

**Глава 6. Структура и организация генома**

<b>6.1. Роль ДНК в наследственности</b>	<b>2</b>
<b>6.2. Структура ДНК</b>	<b>3</b>
<b>6.3. Репликация ДНК</b>	<b>3</b>
<b>6.4. Генетический код</b>	<b>12</b>
<b>6.5. Структура генома эукариот</b>	<b>15</b>
<b>6.6. Мобильные элементы генома</b>	<b>19</b>
<b>6.6.1. Мобильные элементы геномов растений</b>	<b>19</b>
<b>6.6.2. Мобильные элементы у дрозофилы</b>	<b>23</b>
<b>6.6.3. Ту-элементы у дрожжей</b>	<b>26</b>
<b>6.6.4. Транспозоны млекопитающих</b>	<b>27</b>
<b>6.6.5. Функциональное значение мобильных элементов</b>	<b>28</b>
<b>6.7. Мобильные элементы прокариот</b>	<b>31</b>
<b>6.7.1. IS-элементы</b>	<b>31</b>
<b>6.7.2. Транспозоны</b>	<b>31</b>
<b>6.7.3. IS-элементы и транспозоны в плазмидах</b>	<b>33</b>
<b>6.7.4. Бактериофаг <i>Mu</i></b>	<b>35</b>

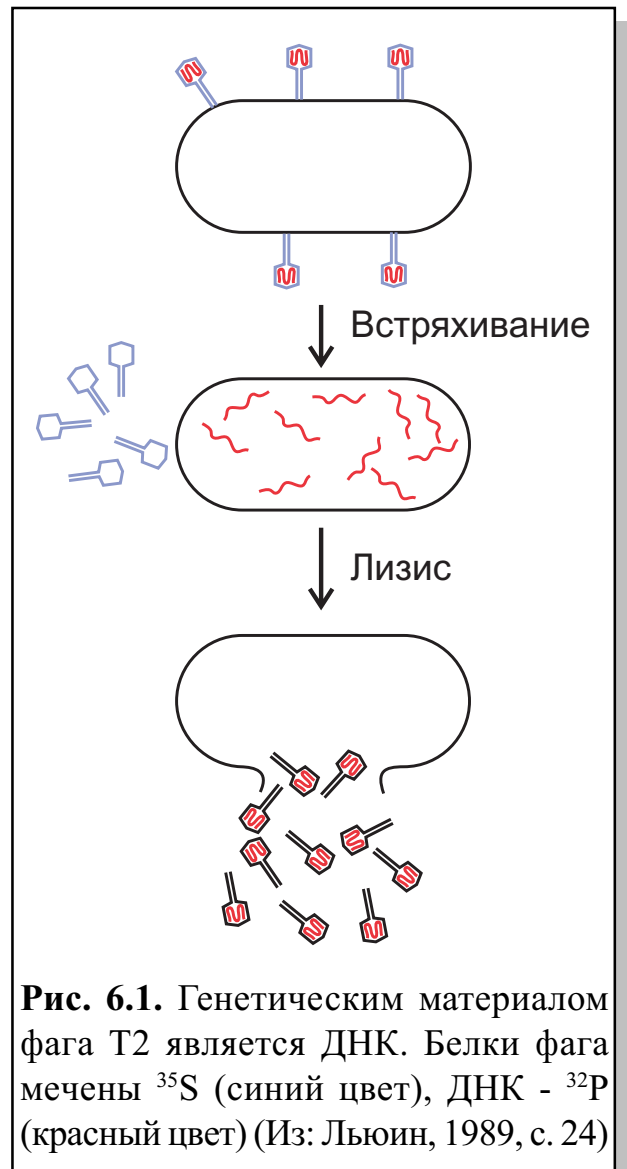
## 6. Структура и организация генома

### 6.1. Роль ДНК в наследственности

Как уже упоминалось в разделе 5.10., в результате изучения явления трансформации у бактерий было впервые показано, что именно ДНК может служить генетическим материалом. В 1952 году были получены новые доказательства этого в экспериментах другого типа. Фаг Т2 является вирусом, инфицирующим бактерию *E. coli*. Фаговые частицы абсорбируются на наружной поверхности клетки, их материал проникает внутрь и примерно через 20 минут бактерия лизируется, освобождая большое количество фаговых частиц - потомков. В 1952 году Альфред Херши и Марта Чейз инфицировали бактерий фагами Т2, которые были мечены радиоактивными соединениями: ДНК - с помощью  $^{32}\text{P}$ , белковая часть фага -  $^{35}\text{S}$  (Рис. 6.1.). После инфекции бактерий фагами, с помощью центрифугирования удалось выделить две фракции: пустые белковые оболочки фага и бактерий, инфицированных фаговой ДНК. Оказалось, что 80% метки  $^{35}\text{S}$  осталась в пустых фаговых оболочках, а 70% метки  $^{32}\text{P}$  - в инфицированных бактериях. Фаги-потомки получили только около 1% исходного белка, меченого  $^{35}\text{S}$ , однако они же обнаружили около 30% метки  $^{32}\text{P}$ .

Результаты этого эксперимента прямо показали, что ДНК родительских фагов проникает в бактерии и затем становится составляющей развившихся новых фаговых частиц.

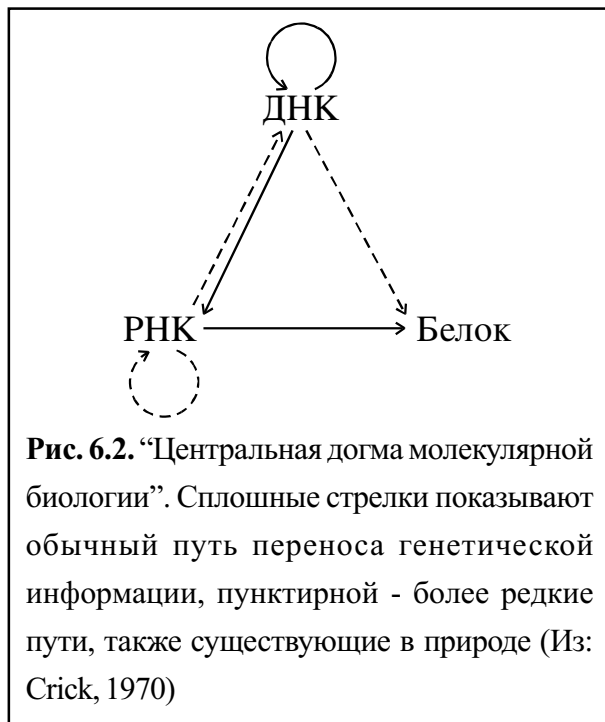
Современные представления о роли ДНК в передаче наследственной информации лучше всего отражает



**Рис. 6.1.** Генетическим материалом фага Т2 является ДНК. Белки фага мечены  $^{35}\text{S}$  (синий цвет), ДНК -  $^{32}\text{P}$  (красный цвет) (Из: Льюин, 1989, с. 24)

“Центральная догма молекулярной биологии”, сформулированная Ф. Криком в 1970 году (Рис. 6.2.).

Автор предложил разделить все виды переноса биологической информации в клетке на три группы: 1. Процессы, существование которых уже показано: ДНК  $\rightarrow$  ДНК, ДНК  $\rightarrow$  РНК, РНК  $\rightarrow$  белок, РНК  $\rightarrow$  РНК. 2. Процессы, которые не были экспериментально выявлены и с теоретической точки зрения не казались строго необходимыми: РНК  $\rightarrow$  ДНК, ДНК  $\rightarrow$  белок. 3. Невозможные переносы: белок  $\rightarrow$  белок, белок  $\rightarrow$  РНК, белок  $\rightarrow$  ДНК. Таким образом,



информация во всех случаях в клетке переносится однонаправленно по цепи: ДНК → РНК → белок. Белок не может служить матрицей для синтеза ДНК или РНК, поскольку у молекул белка нет свойства комплементарности отдельных частей молекулы, что бы позволяло использовать её как матрицу.

## 6.2. Структура ДНК

(Смотреть дополнительно: Чолаков, 1987, стр. 233).

Генетическая информация в молекуле ДНК записана в виде последовательности нуклеотидных остатков, которые содержат одно из четырех азотистых оснований: аденин (А), гуанин (G), цитозин (С) и тимин (Т) (Рис. 6.3.).

Модель ДНК в форме регулярной двойной спирали была предложена Дж. Д.

### Дополнение 6.2.

За исследования нуклеотидов и нуклеозидов Нобелевскую премию в 1957 году получил Александр Тодд (А. Todd) (см. детали в: Чолаков, 1987, стр. 231-248).

### Дополнение 6.1.

В 1969 году Альфред Д. Херши (А.Д. Hershey) получил Нобелевскую премию за открытие генетической структуры вирусов.

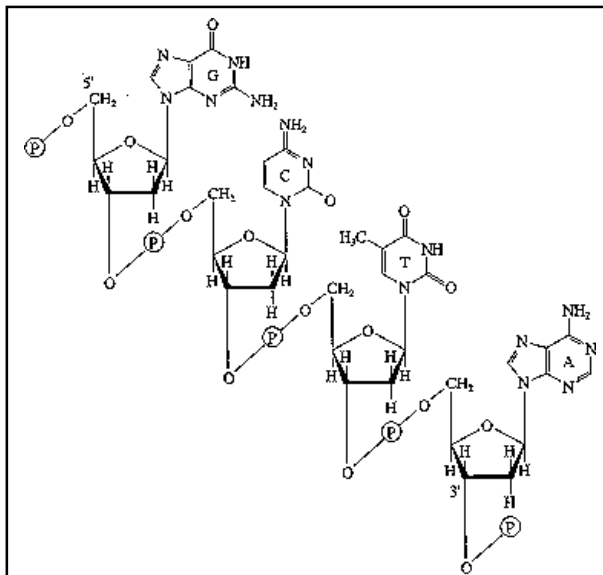
Уотсоном и Ф. Криком в 1953 году (Рис. 6.4.).

Две спиральные полинуклеотидные цепи закручены вправо вокруг общей оси. Пуриновые остатки заштрихованы. Против каждого из них находится остаток пиримидинового основания другой цепи. На схеме показаны размеры спирали, наличие большой и малой бороздок и антипараллельность двух цепей ДНК. Вначале предполагали, что на виток спирали приходится 10 пар нуклеотидов или 3.4 нм. Последующие измерения показали, что виток соответствует 10.5 пар нуклеотидов, или 3.6 нм.

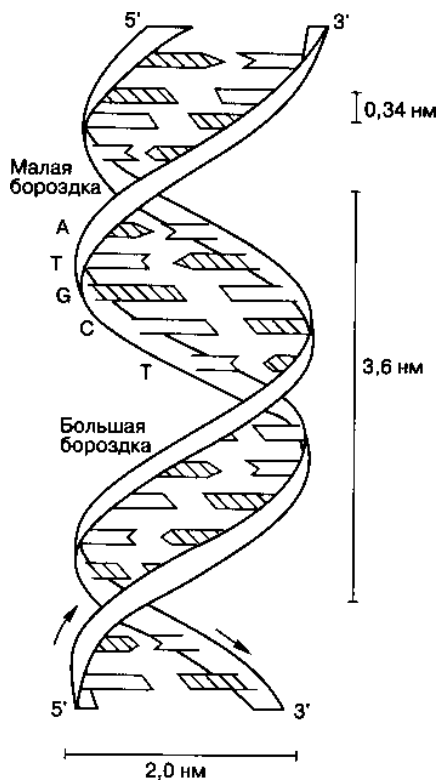
Каждая цепь содержит последовательность нуклеотидов, строго соответствующую последовательности другой цепи. Это соответствие достигается наличием водородных связей между направленными навстречу друг другу основаниями двух цепей: G и С или А и Т. Таким образом, цепи комплементарны. Поскольку цепи имеют противоположную направленность в расположении 5' и 3' свободных концов в молекуле пентозы, их называют антипараллельными (см. детали у Lewin, 1994, pp. 84-98).

## 6.3. Репликация ДНК

Уотсон и Крик уже во второй своей работе 1953 года предположили возможный механизм копирования наследственного материала. Легко представить, что цепи молекулы ДНК расходятся и каждая из них становится матрицей, на которой синтезируется новая комплементарная цепь



**Рис. 6.3.** Фрагмент одной цепи ДНК. Пуриновые основания аденин (А) и гуанин (G) и пиримидиновые основания тимин (Т) и цитозин (С) прикреплены к полимерному остову, состоящему из чередующихся остатков фосфата (Р) и сахара дезоксирибозы



**Рис. 6.4.** Модель структуры ДНК по Уотсону и Крику (Из: Фаворова, 1996, стр. 13)

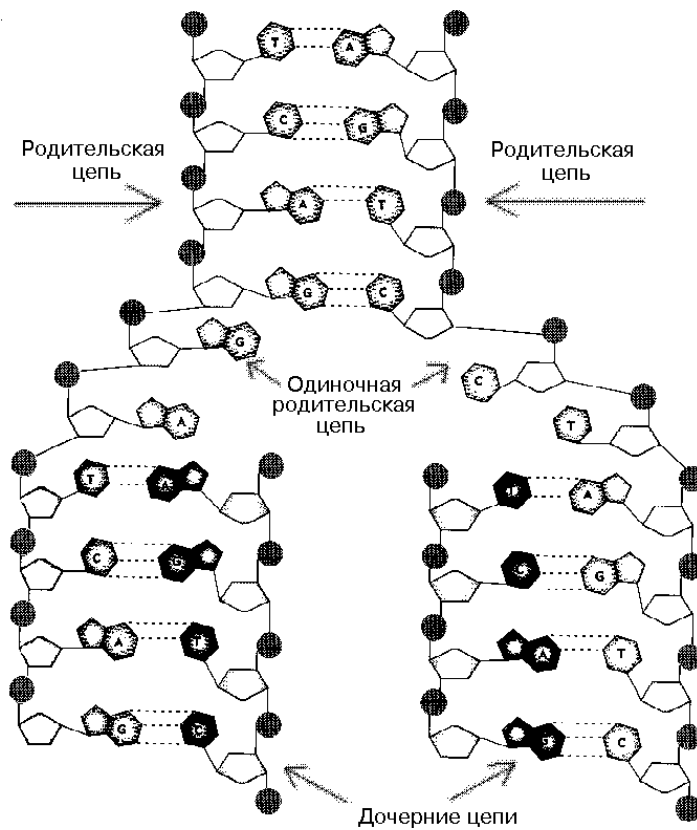
### Дополнение 6.3.

За открытие структуры нуклеиновых кислот Френсис Крик, Джеймс Уотсон и Морис Уилкинс (F.H.C. Crick, J.D. Watson, M.H.F. Wilkins) в 1962 году были награждены Нобелевской премией.

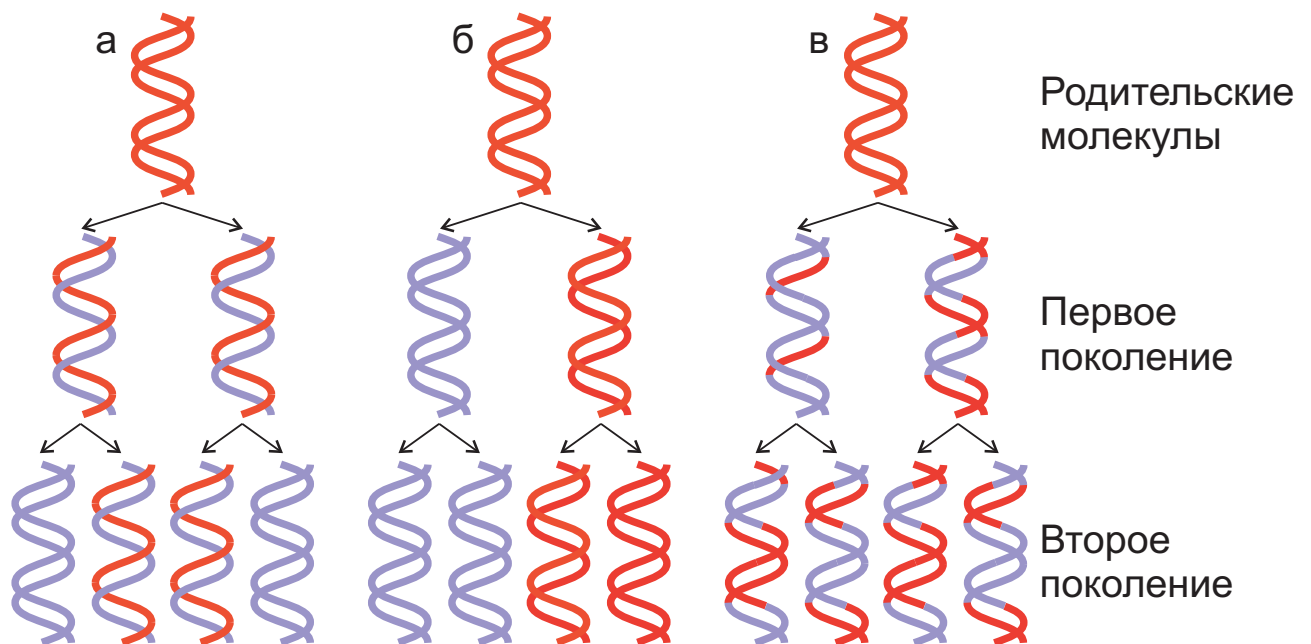
(Рис. 6.5.). В результате образуются две дочерние двуспиральные молекулы ДНК, неотличимые от родительской молекулы.

Каждая молекула ДНК состоит из одной цепи исходной родительской молекулы и одной вновь синтезированной цепи. Такой механизм копирования называется полуконсервативным. В то же время обсуждались две другие модели, одна из них “консервативная” другая “дисперсионная” (Рис. 6.6.). Доказали существование полуконсервативной модели М. Мезелсон и Ф. Сталь в 1958 году. Авторы выращивали бактерии *E. coli* несколько поколений на минимальной среде, в которой единственным источником азота был  $^{15}\text{NH}_4\text{Cl}$  (хлорид аммония). В этом соединении нормальный изотоп  $^{14}\text{N}$ , был заменен на  $^{15}\text{N}$ . В результате все клеточные компоненты бактерий, включая пурины и пиримидины в молекулах ДНК содержали более тяжелый азот  $^{15}\text{N}$ . Затем клетки переносили на среду, содержащую изотоп  $^{14}\text{N}$ . Через 1 или 2 поколения выделяли ДНК и центрифугировали в градиенте  $\text{CsCl}$ . Фракции, содержащие легкие или тяжелые цепи, а так же гибридные  $^{15}\text{N}/^{14}\text{N}$ , легко отделялись одна от другой (Рис. 6.7.).

В 1957 году Артур Корнберг обнаружил у бактерии *E. coli* фермент, катализирующий процесс полимеризации ДНК из нуклеотидов - ДНК-полимеразу. Открытие Корнберга показало, что в



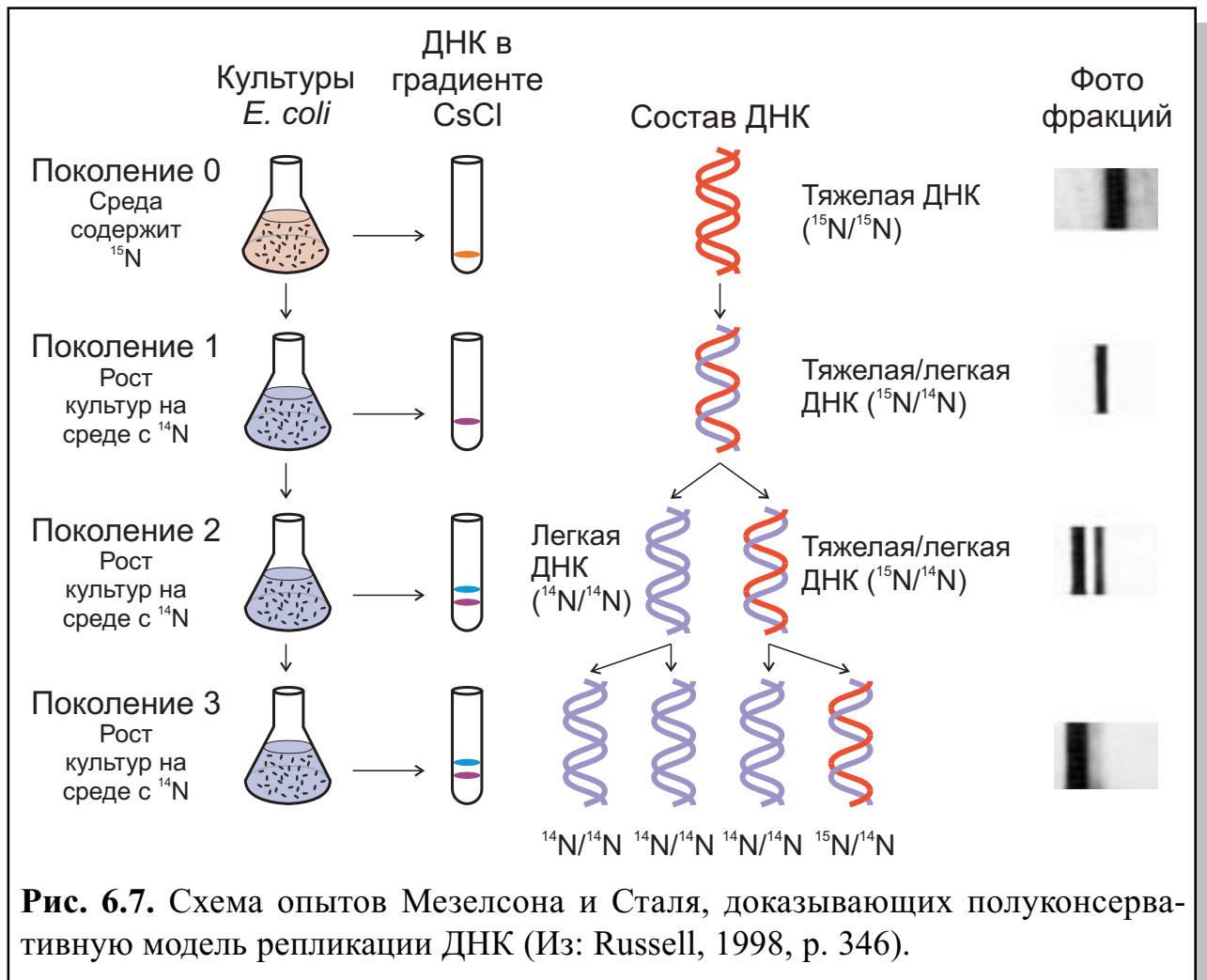
**Рис. 6.5.** Схема полуконсервативной репликации ДНК (Из: Lewin, 1994, стр. 95)



**Рис. 6.6.** Модели репликации ДНК:

- а - полуконсервативная,
- б - консервативная,
- в - дисперсионная.

Родительские цепи изображены в виде красных лент, вновь синтезированные показаны синим цветом. (Из: Russell, 1998, p.345).



**Рис. 6.7.** Схема опытов Мезелсона и Сталя, доказывающих полуконсервативную модель репликации ДНК (Из: Russell, 1998, p. 346).

основе удвоения молекул ДНК лежат обычные биохимические реакции. По современным представлениям в репликации ДНК у прокариот выделяют следующие этапы:

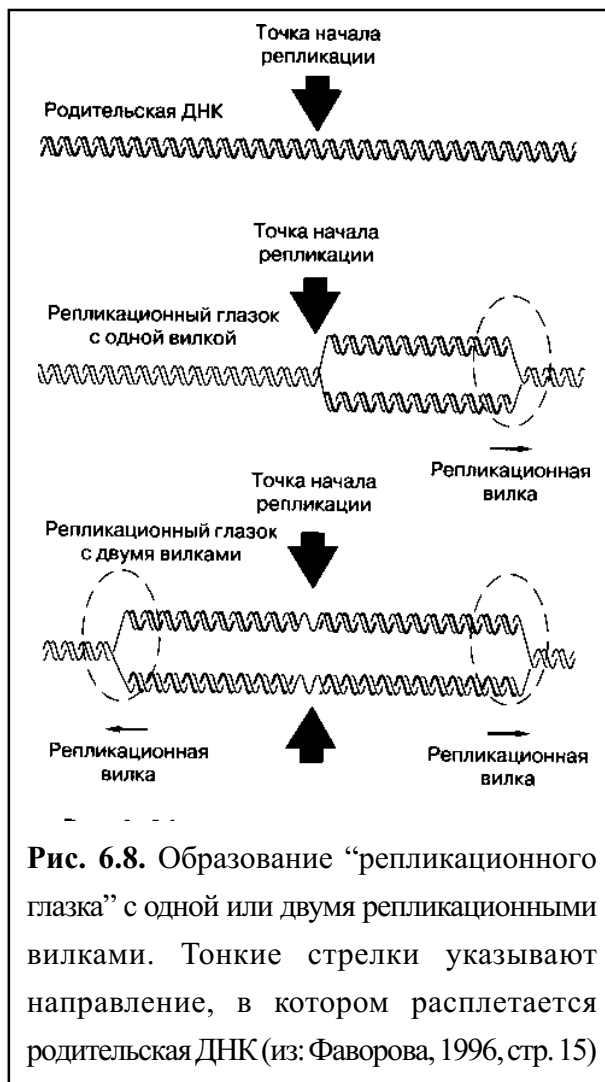
1. Релаксация суперспирализованной ДНК. Этот процесс катализируется ферментом топоизомеразой.
2. Денатурация двойной спирали ДНК. Поскольку синтез ДНК происходит на одноцепочечной матрице, ему должно предшествовать обязательное разделение двух цепей ДНК. Участок начала расхождения цепей называется репликационной вилкой из-за характерной Y-образной формы (Рис. 6.8.).

Именно в этой репликационной вилке ДНК-полимеразы синтезируют

дочерние молекулы ДНК. Участок ДНК, в котором репликация уже завершилась, выглядит как пузырек или “глазок” в нереплицированной ДНК. Репликационные глазки образуются в тех местах, где находятся специфические последовательности - точки начала репликации (origin of replication). Они состоят примерно из 300 нуклеотидов. С ориджинами связываются инициаторные белки репликации.

#### Дополнение 6.4.

В 1959 году Артуру Корнбергу (Arthur Kornberg) была присуждена Нобелевская премия за открытие механизма биосинтеза ДНК.



**Рис. 6.8.** Образование “репликационного глазка” с одной или двумя репликационными вилками. Тонкие стрелки указывают направление, в котором расплетается родительская ДНК (из: Фаворова, 1996, стр. 15)

Для того, чтобы цепи ДНК разъединились, функционирует особый фермент - ДНК-хеликаза, который связывается с инициаторными белками. Этот фермент движется по одиночной цепи ДНК и, встречая участок двойной спирали, он разрывает водородные связи между основаниями, разделяет цепи и продвигает репликационную вилку.

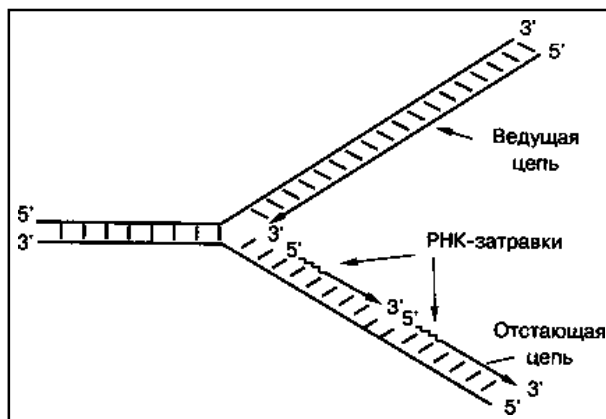
Субстратом для ДНК-полимеразы являются дезоксирибонуклеозидтрифосфаты (дНТФ), полимеризующиеся на одноцепочечной матрице. ДНК-полимеразы последовательно наращивают одну нить ДНК, шаг за шагом присоединяя к ней следующие звенья в направлении от 5' к 3' концу, причем выбор

очередного дНТФ диктуется матрицей. В клетках присутствуют несколько разных типов ДНКполимераз, выполняющих различные функции и имеющих разное строение: они могут быть построены из различного количества белковых цепей (субъединиц), от одной до десятков. Однако, все они работают на любых последовательностях нуклеотидов матрицы; задача этих ферментов - сделать точную копию каждой матрицы (см. Фаворова, 1996, стр. 13).

Генетический материал живых организмов имеет огромные размеры и реплицируется с высокой точностью. В среднем, в процессе воспроизведения генома млекопитающего, состоящего из ДНК длиной 3 млрд. пар нуклеотидов, возникает не более трех ошибок. При этом ДНК синтезируется чрезвычайно быстро (скорость ее полимеризации колеблется в пределах от 500 нуклеотидов в секунду у бактерий до 50 нуклеотидов в секунду у млекопитающих).

Высокая точность репликации, наряду с ее высокой скоростью, обеспечивается наличием специальных механизмов, осуществляющих коррекцию, т.е. устраняющих ошибки.

Суть механизма коррекции заключается в том, что ДНК-полимеразы дважды проверяют соответствие каждого нуклеотида матрице: один раз перед включением его в состав растущей цепи, второй раз перед тем, как включить следующий нуклеотид. Очередная фосфодиэфирная связь синтезируется лишь в том случае, если последний нуклеотид растущей цепи ДНК образовал правильную уотсон-криковскую пару с соответствующим нуклеотидом матрицы (см. Фаворова, 1996, стр. 13).



**Рис. 6.9.** Строение репликационной вилки. Направление синтеза ДНК совпадает с направлением расплетания двойной спирали лишь для одной из новосинтезированных цепей - ведущей. Вторая цепь - отстающая - синтезируется прерывисто, в виде коротких фрагментов Оказаки. В результате обе дочерние цепи растут в направлении от 5' к 3' (из: Фаворова, 1996, стр. 15)

ДНК-полимеразы не могут начинать синтез ДНК на матрице, а способны только добавлять новые дезоксирибонуклеотидные звенья к 3'-концу уже имеющейся полинуклеотидной цепи. Таковую заранее образованную цепь, к которой добавляются нуклеотиды, называют затравкой (или праймером). Короткую РНК-затравку синтезирует из рибонуклеозидтрифосфатов фермент, не обладающий корректирующей активностью и называемый ДНК-праймазой. Праймазная активность может принадлежать либо отдельному ферменту, либо одной из субъединиц ДНК-полимеразы.

Установлено, что дочерние цепи ДНК растут только в направлении 5' → 3', т.е. всегда удлиняется 3'-конец затравки, а матрица считывается ДНК-полимеразой в направлении 3' → 5'.

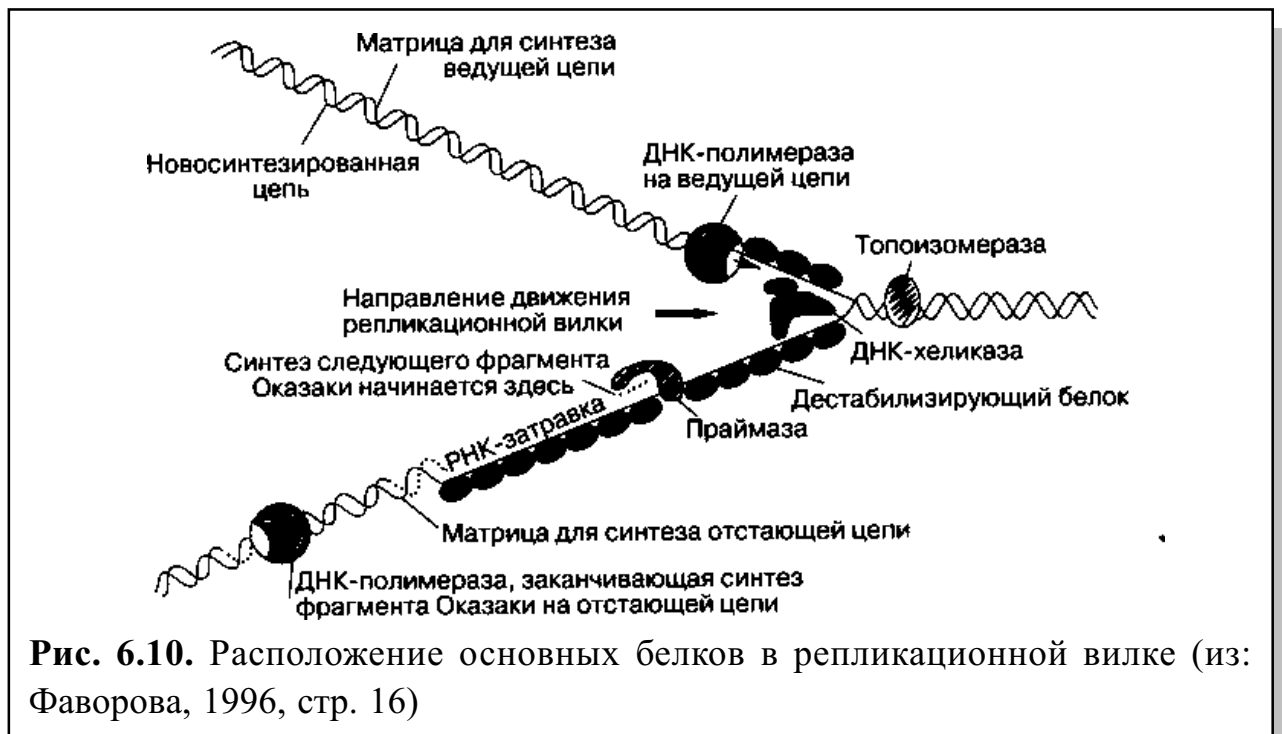
Синтез ДНК происходит непрерывно только на одной из матричных цепей. На второй цепи ДНК синтезируется сравнительно короткими фрагментами - от 100 до 1000 нуклеотидов, названными "фрагментами Оказаки" по имени открывшего их ученого - Тунеко Оказаки (Рис. 6.9.).

С одиночными цепями ДНК связываются специальные белки, дестабилизирующие спираль (SSB - single strand binding proteins). Они не позволяют им сомкнуться. Для того, чтобы репликационная вилка могла продвигаться вперед, вся еще не удвоенная часть ДНК должна была бы очень быстро вращаться. Белки ДНК-

#### Дополнение 6.5

Все живые организмы на Земле обычно делят на прокариот и эукариот (от греч. карион - ядро). Главной особенностью прокариот является отсутствие у них в отличие от эукариот (эу - по-гречески - истинный) полноценного клеточного ядра, покрытого оболочкой. Генетический материал прокариот расположен в нуклеоиде - примитивном эквиваленте ядра эукариот. Клетки прокариот имеют очень небольшие размеры - около 1 мкм. Объем эукариотических клеток в 800-1000 раз больше объема клеток прокариот. К прокариотам относятся бактерии и археи (или археобактерии), предки которых возникли около 4 млрд. лет назад. Эукариоты могут быть как одноклеточными так и многоклеточными. Они появились на Земле примерно через 500 млн. лет после прокариот (см. Кулаев, 1998; Алиханян и др., 1985, с.435,438).





**Рис. 6.10.** Расположение основных белков в репликационной вилке (из: Фаворова, 1996, стр. 16)

топоизомеразы вносят одноцепочечные или двуцепочечные разрывы, позволяющие цепям ДНК разделиться, а затем заделывают эти разрывы.

На Рис. 6.10. показано расположение цепей и молекул ферментов во время репликации.

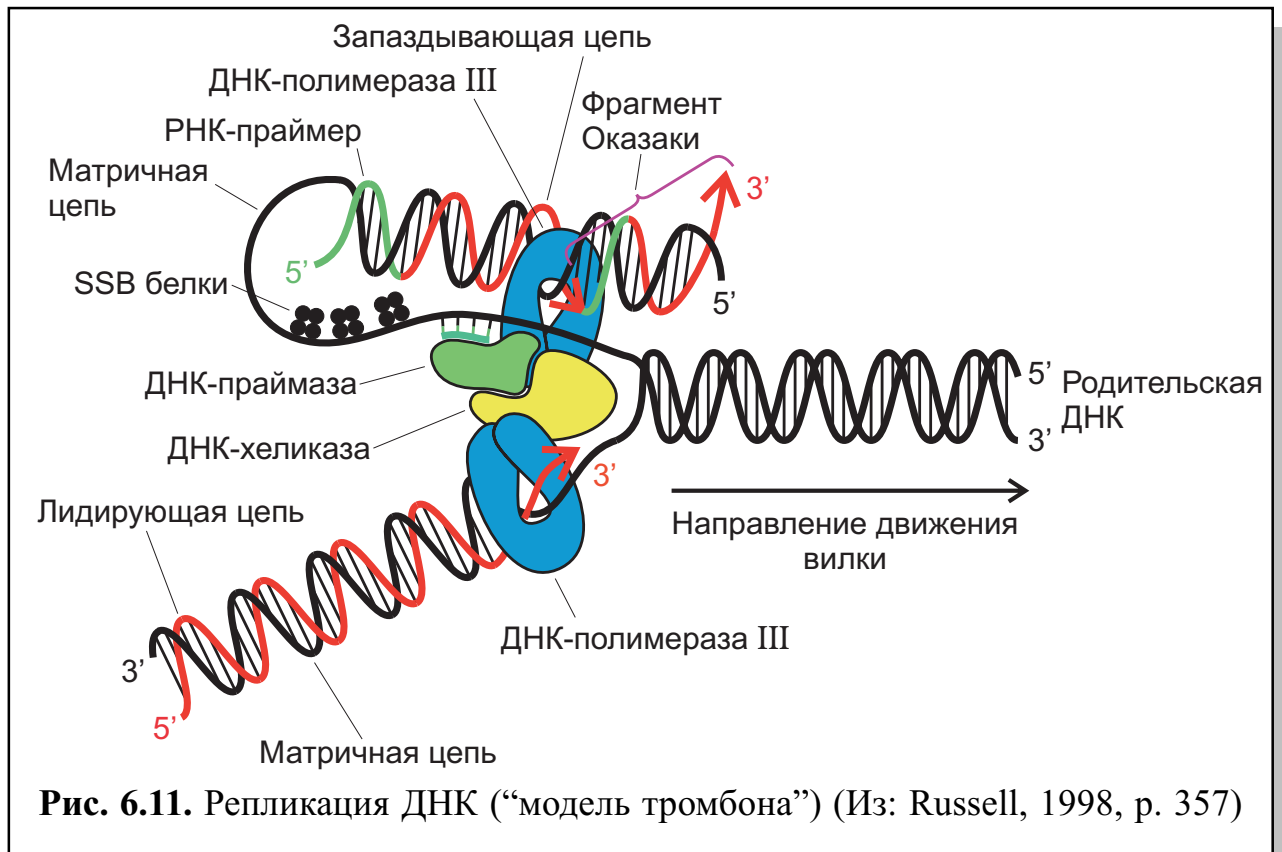
Спираль расплетается ДНК-хеликазой; этому процессу помогают ДНК-топоизомераза, раскручивающая цепь ДНК, и множество молекул дестабилизирующего белка (SSB), связывающихся с обеими одиночными цепями ДНК. В области вилки действуют две ДНК-полимеразы - на ведущей и отстающей цепи. На ведущей цепи ДНК-полимераза работает непрерывно, а на отстающей фермент время от времени прерывает и вновь возобновляет свою работу, используя короткие РНК-затравки, синтезируемые ДНК-праймазой. Молекула ДНК-праймазы непосредственно связана с ДНК-хеликазой, образуя структуру,

называемую праймосомой. Праймосома движется в направлении раскрывания репликационной вилки и по ходу движения синтезирует РНК-затравку для фрагментов Оказаки. В этом же направлении движется ДНК-полимераза ведущей цепи и ДНК-полимераза отстающей цепи. Для этого, как полагают, последняя накладывает цепь ДНК, которая служит ей матрицей, саму на себя, что и обеспечивает разворот ДНК-полимеразы отстающей цепи на 180 градусов (Рис. 6.11.). Согласованное движение двух молекул ДНК-полимераз обеспечивает координированную репликацию обеих нитей.

[Лигаза].

Всего в репликационной вилке одновременно работает около двадцати разных белков (на Рис. 6.10. и 6.11. показана только часть из них).

Процесс репликации хромосомы бактерий начинается в точке начала репликации и продолжается до тех пор, пока не удвоится вся ДНК хромосомы.



Молекулярно-биологические процессы, происходящие во время репликации ДНК, в основном похожи у эукариот и прокариот. Если бактериальная хромосома представляет собой единицу репликации - репликон, то репликация ДНК эукариотической хромосомы осуществляется посредством разделения ее на множество отдельных репликонов. Полагают, что у эукариот гомологами ориджинов начала репликации являются автономно реплицирующиеся последовательности или ARS (autonomously replicating sequences). Сначала у дрожжей *Saccharomyces cerevisiae* были выделены особые последовательности, которые, будучи включенными в экстрахромосомальную ДНК, обеспечивали репродукцию этих ДНК в дрожжевой клетке. Позднее такие последовательности были выделены у многих других организмов.

По эукариотической хромосоме в каждый момент времени может двигаться независимо друг от друга множество репликационных вилок. Остановка продвижения вилки происходит только при столкновении с другой вилкой, движущейся во встречном направлении, или по достижении конца хромосомы. В результате вся ДНК хромосомы в короткий срок оказывается реплицированной (все из Фаворовой, 1996).

Насколько велики репликоны и как много их в геноме? Трудность определения размеров и числа репликонов заключается в том, что трудно выделить индивидуальные “глазки” репликации. Всегда остается возможность, что наблюдаемый глазок является результатом слияния двух соседних репликонов. Чтобы обойти это препятствие, подбирают стадию

**Табл. 6.1.** Параметры репликации ДНК в геномах эу- и прокариот (Из: Lewin, 1994, p. 536)

Организм	Число репликонов	Средняя длина репликона (в т.п.н.)	Скорость движения вилки репликации (в т.п.н./мин.)
Бактерии ( <i>Escherichia coli</i> )	1	4200	50
Дрожжи ( <i>Saccharomyces cerevisiae</i> )	500	40	3.6
Насекомые ( <i>Drosophila melanogaster</i> )	3500	40	2.6
Лягушка ( <i>Xenopus laevis</i> )	15000	200	0.5
Мышь ( <i>Mus musculus</i> )	25000	150	2.2
Растения ( <i>Vicia faba</i> )	35000	300	-

репликации, когда число “глазков” максимально и они еще не начали сливаться. Затем в участке ДНК, содержащем несколько глазков, измеряют расстояние между начальными точками начала репликации (т.е. между средними точками смежных репликонов). Скорость движения вилки репликации определяют по максимальной длине радиоактивно меченого следа реплицирующейся ДНК в единицу времени.

Сведения о параметрах репликации представлены в табл. 6.1.

Как свидетельствуют данные табл. 6.1., репликоны у эукариот имеют существенно меньшие размеры, чем у прокариот (хотя в пределах генома одного вида они могут варьировать в размерах в 10 раз). Скорость репликации существенно ниже у эукариот. (См. более подробно в книге Lewin, 1994, pp. 527-569 и 571-603!).

### **Литература к разделам 6.1 - 6.3.**

- Альбертс Б., Брей Д., Льюис Дж., Рэфф М., Робертс К., Уотсон Дж. Молекулярная биология клетки. Москва, Мир, Том. 2, стр. 287-301, 1994.
- Докинз Р. Эгоистичный ген. Москва, Мир, стр. 1-316, 1993.
- Кулаев И.С. Происхождение эукариотических клеток. Соросовский образовательный журнал, N5 (30), стр.17-22, 1998 .
- Льюин Б. Гены. Москва, Мир, стр.23-36, 1987.
- Фаворова О.О. Сохранение ДНК в ряду поколений: репликация ДНК. Соросовский образовательный журнал, N4, стр. 11-17, 1996.
- Чолаков В. Нобелевские премии. Ученые и открытия. Москва, Мир, 1-363, 1987.

- Crick F. Central dogma of molecular biology. Nature 227, N5258, pp. 561-563, 1970.
- Russell P.J. Genetics. Fifth edition. Menlo Park, California. Addison Wesley Longman, inc., pp.344-366, 1998.
- Watson, J.D., Crick, F.H.C. A structure for deoxyribose nucleic acid. Nature 171, N4356, pp. 737-738, 1953.
- Watson, J.D., Crick, F.H.C. General implications of the structure of deoxyribonucleic acid. Nature 171, pp. 964-967, 1953.
- Wilkins M.H.F., Stikes A.R., Wilson H.R. Molecular structure of deoxypentose nucleic acids. Nature 171, pp. 738-740, 1953.

#### 6.4. Генетический код

В любом данном участке ДНК только одна из двух нитей ДНК кодирует аминокислоты, поэтому код - это последовательность нуклеотидов, а не пар нуклеотидов. Генетический код имеет следующие свойства:

1. Генетический код читается группами по три нуклеотида, т.е. код триплетный. Каждый триплет кодирует аминокислоту, каждый триплет называется кодоном (Табл. 6.2).

2. Основные закономерности организации генетического кода были открыты с помощью генетического анализа района II фага Т6. В 1961 г. Ф. Крик и его коллеги показали, что код должен читаться неперекрывающимися триплетами с фиксированной стартовой точки.

а) неперекрывание подразумевает, что каждый кодон состоит из трех нуклеотидов и каждый последующий кодон представлен следующими тремя нуклеотидами.

б) Фиксированная стартовая точка означает, что считывание начинается на одном конце и завершается на другом; различные части кодирующей последовательности не могут считываться независимо друг от друга. Началом трансляции любого гена является кодон AUG. В конце гена обязательно стоят кодоны UAA, UAG или UGA, которые не кодируют аминокислот и являются сигналами на окончание синтеза белка - стоп-кодона. Для повышения надежности процесса терминации стоп-кодона обычно дублируются. Первым при этом, как правило, выступает кодон UAA (основной терминирующий триплет), а вслед за ним на очень близком расстоянии в той же рамке считывания следует один из запасных терминирующих триплетов - UAG или UGA.

в) Если генетический код считывается неперекрывающимися триплетами, есть только три возможности транслирования нуклеотидной последовательности в аминокислотную, в зависимости от стартовой точки.

```
ACG ACG ACG ACG ACG ACG
CGA CGA CGA CGA CGA CGA
GAC GAC GAC GAC GAC GAC
```

Эти три возможности называют рамками считывания.

Мутация, в результате которой инсертируется или делетируется один нуклеотид и изменяется рамка считывания всей последующей последовательности, называется сдвигом рамки. Поскольку последовательность новой рамки считывания полностью отлична от первоначальной, вся аминокислотная последовательность будет измененной ниже мутации. Функция такого белка

**Табл. 6.2.** Соответствие кодонов генетического кода аминокислотам белка

Первая буква в кодоне (5')	Вторая буква в кодоне				Третья буква в кодоне (3')
	U	C	A	G	
U	Фен (F)	Сер (S)	Тир (Y)	Цис (C)	U
	Фен (F)	Сер (S)	Тир (Y)	Цис (C)	C
	Лей (L)	Сер (S)	Stop	Stop	A
	Лей (L)	Сер (S)	Stop	Трп (W)	G
C	Лей (L)	Про (P)	Гис (H)	Арг (R)	U
	Лей (L)	Про (P)	Гис (H)	Арг (R)	C
	Лей (L)	Про (P)	Глн (Q)	Арг (R)	A
	Лей (L)	Про (P)	Глн (Q)	Арг (R)	G
A	Иле (I)	Тре (T)	Асн (N)	Сер (S)	U
	Иле (I)	Тре (T)	Асн (N)	Сер (S)	C
	Иле (I)	Тре (T)	Лиз (K)	Арг (R)	A
	Мет (M)	Тре (T)	Лиз (K)	Арг (R)	G
G	Вал (V)	Ала (A)	Асп (D)	Гли (G)	U
	Вал (V)	Ала (A)	Асп (D)	Гли (G)	C
	Вал (V)	Ала (A)	Глу (E)	Гли (G)	A
	Вал (V)	Ала (A)	Глу (E)	Гли (G)	G

**Примечание**

U - урацил, C - цитозин, A - Аденин, G - гуанин, F, L, I и т.д. - однобуквенные сокращения названий аминокислот, Фен, Сер и т.д. - трехбуквенные сокращения. Кодоны UAA, UAG, UGA - не кодируют аминокислот. Эти кодоны являются сигналами терминации трансляции - стоп-кодонами

**Дополнение 6.6.**

Проблема кодирования в молекулярной биологии была впервые поставлена Г. Гамовым еще в начале 50х годов, т.е. задолго до открытия самой мРНК. Размышляя над тем, как линейная последовательность четырех различных нуклеотидов может определить последовательность двадцати разных аминокислот в белке, Гамов предположил, что генетический код является триплетным. Он же поставил вопросы и о других свойствах генетического кода: перекрываемости, запятых между кодонами, вырожденности.

полностью утрачена. Пример:

GCU GCU GCU GCU GCU GCU GCU  
Ala Ala Ala Ala Ala Ala Ala

Вставка А

↓  
GCU GCU AGC UGC UGC UGC UGC  
Ala Ala Ser Cys Cys Cys Cys

Делеция G

↑  
GCU GCU AGC UGC UCU GCU GCU  
Ala Ala Ser Cys Ser Ala Ala

(См. детали в кн. Гершензона, 1983, стр. 286, в кн. Lewin, 1994, pp. 98-100 и Инге-Вечтомов, 1996).

3. Генетический код является вырожденным, в том смысле, что одной аминокислоте может соответствовать несколько кодонов (Табл. 6.3).

Однако, кодоны используются не с одинаковой частотой. Например, у дрожофилы, в результате параллельного изучения последовательностей кодонов в генах и аминокислот, кодированных ими, при этом суммарная длина генов соответствовала 269 т.п.н., было показано, что кодоны используются с разными частотами (Табл. 6.4).

#### Дополнение 6.7.

В 1968 году за открытие и интерпретацию генетического кода и его функции в белковом синтезе Нобелевская премия была присуждена Р. Холли, Х. Г. Хоране и М.В. Ниренбергу (Robert W. Holley, H. Gobind Khorana, Marshall W. Nirenberg).

Табл. 6.3. Аминокислоты и соответствующие им кодоны

A	Ala Аланин	GCA GCC GCG GCU
C	Cys Цистеин	UGC UGU
D	Asp Аспарагиновая кислота	GAC GAU
E	Glu Глутаминовая кислота	GAA GAG
F	Phe Фенилаланин	UUC UUU
G	Gly Глицин	GGA GGC GGG GGU
H	His Гистидин	CAC CAU
I	Ile Изолейцин	AUA AUC AUU
K	Lys Лизин	AAA AAG
L	Leu Лейцин	UUA UUG CUA CUC CUG CUU
M	Met Метионин	AUG
N	Asn Аспарагин	AAC AAU
P	Pro Пролин	CCA CCC CCG CCU
Q	Gln Глутамин	CAA CAG
R	Arg Аргинин	AGA AGG CGA CGC CGG CGU
S	Ser Серин	AGC AGU UCA UCC UCG UCU
T	Thr Треонин	ACA ACC ACG ACU
V	Val Валин	GUA GUC GUG GUU
W	Trp Триптофан	UGG
Y	Tyr Тирозин	UAC UAU

**Табл. 6.4.** Частоты использования разных кодонов, кодирующих лейцин и аланин у дрозофилы (часть таблицы из книги: Ashburner, 1989, p. 75.)

Аминокислота	Кодоны	Число случаев
Лейцин	СТТ	610
	СТС	1096
	СТА	538
	СТГ	3425
Аланин	GCC	3534
	GCA	926
	GCT	1397
	GCG	1159

4. Генетический код универсален, в том смысле что определённому кодону соответствует определённая аминокислота. Например, AUG-кодон кодирует метионин у любого организма. Однако, по мере расширения круга объектов молекулярной генетики стали накапливаться исключения, сделавшие код "квазиуниверсальным" (см. Инге-Вечтомов, 1996, стр. 7). Касается это прежде всего митохондриальных геномов (Табл. 6.5).

#### **Литература к разделу 6.4.**

Гершензон С.М. Основы современной генетики. Киев, Наукова думка, 1-558, 1983.

Инге-Вечтомов С.Г. Трансляция как способ существования живых систем, или в чём смысл "бессмысленных" кодонов". Сорос. образ. журн. N12, 2-10, 1996.

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

Crick F.C.H., Barnett L., Brenner S., Watts-Tobin R.J. General nature of the genetic code for proteins. *Nature* 192, N4809, 1227-1232, 1961

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

#### **6.5. Структура генома эукариот**

Главной особенностью генетического материала эукариот в сравнении с прокариотами является наличие избыточной ДНК.

Если средний размер гена бактерий - 1500 п.н., а длина кольцевой молекулы ДНК хромосомы *E. coli* и *Bacillus subtilis* составляет около 1,1 мкм, то в такой хромосоме могут разместиться около 3000 генов. Примерно такое число генов было экспериментально определено у бактерий по числу типов мРНК. Если это число умножить на средний размер гена, то получится, что около 95% генома бактерий состоит из кодирующих (генных) последовательностей.

Иная картина наблюдается у эукариотических организмов. Например, у человека насчитывают приблизительно  $5 \times 10^4$  генов (Алиханян и др., 1985). В то же время размер генома человека -  $3 \times 10^9$  п.н. Это означает, что кодирующая часть его генома составляет всего 15-20% от всей ДНК. Существуют виды, геном которых в десятки раз больше генома человека, например некоторые рыбы, хвостатые амфибии, лилейные растения. Избыточная ДНК характерна для всех эукариот. В связи с этим необходимо

**Табл. 6.5.** Отклонения от универсального генетического кода (Из: Lewin, 1994, p. 217-219)

Геном	Организм	Кодон	Универсальное значение	Необычное значение
Митохондри	Позвоночные, дрозофила, дрожжи, плесени, трипаносомы	UGA	Stop	Trp
	Сахаромицеты	CUU CUC CUA CUG	Leu	Thr
		CGG	Arg	Trp
	Позвоночные, дрозофила, сахаромицеты	AUA	Ile	Met
	Морская звезда	AAA	Lys	Asn
	Позвоночные	AGA AGG	Arg	Stop
	Морская звезда, дрозофила	AGA AGA*	Arg	Ser
	Аскарида, нематода	UUG	Leu	Start
		AUU	Ile	Start
	Нематода	AUA	Ile	Start
Млекопитающие	AUU AUC AUA	Ile	Start	
Ядро	Микоплазма	UGA	Stop	Trp
	Цилиаты	UAA UAG	Stop	Gln
	Гриб кандиды цилиндрика	CUG	Leu	Ser

Примечание. А\* - модифицированный аденин

разграничить понятия геном и генотип. Генотип - это совокупность генов, имеющих фенотипическое проявление. Геном - это количество ДНК в гаплоидном наборе хромосом данного вида (Из: Алиханян и др., 1985, стр. 104).

В конце 60-х годов американские ученые Р. Бриттен и Э. Дэвидсон открыли фундаментальную особенность молекулярной структуры генома эукариот - последовательности нуклеотидов разной степени повторяемости. Это открытие было сделано с помощью молекулярно-



биологического метода изучения кинетики ренатурации денатурированной ДНК. Различают следующие фракции в геноме эукариот:

1. Уникальные, т.е. представленные в одном экземпляре.
2. Промежуточные (или среднечастотные) повторы. Это последовательности, повторяющиеся десятки и сотни раз.
3. Высокочастотные повторы, число которых в геноме достигает  $10^6$  копий.

Уникальные последовательности чаще всего представлены генами. Число генов у эукариот определяют одним из двух способов. Первый способ прямой, т.е. экспериментально определяют последовательности нуклеотидов во всем геноме, число последовательностей, содержащих длинные рамки считывания или кДНК-клоны (см. ниже). Понятно, что такой анализ можно провести пока на очень ограниченном числе видов, и эти исследования исключительно интенсивно проводятся в настоящее время (программы “Геном человека”, “Геном дрозофилы”, “Геном дрожжей” - всего около 30 таких программ). Выполнение подобных программ требует огромных финансовых затрат и скоординированных усилий большого числа ученых из всех развитых стран мира. Например, для того, чтобы расшифровать последовательность нуклеотидов в геноме дрожжей *Saccharomyces cerevisiae* потребовались усилия более чем 600 ученых из 96 лабораторий мира (Из: Clayton et al., 1997). Данные по анализу генома дрожжей представлены в табл. 6.6.

Другой подход используют уже 2-3 десятка лет. С помощью довольно простых процедур *рассчитывают*

возможное число генов у того или иного вида. Сначала определяют общий размер генома этого вида, затем, зная средний размер гена у этого вида и добавив к этому значению половину размера собственно гена (межгенный промежуток), делят значение размера генома на значение размера гена + межгенного промежутка и получают число генов. Все эти оценки в какой-то степени субъективны, поэтому варьируют в довольно широких пределах (см. Табл. 6.6.). Варьирование связано с тем, что разные авторы берут в расчет несколько различающиеся значения длин генов, межгенных промежутков, да и общий объем генома.

Повторы образуют семейства - совокупность последовательностей, полностью или по большей части гомологичных друг другу.

Нередко из-за существенных различий в нуклеотидном составе высокочастотных повторов и остальной ДНК первые образуют при центрифугировании в градиенте плотности хлористого цезия так называемые сателлитные пики, которые имеют большую или меньшую плавучую плотность, чем остальная ДНК. (см разд. 9.4.9.). Эта фракция генома представлена небольшим (10-15) числом семейств коротких (5-12 п.н.) повторов, образующих протяженные блоки. У огромного большинства видов эта фракция занимает не более 10% генома. Близкие виды, например мышь и крыса имеют совершенно различные высокочастотные последовательности: у крысы их нуклеотидный состав не отличается от основной ДНК, тогда как геном мыши содержит четкий АТ-богатый сателлит. Это означает, что высокочастотные повторы

**Табл. 6.6.** Данные о числе генов у разных модельных объектов, полученные на основе расчетов или в результате определения последовательности нуклеотидов

Таксон	Вид	Число генов по данным авторов		
		Lewin, 1994	Miklos, Rubin, 1996	Другие авторы
Prokariota	<i>Mycoplasma genitalium</i>		473	
	<i>Haemophilus influenzae</i>		1760	
	<i>Bacillus subtilis</i>		3700	4200
	<i>Escherichia coli</i>	2350	4100	4290
	<i>Mycococcus xanthus</i>		8000	
Fungi	<i>Saccharomyces cerevisiae</i>	5200	5800	6200 <sup>4</sup>
Protoctista	<i>Cyanidioschyzon merolal</i>		5000	
	<i>Oxytricha similis</i>		12000	
Arthropoda	<i>Drosophila melanogaster</i>	8000	12000	8000-20000 <sup>2</sup> , 20000-30000 <sup>3</sup>
Nematoda	<i>Caenorhabditis elegans</i>		14000	~17000
Mollusca	<i>Loligo peali</i>		>35000	
Chordata	<i>Fugu rubripes</i>		70000	
	<i>Mus musculus</i>	125000	70000	
	<i>Homo sapiens</i>		70000	50000 <sup>1</sup>
Plantae	<i>Nicotiana tabacum</i>		43000	
	<i>Arabidopsis thaliana</i>		16000-33000	

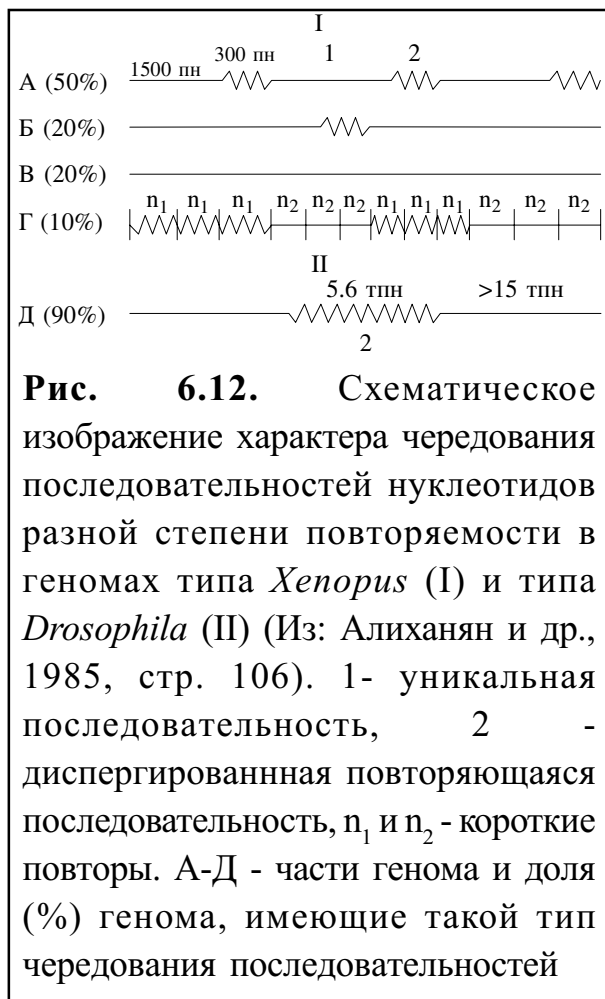
Другие авторы: <sup>1</sup>Алиханян и др., 1985; <sup>2</sup>Nusslein-Volhard, 1994; <sup>3</sup>Zhimulev, Belyaeva, 1974; <sup>2</sup> и <sup>3</sup> из книги: Zhimulev, 1998; <sup>4</sup>Clayton et al., 1997

способны к быстрым изменениям в ходе видообразования.

Остальные 90% генома эукариот построены по принципу чередования (интерсперсии) уникальных и повторяющихся последовательностей. Условно выделяют два основных типа интерсперсии, получивших названия по тем видам, у которых они впервые были описаны: интерсперсия типа “ксенопус” (обнаружена у *Xenopus laevis*) и типа “дрозофила” (впервые описана у *D. melanogaster*). Примерно в 50% генома *Xenopus laevis* (Рис. 6.12.) уникальные последовательности из 800-1200 п.н. чередуются с повторяющимися, средний

размер которых 300 п.н. В остальной части геномов типа “ксенопус” расстояния между соседними повторами значительно превышают 1-2 т.п.н.

Структура генома типа “ксенопус” широко распространена, особенно среди животных. Млекопитающие и человек также относятся к этому типу организации генома. Особенность генома человека и других приматов составляют интерсперсные высокочастотные повторы длиной около 300 п.н. У человека эти повторы содержат сайт, разрезаемый ферментом рестрикции *Alu* I. Число *Alu*-подобных повторов достигает  $5 \times 10^5$  -  $10^6$  копий.



У дрозофилы параметры интерсперсии резко отличаются от видов с типом генома “ксенопус”. Повторяющиеся последовательности длиной 5600 п.н. чередуются с уникальными, длина которых не менее 13000 п.н. (см. Рис. 6.12.). Интересно отметить, что у *Musca domestica*, - вида, близкого *D. melanogaster*, геном устроен по типу “ксенопус”. Этот факт прямо указывает на то, что в ходе эволюции возможны очень быстрые преобразования характера чередования последовательностей.

Птицы по параметрам интерсперсии занимают промежуточное положение между типом “ксенопус” и типом “дрозофилы”. Многие виды не могут быть отнесены ни к какому типу (Из: Алиханян и др., 1985, стр. 104-107).

## Литература к разделу 6.5.

- Алиханян С.И., Акифьев А.П., Чернин Л.С. Общая генетика. Москва, Высшая школа, 1-446, 1985.
- Clayton R.A., White O., Ketchum K.A., Venter J.C. The first genome from third domain of life. *Nature*, V. 387, P. 459-462, 1997.
- Lewin B. *Genes*. V. Oxford, New York, Tokyo; Oxford University Press, 1-1272, 1994.
- Miklos G.L.G., Rubin G.M. The role of the genome project in determining gene function: insights from model organisms. *Cell*, V. 86, P. 521-529, 1996.
- Zhimulev I.F. Genetic organization of polytene chromosomes. *Advances in Genetics*, V. , 1- , 1998.

## 6.6. Мобильные элементы генома

### 6.6.1. Мобильные элементы геномов растений

В начале 1940-х годов американская исследовательница Барбара МакКлинток (Рис. 6.13.) открыла существование гена или локуса, который вызывал повышенные частоты хромосомных перестроек у кукурузы. Среди потомков от скрещивания, в котором оба родителя несли такие перестройки, появлялись нестабильные мутации с неожиданно высокой частотой. В 1948 году она опубликовала результаты исследований этого локуса, вызывающего разрывы хромосом, сделав вывод, что он был совершенно необычным поскольку мог перемещаться из одного участка хромосомы в другой. Б. МакКлинток назвала феномен перемещения



**Рис. 6.13.** Барбара МакКлинток  
1902-1992

#### **Дополнение 6.8.**

В 1983 году за открытие мобильных генетических элементов Нобелевская премия была присуждена Барбаре МакКлинток (В. McClintock).

транспозицией (см. Federoff, 1994), а сами локусы - “контролирующими элементами” (КЭ). Эти элементы характеризуются следующими свойствами:

1. Они могут перемещаться из одного сайта в другой.
2. Их встраивание в данный район влияет на активность генов, расположенных рядом.
3. Утрата КЭ в данном локусе превращает прежде мутабельный локус в стабильный.
4. В сайтах, в которых присутствуют КЭ, могут возникать делеции, транслокации, транспозиции,

инверсии, а также разрывы хромосом.

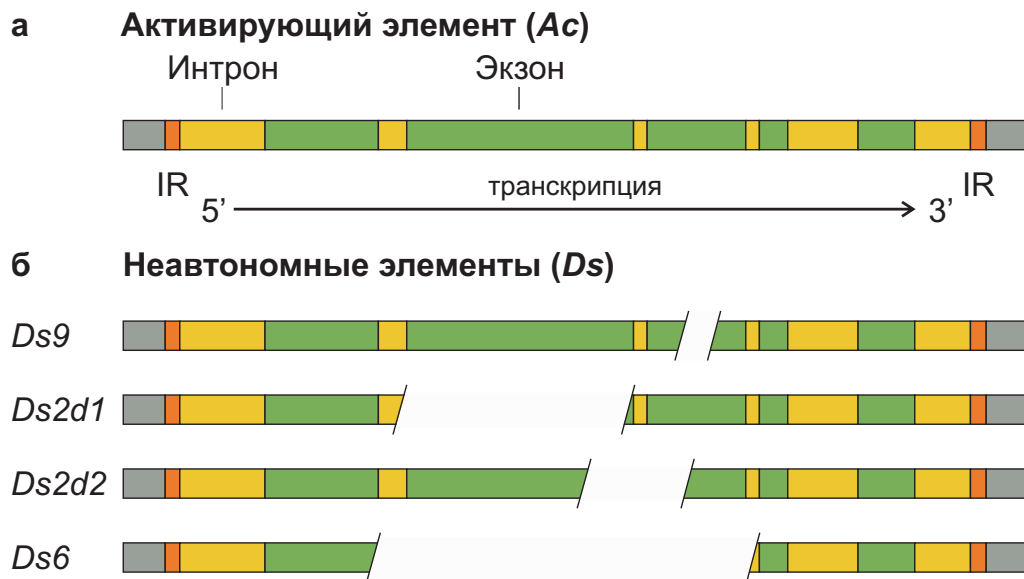
Геном кукурузы содержит несколько семейств КЭ. Члены каждого семейства могут быть подразделены на два класса:

1) Автономные элементы, которые способны вырезаться и транспозироваться. Их внедрение ведет к появлению нестабильных аллелей. Неавтономные элементы теряют свою стабильность только в том случае, если в какой-то области генома присутствует автономный член того же семейства.

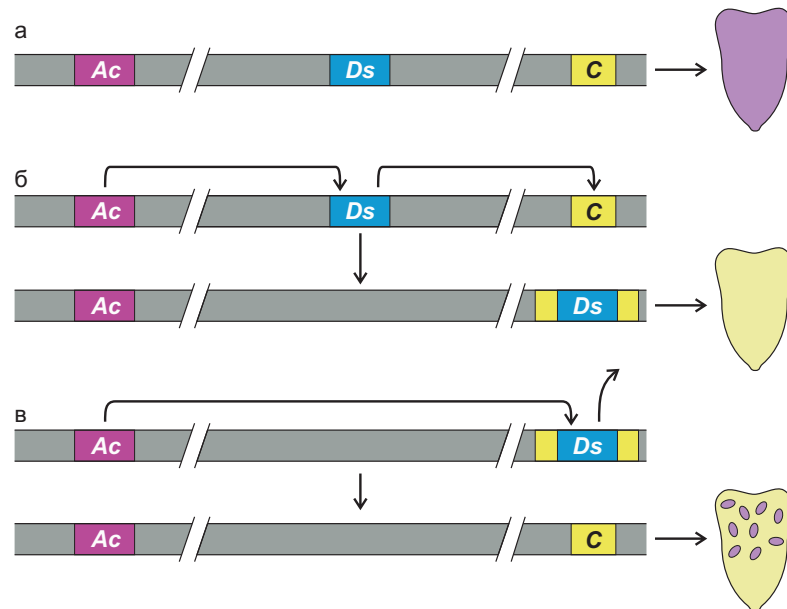
2) Неавтономные элементы могут быть активированы только определенными автономными элементами (членами того же семейства). У кукурузы изучены лучше всего *Ac-Ds*, *Spm* (супрессор-мутатор) и *Dt* семейства (см. более подробно: Льюин, 1987, стр. 841).

При встраивании автономного элемента в ген, последний мутирует, однако мутация эта будет нестабильной, поскольку элемент может выйти из данного гена и переместиться в другой участок генома. Поскольку частота транспозиции значительно выше, чем частота обратного спонтанного мутирования, то аллель, индуцированный автономным элементом, называется мутабельным.

Система *Ac-Ds* у кукурузы была изучена в деталях. *Ac*-элемент имеет длину 4563 п.н. с инвертированными повторами (IR) на концах. Он содержит единственную единицу транскрипции (5 экзонов и 6 нитронов), кодирующего фермент транспозазу. Элементы *Ds* происходят из *Ac* в результате делеций внутренних участков *Ac*-элемента (Рис. 6.14.).



**Рис. 6.14.** Структура автономного *Ac*-элемента (а) и неавтономных *Ds*-элементов (б) у кукурузы (Из: Russell, 1998, p. 669)



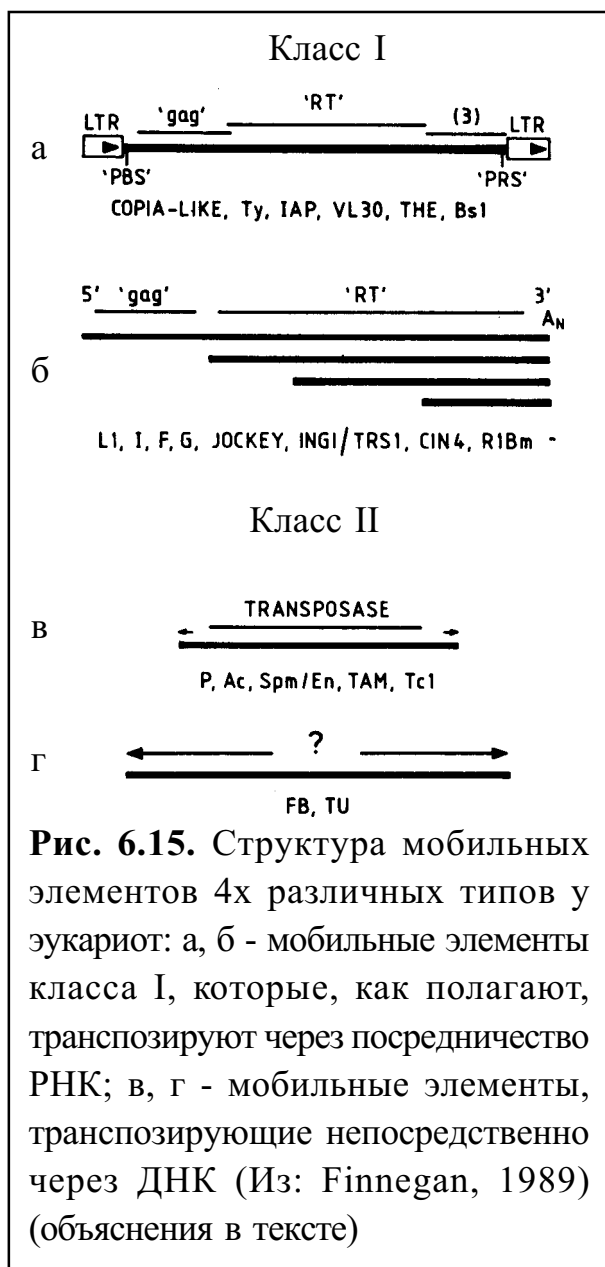
**Рис. 6.15.** Изменение окраски кукурузного зерна под влиянием перемещений элементов *Ac-Ds* (Из: Russell, 1998, p. 668)

Если растение имеет аллель гена *C* дикого типа, зерно будет иметь пурпурную окраску (Рис. 6.15а.), если *Ac* элемент индуцировал инсерцию *Ds* в ген *C*, возникает мутантный аллель *c*, и окраска зерна будет бесцветной (Рис. 6.15б.). В ходе развития *Ds* элемент может в некоторых клетках выйти из гена *C*, в результате чего зерна вновь приобретут пурпурную окраску. Таким

образом, возникает мозаичность (Рис. 6.15в.).

К настоящему времени мобильные элементы открыты у множества видов растений, животных и микроорганизмов. Ниже дается их классификация и подробное описание.

По механизмам транспозиции мобильные элементы делятся на две большие группы: элементы класса I



перемещаются, используя обратную транскриптазу, т.е. на ДНК-матрице мобильного элемента синтезируется РНК. Фермент, осуществляющий эту реакцию синтеза ДНК на РНК, называют обратной транскриптазой или (в русской литературе) - ревертазой. Обратная транскриптаза не только ведет синтез нити ДНК на РНК, но и осуществляет синтез второй комплементарной нити ДНК, а РНК-матрица распадается и удаляется. Двунитевая ДНК синтезируется в цитоплазме, а затем перемещается в ядро и может встроиться в геном, образуя

провирус. Такие мобильные элементы называют ретротранспозонами (или ретропозонами). Элементы класса II перемещаются непосредственно как ДНК-овые элементы (Рис. 6.16.). Такие элементы называются транспозонами.

На Рис. 6.16а. показаны ретровирусоподобные элементы или ретротранспозоны. Они обладают способностью ретровирусов встраивать (в форме провирусов) ДНК-копии своих РНК-геномов в хромосомы клеток-хозяев. Ретропозоны имеют прямые повторы (LTR) на каждом конце. Есть потенциальный тРНК праймер-связывающий сайт (PBS), сразу после левого LTR, и обогащенная пуринами последовательность непосредственно перед правым LTR. ДНК между LTR-ами содержит открытые рамки считывания. Первая из них имеет гомологию с геном *gag* ретровирусов, который кодирует белковые компоненты нуклеопротеиновой сердцевины вириона. Вторая рамка напоминает вирусный ген *pol* и кодирует потенциальную обратную транскриптазу (RT). У некоторых мобильных элементов эти открытые рамки сливаются. Некоторые элементы имеют три рамки. Третья рамка находится в похожем положении с вирусными генами *env*, но имеет другую последовательность нуклеотидов. У ретровирусов ген *env* кодирует компоненты оболочки вирусной частицы. Элементы этого типа встречаются у дрозофилы (*copla-like*), у дрожжей (*Ty*), у грызунов (*IAP* и *VL30*), у человека (*THE*), у кукурузы (*BS1*). (Более детально описание ретропозонов представлено в книге: Lewin, 1994, P. 1034-1045).

На Рис. 6.16б. показаны элементы, не имеющие концевых повторов (не вирусные ретропозоны). Эти элементы обычно имеют две открытые рамки считывания. Первая напоминает ген *gag*, а вторая кодирует потенциальную обратную транскриптазу (RT). У этих элементов есть последовательность, обогащенная аденином (A<sub>n</sub>) на 3' конце одной из нитей (ее нет у R1Bm-элемента, встраивающегося в некоторые гены 28S РНК у *Bombix mori*). У них часто делетирован определенный участок с 5' конца, но они имеют фиксированный 3' конец. Транспозоны такого типа встречаются у млекопитающих (*L1*), у дрозофилы (*I, F, G, jockey*), у трипаномы (*INGI/TRS1*), у кукурузы (*Cin4*), а также инсерции в 28S рРНК у *Bombix mori*, *D. melanogaster*, *Ascaris lumbricoides* (Из: Finnegan, 1989).

Второй класс элементов (см. Рис. 6.16в, г.) объединяет представителей которые перемещаются в геноме, как ДНК-элементы. В этот класс входят транспозоны бактерий (IS-элементы, см. ниже и Lewin, 1994, Р. 1001), *P* и *hobo* у дрозофилы, *Ac/Ds* и *Spm/En* у кукурузы, *Tam* у *Antirrhinum majus* и *Tc1* у нематоды *C. elegans*. Все они имеют короткие инвертированные повторы на концах (Рис. 6.16в.). *P*, *Ac* и *Spm/E* кодируют по крайней мере одну функцию транспозазы, поскольку элементы с внутренними делециями могут перемещаться только в присутствии полных элементов.

Элементы с длинными концевыми инвертированными повторами (Рис. 6.16.) составляют вторую группу класса II. Это *fold-back* (или *FB*) элементы у дрозофилы, *TU* у морского ежа. О механизмах их перемещений известно мало (Из: Finnegan, 1989).

## 6.6.2. Мобильные элементы у дрозофилы

Скрещивания определенных линий дрозофилы приводит к образованию потомства с “дисгенетическими признаками”. Это выражается в появлении у них серии генетических дефектов, таких как мутации, хромосомные aberrации, нарушение расхождения хромосом в мейозе и стерильность. Комплекс этих генетических аномалий характеризует явление, получившее название гибридного дисгенеза.

У дрозофилы были выделены несколько систем, обуславливающих гибридный дисгенез: например, I-R, P-M. При скрещиваниях самцов из линий I (*inducer*) с самками R (*reactive*) наблюдается уменьшение плодовитости потомства, однако реципрокное скрещивание проходит нормально.

Скрещивание между самцом P (*paternal*) и самкой M (*maternal*), вызывает дисгенез, а в реципрокном скрещивании этого эффекта нет (Рис. 6.17.).

Дисгенез проявляется преимущественно в зародышевых клетках. Морфологический дефект в развитии зиготы проявляется начиная со стадии, на которой в зародышевой



**Рис. 6.17.** Гибридный дисгенез асимметричен. Он проявляется только в одном из реципрокных скрещиваний

линии начинается быстрое клеточное деление. Дисгенез вызывается фактором Р, находящемся в хромосомах Р-линии, в линии М нет Р-фактора. Показано, что Р-фактор активируется под действием М-цитоплазмы, унаследованной по материнской линии. Материнская М-цитоплазма названа М-цитотипом. Внедрения Р-элемента стабильны, если хромосома находится в Р-цитотипе; если хромосома попадает в М-цитотип, Р-элементы начинают перемещаться.

Любая хромосома Р-самца может вызвать гибридный дисгенез в скрещиваниях с М-самкой. В пределах же одной хромосомы довольно много районов способно вызвать дисгенез. Это предполагает, что Р-самец в своем геноме имеет большое число Р-факторов, число которых варьирует от 30 до 50, и в разных линиях сайты их локализации различны. (Льюин, 1987, стр. 480-481; Lewin, 1994, Р. 1026).

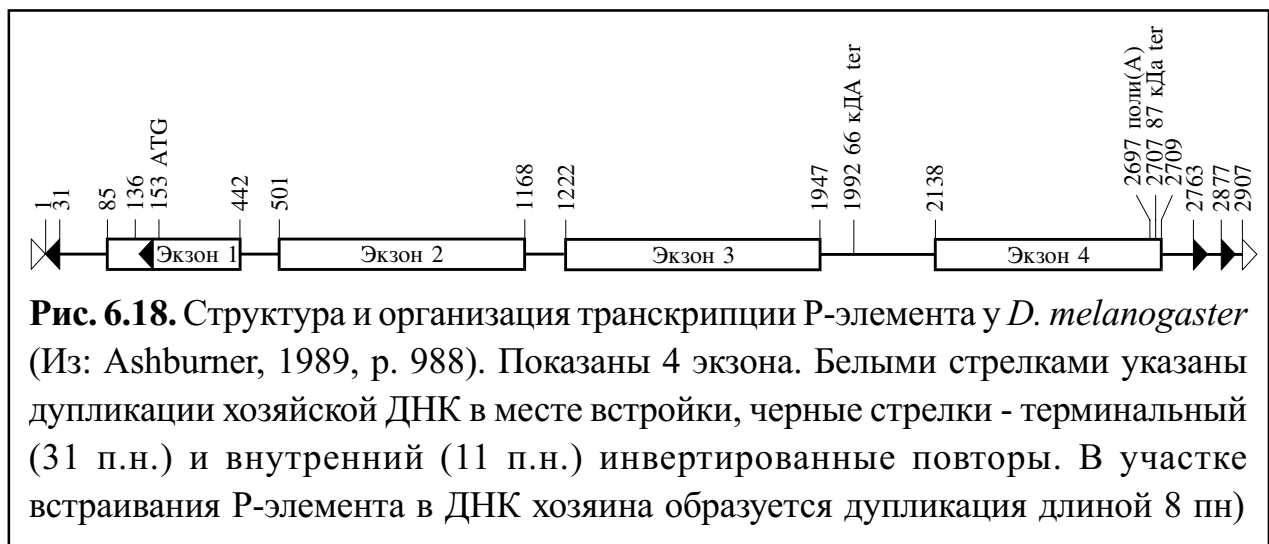
ДНК Р-элемента выделена и охарактеризована (Рис. 6.18.). Полный Р-элемент имеет длину 2907 п.н., и ограничен терминальными повторами размером 31 п.н. Функционально это один ген, дающий транскрипт размером 2,7 т.п.н., кодирующий белок с

молекулярным весом 87кДа - транспозазу. В половых клетках все три интрона процессируются. В соматических клетках третий интрон не процессируется, что определяет специфичность активирования перемещений в половых клетках (Из: Ashburner, 1989, р. 987).

В результате трансляции РНК, синтезированной в соматических клетках, формируется белок размером 66 кДа, который является репрессором активности транспозона.

В половых клетках удаляются все три интрона, в результате чего четыре экзона входят в состав мРНК, которая в результате трансляции дает белок размером 87 кДа, являющийся катализатором транспозиции - транспозазой.

Для осуществления транспозиции Р-элемента необходимы примерно 150 пн на терминальных концах элемента. Транспозаза связывается с последовательностью длиной 10 пн, расположенной рядом с инвертированным повтором 31 пн (см. Рис. 6.18.). Транспозиция происходит по принципу "вырезание - встраивание".



**Рис. 6.18.** Структура и организация транскрипции Р-элемента у *D. melanogaster* (Из: Ashburner, 1989, р. 988). Показаны 4 экзона. Белыми стрелками указаны дубликации хозяйской ДНК в месте встройки, черные стрелки - терминальный (31 п.н.) и внутренний (11 п.н.) инвертированные повторы. В участке встраивания Р-элемента в ДНК хозяина образуется дубликация длиной 8 пн)



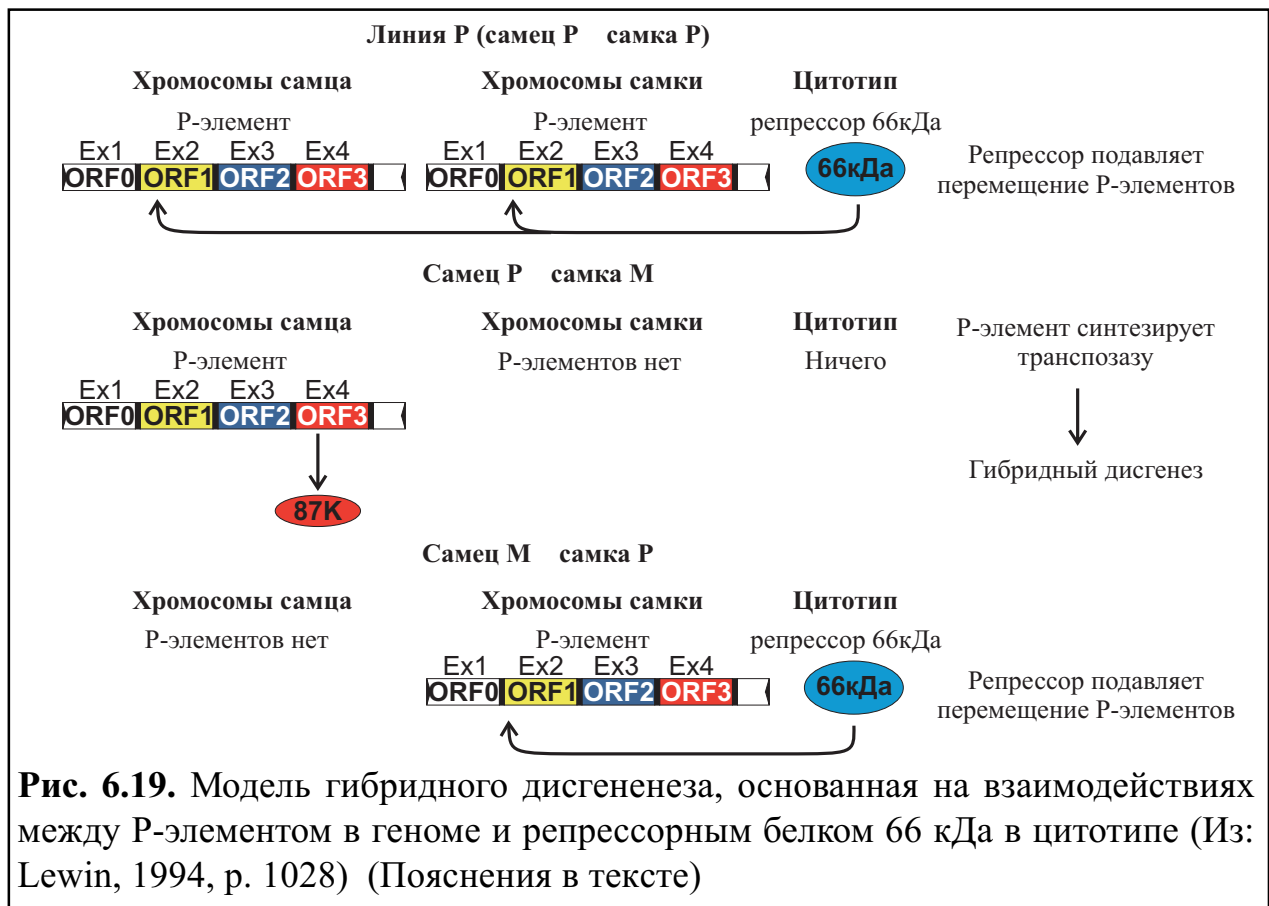
Эффект цитотипа при гибридном дисгенезе у дрозофилы объясняется моделью, представленной на Рис. 6.19. Белок 66кДа, который репрессирует транспозицию, в большом количестве представлен в яйце как материнский фактор. В линии Р должно быть достаточно этого белка, чтобы полностью предотвратить транспозицию, хотя Р-элементы и присутствуют. В любом скрещивании, в котором принимает участие Р-самка, транспозаза не синтезируется. В том случае, когда самка М-цитотипа, в яйце не накапливается белок репрессор, и внесение Р-элемента из генома самца приводит к наработке транспозазы в клетках зародышевого пути. (см. Рис. 6.19.).

Интересно, что линии *D. melanogaster*, выделенные из диких популяций более 30 лет назад, всегда

имеют М-цитотип. В последние 10 лет почти все дикие популяции имеют Р-элементы. Полагают, что повсеместное распространение Р-элемента связано с инвазией и что источником его являются какие-то другие виды (Из: Lewin, 1994, Р. 1027-1029).

(Современные представления об инвазиях Р-элемента см. у Engels, 1997).

Кроме Р-элемента у дрозофилы известно множество других мобильных элементов. Впервые они были выделены и охарактеризованы в лабораториях Г.П. Георгиева (Рис. 6.20.) и В.А. Гвоздева в России, а также Д. Хогнеса (Рис. 6.21.) в США в 1975-1976 гг. Около 12% генома дрозофилы приходится на умеренные повторы, примерно четверть от этого количества занято умеренно повторенными генами (рРНК, гистоны). Остаточные 15000 т.п.н. (9% генома)





**Рис. 6.20.** Георгий Павлович  
Георгиев  
род. 1933



**Рис. 6.21.** Дэвид Хогнесс  
род. 1925

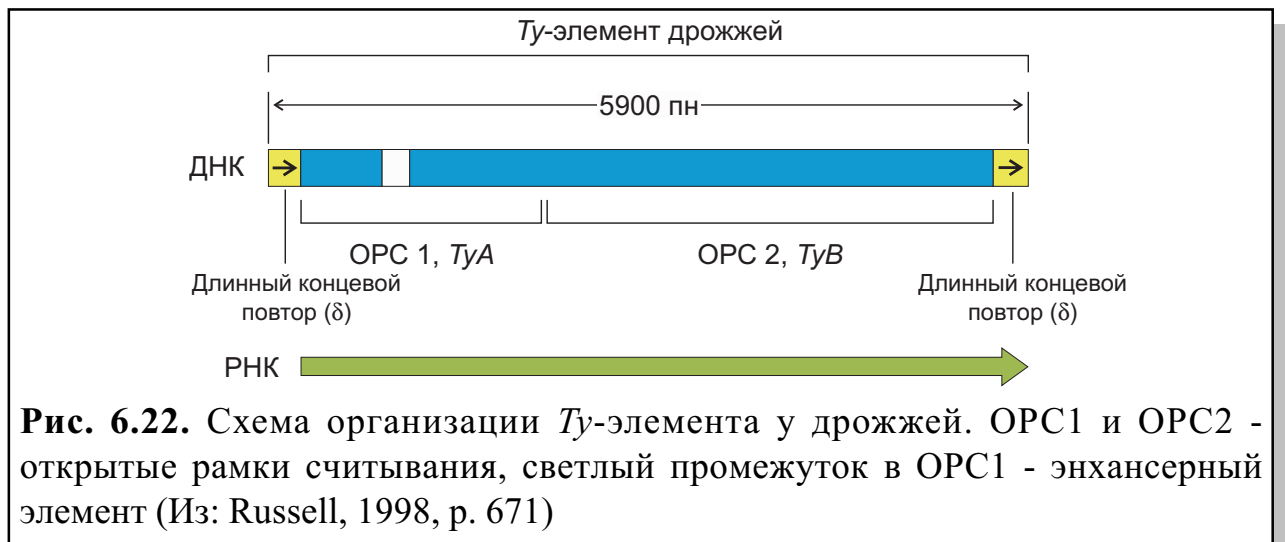
организованы примерно в 50 семейств мобильных элементов (Ashburner, 1989, pp. 91-97). Мобильные элементы часто получают названия, отражающие их способность к перемещению: Магеллан, “Бигль”, *hobo* - бродяга, *gypsy* - цыган, *flea* - блоха, *burdock* - репейник, *jockey* - наездник и т.д. Они отличаются друг от друга по следующим характеристикам:

1. По размерам - средние размеры - 5, т.п.н., причем самый маленький - “элемент 1360” - 1.176 т.п.н., самый большой - “17,6” - имеет размер 7,4 т.п.н.
2. По числу копий: от 1 до 120 копий на геном.
3. По наличию и размерам длинных концевых повторов (ДКП-LTR). Они могут иметь длину 270-840 п.н., быть прямыми или обратными.
4. По индукции дупликаций в сайте встраивания - 4-6 п.н.

### 6.6.3. Ту-элементы у дрожжей

На Рис. 6.22. показана схема организации транспозона *Tu*. Он ограничен на концах длинными концевыми повторами (LTR) или дельта ( $\delta$ ). Элементы дельта на 70% состоят из АТ последовательностей. Каждый из них имеет промотор и последовательность, опознаваемую ферментами транспозиции. *Tu*-элемент имеет длину 5,9 т.п.н. и кодирует единственную мРНК длиной 5700 н, старт транскрипции которой находится в промоторе элемента дельты. Матричная РНК имеет открытые рамки считывания, т.е. рамку, начинающуюся со стартового кодона и кончающуюся терминирующим кодоном. Эти две рамки называют *TuA*, *TuB*.

Генетических маркеров у этого транспозона нет, поэтому следить за его перемещениями трудно.



Tu - имеет много черт сходства с ретровирусами: в частности, транспозиции у них происходят как у ретротранспозонов, уже установлено, что Ту-элемент кодирует обратную транскриптазу.

#### 6.6.4. Транспозоны млекопитающих

У млекопитающих в составе генома обнаружено несколько классов умеренно повторенных последовательностей: SINE (short interspersed nuclear sequences) и LINE (long interspersed nuclear sequences). Элементы SINE - это фрагменты длиной 100-300 п.н., чередующиеся с уникальными последовательностями от 1000 до 2000 п.н. длиной. Элементы LINE имеют длину более 5 т.п.н., они чередуются с уникальными последовательностями до 35 т.п.н. длиной. Как SINE, так и LINE представлены семействами, состоящими из одинаковых элементов.

В геноме человека SINE широко представлено семейством элементов Alu. Члены этого семейства имеют длину 300 п.н. и повторены от 300000 до 500000 раз в геноме. Около 3% генома человека приходится на долю этих повторов.

Наименование Alu этот элемент получил поскольку содержит сайт рестрикции рестриктазы AluI. Каждая Alu-последовательность фланкирована прямыми повторами длиной от 7 до 20 п.н. По этой причине полагают, что Alu-повторы являются мобильными элементами, скорее всего ретротранспозонами.

Одно из семейств, принадлежащих LINE, это LINE-1 (или L1-элемент). Полагают, что в геноме человека присутствует от 50 до 100 тыс. копий L1-элемента, т.е. он представляет около 5% генома. Максимальная длина этих элементов составляет 6500 п.н., хотя таких элементов в геноме не более 3500. Остальные же копии по аналогии с Ds элементами кукурузы имеют внутренние делеции различной длины. Полноразмерные L1 элементы содержат большие открытые рамки считывания, имеющие гомологию с известными обратными транскриптазами (См. Russell, 1998; Lewin, 1994, p. 1051-1055).

### 6.6.5. Функциональное значение мобильных элементов

Наличие мобильных элементов в геномах имеет разнообразные генетические последствия.

1. Перемещения и внедрение мобильных элементов в гены может вызывать мутации. Так, большинство “спонтанных” мутаций в локусе *white* у дрозофилы индуцированы инсерциями мобильных генетических элементов (см. Юрченко, Голубовский, 1988).

Около 80% спонтанных мутаций, изученных в разных локусах дрозофилы, вызвано инсерциями мобильных элементов. Внедряясь в ген, мобильный элемент может повредить экзон, разорвав его. В таком случае ген будет лишен возможности кодировать белок. Попадая в район промоторов или энхансеров, мобильный элемент может повредить регуляторную зону гена. Наконец, инсерция в район интрона может оказаться безвредной, поскольку вся последовательность интрона вместе с мобильным элементом будет вырезана

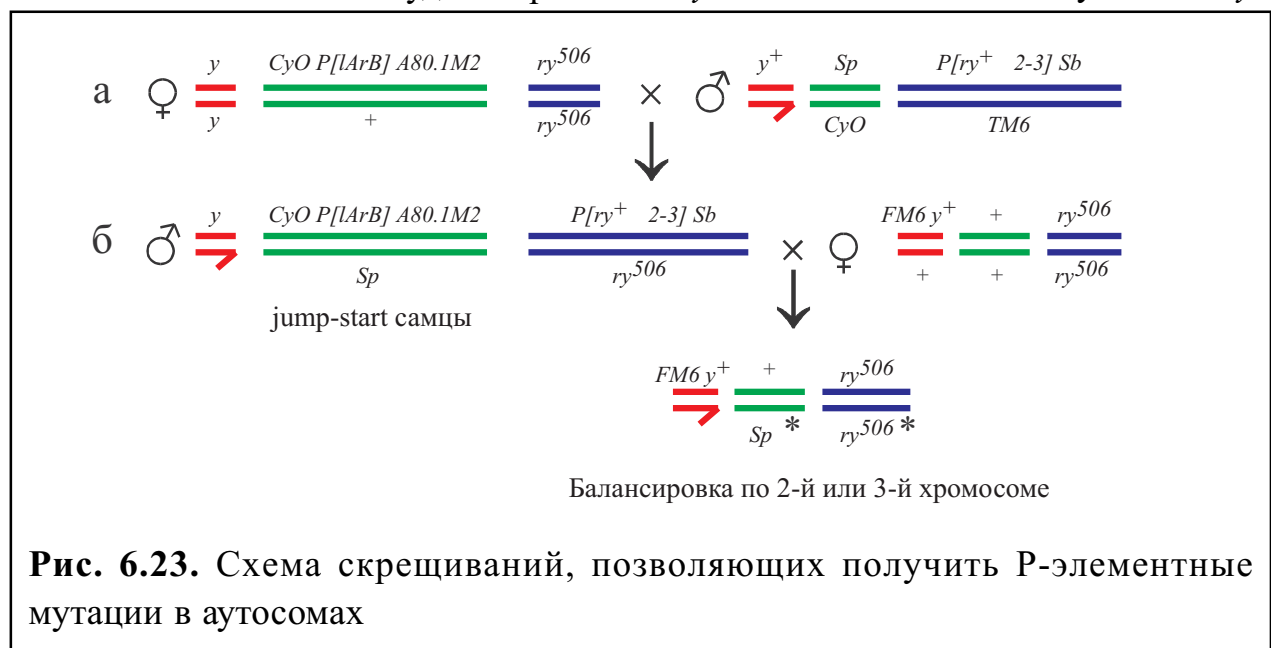
во время процессинга мРНК, а соседние экзоны беспрепятственно сплайсируются.

Синтезированы экспериментальные схемы, позволяющие получать большое число мутаций у дрозофилы, индуцируемых мобильными Р-элементами. Получают гетерозиготу по двум Р-элементам:

#### **P[ry<sup>+</sup>Δ2-3](99B) и P[ArB]**

Элемент P[ry<sup>+</sup>Δ2-3] (99B) имеет два полезных свойства: он синтезирует транспозазу, которая может катализировать транспозицию неавтономного Р-элемента, но не может перемещаться сам. Элемент P[ArB] - способен перемещаться, но не имеет своей транспозазы. Плазмида P[ArB] содержит гены *rosy*<sup>+</sup> и *Adh*<sup>+</sup>, которые облегчают обнаружение трансформантов на фоне мутаций *ry* и *Adh*. Этот метод получил название “Jumpstart” (см.: Lai, 1994, p. 520). Ставится следующее скрещивание (Рис. 6.23.).

Самок, содержащих встроенную плазмиду P[ArB] в хромосому, маркированную доминантной мутацией *CyO* и гомозиготных по мутации *rosy*<sup>506</sup>,



**Рис. 6.23.** Схема скрещиваний, позволяющих получить Р-элементные мутации в аутосомах

скрещивают с самцами, у которых по одной доминантной мутации (*Sp* и *CyO*) во второй хромосоме и инсерция  $P[ry^+\Delta 2-3]$  транспозона в третью хромосому, меченную мутацией *Sb* (Рис. 6.23а.). В потомстве отбирают самцов *rosy<sup>+</sup>/Sb; CyO/ Sp*. Они имеют инсерции транспозонов обоих типов. Это самцы jump-start. Их скрещивают с самками *ry<sup>506</sup>/ry<sup>506</sup>* (Рис. 6.23б.). В потомстве отбирают самцов *CyO<sup>+</sup>* (т.е. без  $P[1ArB]$  в исходной хромосоме), *Sb<sup>+</sup>* (т.е. без  $P[ry^+\Delta 2-3]$ ), но *rosy<sup>+</sup>* (т.е. содержащих  $P[1ArB]$  в новой позиции во второй или третьей хромосоме - отмечены звездочками на Рис. 6.23б.).

2. Изменение состояния активности генов.

ДКП мобильного элемента являются промоторами ретротранспозона, причем как ДКП, так и сам ретротранспозон содержит нуклеотидные последовательности, являющиеся энхансерами транскрипции. Поэтому перемещение этих сигналов в геноме может изменять регуляцию активности генов. Например, если мобильный элемент оказался около протоонкогена, то результатом может быть сверхпродукция белка и злокачественное перерождение клетки (Гвоздев, 1998б).

В случае ретротранспозонов особые возможности для перенесения и приобращения регуляторных сигналов возникают в том случае, когда элемент удаляется за счет кроссинговера между ДКП с идентичными последовательностями (Рис. 6.24а.), в результате сохраняется лишь один ДКП на месте внедрения ретротранспозона. Это явление широко распространено в клетках дрожжей. Установлено, что такие

одинокие ДКП оказывают серьезное влияние на изменчивость регуляторных систем дрожжевой клетки.

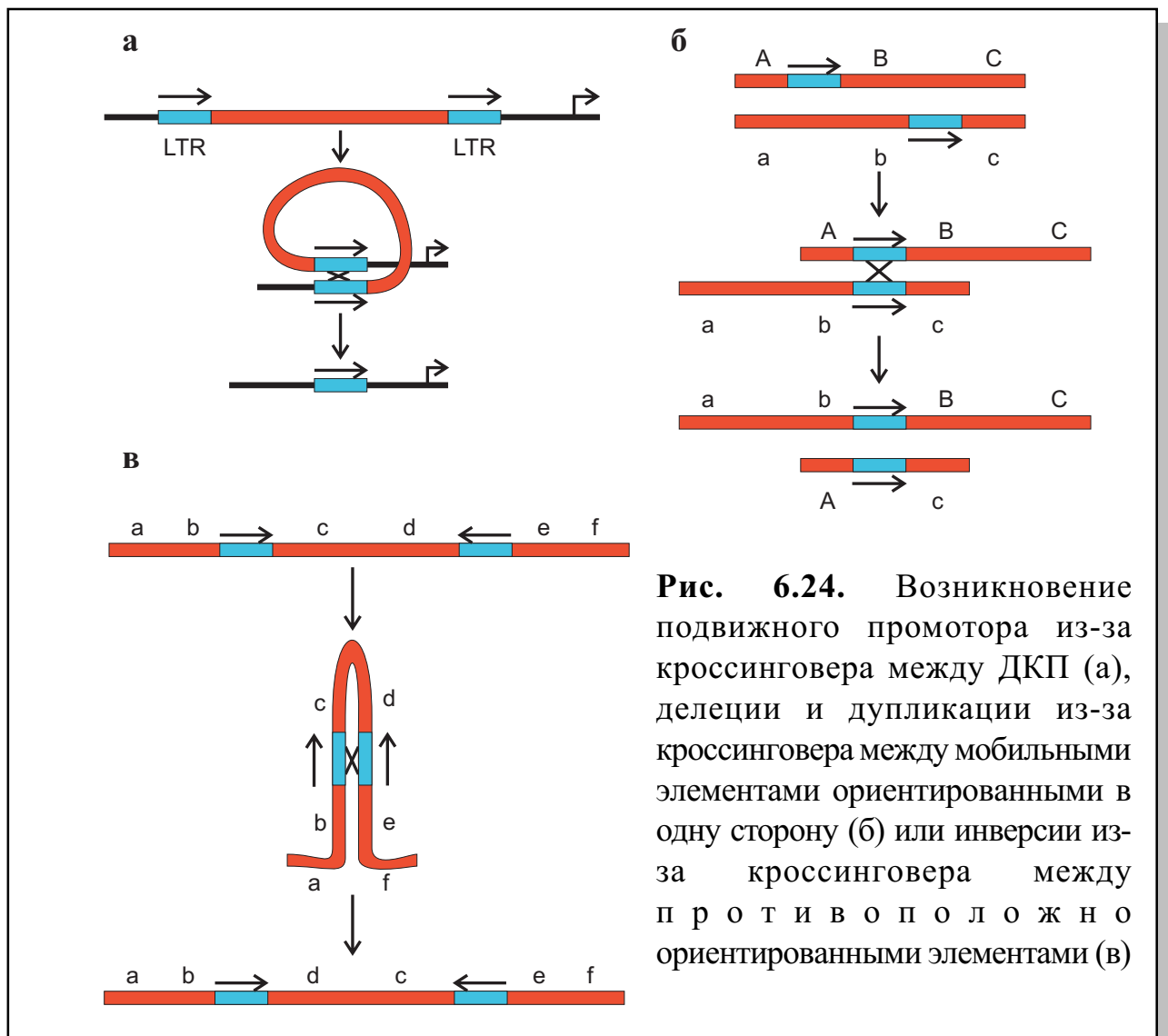
3. Формирование хромосомных перестроек.

В результате кроссинговера между одинаково ориентированными элементами возникает делеция и дупликация материала, расположенного между инсерциями (Рис.6.24б.). Если инсерции ориентированы в противоположном направлении, возникает инверсия (Рис.6.24в.).

4. У дрозофилы отсутствует теломеразная машина, но концы ДНК удлиняются за счет перемещений ретротранспозонов (см. раздел 9.5.).

5. Участие в горизонтальном переносе генов.

Инфекционные ретровирусы способны заражать организмы, принадлежащие к разным видам и переносить собственный генетический материал, образуя копии ДНК, встраивающиеся в геном. Таким образом могли распространяться ретротранспозоны. Подобный способ передачи генов получил название горизонтального в отличие от вертикального наследования генов из поколения в поколение. Широкое распространение транспозона *mariner* среди филогенетически отдаленных групп организмов может свидетельствовать о повторных переходах данного элемента от вида к виду (Гвоздев, 1998б). Так, один из ретротранспозонов дрозофилы (*gypsy*), как оказалось, является настоящим ретровирусом: путем инъекции или скармливанием вирусных частиц удается заразить мух, не несущих этих ретротранспозонов (Гвоздев, 1998а).



**Рис. 6.24.** Возникновение подвижного промотора из-за кроссинговера между ДКП (а), делеции и дупликации из-за кроссинговера между мобильными элементами ориентированными в одну сторону (б) или инверсии из-за кроссинговера между противоположно ориентированными элементами (в)

6. Транспозоны на основе Р-элемента используют для трансформации у эукариот, клонирования генов, поиска энхансеров (см. более подробно в разделе 7).

### Литература к разделу 6.6.

Гвоздев В.А. Подвижная ДНК эукариот. Часть 1. Структура, механизмы перемещения и роль подвижных элементов в поддержании целостности хромосом. Соросовский образовательный журнал. N8, С.8-14, 1998а.

Гвоздев В.А. Подвижная ДНК эукариот. Часть 2. Роль в регуляции активности генов и эволюции

генома. Соросовский образовательный журнал. N8, С.15-21, 1998б.

Льюин Б. Гены. Москва, Мир, 1-544, 1987.

Хесин Р.Б. Непостоянство генома. Москва, Наука, 1-472, 1984.

Юрченко Н.Н., Голубовский М.Д. Современная генетика локуса *white* у *Drosophila melanogaster*. Генетика 29, N4, 581-591, 1988.

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

Cooley L., Kelley R., Spradling A. Insertional mutagenesis of the

- Drosophila* genome with single P elements. Science, V. 239, P. 1121-1128, 1988.
- Engels W.R. Invasions of P elements. Genetics, V. 145, P. 11-15, 1997.
- Federoff N.V. Barbara McClintock (June 16, 1902 - September 2, 1992). Genetics 136, 1-10, 1994.
- Finnegan D.J. Eukaryotic transposable elements and genome evolution. Trends in Genetics 5, N4, 103-107, 1989.
- Lai C. Genetic application of transposable elements in eukaryotes. Genome 37, 519-525, 1994.
- Lewin B. Genes V. Oxford, New York, Tokyo, Oxford University Press, 1-1272, 1994.

## 6.7. Мобильные элементы прокариот

У прокариот существуют три типа мобильных элементов - *IS*-элементы (insertion sequence), транспозоны (*Tn*) и некоторые бактериофаги.

### 6.7.1. *IS*-элементы

Эти элементы содержат только минимальное число генов, необходимых для мобилизации элемента и его инсерции в новый участок хромосомы. *IS*-элементы являются обычным компонентом бактериальных хромосом и плазмид. Три инсерционных элемента - *IS1*, *IS2*, и *IS10R* представлены в геноме *E. coli* в 0-30 копиях. Их размеры варьируют от 768 до 1329 п.н. Некоторое число копий этих элементов встречается в плазмидах. Они содержат ген транспозазы. На концах *IS*-элементов находятся инвертированные повторы IR, длина которых варьирует от 22 до

41 п.н. В участке встраивания *IS*-элементов в геномной ДНК образуется дупликация размером от 5 до 9 п.н. (Рис. 6.25.).

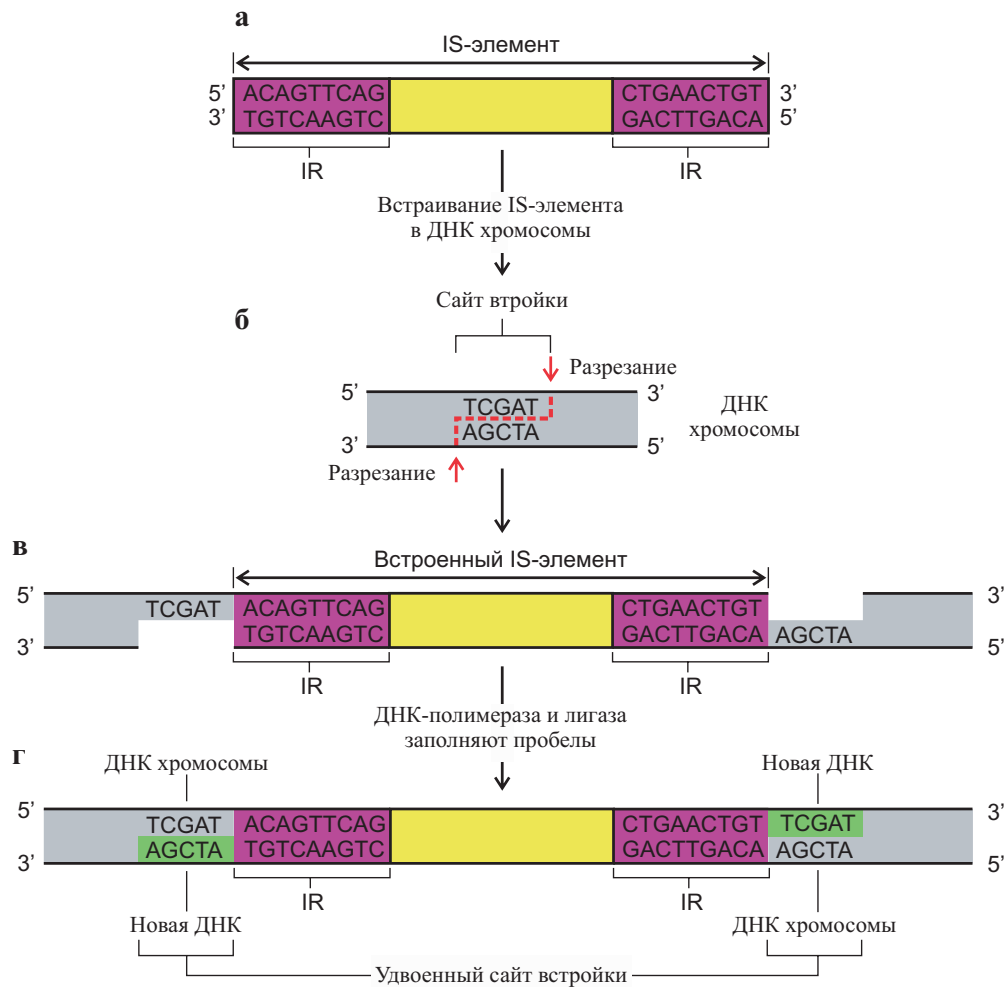
Поскольку *IS*-элементы встраиваются в любой участок ДНК, они часто вызывают мутации, разрушая либо кодирующие или регуляторные последовательности.

Промоторы в самом *IS*-элементе могут влиять на экспрессию соседних генов. Из-за присутствия *IS*-элементов в хромосоме кроссинговер между ними может вызывать делеции и инверсии.

В процессе встраивания *IS*-элементов происходит точное копирование уже встроенного элемента, затем старая копия остается на месте, а вновь синтезированная внедряется в новый сайт. Репликация новой копии происходит с использованием энзимов репликационной машины клетки хозяина. Транспозиция происходит с использованием транспозазы, которая опознает IR последовательности, где и инициируется транспозиция. Мутации в IR фрагментах влияют на частоту транспозиции. Частота транспозиций варьирует между  $10^{-5}$  и  $10^{-7}$  на одно поколение.

### 6.7.2. Транспозоны

Эти элементы (*Tn*) устроены значительно сложнее. Известны два типа транспозонов эукариот: сложные и не сложные. Сложные (composite) транспозоны имеют центральный район, содержащий гены. На обоих концах транспозона расположены *IS*-элементы (Рис.6.26.). Оба *IS*-элемента в пределах одного транспозона



**Рис. 6.25.** Схема строения и инсерции *IS*-элемента в хромосомную ДНК бактерии.

а - *IS*-элемент с *IR*-повторами длиной 9 п.н.

б - образование двуцепочечного надреза в ДНК бактерии.

в - инсерция *IS*-элемента в участке надреза.

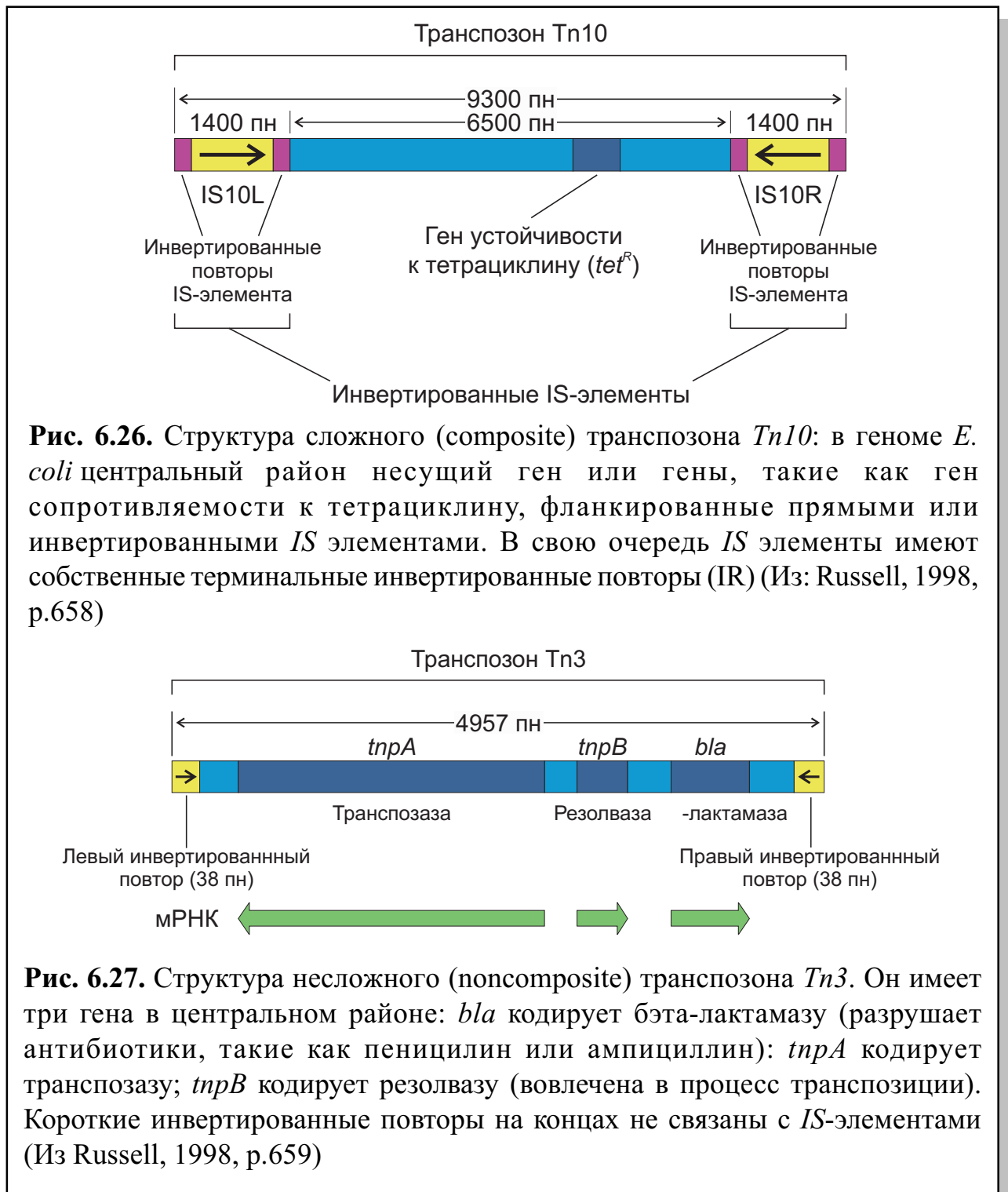
г - достраивание фрагмента ДНК в участке надреза и образование дупликации геномной ДНК бактерии (Из: Russell, 1998, p.657)

принадлежат к одному типу и называются “левый” и “правый”.

Транспозоны *E. coli* варьируют по длине от 2638 п.н., до 9300 п.н., имеют гены чувствительности к антибиотикам, *IS*-элементы на концах. В участках встраивания в хромосоме хозяина образуется дупликация длиной 9 п.н. Транспозиции происходят, поскольку один или оба *IS*-элемента представляют транспозазу.

Не сложные (noncomposite) транспозоны также содержат гены чувствительности к антибиотикам, однако эти транспозоны не терминируются *IS*-элементами. На своих концах они все-таки имеют повторенные последовательности, необходимые для транспозиций. (Рис. 6.27.).

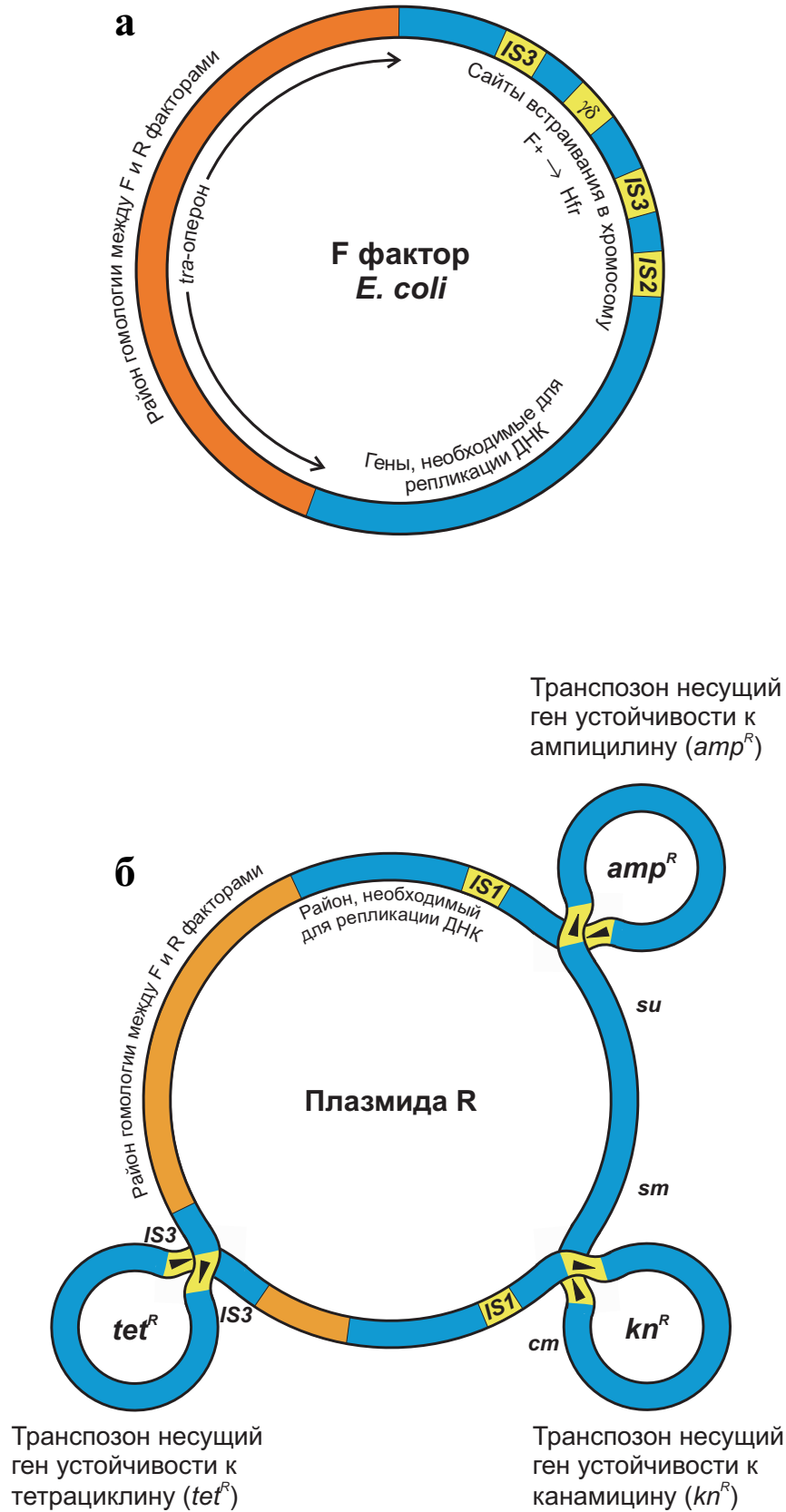




### 6.7.3. IS-элементы и транспозоны в плазидах

Как уже упоминалось выше (раздел 5), одной из первых открытых плазмид была плазида *F* (*F*-фактор). Эта кольцевая структура имеет длину 94,5 т.п.н. и содержит ряд генов: а) *tra*

(transfer) гены, необходимые для переноса ДНК из бактерии-донора в бактерию-реципиента, б) гены, необходимые для репликации плазмиды, в) четыре *IS*-элемента: две копии *IS3*, одну копию *IS2* и один элемент –  $\gamma\delta$  (Рис. 6.28.).



**Рис. 6.28.** Карты *F*-фактора *E. coli* (а) и плазмиды *R* (б). Объяснения в тексте (Из: Russell, 1998, p. 661)

Поскольку в различных позициях в хромосоме *E. coli* имеются копии этих же инсерционных элементов, *F*-фактор может встраиваться в различные участки бактериальной хромосомы и в различной ориентации. Встраивание происходит за счет кроссинговера между гомологичными последовательностями инсерционных элементов.

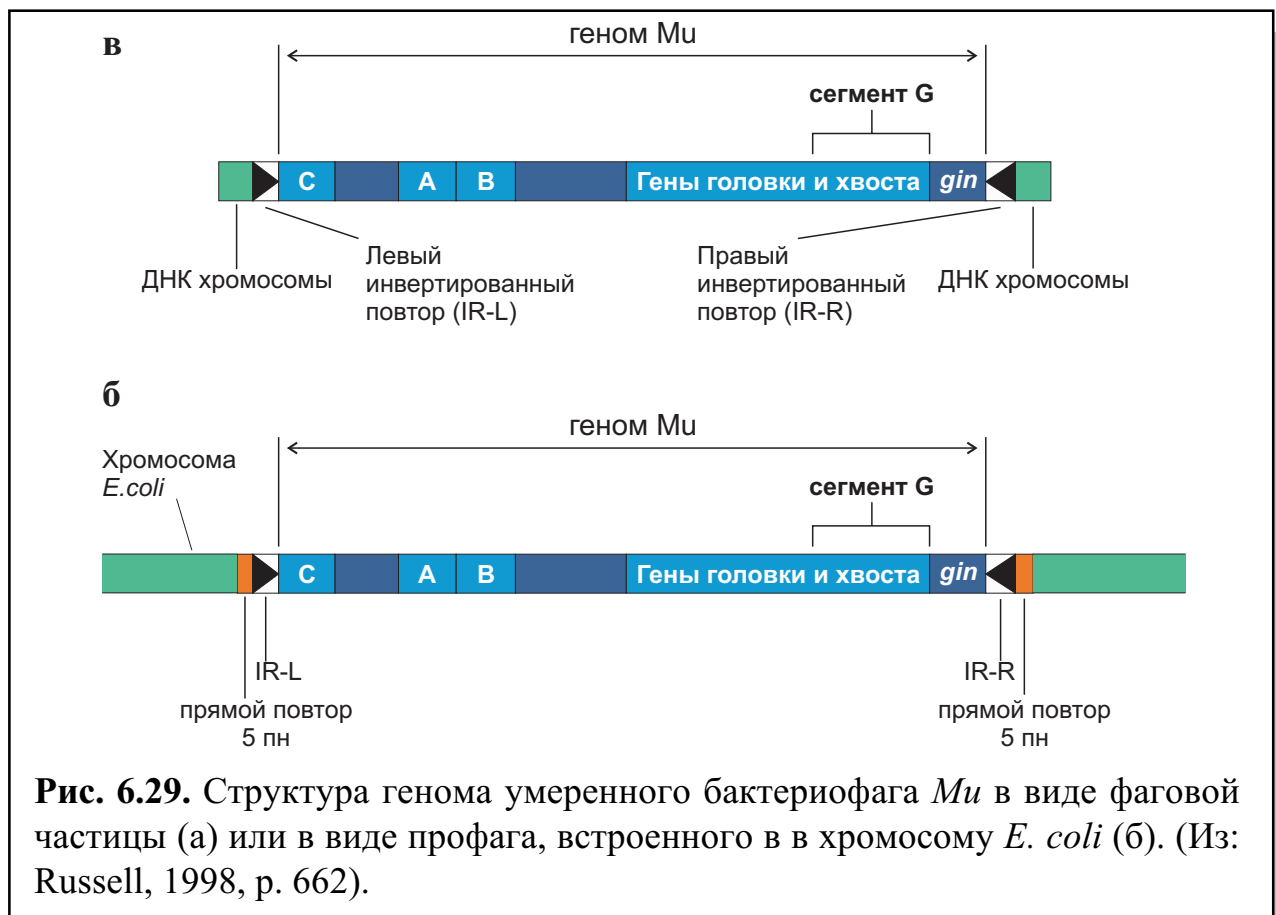
Плазмида *R* имеет те же *tra* гены, что и в *F* факторе, гены, необходимые для репликации ДНК, затем есть две копии *IS1* элемента, кроме того, есть три гена резистенции к антибиотикам - *amp* (ампициллин), *kan* (канамицин) и *tet* (тетрациклин) (Рис. 6.28.). Эти гены фланкированы терминальными повторенными последовательностями и, следовательно сами являются истинными транспозонами. Плазмида

*R*, обычно находящаяся у патогенной бактерии *Shigella*, передается другим (нерезистентным к антибиотикам) штаммам этого вида, а также другим бактериям, населяющим кишечник человека, в результате чего они приобретают резистентность.

#### 6.7.4. Бактериофаг *Mu*

Это умеренный бактериофаг, инфицирующий *E. coli*, “умеренный” - значит может проходить через литический цикл или лизогенную стадию. В то же время, это транспозон и он может вызывать мутации в результате встраивания.

В фаговой частице *Mu* геном содержит 37 т.п.н. линейной молекулы ДНК, главным образом фаговой с небольшими фрагментами ДНК из хромосомы *E. coli* на каждом из концов (Рис. 6.29а).



***Литература к разделу 6.7.***

Хесин Р.Б. Непостоянство генома.

Москва, Наука, 1984.

Russell P.J. Genetics. Fifth edition, Menlo

Park, California, 655-678, 1998.