное выше, эксперимент по трансформации

у дрозофилы схематически выглядит следующим образом: готовят ДНК двух Р-элементов, один из которых может синтезировать транспозазу, но не способен встраиваться в геном, так как у него поврежден один из концевых (с 3'-конца) повторов - рп25.7wc на Рис. 7.27. - в то время как во втором Р-элементе удалена значительная часть гена транспозазы, но он может перемещаться по геному, поскольку имеет нормальные концевые повторы - это $P[(ry^+)A]$ на Рис. 7.27.

В Р-элемент $P[(ry^+)A]$, являющийся вектором и имеющий дефективный ген транспозазы, обычно встраивают дополнительный фрагмент ДНК длиной до 20 т.п.н., который хотят трансформировать в геном дрозофилы. Р-элемент $P[(ry^+)A]$ называют транспозоном, а рп25.7wc - элементом-хелпером или помощником.

Затем смесь этих двух сортов ДНК инъецируют в яйцо дрозофилы. Там некоторая ее часть встраивается в ДНК полярных клеток, из которых позднее образуются клетки зародышевого пути. Таким образом, чужеродная ДНК, инъецированная в эмбрион на ранних стадиях развития, передается следующим поколениям потомков.

Многие факторы, в том числе и размер встроеного в Р-элемент фрагмента чужеродной ДНК влияют на успех трансформации, в среднем только около 14% инъецированных эмбрионов развиваются в плодовитое потомство. Поэтому встает вопрос о том как различить мух, у которых инъецированная ДНК встроилась в геном и мух без встройки. Для этого в Р-элемент $P[(ry^+)A]$ вводят специальные маркерные гены. Это могут быть либо гены дрозофилы, контролирующие окраску глаз, например, ген *rosy* (ry^+ на Рис. 7.27.),



Рис. 7.27. Схема скрещиваний для осуществления трансформации у дрозофилы (Из: Spradling, 1986, p. 177)

гены дрозофилы, кодирующие синтез ферментов, например, алкогольдегидрогеназы или даже ген, выделенный из генома медузы, кодирующий белок - GFP, обнаруживающий при определенных условиях зеленую флюоресценцию.

Разберем сначала схему, в которой используется ген *rosy* (Рис. 7.27.). В этом случае в Р-элемент $P[(ry^+)A]$ введен фрагмент ДНК, содержащий нормальный аллель гена *rosy*. ДНК этого транспозона вместе с ДНК-хелпера инъецируют в эмбрионы, гомозигонные по мутации *rosy* - ry^{506} . Если транспозон $P[(ry^+)A]$ встроился в геном, то в поколении G1 появятся мухи с нормальным цветом глаз, поскольку у них - гомозигот по мутации ry^{506} - появится ДНК с нормальным аллелем гена ry^+ в транспозоне.

Недавно для маркировки произошедшего встраивания транспозона в качестве маркера стали использовать кДНК гена GFP, подшитого вместе с подходящим промотором. В случае инсерции транспозона в различных клетках дрозофилы синтезируется белок GFP. В результате облучения этих клеток длинноволновым ультрафиолетовым светом белок GFP начинает флуоресцировать зеленым светом.

С какими целями используют трансформацию в экспериментальной генетике?

Можно выделить несколько направлений.

1. В процессе клонирования и выделения генов всегда остается вопрос о том, весь ли ген уже в руках экспериментатора, т.е. полностью ли

выделены его регуляторная и структурная части. Для выяснения этого вопроса в состав транспозона вводят фрагмент ДНК, предположительно содержащий ген, это фрагмент "А" в транспозоне $P[(ry^+)A]$ на Рис. 7.27. Затем хромосому со встроенным транспозоном $P[(ry^+)A]$ путем скрещивания вводят мухам, гомозиготным по мутации испытуемого гена. Если в транспозоне присутствуют все части гена "А", необходимые для его функционирования, то такая особь будет иметь нормальный фенотип, поскольку мутантное действие гена *A* полностью перекрывается нормальным аллелем A^+ , расположенным в транспозоне.

Нетрудно заметить, что, хотя эти работы и выполнены на дрозофиле, по сути здесь речь идет о возможности исправления генетических дефектов посредством вмешательства человека. Если учесть, что трансформация у дрозофилы за последние 16 лет применялась тысячи раз, становится очевидным, что по крайней мере на этом модельном объекте проблема исправления генетических дефектов уже в наше время эффективно решается достаточно рутинными методами.

2. Другая цель использования трансформации заключается в исследовании структуры самого гена, т.е. анализа его структурной и регуляторной частей (см. раздел 7.7.).

3. Трансформацию на дрозофиле используют для экспериментальной экспрессии генов, из генов других организмов, у которых еще не открыты мобильные элементы, пригодные для создания транспозонов, и не

разработаны методы трансформации. Например, можно выделить какой-то ген из генома бабочки-шелкопряда, комара или человека, ввести в геном дрозофилы и проанализировать экспрессию.

Созданы еще более уникальные конструкции, связанные с пересадками генов. Можно привести один пример. Известно, что в геноме комара *Chironomus thummi thummi* присутствуют фракции, повторенные множество раз. Единицей этого повтора является фрагмент длиной 120 п.н., обогащенный А-Т нуклеотидами. Это так называемый *Cla*-повтор. Он рассеян по всей длине хромосом *Ch. th. thummi*, но почти полностью отсутствует у очень близкого вида *Ch. th. piger*, т.е. этот повтор является варьирующей фракцией генома. Считается, что повторенные последовательности оказывают инактивирующее влияние на гены, расположенные рядом.

Для того, чтобы проверить предположение об инактивирующем влиянии *Cla*-повтора, был создан транспозон, содержащий кроме маркерного гена *rosy*⁺, еще последовательность нуклеотидов, состоящую из нескольких копий *Cla*-повтора из ДНК хирономуса и расположенного рядом с ним гена *white* (*w*), выделенного из генома дрозофилы. Этот транспозон был введен в эмбрионы дрозофилы, мутантные по гену *w*. Дело в том, что нормальный аллель гена *w*⁺, обеспечивает формирование красного цвета глаз у мухи, в то время как у гомозигот по мутации *w* глаза белого цвета. Если *Cla*-повтор не оказывает инактивирующего влияния на ген *w*⁺ в транспозоне, глаза

у мух *w/w/Tw*⁺ будут нормального красного цвета. Если же *Cla*-повтор инактивирует ген *w*⁺ в транспозоне во всех клетках или только в некоторых, то глаза у мух будут соответственно полностью белыми или мозаичными, т.е. в некоторых клетках ген *w*⁺ сохраняют активность, в других клетках *w*⁺ инактивирован.

4. Поиски энхансеров генов (см. раздел 7.8.3.).

5. Создание систем трансформации для направленной экспрессии генов (см. раздел 7.8.3.).

Литература к разделу 7.5.10.

- Жимулев И.Ф. Трансформация у дрозофилы - новый подход в генетике. Сорос. образов. журнал, 1999, в печати.
- Ashburner M. *Drosophila*. A laboratory handbook. Cold Spring Harbor, Cold Spring Harbor Laboratory Press, 1989, 1-1331.
- Brand A.H., Perimon N. Targeted gene expression as a means of altering cell fates and generating dominant phenotypes. *Development*, V. 118, 401-415, 1993.
- Gehring W.J. The master control gene for morphogenesis and evolution of the eye. *Genes to Cells*, V. 1, 11-15, 1993.
- Halder G., Callaerts P., Gehring W.J. New perspectives on eye evolution. *Current Opinion in Genetics & Development*, V.5, 602-609, 1995.
- Lawrence P.A. *The making of a fly*. Oxford, Blackwell Scientific Publications, 1-228, 1992.
- Prasher D.C. Using GFP to see the light. *Trends in Genetics*, V. 11, N8, 320-323, 1995.

Spradling A.C. P-element mediated transformation. In: “*Drosophila: a practical approach*” (D.B. Roberts, ed.), Oxford-Washington, IPL Press, p. 175-197, 1986.

Wilson C. Pearson R.K., Bellen H.J., O’Kane C.J. Grossniklaus U., Gehring W.J. P-element-mediated enhancer detection: an efficient method for isolating and characterizing developmentally regulated genes in *Drosophila*. *Genes & Development*, V. 3, 1301-1313, 1989.

7.6. Расположение генов в хромосомах эукариот

У эукариот не распространен оперонный тип расположения генов, т.е. объединение в блоки генов, находящихся под общим контролем. Гены, контролирующие даже последовательные биохимические реакции, расположены в различных районах хромосомы и даже в различных хромосомах. Например, у дрозофилы гены, кодирующие ферменты, под контролем которых происходит превращение триптофана в бурый глазной пигмент, разбросаны во множестве участков генома (табл. 7.1. - делать)

Вместе с тем, известны некоторые примеры кластерной организации генов. У человека известно несколько типов гемоглобинов. Каждый из них синтезируется в определенном органе и на определенной стадии развития. Например, гемоглобины ζ и ϵ синтезируются в клетках эмбрионального желточного мешка на ранних этапах развития. В это время молекула белка состоит из двух цепей ζ -подобный α -гемоглобин и двух цепей ϵ (β -подобный). К концу третьего месяца развития синтез

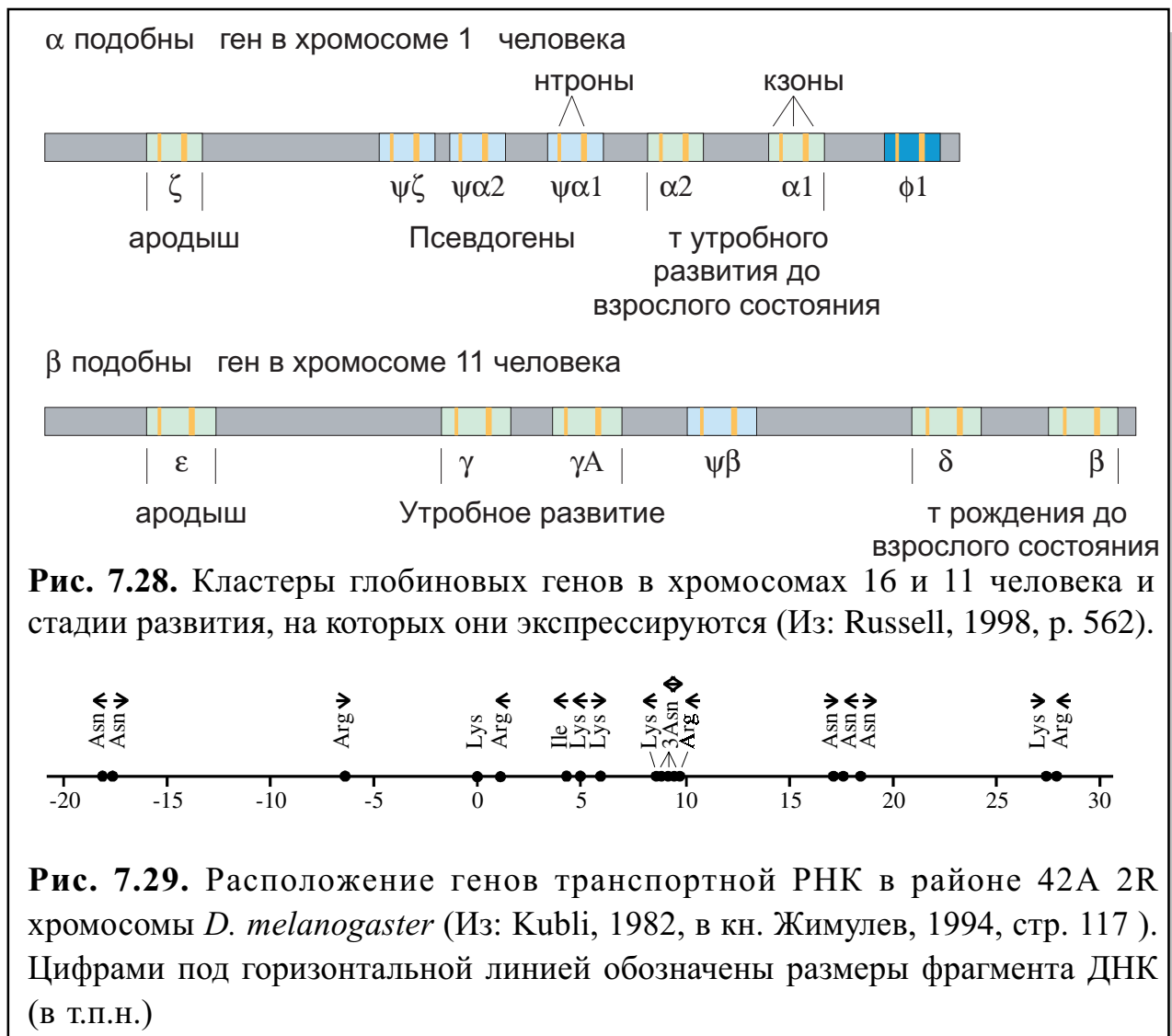
гемоглобина этого типа заканчивается и начинается синтез фетального (утробного) гемоглобина (Hb-F) в клетках печени и селезенки. Фетальный гемоглобин состоит из двух α -полипептидов и двух γ (β -подобных): двух γ^A или двух γ^G . Во время постэмбрионального развития гемоглобин синтезируется в клетках костного мозга, и состоит из α , β полипептидов и некоторого количества β -подобного δ полипептидов. Большая часть гемоглобина в крови взрослого человека состоит из тетрамера $\alpha_2\beta_2$ (Hb-A).

В геноме человека гены гемоглобина расположены двумя кластерами: в хромосоме 16 расположены все α -подобные гены, в то время как все β -подобные гены расположены в хромосоме 11 (Рис. 7.28.). В каждом кластере есть псевдогены.

Интересно, что гены расположены в том порядке, вдоль по хромосоме, в какой очередности они включаются в работу в ходе онтогенеза. Однако, каких-либо данных, свидетельствующих об их функциональной сцепленности или общем контроле по принципу оперонной организации, не получено. Тот факт, что эти гены функционируют в разных тканях и на разных этапах онтогенеза, скорее свидетельствует о независимом контроле экспрессии этих генов.

Гены транспортной РНК у дрозофилы расположены во многих районах по 1-2 гена в каждом. Однако, в одном из районов 16 генов занимают небольшой участок длиной 50 т.п.н. (Рис. 7.29.), хотя их транскрипция не контролируется одним промотором.

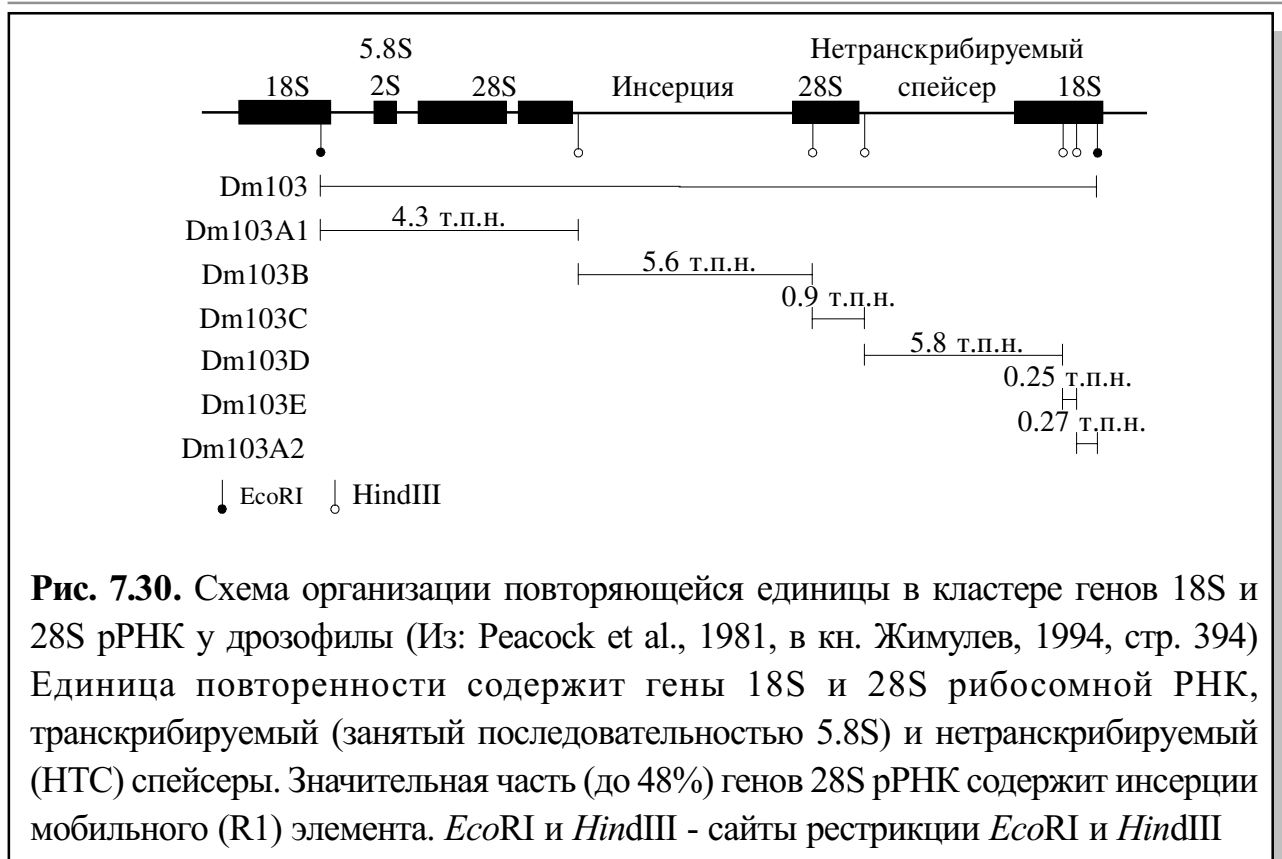
Организация генов 18S и 28S рРНК у всех эукариот в общих чертах одинаковая. Последовательности генов



18S и 28S рРНК, лидерной последовательности, а также транскрибируемого и нетранскрибируемого спейсеров составляют единицу длиной около 11 т.п.н. (Рис. 7.30.), которая повторена несколько сот раз. Как правило число копий варьирует в пределах от 100 до 1000: у дрожжей - 140 единиц повторности, у дрозофилы - 200-250, у человека - 1250. Лидерная последовательность расположена перед геном 18S, и с неё начинается транскрипция. Каждая единица повтора генов рРНК считывается РНК-полимеразой I от лидерной последовательности до конца гена 28S рРНК в виде одного цистрона.

Участок ДНК между генами 18S и 28S транскрибируется вместе с этими генами и называется транскрибируемым спейсером. В нём расположены короткие последовательности, которые выделяются из общего транскрипта в ходе процессинга РНК. У млекопитающих и амфибий в коротком спейсере формируется 5.8S РНК - небольшая молекула, образующая водородную связь с 28S рРНК в рибосоме.

Транскрипционная единица рРНК самая короткая у бактерий, у которых при общей длине транскрипта 6 т.п.н. 80% приходится на кодирующую часть. У эукариот длина транскрипта варьирует от 7 до 8 т.п.н. и кодирующая часть



составляет 70-80%. В ходе созревания рРНК лидирующая последовательность и часть транскрибируемого спейсера, не кодирующая 5.8S РНК, деградируют до нуклеотидов.

Транскрибируемые единицы разделяются участком ДНК, называемым нетранскрибируемым спейсером (НТС). Его длина варьирует в широких пределах: от 1750 п.н. у дрожжей до 30000 п.н. у мыши, у дрозофилы он имеет длину 3750-6450 п.н. В самом НТС обнаружили внутренние повторы, которые в частности у дрозофилы, играют существенную роль в мейотическом спаривании X и Y хромосом (По: Lewin 1994, pp. 726-728). Существенная часть генов 28S рРНК содержит инсерции (Рис. 7.30.).

Каждая из единиц повтора может функционировать независимо от других. Если один из повторов заключить в Р-элемент и с помощью трансформации

ввести его опять в геном дрозофилы, в участке встраивания транспозона образуется маленькое ядрышко, возникшее в результате транскрипционной активности этого повтора.

У дрозофилы гены, кодирующие 5S рРНК, представлены блоком, состоящим из 160-200 идентичных последовательностей длиной 385 п.н. каждая. Общая длина кластера составляет 60-80 т.п.н. Повторяющаяся единица состоит из кодирующей части (~135 п.н.) и спейсера (250 п.н.). Отдельные гены или группы генов в пределах кластера транскрибируются независимо друг от друга: делеции части генов не предотвращают активности остальных. Клоны генов 5S рРНК, выделенные из разных частей кластера, нормально транскрибируются в молекулы 5S рРНК в системе ооцита ксенопуса (*Xenopus*). Таким образом, последовательности ДНК не составляют

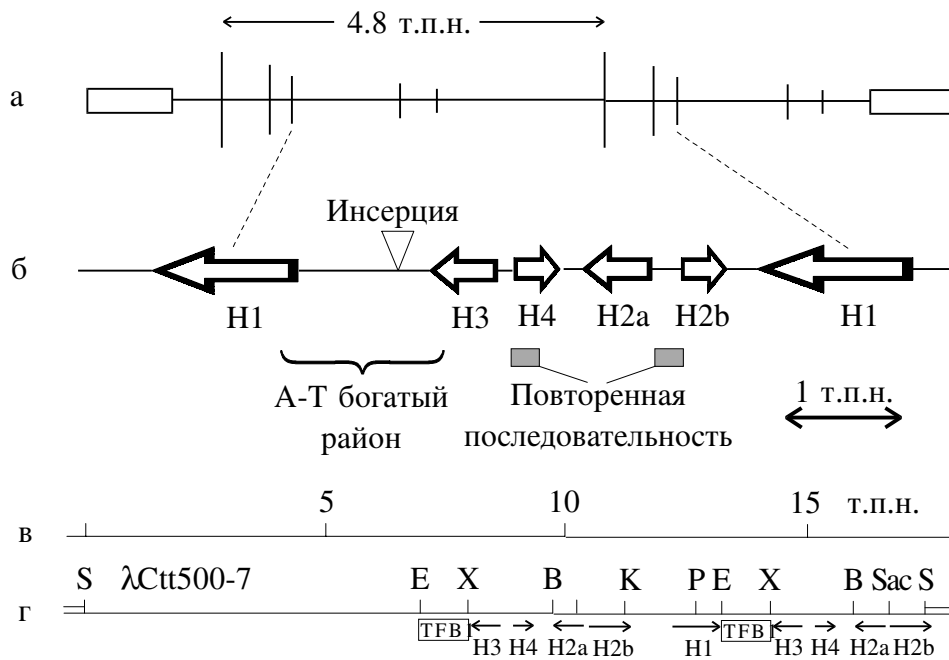


Рис. 7.31. Схема организации генов, кодирующих гистоны в повторяющейся единице у *D. melanogaster* (а-б) и *Chironomus thummi* (в-г) (Из: Lifton et al., 1978; Hankeln, Schmidt, 1990, в кн. Жимулев, 1994, стр. 124); а - рестрикционная карта клона сDm500. Вертикальными линиями разной высоты обозначены слева направо *Bgl*II, *Bam*HI, *Hind*III, *Hpa*I, *Sst*I; б - расположение генов (стрелками указано направление транскрипции), в - расстояния на карте; г - сайты рестрикции и локализация генов, TFB1 - мобильный элемент

единицу транскрипции и функционируют независимо одна от другой (Текст по: Жимулев, 1994, стр. 122-123).

Общая длина кластера гистоновых генов у дрозофилы составляет примерно 500 т.п.н. (100 повторов единицы около 5 т.п.н. с пятью генами в каждой). Отдельные гены в пределах повторенной единицы транскрибируются в противоположных направлениях (т.е. с разных цепей ДНК), что свидетельствует о независимости их функционирования (Рис. 7.31.).

У другого вида насекомых - хирономуса (*Chironomus thummi*) эти же пять гистоновых генов составляют кластер длиной 6262 п.н., однако, и у этого вида гены, расположенные в

кластере, могут считываться с разных цепей (см. Рис. 7.31.).

Иногда в кластеры объединяются и неповторенные гены. Семь генов, кодирующих белки теплового шока (см. раздел 12), располагаются во фрагменте ДНК длиной 12 т.п.н., картируемом с помощью гибридизации *in situ* в пуфе теплового шока 67В. Для генов этого района также характерна разная направленность транскрипции (см. Рис. 12..).

Гены, контролирующие развитие крупных частей тела (гомейозисные гены) дрозофилы, собраны в кластеры. Их называют комплексами. Головная капсула дрозофилы формируется в результате активности генов комплекса *Antennapedia* (*Antp-C*), брюшная часть -

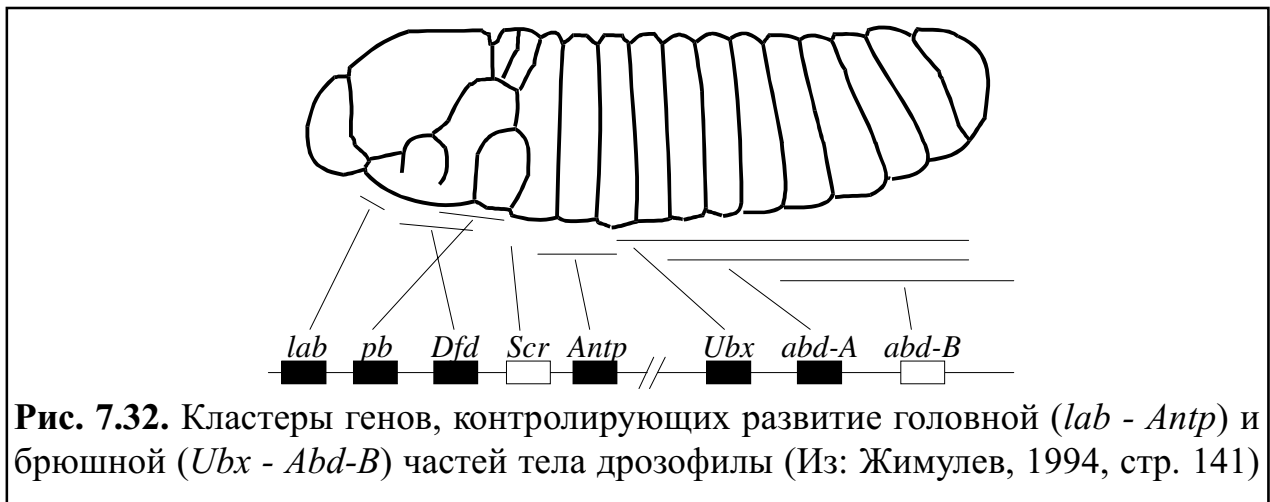


Рис. 7.32. Кластеры генов, контролирующих развитие головной (*lab - Antp*) и брюшной (*Ubx - Abd-B*) частей тела дрозофилы (Из: Жимулев, 1994, стр. 141)

генами комплекса *Bithorax (BX-C)* (Рис. 7.32.).

Литература к разделу 7.6.

- Жимулев И.Ф. Хромомерная организация политенных хромосом. Наука, Новосибирск, 1-564, 1994.
- Lewin B. Genes. Oxford, New York, Tokyo; Oxford University Press, p. 726-728, 1994.
- Russell, P.J. Genetics Fifth edition. Addison Wesley Longman, Ins, Menlo Park, California, p. 400-408, 1998.

7.7. Структура транскрипта: структурная и регуляторная части гена

Синтез РНК, кодируемой данным геном, называется экспрессией гена. У прокариот единственная РНК-полимераза, состоящая из белкового комплекса - собственно РНК-полимеразы и σ -фактора, синтезирует все виды РНК: мРНК, тРНК и рРНК.

У эукариот три разных РНК-полимеразы транскрибируют РНК с трех разных типов генов. РНК-полимераза I синтезирует 18, 28 и 5.8SРНК, РНК-полимераза II считывает мРНК с генов, кодирующих белки и некоторые snРНК, РНК-полимераза III

транскрибирует гены 5S рРНК, тРНК и остальные snРНК.

Схема организации типичного гена представлена на Рис. 7.33. Каждый ген состоит из регуляторной части, с которой начинается транскрипция кодирующей части, в которой записана информация о структуре белка, и терминирующей части, в которой завершается транскрипция. Двойная спираль ДНК в области промотора денатурируется под действием РНК-полимеразы у прокариот или транскрипционного комплекса у эукариот (Рис. 7.34.) с образованием "транскрипционного пузырька". РНК синтезируется в направлении от 5' к 3' концу одной из цепей ДНК, которая называется матричной. Вторая цепь называется кодирующей.

Транскрипция заканчивается, когда молекула РНК-полимеразы достигнет терминирующего участка или терминатора (Рис. 7.33.). Обнаружено два типа терминирующих последовательностей, и каждый ген имеет один из них. Один тип последовательностей опознается непосредственно РНК-полимеразой, другой тип - РНК-полимеразой в ассоциации с ρ -фактором.

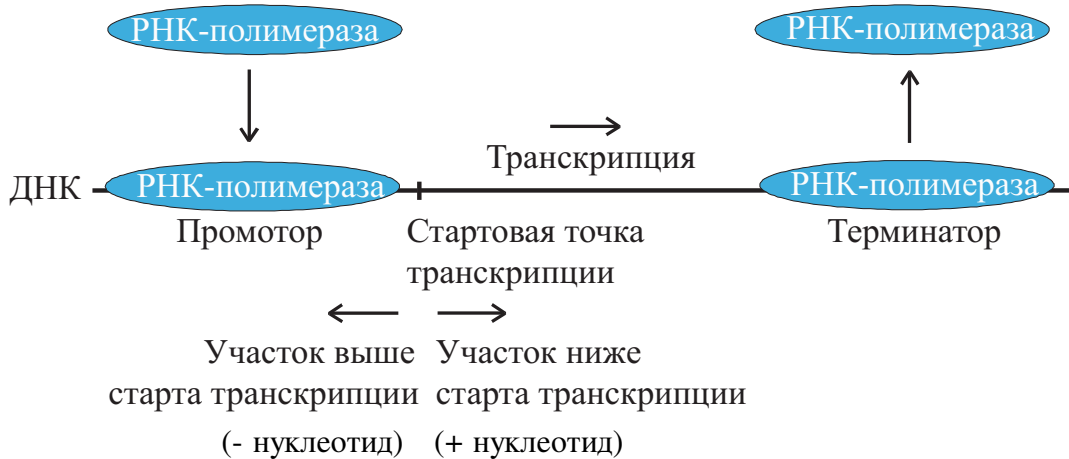


Рис. 7.33. Единица транскрипции, содержащая различные элементы гена (Из: Lewin, 1994, p. 378)

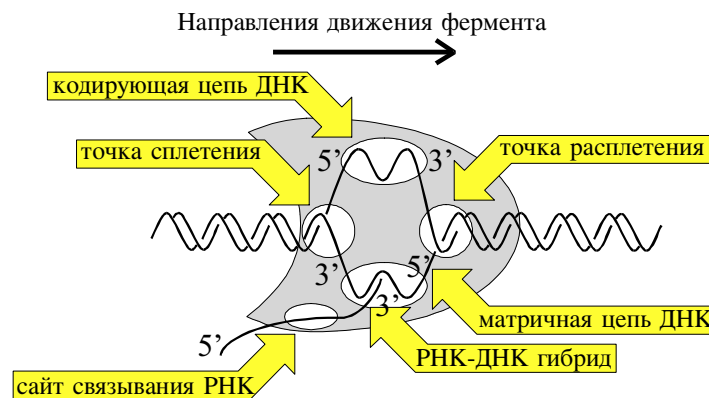


Рис. 7.34. Транскрипционный “пузырек”. (Из: Lewin, 1994, p. 380). В ходе транскрипции РНК-полимераза денатурирует двуцепочечную спираль ДНК и вновь восстанавливает ее, поддерживает кодирующую и матричную цепи ДНК в выпрямленном состоянии и синтезирует РНК

Исследование процессов, происходящих при синтезе мРНК, показало, что продукты транскрипции, синтезируемые в ядрах эукариот на ДНК хромосом, гораздо крупнее, чем образуемые из них и выходящие в цитоплазму матричные РНК, поступающие в рибосомы и участвующие в трансляции. Такие транскрибированные с ДНК и находящиеся в клеточном ядре предшественники мРНК, очень разнообразные по своему нуклеотидному составу, называются гетерогенными ядерными РНК или про-мРНК.

Дополнение 7.6.

Нобелевская премия за 1959 год присуждена С. Очоа (Severo Ochoa) за открытие механизма биологического синтеза рибонуклеиновой кислоты.

Выяснилось, что в гетерогенных ядерных РНК присутствуют протяженные участки, считанные с участков генов, не кодирующие информацию о структуре белка - интронов. В результате сложного процесса созревания РНК, называемого процессингом, образуется зрелая молекула матричной

РНК уже существенно меньшего размера. В ходе процессинга (или постраскрипционных модификаций РНК) удаляются участки, соответствующие интронам, происходит сращивание экзонов (сплайсинг), добавляется короткий фрагмент на 3' конце (5'- кэп) и поли (А)⁺-хвост на 3' конце молекулы РНК.

На Рис. 7.35. представлена схема строения зрелой биологически активной мРНК из клеток эукариот. Аналогичная мРНК обнаруживается и в клетках прокариот, но в ней нет поли(А)⁺ последовательности.

Литература к разделу 7.7.

Гершензон С.М. Основы современной генетики. Киев, Наукова Думка, стр. 349-356, 1983.

Lewin B. Genes V. Oxford University Press. Oxford, New York, Tokyo, 377-401, 1994.

Russell, P.J. Genetics Fifth edition. Addison Wesley Longman, Ins, Menlo Park, California, p. 379-417, 1998.

Георгиев Г.П. Гипотеза о структурной организации оперона и регуляции синтеза РНК в животной клетке. Молекулярная биология т.4, вып. 1, стр 17-29, 1970.

Георгиев Г.П. Гены высших организмов и их экспрессия. Москва, Наука, 1-254, 1989.

Овчинников Л.П. Что и как закодировано в мРНК. Соросовский образовательный журнал, N 4, с. 10-18, 1998.

7.8. Регуляторная часть гена

7.8.1. Промоторы и регуляторы

У прокариот процесс транскрипции осуществляется с помощью голоэнзима (или “полного энзима”) РНК-полимеразы, состоящей



У эукариот процесс транскрипции значительно усложнён в связи с тремя обстоятельствами. Во-первых, у эукариот транскрипция осуществляется под действием трех разных РНК-полимераз.

Во-вторых, РНК-полимераза эукариот не может самостоятельно инициировать транскрипцию. Для ее активирования необходимо большое число белков, называемых общими факторами транскрипции, которые должны объединяться в комплекс, прежде чем транскрипция начнется. Формирование комплекса - это многоступенчатый процесс, от прохождения этапов которого будет в конечном счете зависеть скорость инициации транскрипции. Во многих случаях регуляторные белки действуют, влияя главным образом на процесс сборки транскрипционного комплекса.

И, наконец, в-третьих, большинство регуляторных белков у эукариот могут влиять на скорость транскрипции даже если они

связываются с участками ДНК, расположенными за тысячи пар нуклеотидов от промотора, что означает, что любой конкретный промотор может находиться под контролем неограниченного числа регуляторных последовательностей, разбросанных по геному.

Рассмотрим сначала организацию контролирующей зоны для действия РНК-полимеразы II. Контролирующей областью называют последовательности ДНК, необходимые как для инициации транскрипции, так и для регулирования ее скорости и интенсивности.

Поэтому контролирующий район состоит из промотора, на котором образуется комплекс из РНК-полимеразы II и общих факторов транскрипции, а также многочисленные регуляторные последовательности, с которыми связываются различные регуляторные белки (Рис. 7.38.). Регуляторные элементы гена могут быть расположены как выше, так и ниже

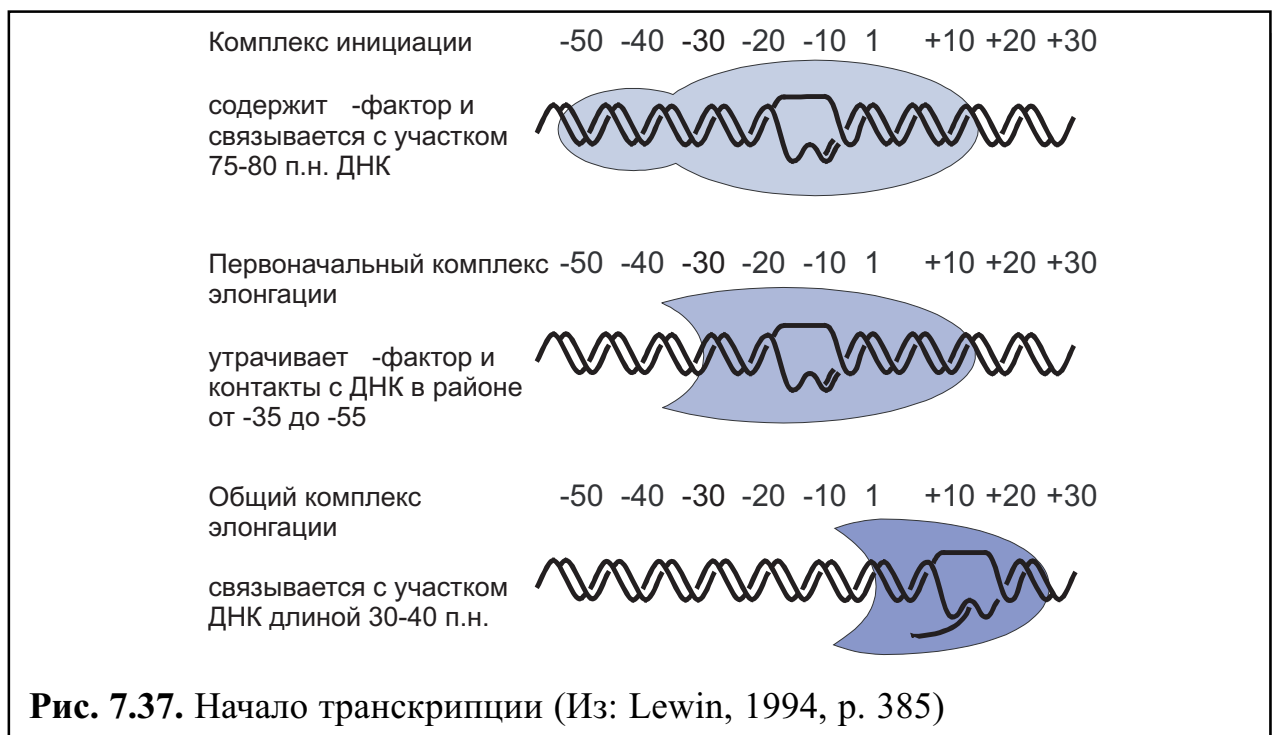


Рис. 7.37. Начало транскрипции (Из: Lewin, 1994, p. 385)

точки начала транскрипции. Одним из наиболее известных регуляторных элементов гена является промотор, который состоит из нескольких промоторных элементов. Самым близким к точке начала транскрипции является ТАТА-домен (ТАТААА) называемый по-другому доменом Хогнесса или доменом Голдберга-Хогнесса). Затем следуют домены СААТ (GGССААТСТ) и GG (GGССGG). Промоторы могут содержать различные комбинации промоторных элементов, но ни один элемент не встречается во всех промоторах, например у *D. melanogaster* были проанализированы данные о первичной структуре 252 независимых промоторов. ТАТА-домен находится в положении -25/-30. Неожиданно было найдено, что половина промоторов у дрозофилы не имеют ТАТА-домена (Arkhipova, 1995). Некоторые промоторы имеют более одной копии СААТ или GC. Домен СААТ играет существенную роль в инициации транскрипции, ТАТА и GC, по-видимому, выполняют вспомогательные роли.

Общие факторы транскрипции к настоящему времени очищены и выделены (Рис. 7.39.). Их шесть: TFIIA, TFIIB, TFIID, TFIIE, TFIIF и TFIIH.

Вместе с РНК полимеразой II они могут инициировать транскрипцию во внеклеточных системах *in vitro*. Некоторые из общих факторов транскрипции в свою очередь состоят из многих полипептидов. Наиболее интересны TFIID, которые в своем составе содержат белки ТВР (ТАТА-box binding protein) и около 8 связанных с ними других молекул, т.н. ТАФ (ТВР-associated factors). Сложную организацию имеет комплекс TFIIH. У млекопитающих РНК-полимераза II состоит из 12-14 полипептидов, в результате чего молекулярная масса этого комплекса доходит до 600 килодальтон (кДА). В состав транскрибирующего комплекса входят недавно открытые белки, называемые SRB (suppressors of RNA PolII) которые связываются с большой субъединицей РНК-полимеразы (Рис. 7.39.). Эти белки помогают РНК-полимеразе разрушить нуклеосомы и декомпактизовать молекулу ДНК.

Довольно давно было показано, что молекула фактора транскрипции TFIIH связана с белками, участвующими в репарации ДНК (т.н. эксцизия нуклеотидов). РНК полимеразы II связана еще с группой белков, которые могут разрушать

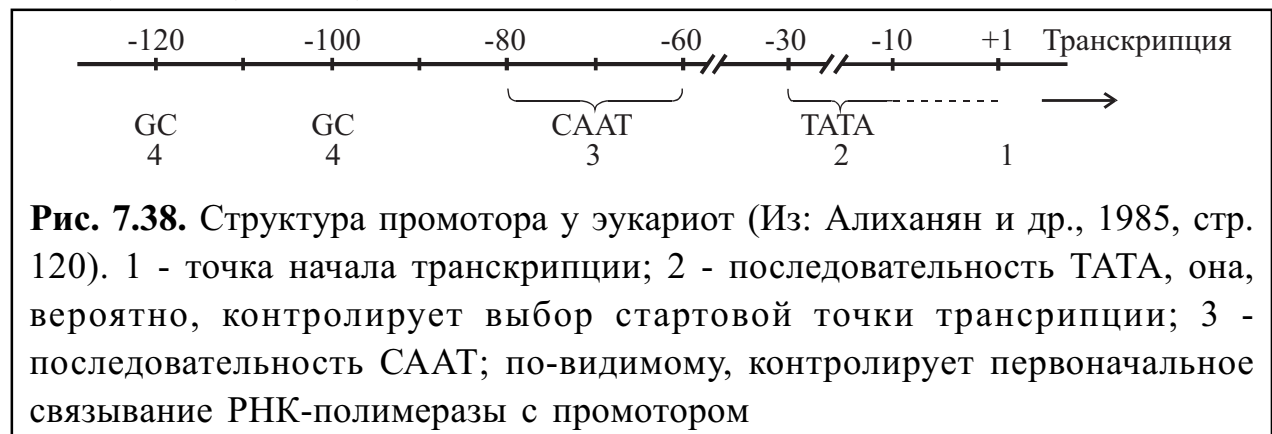


Рис. 7.38. Структура промотора у эукариот (Из: Алиханян и др., 1985, стр. 120). 1 - точка начала транскрипции; 2 - последовательность ТАТА, она, вероятно, контролирует выбор стартовой точки транскрипции; 3 - последовательность СААТ; по-видимому, контролирует первоначальное связывание РНК-полимеразы с промотором

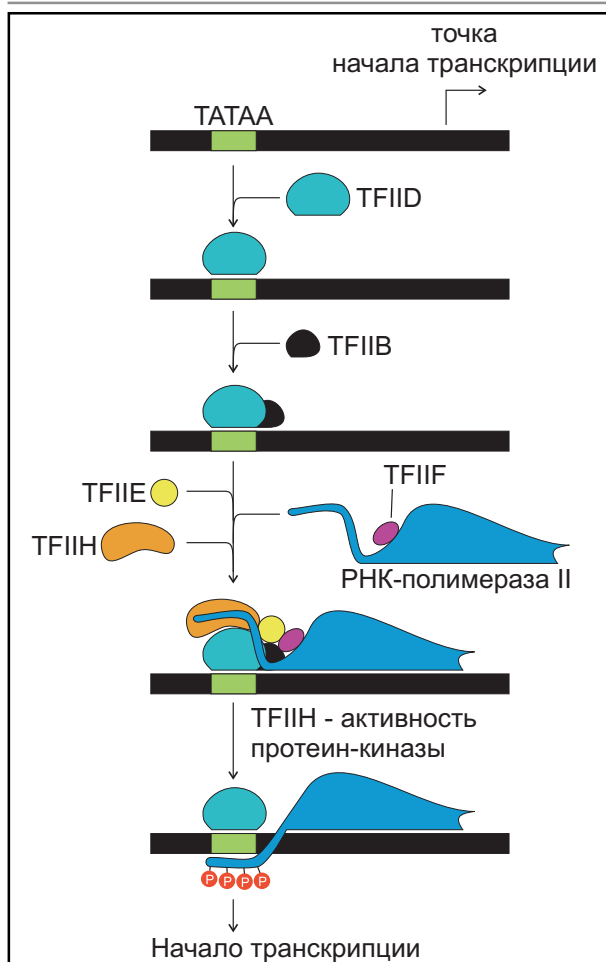


Рис. 7.39. Белковый состав регуляторных зон генов эукариот. Образование комплекса из РНК-полимеразы II и общих факторов транскрипции (TFIID, TFIIB, F, E, и H). В присутствии АТФ TFIIH фосфорилирует РНК-полимеразу II. Фосфорилирование происходит в зоне длинного полипептидного “хвоста”, в результате чего молекула полимеразы изменяет конформацию, становится готовой к транскрипции и освобождается от всех факторов транскрипции, кроме TFIID (Из: Alberts et al., 1994, p. 421)

нуклеосомы и называются семейством SWI/SNK.

Процесс сборки комплекса показан на Рис. 7.39.

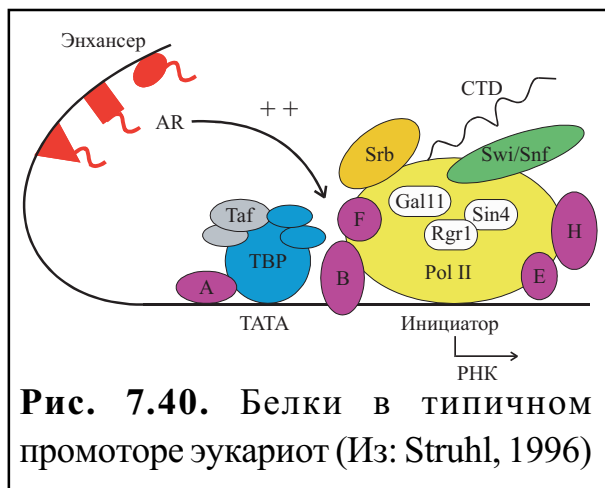
Формирование белкового комплекса на промоторной

последовательности начинается с того, что фактор транскрипции TFIID связывается с ТАТА-доменом, который расположен выше точки инициации транскрипции примерно на 25 нуклеотидов. TFIID состоит в свою очередь из многих субъединиц. Те из них, которые ответственны за распознавание ТАТА-последовательности, называются ТВР (TATA-binding proteins). Теперь для начала транскрипции РНК-полимераза II должна освободиться от комплекса факторов транскрипции. Ключевым в процессе инициации транскрипции является присоединение фактора транскрипции TFIIH. Одна из его субъединиц, обладающая протеинкиназной активностью, фосфорилирует молекулу РНК-полимеразы II. Установлено на примере по крайней мере нескольких генов, что это фосфорилирование освобождает РНК-полимеразу II и позволяет начать транскрипцию (Рис. 7.39.).

В состав комплексов общих факторов транскрипции и РНК-полимеразы II входит до 50 белков, иногда эти комплексы называют транскриптосомами (Рис. 7.40.)

Две другие РНК-полимеразы, I и III, найденные у эукариот, также требуют для активирования набора общих факторов транскрипции. С ними не все ясно. Установлено лишь, что белки ТВР требуются для всех трех полимераз. Другие факторы отличаются от тех, что были описаны в комплексах с РНК-полимеразой II.

Кроме промотора в контрольной зоне находятся регуляторные последовательности (Рис. 7.41.), с которыми связываются регуляторные



белки, необходимые для контроля процесса образования белкового комплекса на промоторе.

Ниже приведены примеры организации регуляторных участков, обнаруженных у генов эукариот.

Гены, отвечающие на тепловой шок (см. раздел 12.), имеют в составе регуляторной зоны от трех до шести “элементов теплового шока” (HSE) (Рис. 7.42.), которые представлены 14-членными последовательностями нуклеотидов (Рис. 7.43.).

Известно, что гены теплового шока очень быстро реагируют на индуцирующее воздействие (см. раздел 12.). В соответствии с таким типом

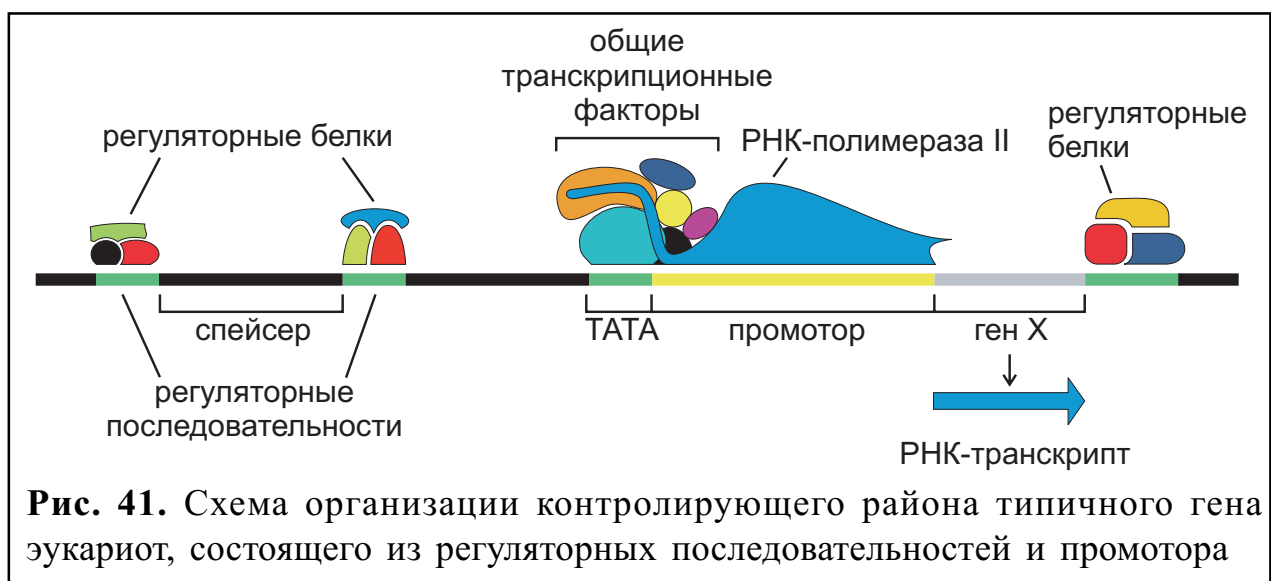
индукции у них устроены регуляторные районы (Рис. 7.44.).

В неиндуцированной тепловым шоком клетке (Рис. 7.44а.) находятся белковые факторы HSF, они неактивны и не связываются с промоторными HSE-элементами. РНК-полимераза II занимает участок гена в районе точки инициации транскрипции. Под воздействием теплового шока белок HSF модифицируется и связывается с промотором гена *hsp*. Это активирует РНК-полимеразу II, либо воздействуя на TATA-домен, либо освобождая гипотетический белок “X”, блокирующий РНК-полимеразу II (см. Рис. 7.44.).

Некоторые из генов *hsp* индуцируются также гормонами. Соответственно, промоторы этих генов организованы сложнее (Рис. 7.45.). В них перемежаются контролирующие элементы обоих типов.

7.8.2. Метод репортерных генов для изучения промоторных участков генов

Одним из подходов к изучению организации регуляторной части гена является методика репортерных генов,



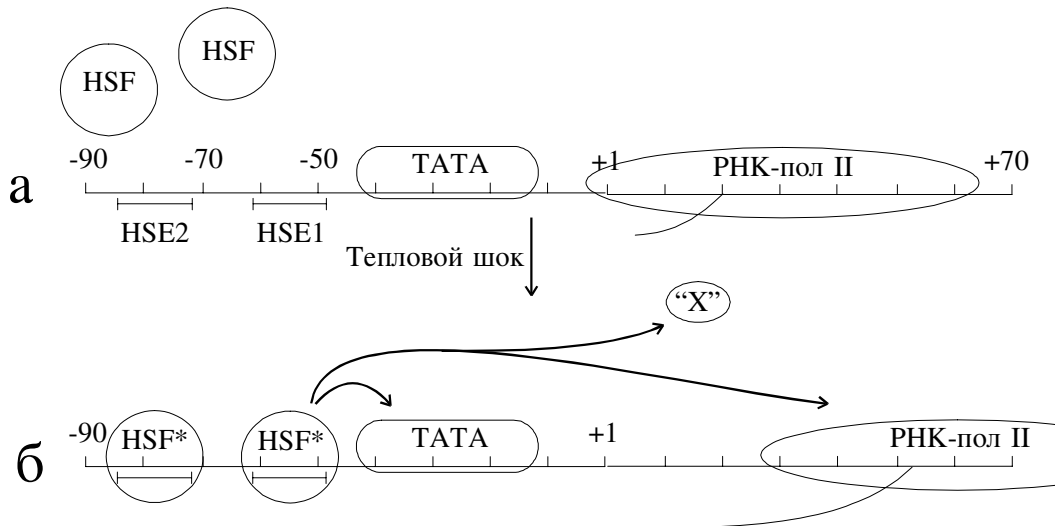


Рис. 7.44. Схема изменений в компонентах промоторной зоны гена *hsp70* дрозофилы во время индукции тепловым шоком (Из: Rougvie, Lis, 1988 в кн. Жимулёв, 1994, стр. 354)

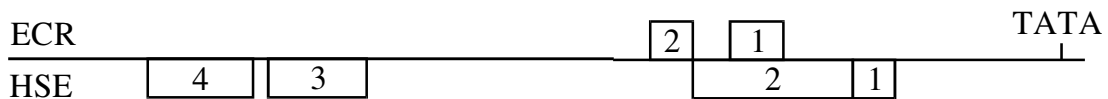


Рис. 7.45. Регуляторные последовательности в 5' нетранскрибируемом районе гена *hsp23* у дрозофилы (Из: Mestril et al., 1986, в кн. Жимулёв, 1994, стр. 343), ECR - участки связывания с эктистерон-рецепторным комплексом, HSE - элементы связывания фактора теплового шока

разработанная в лаборатории В. Геринга в 1985 году. Впервые этот метод был применен на гене *fushi tarazu* (*ftz*) и состоит в следующем: авторы имели фрагмент ДНК, в котором располагалась кодирующая часть гена *ftz*. Более того, выше 5' конца кодирующей части гена был расположен еще один отрезок ДНК длиной 6,1 т.п.н., содержащий регуляторную часть гена. Если трансформировать мутантов *ftz* транспозоном, содержащим весь фрагмент ДНК, у потомков формируется нормальный фенотип. Это значит, что и регуляторная, и структурная части гена функционируют нормально. После этого авторы удалили кодирующую часть *ftz*, присоединили на ее место кодирующую

часть гена *lacZ*, выделенного из бактерии *E. coli*, и трансформировали этим транспозоном эмбрионов дрозофилы. Следует отметить, что ген *lacZ* кодирует фермент β -галактозидазу, расщепляющую β -галактозиды. В условиях эксперимента можно подобрать особый галактозид, например, вещество, называемое X-gal, в результате расщепления которого образуются соли, имеющие синюю окраску и нерастворимые в воде (Рис. 7.46.). Теперь, если искусственно созданный "ген", состоящий из регуляторной части *ftz* и гена *lacZ*, проработает, а затем органы мухи проинкубировать с X-gal, клетки, в которых такой "ген" был активен, окрасятся в синий цвет - цвет



Рис. 7.46. Локализация клеток в имагинальных дисках дрозофилы, в которых экспрессируется ген *lacZ*

водонерастворимого продукта расщепления X-gal. В данном случае ген *lacZ* ведет себя как репортер, сообщающий о работе регуляторной части гена *ftz*. Оказалось, что у трансформантов, содержащих в геноме транспозон, “ген” *ftz/lacZ* функционирует в тех же самых клетках, что и нормальный ген *ftz*.

Теперь можно попытаться узнать какие именно участки регуляторной части гена более необходимы для нормального функционирования. С помощью особых приемов можно удалять разные части энхансерного фрагмента, объединять их с *lacZ*, трансформировать и следить за экспрессией *lacZ*.

На Рис. 7.47. показаны результаты такого эксперимента с геном *hunchback*. В нормальном развитии белковый продукт гена *hunchback*⁺ распределяется в передней части эмбриона (верхний ряд на Рис. 7.47.).

Используя трансформацию, установили, что регуляторная последовательность гена *hunchback* длиной 747 п.н. выше репортера *lacZ* (второй ряд), вполне достаточна для почти нормального распределения продукта *lacZ*. В последующих экспериментах удалось показать, что фрагмент длиной 263 п.н. (между +55 и +318) абсолютно необходим для экспрессии *lacZ* (третий ряд). Если добавить только этот фрагмент, его достаточно для почти нормальной экспрессии *lacZ* (четвертый ряд). Другие варианты состава регуляторной зоны дают разные промежуточные уровни активности “репортера” (ряды 5 и 6 на Рис. 7.47.).

7.8.3. Энхансерные участки гена

Хотя промоторные элементы являются решающими в осуществлении транскрипции в целом для того, чтобы она осуществлялась максимально эффективно, необходимы энхансерные элементы.

В 1979 было установлено, что последовательности ДНК, расположенные в тысячах пар нуклеотидов от промотора эукариотического гена, могут активировать его транскрипцию. Сейчас известно, что эти энхансерные (т.е. усиливающие) последовательности служат в качестве специфических участков (сайтов) связывания особых регуляторных белков, усиливающих или активирующих процесс транскрипции. Этот тип контроля генной активности на расстоянии является скорее правилом, чем исключением. Как эти белки могут

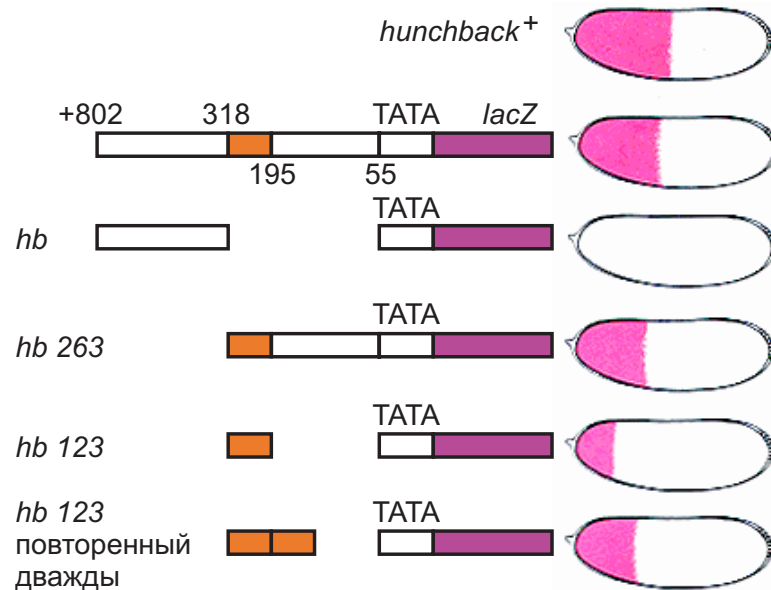


Рис. 7.47. Локализация регуляторной зоны гена *hunchback* с помощью гена-репортера *lacZ*. (Из: Lawrence, 1992, стр. 53). В верхней строке красным цветом показано распределение продукта гена *hunchback*⁺ в передней части эмбриона. Строки 2-6 показывают, как изменяется распределение продукта гена *lacZ* при удалении той или иной части регуляторной зоны. Цифры +802, 318, 195, 55 обозначают порядковый номер нуклеотида, считая нулевым первый нуклеотид гена *lacZ*

действовать на больших расстояниях?

Согласно самой простой модели, ДНК между энхансером и промотором образует петлю, в результате чего белки, связанные с энхансером, непосредственно взаимодействуют с одним из общих факторов транскрипции или с молекулой самой РНК-полимеразы (Рис. 7.40.).

Число общих факторов транскрипции невелико, хотя они очень обильно представлены в клетке, поскольку связываются с промоторами всех генов, транскрибируемых РНК-полимеразой II. Кроме этого, в клетке существуют тысячи различных других регуляторных белков, связывающихся с регуляторными сайтами контролирующей зоны. Их наборы различаются в разных клетках и у разных генов. Каждый из этих белков

представлен малым числом молекул. Большинство из этих белков распознают особую, специфическую только для них, последовательность нуклеотидов в регуляторных сайтах генов. С помощью белков-регуляторов каждый ген специфически включается или выключается.

Очень часто для понимания процессов развития необходимо выявить максимальное число генов, функционирующих в той или иной ткани. В лаборатории В. Геринга в 1987-1989 г.г. был разработан метод трансформации, позволяющий выявлять гены, вовлеченные в развитие. Геринг и его коллеги создали транспозон, в котором ген *lacZ* помещен под контроль слабого промотора, имеющегося в Р-элементе. Если такой транспозон встраивается в окрестности энхансерного элемента какого-то гена, функционирующего у

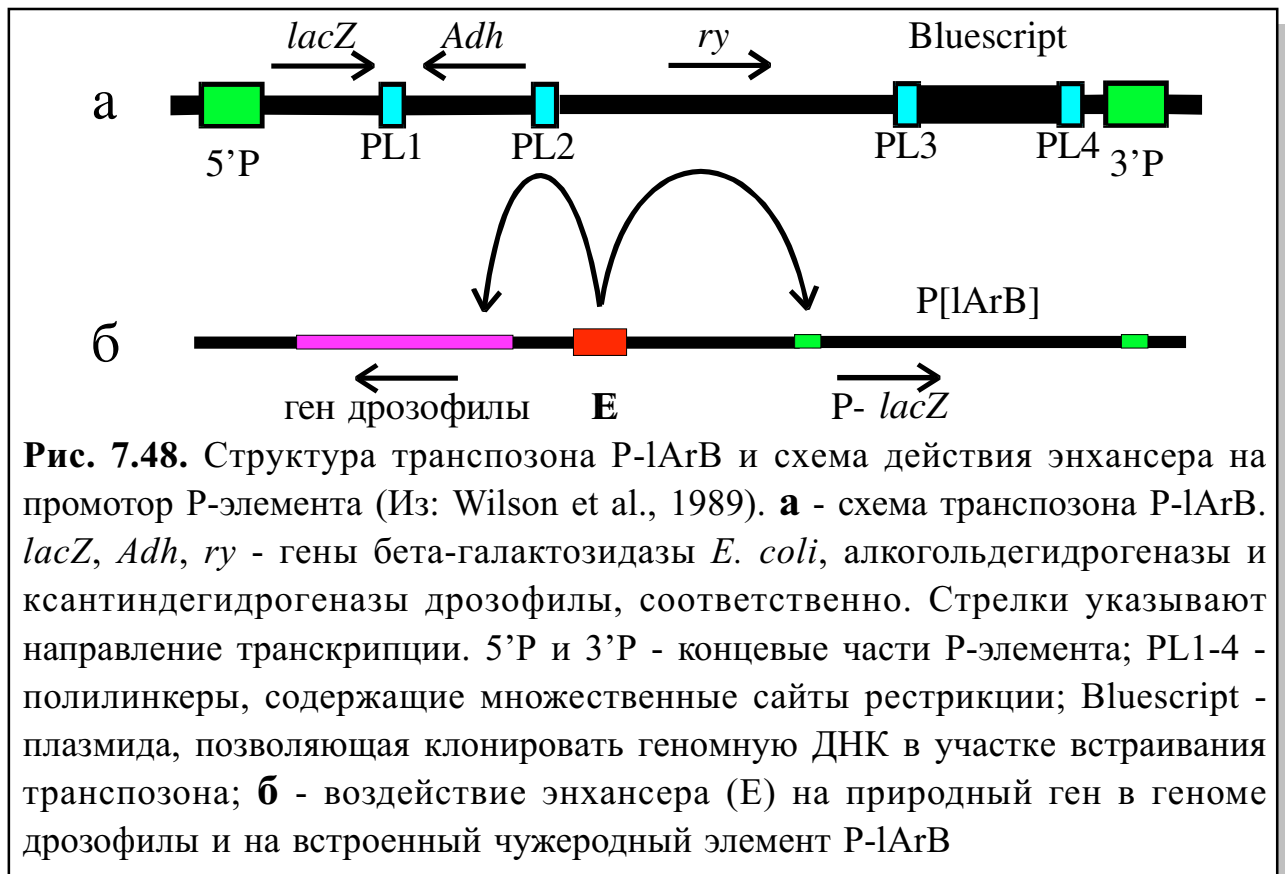


Рис. 7.48. Структура транспозона P-lArB и схема действия энхансера на промотор P-элемента (Из: Wilson et al., 1989). **а** - схема транспозона P-lArB. *lacZ*, *Adh*, *ry* - гены бета-галактозидазы *E. coli*, алкогольдегидрогеназы и ксантиндегидрогеназы дрозофилы, соответственно. Стрелки указывают направление транскрипции. 5'P и 3'P - концевые части P-элемента; PL1-4 - полилинкеры, содержащие множественные сайты рестрикции; Bluescript - плаزمид, позволяющая клонировать геномную ДНК в участке встраивания транспозона; **б** - воздействие энхансера (E) на природный ген в геноме дрозофилы и на встроенный чужеродный элемент P-lArB

дрозофилы в ходе нормального развития (элемент E на Рис. 7.48б.), этот энхансер может активировать слабый промотор P-элемента и соединенного с ним гена *lacZ* (Рис. 7.48а.). В результате активирования гена *lacZ* и окраски внутренних органов дрозофилы реактивом X-gal можно обнаружить клетки, окрашенные в синий цвет и, таким образом, установить, что в зоне встраивания транспозона в молекуле ДНК находится ген, активно функционирующий в данной клетке. Используя перемещения транспозона и собирая те линии, в которых *lacZ* активен, можно выявить все или существенную часть генов, активно функционирующих в той или иной дифференцированной ткани. Наличие в составе транспозона плазмиды *Bluescript* (см. Рис. 7.48а.) дает возможность клонировать и ген, находящийся под контролем энхансера,

и сам энхансер. Этот метод имеет название *enhancer trapping*.

У дрожжей известны элементы, функционально похожие на энхансеры, которые называются *upstream activator sequences* (UAS). Подобно энхансерам UAS могут функционировать в любой ориентации и на различных расстояниях до промотора. Однако, в отличие от энхансера, UAS не могут функционировать когда локализованы ниже промотора. UAS-последовательности активируются, когда с ними связываются, белки GAL4 (рис. 7.49.).

В последние несколько лет появился новый метод активирования направленной экспрессии генов у дрозофилы. Суть его заключается в том, что исследователи создают две трансформированные линии мух, в одну из которых вводится GAL4-транспозон, а в другую UAS-транспозон (Рис. 7.50.).

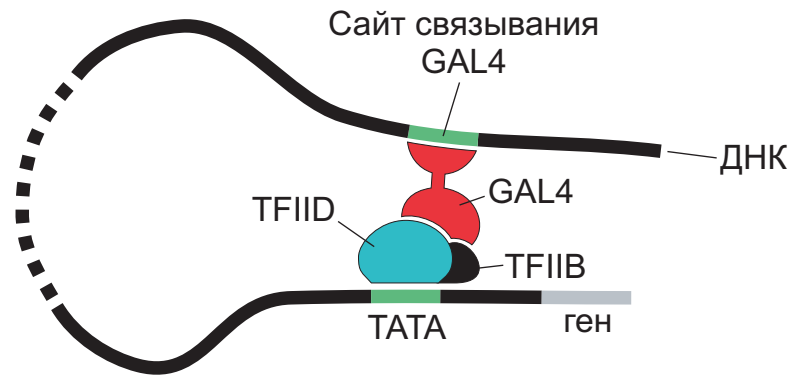


Рис. 7.49. Взаимодействие белка GAL4 с участком UAS и транскрипционным комплексом, в результате чего облегчается добавление к комплексу фактора транскрипции TFIIB. Это усиливает скорость транскрипции примерно в 1000 раз (Из: Alberts et al., 1994, p. 425)

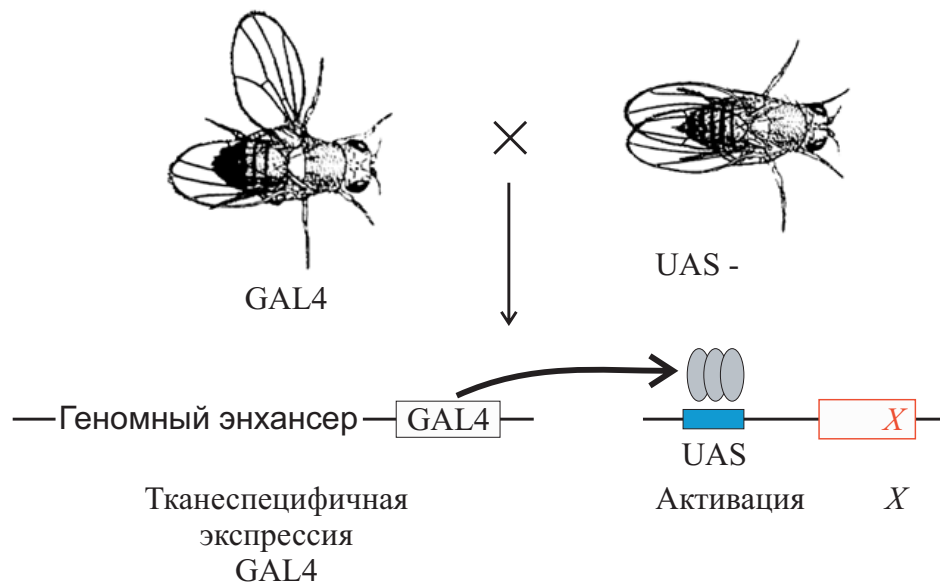


Рис. 7.50. Схема активирования генов в транспозонах GAL4-UAS. Ген *GAL4* встраивается под любой энхансер в геноме дрозофилы. В результате экспрессии гена *GAL4* считывается белок, который связывается с промотором UAS, активируя встроенный ниже ген *X* (Из: Brand, Perrimon, 1993)

Для создания первой линии в Р-элемент, наряду с обычными маркерами вводится ген, кодирующий белок GAL4. Такой транспозон встраивается в случайные районы хромосом дрозофилы, в том числе может встроиться под какой-нибудь энхансер, например, под энхансер гена, работающего в клетках формирующегося крыла или ноги. Поэтому GAL4 будет экспрессироваться

в клетках крыла или ноги, однако никаких последствий для этих клеток не будет, поскольку белок GAL4 может активировать транскрипцию другого гена лишь в том случае, если он свяжется с промотором UAS, в нормальном развитии функционирующий также только в геноме дрожжей. Поэтому и создают вторую линию дрозофилы, которую

трансформируют транспозоном, содержащим испытуемый ген (X - на рис. 7.50.), сшитый с промотором UAS. Когда в результате скрещивания двух линий оба транспозона объединяются в генотипе потомка, ген *GAL4* начинает функционировать в тех клетках, в которых активен его энхансер. Белок GAL4 связывается с промотором UAS, который и активирует ген "X". При этом активация "X" будет совсем не в той ткани, в которой он активируется своим собственным энхансером, а в той, в которой активен энхансер гена *GAL4* (см. рис. 7.50.). Такая экспрессия называется эктопической, т.е. "не на своем месте". Рассмотрим один из наиболее интересных примеров, связанный с использованием GAL4-UAS транспозонов.

В лаборатории В. Геринга (Рис. 7.51.) заинтересовались развитием глаз у дрозофилы и некоторых других животных. Мутации некоторых генов приводят к полному отсутствию глаз, например *eyeless* у дрозофилы. В результате сопоставления молекулярной организации гена *Aniridia* у человека, гена *Small eye* у мыши, и *eyeless* у дрозофилы было найдено значительное их сходство. Это тем более удивительно потому, что анатомически глаза у мухи и млекопитающих имеют очень мало черт сходства: всем известный из учебников по анатомии глаз человека и сложный глаз, состоящий из 800 отдельных глазков (фасеток), собранных в один общий орган зрения, у дрозофилы. Оказалось возможным получить гибридный генотип - гомозигот по мутации *eyeless* и дрозофилы, содержащий транспозон



Рис. 7.51. Вальтер Геринг

GAL4 и UAS, в котором *GAL4* находится под контролем энхансера, функционирующего в клетках развивающихся ног, крыльев и органов головного комплекса. В транспозон UAS на место гена "X" встроена ДНК нормального аллеля гена *eyeless*⁺. В результате функционирования гена *eyeless* появляются, во-первых, нормально функционирующие глаза, а также маленькие глаза в разных участках крыльев, ног, на голове - там, где был активен энхансер (Рис. 7.52.). Затем был сделан еще более острый эксперимент: на место гена "X" в транспозон UAS встроили ДНК гена *Small eye*, выделенную из генома мыши. И опять начали формироваться сложные глаза, характерные для дрозофилы в разных участках тела мухи (Рис. 7.52.). Как же могло получиться, что действие гена, контролирующего развитие одного типа глаза у мыши, привело к развитию глаза совершенно другого типа у дрозофилы? Согласно существующим расчетам в процессах формирования глаз у животных принимает участие около 2 500 генов, и их действие

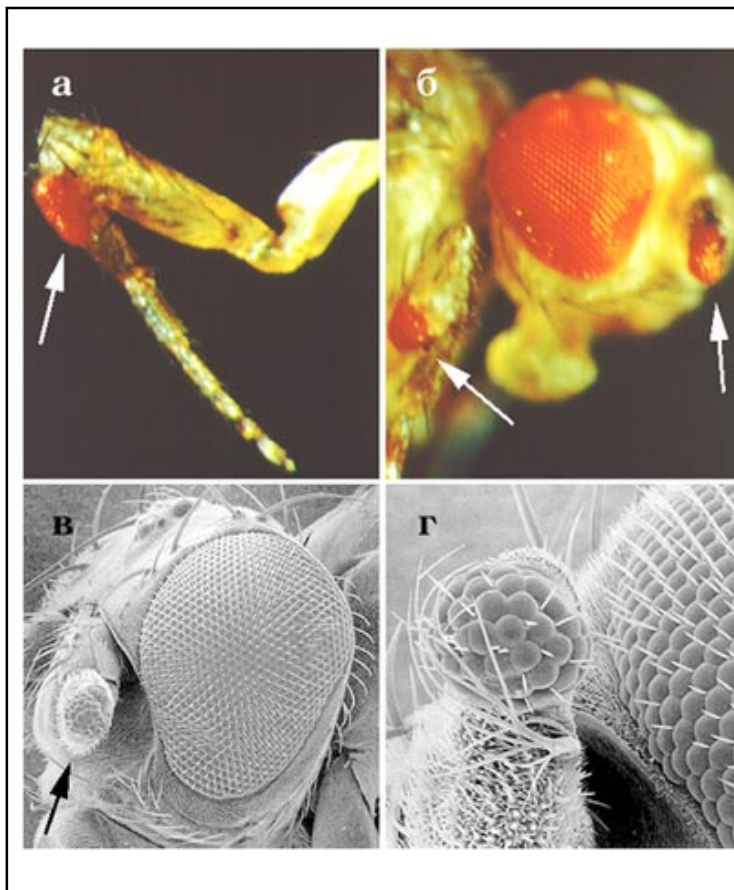


Рис. 7.52. Формирование глаз в обычных участках головы и маленьких эктопических глаз на ногах, под крыльями и антеннах у трансформированных дрозофил: эктопический глаз на ноге (а), под крылом и на антенне (б); дрозофилы трансформированы транспозонами, содержащими ген дрозофилы *eyeless*⁺; в и з - развитие нормального и эктопического глаза на антенне (з - большое увеличение); мухи трансформированы транспозонами, содержащими ген *Small eye*, выделенный из мыши. а, б - из: Halder et al., 1995; 1995; в, з - из: Gehring, 1996

организовано в каскад. Оказалось, что все три рассмотренных выше гена - *Small eyes*, *Aniridia* и *eyeless* находятся в начале каскада, т.е. они только дают команду на начало развития глаза, а остальные гены, функционирующие после них, уже определяют, какой глаз будет сформирован.

Литература к разделу 7.8.

Алиханян С.И., Акифьев А.П., Чернин Л.С. Общая генетика. Москва, Высшая школа, 1-446, 1985.

Гвоздев В.А. Механизмы регуляции активности генов в процессе транскрипции. Соросовский образовательный журнал N1, p. 23-31, 1996.

Жимулев И.Ф. Хромомерная организация политеменных хромосом. Новосибирск, Наука, 1-564, 1994.

Жимулев И.Ф. Трансформация у дрозофилы - новый подход в генетике. Сорос. образов. журнал, 1999, в печати.

Alberts B., Bray D., Lewis J., Raff M., Roberts K., Watson J.D. Molecular biology of the cell (Third edition) Garland Publishing, Inc. New York, London, 1994.

Arkhipova I.R. Promoter elements in *Drosophila melanogaster* revealed by sequence analysis. Genetics V. 139, p. 1359-1369, 1995.

Bellen H.J., O'Kane C.J., Wilson C., Grossniklaus U., Pearson R.K., Gehring W.J. P-element mediated enhancer detection: a versatile method to study development in *Drosophila*. Genes and Development V. 3, p. 1288-1300, 1989.

Bier E., Vaessin H., Shepherd S., Lee K., McCall K., Barbel S., Ackerman L., Carretto R., Uemura T., Grell E., Jan

- L.Y., Jan Y.N. Searching for pattern and mutation in the *Drosophila* genome with a P-*lacZ* vector. *Genes and Development* V. 3, p. 1273-1287, 1989.
- Brand A.H., Perimon N. Targeted gene expression as a means of altering cell fates and generating dominant phenotypes. *Development*, 1993, V. 118, p. 401-415.
- Gehring W.J. The master control gene for morphogenesis and evolution of the eye. *Genes to Cells*, 1993, V. 1, p. 11-15.
- Halder G., Callaerts P., Gehring W.J. New perspectives on eye evolution. *Current Opinion in Genetics & Development*, 1995, V. 5, p. 602-609.
- Halle J.-P., Meisterernst M. Gene expression: increasing evidence for a transcriptosome. *TIG* V. 12, p. 161-163, 1996.
- Lawrence P.A. The making of fly. Oxford, Blackwell Scientific Publications, 1-228, 1992.
- Lewin B. *Genes* V. Oxford University Press. Oxford, New York, Tokyo, 1-1272, 1994.
- O'Kane K., Gehring W.J. Detection in situ of genomic regulatory elements in *Drosophila*. *Proc. Nat. Acad. Sci. U.S.A.* V. 84, p. 9123-9127, 1987.
- Russell, P.J. *Genetics* Fifth edition. Addison Wesley Longman, Inc, Menlo Park, California, p. 379-417, 1998.
- Struhl K. Chromatin structure and RNA polymerase II connection: implications for transcription. *Cell* V. 84, p. 179-182, 1996.
- Wilson C. Pearson R.K., Bellen H.J., O'Kane C.J. Grossniklaus U., Gehring W.J. P-element-mediated enhancer detection: an efficient method for isolating and characterizing

developmentally regulated genes in *Drosophila*. *Genes & Development*, 1989, V. 3, p. 1301-1313.

7.9. Структурная часть гена

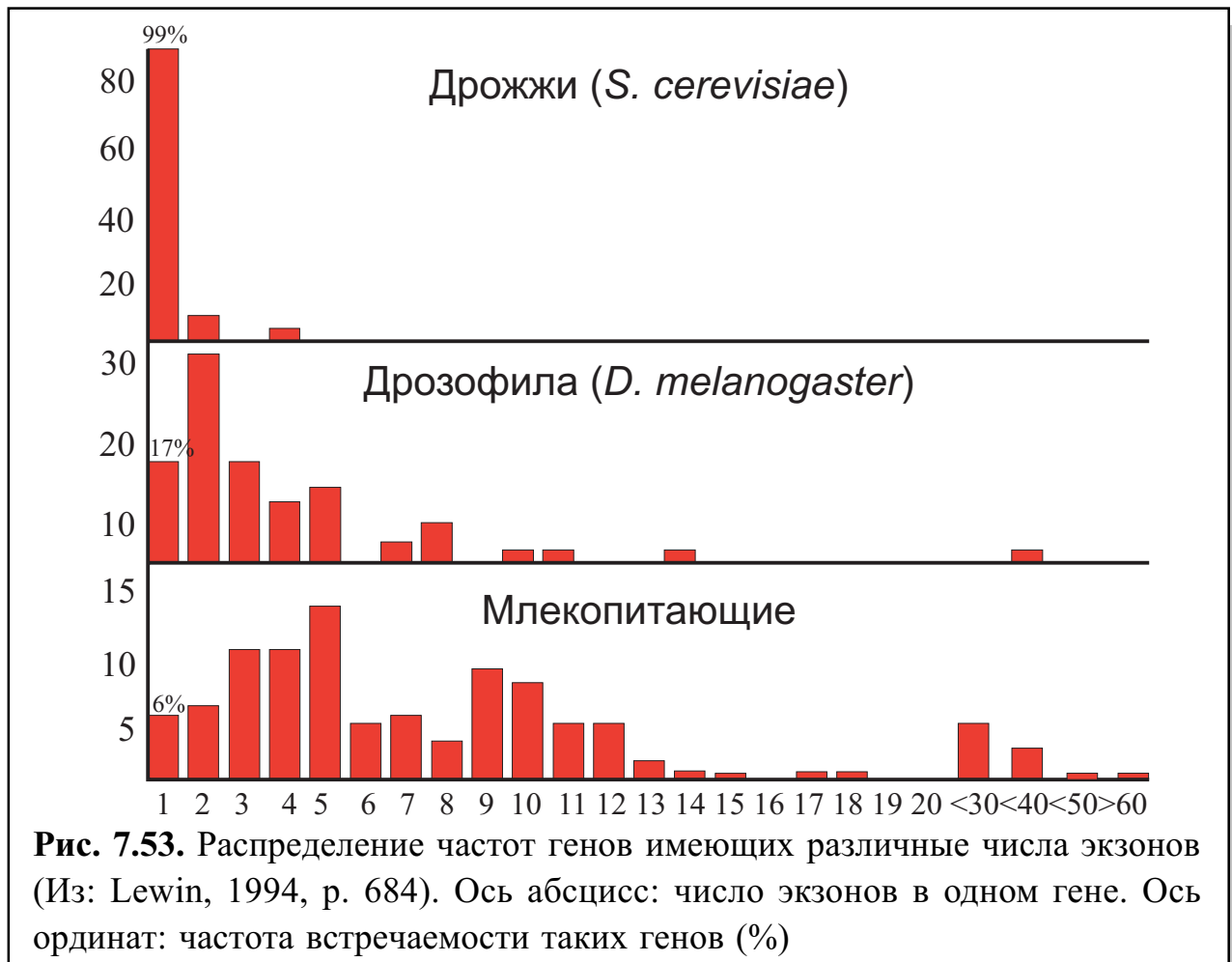
7.9.1. Интроны и экзоны

При изучении первичной структуры, т.е. последовательности нуклеотидов ряда генов выяснилось, что в них, наряду с участками, кодирующими специфичный для этого гена продукт (полипептид, рРНК, тРНК и т.д.) имеются участки, которые ничего не кодируют, т.е. они подобно межгенным спейсерам (участкам между генами) не содержат генетической информации. Группы учёных, возглавляемых Р. Робертсом и П.А. Шарпом обнаружили такие расщеплённые гены у аденовируса 2 в 1977 году.

Некодирующие участки получили название интронов, кодирующие - экзонов. Такой тип структурной организации обнаружен для множества генов, локализованных в хромосомах, некоторых генов внутриклеточных органелл эукариот - пластид и митохондрий, а также для генов нескольких РНК-содержащих и ДНК-содержащих вирусов, поражающих эукариот. У бактерий интронов в генах нет. Нет интронов и в генах вирусов, поражающих бактерии.

Число и внутригенная локализация интронов характерны для каждого гена, что становится очевидным в результате сравнения организации гомологичных генов у разных видов.

Некоторые гены содержат только один-два интрона, но часто их значительно больше. Так, например в гене овальбумина курицы 7 интронов,



в гене сывороточного альбумина крысы их 13, а один из генов коллагена курицы имеет даже 51 интрон.

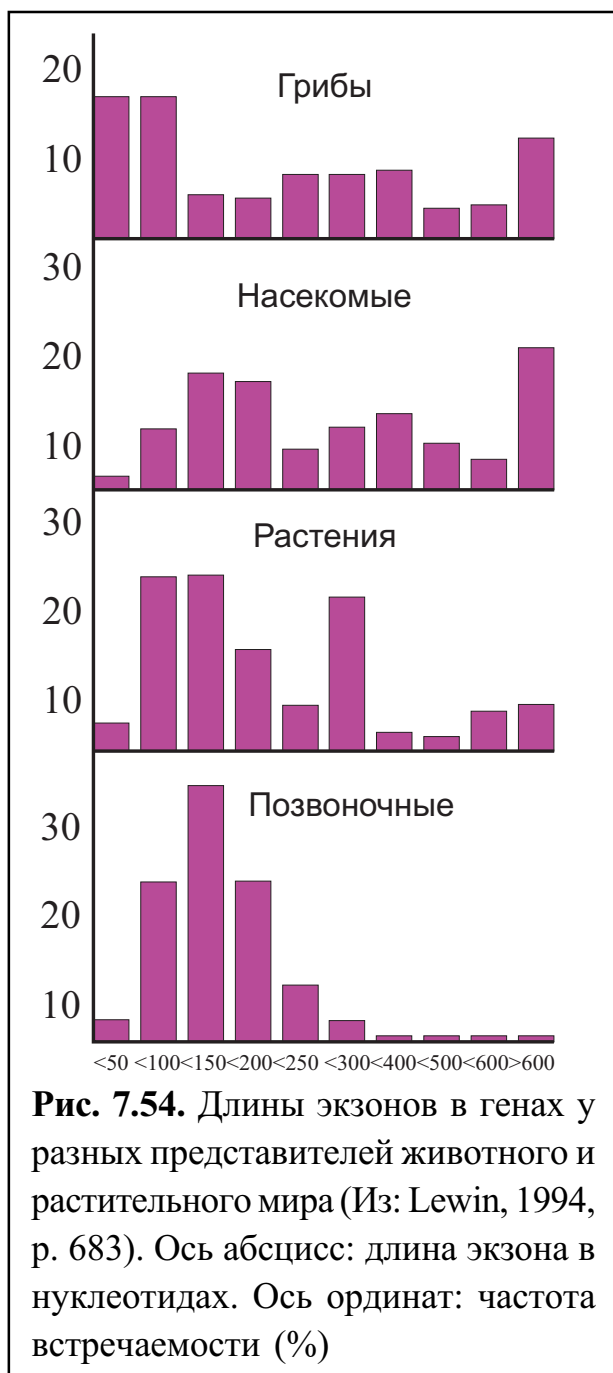
На Рис. 7.53. показано распределение частот экзонов, приходящихся на один ген. Видно, что у низших эукариот, таких как дрожжи, 95% генов содержат только один экзон, значит такие гены не прерываются интронами. У дрозофилы таких генов только 17%, а у млекопитающих - только 6%.

Экзоны имеют, как правило, небольшую длину (Рис. 7.54.).

Длина интрона может быть очень различной - от нескольких десятков пар нуклеотидов до многих тысяч (Рис. 7.55.). Общая длина всех интронов часто значительно превышает суммарную длину экзонов. К примеру, из

приблизительно 7000 пар нуклеотидов, образующих ген овальбумина, на долю экзонов приходится всего 1872 п.н., т.е. почти 3/4 длины составляют интроны.

Интроны транскрибируются наравне с экзонами, так что про-мРНК содержит участки, транскрибированные как с экзонов, так и с интронов. В дальнейшем в ходе процессинга, происходящего в ядре, участки про-мРНК, транскрибированные с интронов, вырезаются, а бывшие разобщенными участки, считанные с экзонов, "сращиваются", и зрелая мРНК содержит только транскрипты экзонов. Эти прежде разобщенные участки соединяются в нужном порядке. Воссоединение участков, транскрибированных с экзонов при образовании зрелой мРНК, называют



сплайсингом (splicing - сращивание морских канатов - англ.) (Рис. 7.56.). Длина гена существенно больше длины мРНК (Табл. 7.2.).

Интроны всегда (установлено для генов, кодирующих белки) имеют на 5'-конце двойку последовательностей гуанин-тимин, а на 3'-конце - аденин-гуанин.

Последовательности нуклеотидов в экзонах консервативны, а в интронах сильно варьируют. Иногда экзон одного

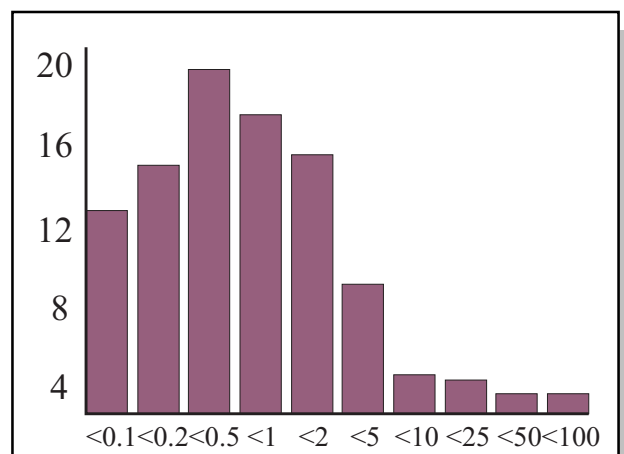


Рис. 7.55. Длины интронов в генах позвоночных (Из: Lewin, 1994, р. 683). Ось абсцисс: длина интрона в т.п.н. Ось ординат: частота встречаемости (%)

Дополнение 7.5.

Нобелевская премия 1993 года была присуждена Р. Дж. Робертсу и Ф. А. Шарпу (Richard J. Roberts, Philip A. Sharp) за открытие расщепленных генов.

гена может быть гомологичным экзону даже другого гена. Например, два β -глобиновых гена мыши имеют по три гомологичных экзона в каждом гене. Между интронами этих генов гомология не найдена. Связано это с тем, что интроны эволюционируют значительно быстрее, чем экзоны. При сравнениях последовательностей нуклеотидов в одних и тех же генах у разных видов, находят большую гомологию в интронах (Из: Lewin, 1994, pp. 688-689).

7.9.2. Альтернативный сплайсинг

В некоторых случаях в аминокислотные последовательности транслируются не все существующие экзоны. В результате с одного гена считывается более одного типа мРНК.

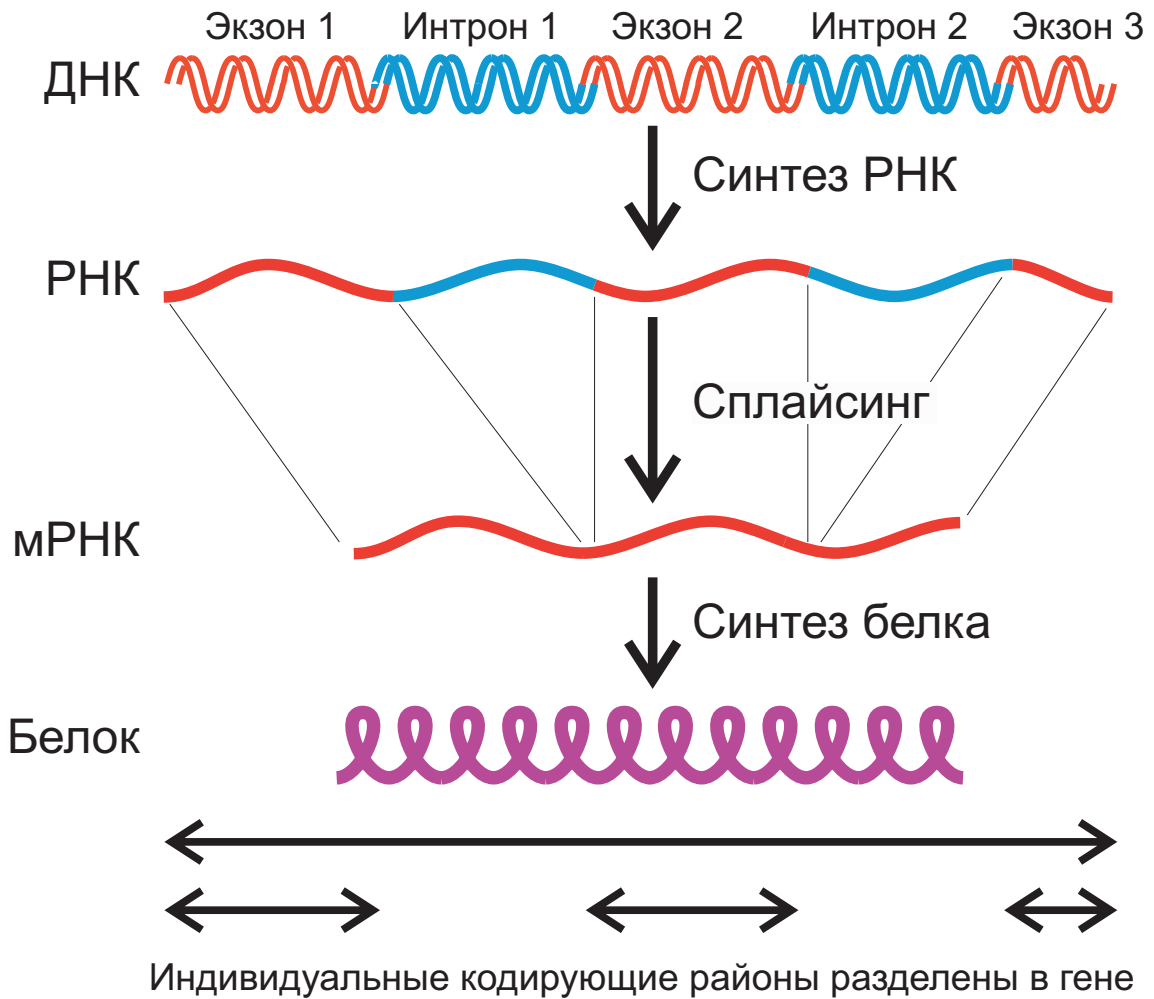


Рис. 7.56. Процесс передачи информации от ДНК до белка, кодируемого расщепленным геном (Lewin, 1994, p. 150)

Табл. 7.2. Соотношение длины гена и матричной РНК в зависимости от числа экзонов (Из: Lewin, 1994, p. 687)

Виды	Среднее число экзонов	Средняя длина гена (т.п.н.)	Средняя длина мРНК (т.п.н.)
<i>Saccharomyces cerevisiae</i>	1	1.6	1.6
Грибы	3	1.5	1.5
<i>Cenorabditis elegans</i>	4	4.0	3.0
<i>Drosophila melanogaster</i>	4	11.3	2.7
Куры	9	13.9	2.4
Млекопитающие	7	16.6	2.2

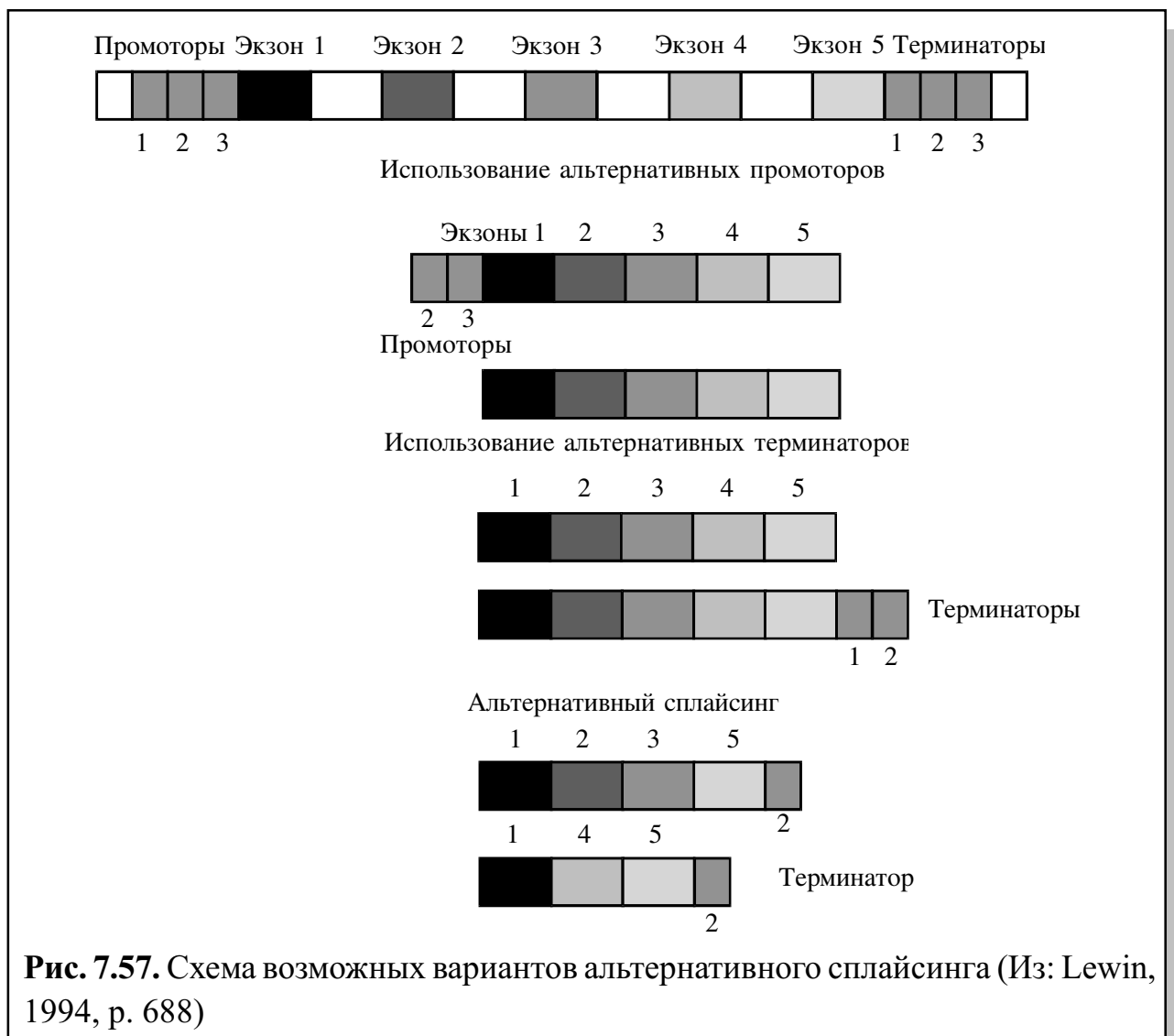
На Рис. 7.57. суммированы возможные механизмы образования многих типов молекул РНК, считываемых с одного гена: за счет изменения инициации и терминации транскрипции, а также сплайсинга. Использование альтернативных промоторов может изменять 5' конец, использование альтернативных терминаторов может изменять 3' конец транскрипта. Эти изменения могут организоваться только иницирующими или терминальными последовательностями мРНК. Иногда изменения в нетранслируемой 5' лидерной последовательности или 3' нетранскрибируемой трэйлерной

последовательности могут иметь регуляторные последствия.

На Рис. 7.58. представлена схема альтернативного сплайсинга в гене *BR-C* у дрозофилы. Весь ген занимает около 120 т.п.н., в нем выделяют 10 экзонов. Обнаружено более 15 типов различных мРНК.

Другой тип альтернативного сплайсинга, когда один ген функционирует фактически как два (Рис. 7.59.). Показаны два транскрипта: E74A и E74B, синтезирующиеся одновременно с двух разных промоторов.

Альтернативный сплайсинг играет огромную роль в генетическом определении пола (см. Рис. 13.6.).



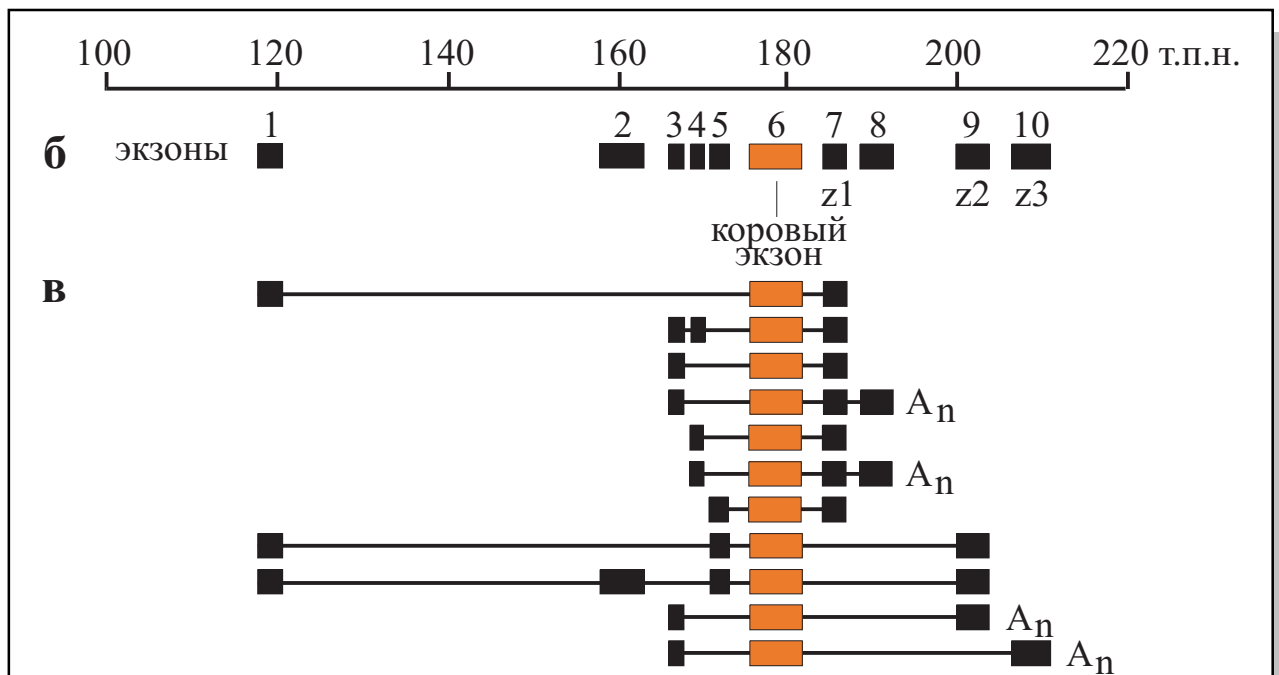


Рис. 7.58. Организация транскрипции гена *BR-C* у *D. melanogaster* (Из: DiBello et al., 1991, в кн. Жимулёв, 1994, стр. 284); **а** - физическая карта гена (от 100 до 220 т.п.н.); **б** - расположение экзонов (1-10) на физической карте; **в** - различные транскрипты, возникающие за счет альтернативного сплайсинга

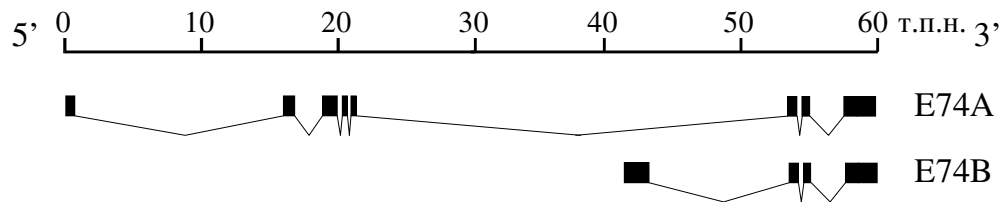


Рис. 7.59. Структура гена *E74*, расположенного в пуфе 74EF (Из: Karim, Thummel, 1991, в кн. Жимулев, 1994, стр. 293)

Во время созревания макронуклеуса у инфузорий происходит удаление внутригенных последовательностей без синтеза РНК. Более того, происходит престановка экзонов. Так, в микронуклеарной ДНК экзоны гена актина I у *Oxytricha nova* расположены в последовательности: 3, 4, 6, 5, 7, 9, 2, 1, 8. После удаления “интронов” экзоны занимают положение в ряду с 1 по 9 и только такая их последовательность даёт функциональный белок (См. детали: Prescott, 1992).

Ещё более впечатляет реорганизация гена α ТВР, который в микронуклеусе

содержит 14 экзонов, расположенных в следующем порядке: 1, 3, 5, 7, 9, 11, 2, 4, 6, 8, 10, 12, 13, 14, а нормальный по нумерации порядок экзонов образуется в созревшем макронуклеусе (Рис. 7.60.).

7.9.3. Локализация генов в интронах

В интронах некоторых генов располагаются другие гены (Рис. 7.61 - 7.63.).

Удивительным по сложности организации и величине интронов является ген *dnc* у *D. melanogaster*. Ген занимает минимум 130 т.п.н. (от -90 до

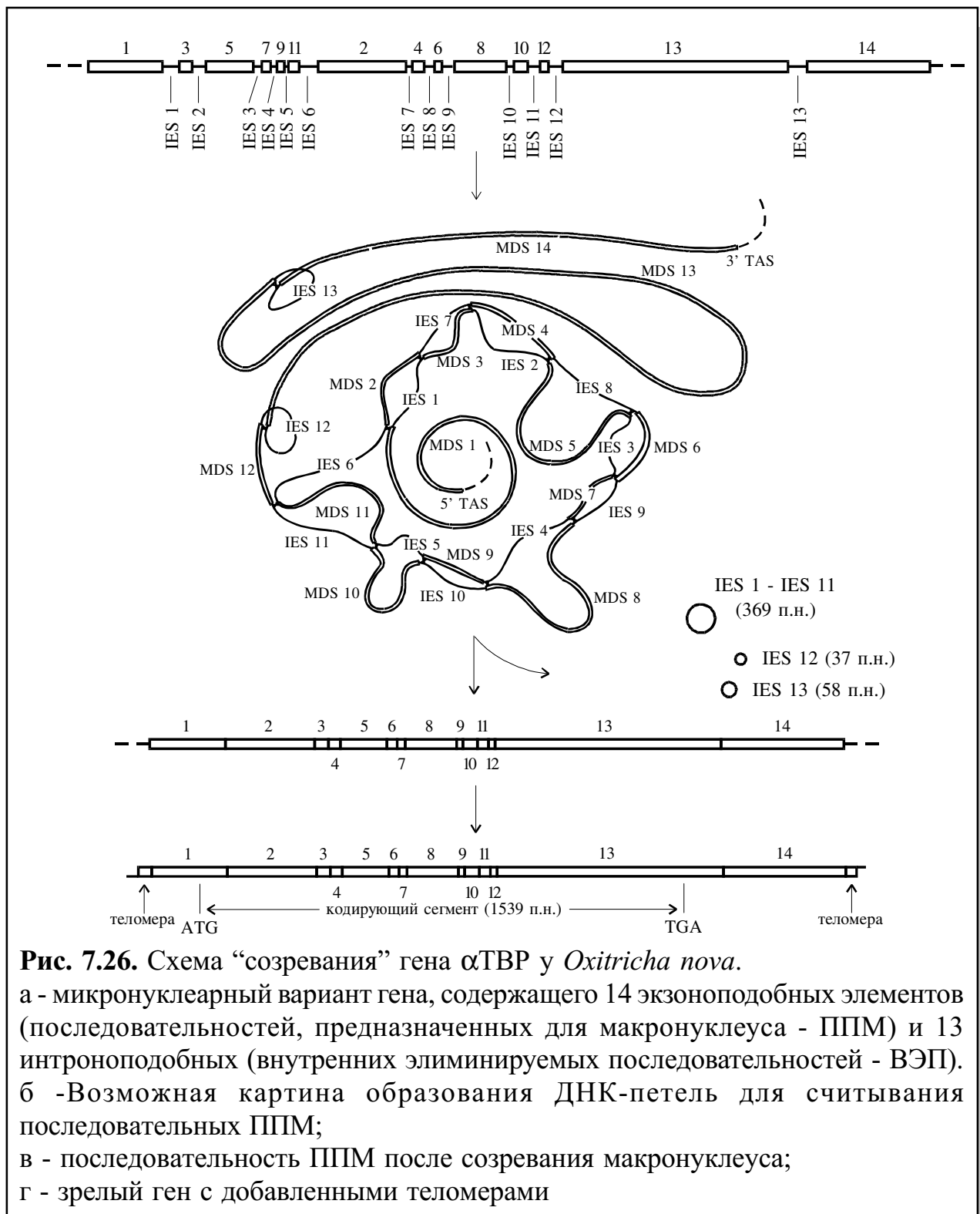


Рис. 7.26. Схема “созревания” гена α TVP у *Oxitricha nova*.

- а - микронуклеарный вариант гена, содержащего 14 экзонотипных элементов (последовательностей, предназначенных для макронуклеуса - ППМ) и 13 интронотипных (внутренних элиминируемых последовательностей - ВЭП).
 б - Возможная картина образования ДНК-петель для считывания последовательных ППМ;
 в - последовательность ППМ после созревания макронуклеуса;
 г - зрелый ген с добавленными теломерами

+40 т.п.н.) и содержит 13 экзонов. Между первым экзонам и экзонам, расположенным еще более дистально (около 40 т.п.н.), расположено несколько генов из семейства *Pig/Sgs*. Между вторым и третьим экзонами (около 70 т.п.н.) располагается еще 4 гена (Рис. 7.63).

7.9.4. Использование промоторов генов теплового шока

Организацию и экспрессию структурной части генов можно изучать, объединяя их в конструкцию с известным промотором, и трансформируя

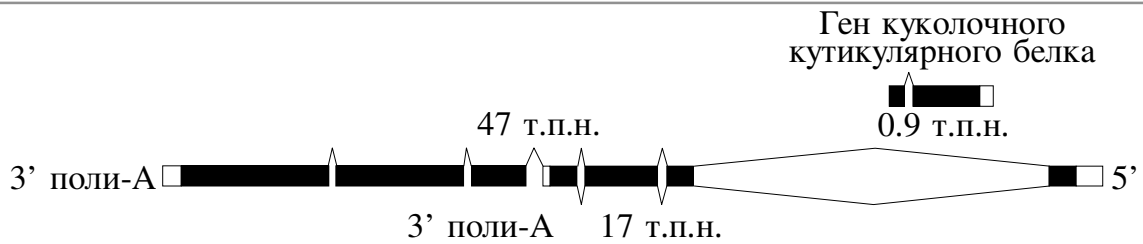


Рис. 7.61. Экзон-интронная карта гена *Gart* и расположенного в его интроне гена кукольного кутикулярного белка дрозофилы (Из: Жимулев, 1994, стр. 147)

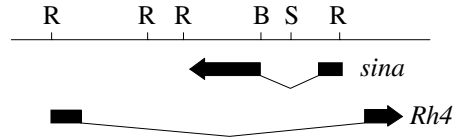


Рис. 7.62. Расположение гена *sina* в интроне гена *Rh4* дрозофилы. Вверху - рестрикционная карта района (Из: Carthew, Rubin, 1990, в кн. Жимулев, 1994, стр. 148)

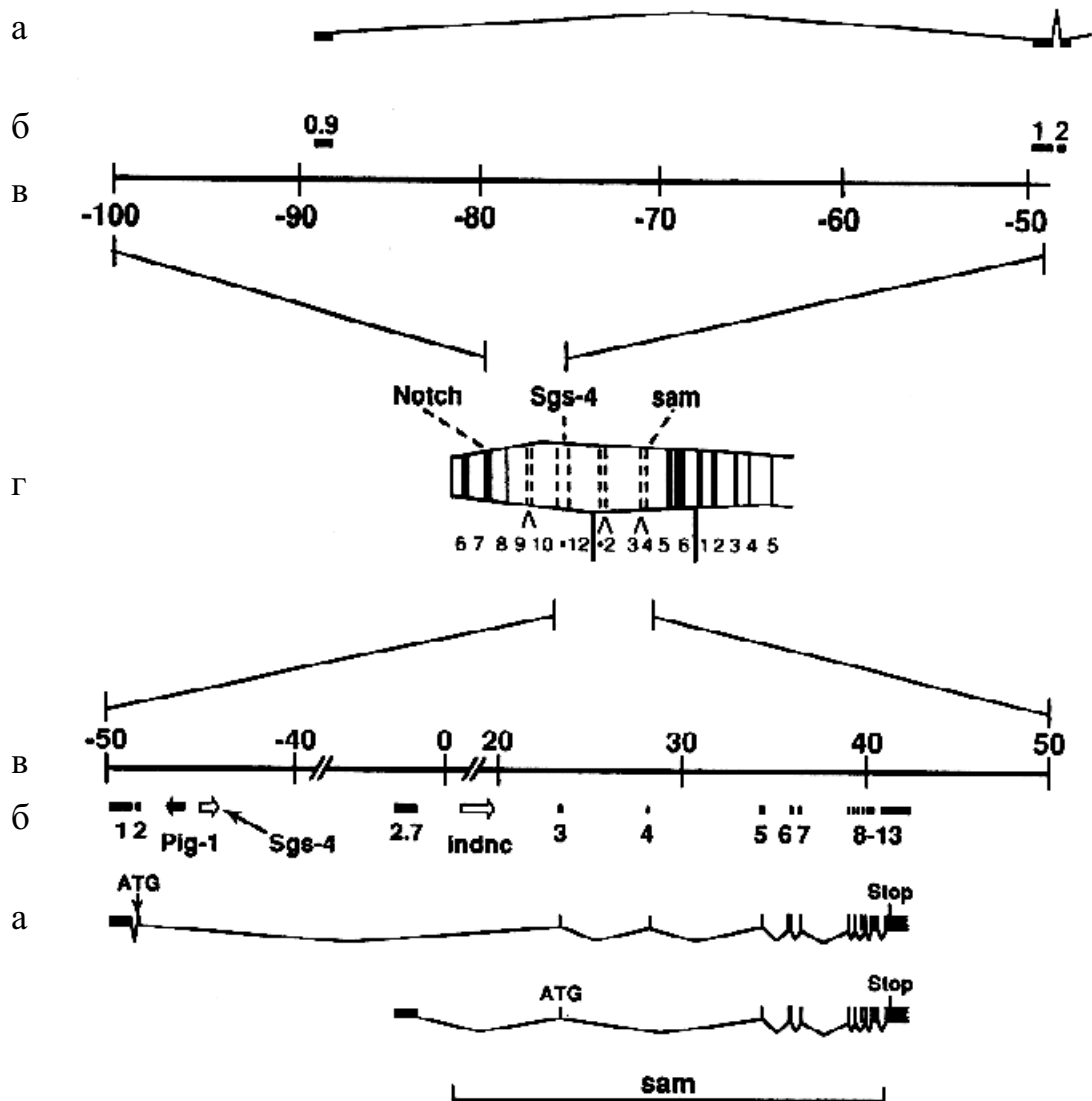


Рис. 7.63. Структура гена *dnc* у *D. melanogaster* (Из: Davis, Dauwalder, 1990, в кн. Жимулев, 1994, стр. 149); а - карта интронов и экзонов гена *dunce*; б - карта других генов, расположенных в интронах гена *dunce*; в - физическая карта ДНК в районе хромосомы 3С; г - участок хромосомы 3С6 - 3Е5

дрозофил, мутантных по этому гену. Для этих целей используют промоторы генов теплового шока. В нормальном развитии эти промоторы не функционируют, но активируются при резком повышении температуры (37°C на 30 мин) (см. раздел 12.).

Ген под действием шока начинает экспрессироваться, о чем можно судить по “исправлению” мутантного фенотипа, появлению нормального белка или РНК (см. детали: Lawrence, 1992, p. 55).

7.9.5. Участки, терминирующие транскрипцию

Терминация транскрипции прокариотических генов обусловлена элементами, называемые терминаторами. Одним из важных белков, вовлеченных в терминацию транскрипции некоторых генов у *E. coli*, является белок ρ . Терминаторы в таких генах называются ρ -зависимыми (или терминаторы II типа). Во многих других терминаторах сердцевинная РНК-полимераза сама

может проводить события терминации. Терминаторы такого типа называются ρ -независимыми терминаторами (терминаторы I типа). ρ -независимые терминаторы состоят из последовательностей, расположенных за 16-20 п.н. от точки терминации и представляющих собой инвертированный повтор (Рис. 7.64.).

Эта последовательность замыкается нитью из 4-8 АТ-последовательностей, на которой синтезируются цепь из U-последовательностей.

ρ -зависимые терминаторы не имеют АТ нити и во многих случаях не формируют шпилечных структур. ρ -фактор - это протеин с двумя доменами: один связывается с РНК, другой домен - с АТФ.

Три ключевые события происходят на терминаторах обоих типов:

1. останавливается синтез РНК,
2. цепь РНК освобождается от ДНК,
3. РНК-полимераза освобождается от ДНК.

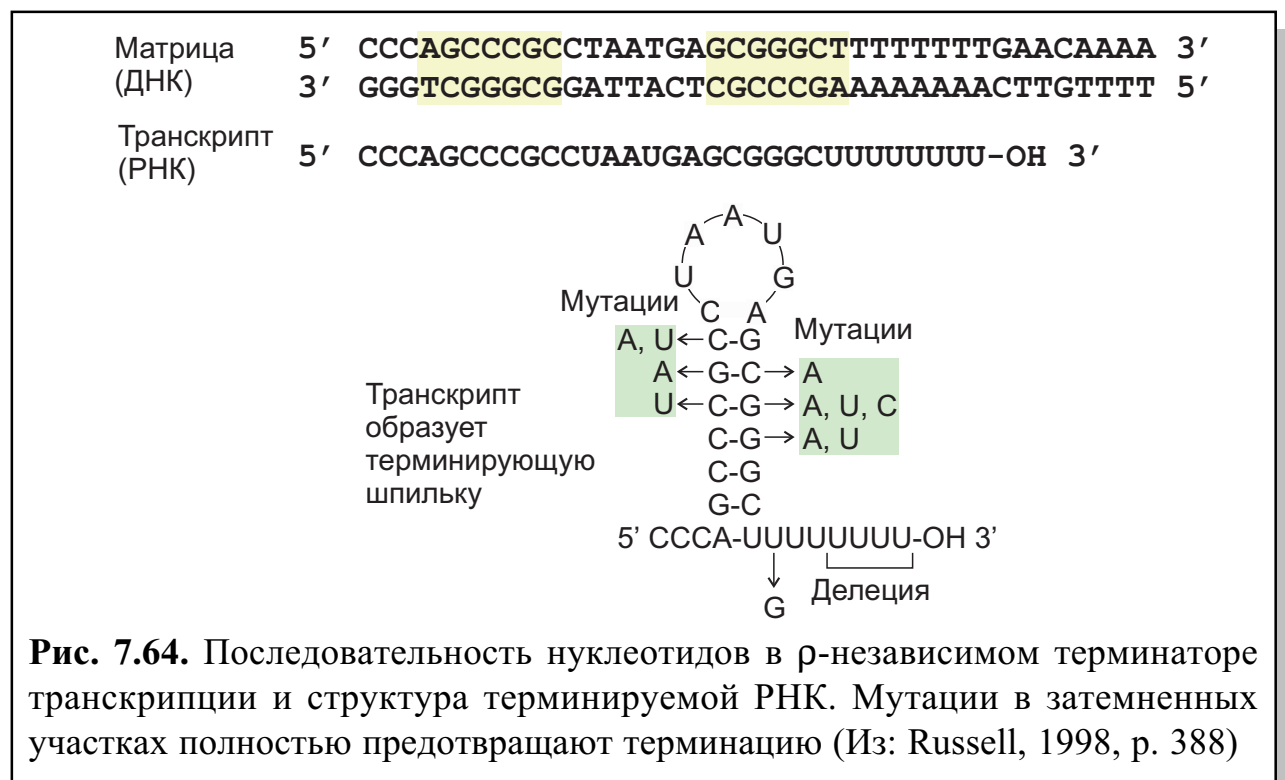


Рис. 7.64. Последовательность нуклеотидов в ρ -независимом терминаторе транскрипции и структура терминируемой РНК. Мутации в затемненных участках полностью предотвращают терминацию (Из: Russell, 1998, p. 388)

7.9.6. Гомология генов

(не расписано)

О гомологии свидетельствует закон гомологических рядов Вавилова. Обладают гомологией:

1. Гены домашнего хозяйства (рРНК, гистонов)

2. Гены, выполняющие одинаковые функции

3. Домены в составе генов (цинковые пальцы, гомео-домен)

4. Гены теплового шока (книга по тепловому шоку Ashburner, Shlesinger)

7.9.7. Псевдогены

(не расписано)

Иногда в геномах встречаются т.н. псевдогены. Они имеют все необходимые черты генов, т.е. полный набор экзонов, характерных для этого гена, поли-А/Т хвост и короткие прямые повторы (как у мобильных элементов) в ДНК мишени.

Псевдоген начинается с 5' точки, эквивалентной 5' точке мРНК, а заканчивается трактом поли-А нуклеотидов, что вероятно имеет происхождение от поли-А конца в молекуле мРНК. (Из: Lewin, 1994, p.1051-1055, см. также главу 24).

У дрозофилы псевдогены встречаются редко, чаще это гены транспортных или малых ядерных РНК (тРНК, snRNA). Еще один известный пример - псевдоген личиночного кутикулярного белка и псевдоген *Adh* у *D. mulleri*, они сохраняют свои интроны. Пример процессированного псевдогена (т.е. без интронов) описан для гена *Adh* у *D. teissieri* и *D. yakuba*.

Литература к разделу 7.9.

Ashburner M. *Drosophila*. A laboratory handbook. Cold Spring Harbor. Cold Spring Harbor Laboratory Press, 1-1331, 1989

Lawrence P.A. The making of fly. Oxford, Blackwell Scientific Publications, 1-228, 1992.

Lewin B. Genes V. Oxford University Press. Oxford, New York, Tokyo, 1-1272, 1994.

Prescott D.M. The unusual organization and processing of genomic DNA in hypotrichous ciliates. TIG 8, N12, 439-445, 1992.

Prescott D.M. The DNA of ciliated protozoa. Microbiol. Rev. 58, N2, 233-267. 1994.

Russell, P.J. Genetics Fifth edition. Addison Wesley Longman, Ins, Menlo Park, California, p. 379-417, 1998.