

6 28

П58

В. В. Попов

# ГЕНОМИКА

## С МОЛЕКУЛЯРНО-ГЕНЕТИЧЕСКИМИ ОСНОВАМИ



URSS

828  
П58

Н

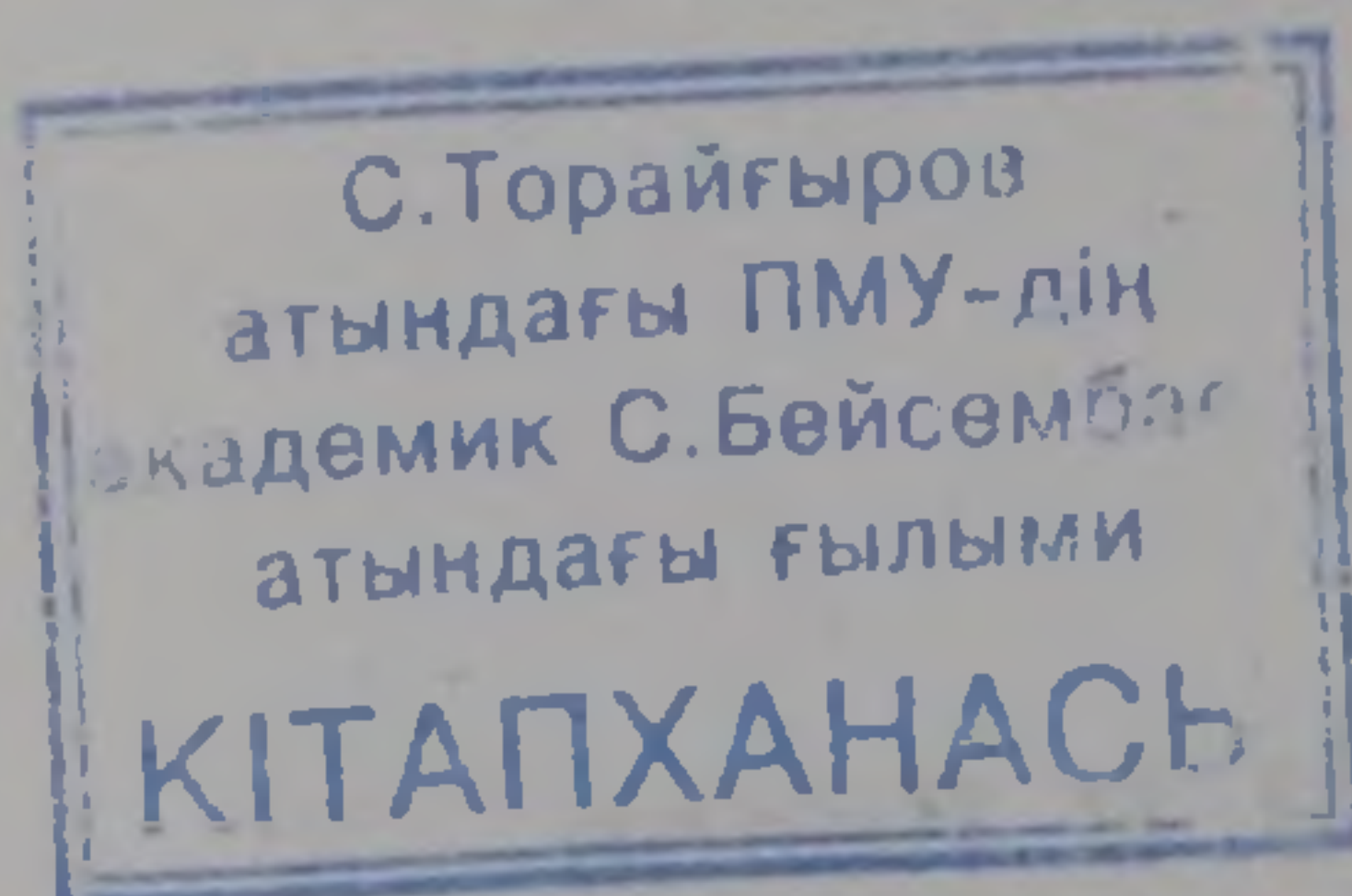
**В. В. Попов**

**ГЕНОМИКА  
С МОЛЕКУЛЯРНО-  
ГЕНЕТИЧЕСКИМИ  
ОСНОВАМИ**



**URSS  
МОСКВА**

ББК 28.04



Попов Вадим Васильевич

606498

Геномика с молекулярно-генетическими основами. — М.: Книжный дом «ЛИБРОКОМ», 2009. — 304 с.

Предлагаемая читателю книга посвящена рассмотрению актуальной проблемы современной генетики — изучению и расшифровке геномов (определению полной структуры ДНК) живых организмов, включая геном человека. Полная информация о фундаментальных открытиях в этой области доступна в основном узким специалистам по публикациям в научных журналах, преимущественно на английском языке. За основу исторического развития генетики автором принята периодизация Н. П. Дубинина: эпоха классической генетики, эпоха неоклассицизма и эпоха молекулярной генетики (с 1953 г.). В первой части монографии рассмотрены структура и функционирование генетического материала как молекулярно-генетические основы геномики. В отдельной главе описана как классическая мегасистематика высших таксонов древа жизни, так и современная мегасистематика на основе молекулярных реконструкций. Во второй части анализируется массив данных по геномике разных организмов — от вирусов до человека. Приведены базы данных по геномам и их адреса в Интернете. Охарактеризованы основные направления геномных исследований в начавшейся постгеномной эпохе, последовавшей после 2001 г., когда были опубликованы данные о структуре генома человека.

Для преподавателей, аспирантов и студентов биологических специальностей вузов, а также для широкого круга читателей, интересующихся современными достижениями молекулярной генетики и геномики.

#### Рецензенты:

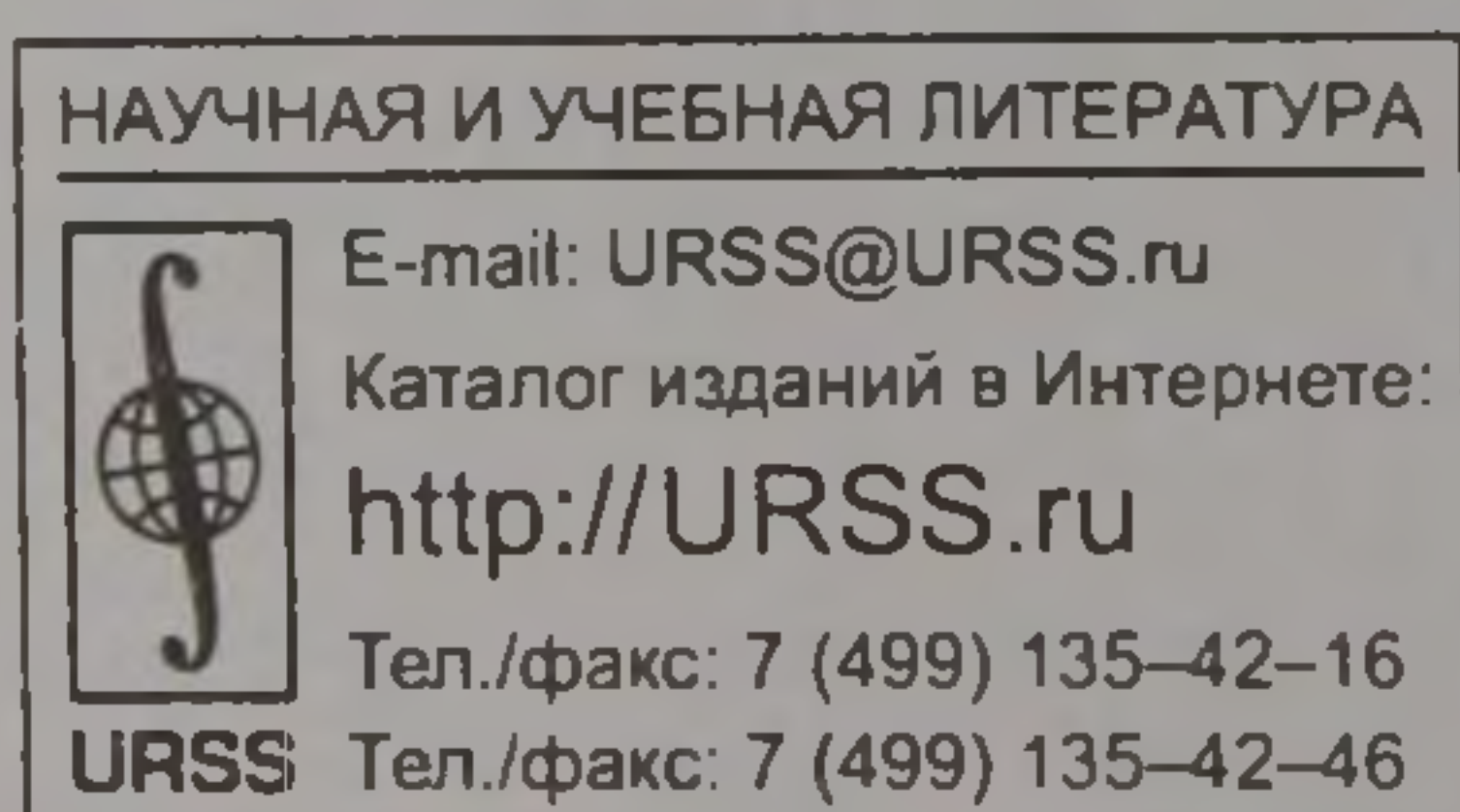
д-р биол. наук, проф. Л. А. Калашникова  
(отдел ДНК-технологий ВНИИ племенного дела);  
д-р биол. наук А. В. Проняев (ГУ «Центрохотконтроль»,  
гл. редактор журнала «Вестник охотоведения»);  
д-р биол. наук А. И. Шаталкин (Научно-исследовательский  
зоологический музей МГУ им. М. В. Ломоносова)

Издательство «Книжный дом «ЛИБРОКОМ»».  
117312, Москва, пр-т Шестидесятилетия Октября, 9.  
Формат 60×90/16. Печ. л. 19. Зак. № 2050.

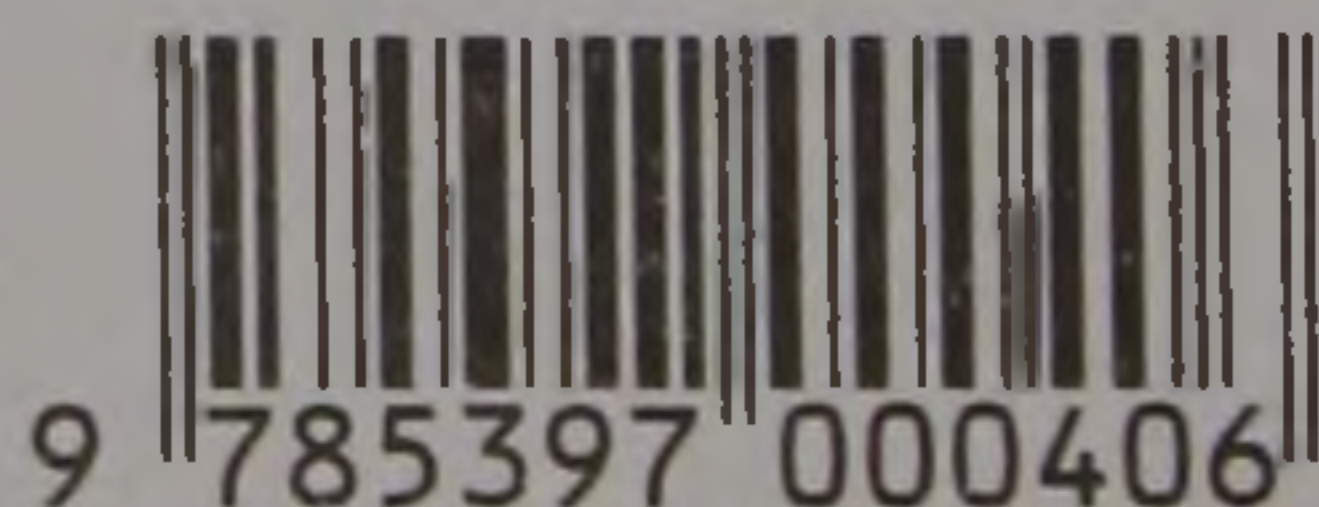
Отпечатано в ООО «ЛЕНАНД».  
117312, Москва, пр-т Шестидесятилетия Октября, 11А, стр. 11.

ISBN 978-5-397-00040-6

© Книжный дом «ЛИБРОКОМ», 2008



5892 ID 73179



# Оглавление

Предисловие .....	4
<i>Введение.</i> Предмет молекулярной генетики и геномика .....	6
<b>Часть первая. Структура и функционирование генетического материала</b> .....	23
<i>Глава 1.</i> Структура нуклеиновых кислот ДНК и РНК .....	24
<i>Глава 2.</i> Репликация ДНК .....	40
<i>Глава 3.</i> Транскрипция ДНК .....	47
<i>Глава 4.</i> Структура гена у прокариот и эукариот .....	60
<i>Глава 5.</i> Трансляция. Биосинтез белка .....	72
<i>Глава 6.</i> Мутации и репарация повреждений ДНК .....	88
<i>Глава 7.</i> Молекулярно-филогенетическая систематика .....	102
<i>Глава 8.</i> Горизонтальный (латеральный) перенос генов как информационный фактор эволюции .....	131
<b>Часть вторая. Геномика</b> .....	155
<i>Глава 9.</i> Становление геномики как самостоятельного раздела молекулярной генетики .....	156
<i>Глава 10.</i> Геномика вирусов и фагов. Вирусы как объект молекулярной генетики .....	162
<i>Глава 11.</i> Геномика прокариот .....	194
<i>Глава 12.</i> Геномика эукариот .....	221
<i>Глава 13.</i> Геном человека .....	242
<i>Заключение.</i> Эра геномики: This is the End of the Beginning (англ. выражение: Это конец начала) .....	256
Именной указатель .....	266
Терминологический словарь .....	269
Основная литература .....	292

# Предисловие

В последние два десятилетия в генетике наблюдался быстрый прогресс в изучении копирования генетической информации и ее молекулярных основ у вирусов, прокариот и эукариот. В области молекулярной генетики выявлены материальные носители генетической информации, изучены закономерности кодирования, принципы управления работой генов, проведен анализ механизмов фундаментальных генетических процессов.

Еще в 1950-е гг. была доказана роль ДНК как материального носителя наследственности и стало понятно значение магистрального пути реализации генетической информации в клетке и основного постулата молекулярной генетики ДНК → РНК → белок. Была установлена роль рибонуклеиновой кислоты (РНК) как прямого продукта гена.

В настоящее время концепция генетической информации и ее материальных носителей является краеугольным камнем генетики и биологии в целом. К началу XXI века установлено далеко не полное знание о том, как именно, по каким общим правилам генетическая информация записана в генах, регуляторных сайтах и функциональных центрах макромолекул, в структурных элементах геномов и как она реализуется в молекулярные функции и свойства. Наши сведения об этой сложной системе пока еще не окончательны.

Френсис Крик, внесший решающий вклад в изучение генетического кода, писал: «...Это ключ к молекулярной биологии, так как он показывает, как два великих языка полимеров — язык полинуклеотидов ДНК и язык полипептидов белков — связаны между собой» (1966). Генетический код универсален и его основная часть одинакова для всех форм жизни на Земле. Этот вывод явился результатом массового секвенирования (определения первичной структуры биополимеров — нуклеиновых кислот и белков) генов и белков. Алфавиты 4 нуклеотидов и 20 аминокислот считаются предзаданными. Набор аминокислот должен удовлетворять принципу достаточного разнообразия свойств для формирования глобулярных белков. В то же время встает вопрос: насколько случаен выбор 20 канонических аминокислот из более чем 200 известных?

Логическим развитием методов молекулярной генетики явилось возникновение *генетической инженерии, биотехнологии и геномики.*

Генетическая инженерия как технология получения рекомбинантных ДНК включает в себя методы: рестрикции, гибридизации, клонирования, определения нуклеотидных последовательностей (секвенирование) и химико-ферментативный синтез полинуклеотидов.

Биотехнология — это многопрофильная комплексная область научно-технического прогресса, включающая микробиологический синтез, генетическую и клеточную инженерию, инженерную энзимологию. Биотехнология возникла на стык микробиологии, биохимии, биофизики, молекулярной генетики, цитологии и иммунологии. Ее уровень развития определяет во многом научно-технический потенциал страны, а прогресс ее в животноводстве в предстоящее десятилетие будет определяться развитием генной, клеточной и эмбриогенетической инженерии.

Одной из главных точек роста современной молекулярной генетики стала *геномика* (Ратнер, 2001). Огромным событием в генетике было секвенирование малого фага φX174 («фи-десять-174») с кольцевой одноцепочечной ДНК размером 5386 п. н. Ф. Сэнджером с соавт. (1977), удостоенное Нобелевской премии 1980 г. Это достижение знаменовало собой начало новой эры в молекулярной генетике — *эры геномики*. В конце 80-х гг. были начаты первые международные программы секвенирования полных клеточных геномов бактерий, грибов, растений, насекомых, млекопитающих и человека. В настоящее время это направление получило развитие в *биоинформатике* — науке о технических средствах исследования и моделирования макромолекул и их систем. Ее основными разделами являются *компьютерная геномика* и *метабономика* (*протеомика*).

Последняя исследует принципиальную сторону организации метаболизма клетки и его управления со стороны генома.

При подготовке монографии использованы материалы ряда авторов, которым выражаю благодарность: Ю. П. Алтухова, В. С. Баранова, С. А. Боринской, С. В. Бокуть, Н. Н. Воронцова, В. И. Глазко, С. Е. Дромашко, Г. А. Заварзина, А. В. Зеленина, В. И. Иванова, Л. А. Калашниковой, М. А. Каменской, Л. Л. Киселева, А. С. Коничева, Л. И. Корочкина, Н. Н. Мушкамбарова, В. И. Назарова, И. Я. Павлинова, В. А. Ратнера, Е. Д. Свердлова, В. З. Тарантула, А. В. Яблокова, Н. К. Янковского, А. И. Зинченко, Е. В. Щеглова, С. Н. Щелкунова. Я благодарю также проф. В. Б. Яковлева и лаборанта каф. генетики и разведения с/х животных Е. В. Мелюшину за помощь при оформлении работы. Выражаю искреннюю признательность профессору А. И. Шаталкину (Зоологический музей МГУ им. М. В. Ломоносова) за прочтение рукописи и важные замечания по улучшению материала.

# Предмет молекулярной генетики и геномика

Молекулярная генетика призвана расшифровывать структуры геномов, генов, механизмы их воспроизводства и реализации информации, закодированной в этих структурах. Эта область науки исследует также механизмы изменчивости генома, лежащие в основе эволюции и биологического разнообразия. С практической точки зрения, молекулярная генетика лежит в основе современной биотехнологии, диагностики и лечения болезней, интенсификации сельского хозяйства и охраны окружающей среды.

Академик *Е. Д. Свердлов*

## 1. Основные понятия. Область молекулярной генетики

Генетика является фундаментальной биологической дисциплиной, изучающей универсальные свойства всех живых организмов — наследственность и изменчивость, а также способы управления ими. Сам термин был введен в 1906 г. английским генетиком Уильямом Бэтсоном.

Генетика включает различные направления исследований:

- 1) биохимическая генетика изучает специфические, генетически контролируемые синтезы;
- 2) физиологическая генетика исследует физиологические функции организма (воспроизведение, перераспределение, изменение и действие наследственных факторов);

- 3) радиационная генетика изучает действие различных видов излучений на генетический материал;
- 4) популяционная и эволюционная генетика направлена на изучение генетической структуры элементарных эволюционных единиц (популяций), закономерностей динамики их генетической структуры, а также направленного изменения их генотипического состава.

По объектам исследования выделяют цитогенетику, генетику растений, генетику микроорганизмов, генетику простейших, генетику животных и человека.

Основными методами генетики можно назвать гибридологический, цитологический и математический. Наряду с этим активно используются методы других естественных наук. Так, методы химии и биохимии применяются для подробной характеристики наследуемых признаков обмена веществ, для изучения свойств молекул белков и нуклеиновых кислот. Широкое применение нашли методы физики — оптические, седиментационные, меченых атомов для маркирования и идентификации различных макромолекул. В последнее десятилетие на базе теории информации возникла биоинформатика и компьютерная геномика.

Методы современной генетики могут быть использованы в любых биологических исследованиях. Поэтому глубоко прав С. Г. Инге-Вечтомов (1989), утверждая: «Генетика — основа современной биологии. Нельзя стать генетиком, не будучи биологом, и нельзя быть современным биологом, не зная генетики».

Логически обоснованную периодизацию развития генетики выдвинул Н. П. Дубинин. Он выделяет в истории генетики три хорошо очерченных этапа — *эпоху классической генетики* (1900–1930), *эпоху неоклассицизма* (1930–1953) и *эпоху молекулярной генетики* (началась в 1953 г.)

### **Первый этап — эпоха классической генетики**

Основные события и открытия этого этапа следующие:

- опыты Г. Менделя и их математический анализ, создание логической модели и установление дискретности гена, обоснование абстрактной модели наследственного фактора;
- В. Иогансен: учение о генотипе и фенотипе организмов;
- У. Бэтсон: явление взаимодействия генов при развитии особи;
- Т. Х. Морган: создание хромосомной теории наследственности, ген как материальная структура хромосомы, метод картирования хромосом;
- Н. И. Вавилов: обосновал учение о мобилизации генетических ресурсов планеты для селекции и создания новых генотипов.



## Второй этап — эпоха неоклассицизма

Открытие влияния радиации на мутационный процесс (Надсон, Филиппов, Мёллер). Роль мутаций, скрещивания и отбора для структуры и эволюции популяций (Четвериков). Теория генетико-автоматических процессов (Дубинин, Ромашов, 1931, 1932), явление дрейфа генов (Райт, 1931). Генетическая теория отбора (Фишер, Холдейн, 1931–1935). Дробимость гена и его линейная модель (Дубинин, Серебровский и др., 1929–1934). Биохимические основы действия генов — афоризм «один ген — один фермент» (Бидл, Татум, 1945). Регуляция пола путем использования цитогенетических методов у тутового шелкопряда (Астауров, 1937). Открытие трансформирующей активности ДНК у пневмококков, что объяснило феномен трансформации (передачи наследственных признаков у бактерий через ДНК) (Эвери и др., 1944). Подтверждение генетической роли ДНК в опытах на бактериофаге Т2 (Херши, Чейз, 1952). Два последних открытия сыграли историческую роль в генетике — сместили центр основных интересов: вместо белка в качестве материального носителя наследственности стали признаваться молекулы ДНК, имеющиеся в хромосомах человека, у всех высших форм, у бактерий, а также многих видов вирусов и фагов.

## Краткие итоги достижений классической генетики

В рассматриваемый период представление о взаимодействии генов и признаков сводилось к схеме:

ген → признак,

по которой гены контролируют формирование признаков. Основой таких представлений явились результаты гибридологического анализа альтернативных признаков морфологии (форма, размер), а также окрасок, чувствительность к патогенам и др. Эти признаки наследуются по Менделю, они необратимы и устойчивы в онтогенезе, а внешние воздействия могли вызывать модификационные изменения фенотипа. Одним из ключевых понятий генетики является представление о генах. На основе данных классической генетики ряд авторов (Лобашев, 1967; Дубинин, 1976; Алиханян, Акифьев, Чернин, 1985; Ратнер, 1983) сформулировали положения теории генов и их основные особенности:

- 1) дискретность и ее проявления — независимость проявления от других генов и мутирования, неделимость и не смешиваемость при кроссинговере, автономность и выделенность в геноме;
- 2) способность к самоудвоению;

- 3) способность мутировать независимо от других генов, изменяя свое состояние как целое;
- 4) определенная локализация относительно других генов;
- 5) способность контролировать развитие признака, что отражено в наименовании генов.

В рамках этой канонической теории, так же как и в любой теории вообще, имелось ряд фактов, в нее не укладывающихся. Именно они и стали предвестниками крупных будущих открытий.

### Третий этап — эпоха молекулярной генетики

Анализ молекулы ДНК Дж. Уотсоном и Ф. Криком и его осмысливание на базе теории гена привели к революционной ситуации в генетике. Взаимодействие генетики, химии, физики и математики при анализе молекулярных структур обусловило рождение молекулярной генетики и молекулярной биологии в целом.

#### Основопологающие открытия молекулярной генетики

**1953** — Дж. Уотсон и Ф. Крик создали модель двойной спирали ДНК на основе рентгенограмм, полученных Р. Франклин и М. Уилкинсом. Открытие удостоено Нобелевской премией 1962 г. совместно с М. Уилкинсом.

**1956** — А. Корнберг открыл фермент ДНК-полимеразу, с помощью которого синтезировал биологически активную молекулу ДНК. Лауреат Нобелевской премии 1959 г. с С. Очоа.

**1958** — М. Мезельсон и Ф. Сталь доказали полуконсервативный механизм репликации ДНК

А. Леван и Д. Тийо получили точные данные о количестве и структуре хромосом человека.

**1959** — описан фермент РНК-полимераза, полимеризующий нуклеиновые кислоты (Вейс, Гурвич, Стивенс).

**1961** — Ф. Жакоб и Ж. Моно разработали модель оперона. Вместе с А. Львовым удостоены Нобелевской премии 1965 г.

**1965–1967** — Р. Холли выяснил первичную структуру аланиновой т-РНК, а А. А. Баев — дрожжевой валиновой т-РНК.

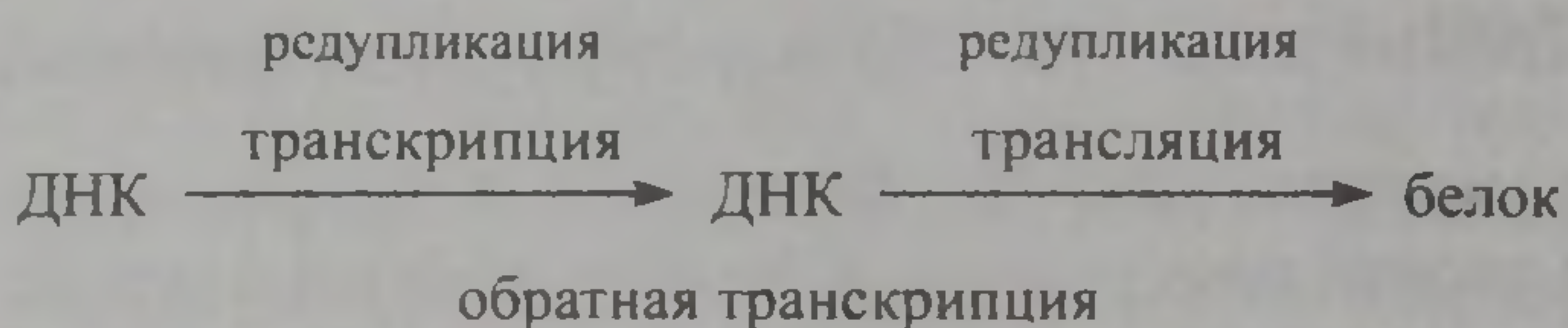
**1966** — А. Маккьюсик опубликовал каталог «Менделевское наследование у человека», содержащий описание 1500 генетических болезней у человека.

**1966** — М. Ниренберг осуществил основополагающие труды по расшифровке генетического кода (частичное участие Г. Кораны и С. Очоа). Синтезировал и испытал все 64 теоретически возможных тринуклеотида,

укомплектовал кодовый словарь. Лауреат Нобелевской премии 1968 г. совместно с Р. Холли и Г. Кораной.

1967 — М. Геллер открыл ДНК-лигазу — фермент, используемый для сшивания фрагментов ДНК.

Ф. Крик сформулировал центральную догму молекулярной генетики: генетическая информация передается по схеме ДНК → РНК → белок (у вирусов может быть несколько изменена), но никогда — от белка к РНК. Позже эта схема была дополнена представлениями об обратной транскрипции и репликации как ДНК, так и РНК. Она приобрела следующий вид:



Изучение деталей репликации, транскрипции и трансляции продолжается как важнейшая проблема молекулярной генетики.

1970 — Г. Темин и Д. Балтимор независимо открыли обратную транскриптазу в вирусе саркомы птиц — фермент, синтезирующий ДНК, используя комплементарную РНК в качестве матрицы в вирионах (вирусных частицах) РНК-содержащих вирусов. Лауреаты Нобелевской премии 1975 г. совместно с Р. Дюльбекко.

Х. Г. Корана впервые синтезировал химическим путем ген аланиновой т-РНК дрожжей. Лауреат Нобелевской премии 1968 г. совместно с Р. Холли и М. Ниренбергом.

В. Арбер, Д. Натанс и Г. Смит впервые выделили рестрикционную нуклеазу Hind II — фермент, разрезающий ДНК в строго определенных местах. С их помощью проведено картирование ДНК вируса SV-40. Лауреаты Нобелевской премии 1978 г.

1972–1973 — Х. Бойер, С. Козн и П. Берг разработали технологию клонирования ДНК. Берг впервые получил рекомбинантные молекулы ДНК бактериального вируса лямбда и обезьяньего вируса SV-40. Первый генно-инженерный эксперимент заключался в том, что в плазмидную ДНК встраивались фрагменты чужеродной ДНК (ДНК дрозофилы), в результате чего получены химерные плазмиды. Показано, что их можно ввести в клетки бактерии *E. coli* в функционально активном состоянии. Продемонстрирована принципиальная возможность клонирования в бактериях любого гена. Авторы удостоены Нобелевской премии 1980 г.

1975 — осуществлено первое клонирование к-ДНК.

1976 — В. А. Гвоздев и Д. Хогнесс открыли подвижные генетические элементы у дрозофилы.

1975–1977 — Ф. Сэнджер, А. Максам, У. Гилберт разработали методы быстрого определения длинных нуклеотидных последовательностей ДНК

(прямое секвенирование). Лауреаты Нобелевской премии 1980 г. Этими же авторами обнаружена мозаичная экзон-интронная структура гена. Впоследствии окажется, что это общее свойство генов эукариот.

1977 — Ф. Сэнджер с соавт. (Кембридж) установили последовательность нуклеотидов в геноме малого фага φX174 («фи-десять-174») (кольцевая одноцепочечная ДНК размером 5386 п. н.). В геноме фага обнаружены перекрывающиеся гены и «гены внутри генов». Это достижение принесло Ф. Сэнджеру вторую Нобелевскую премию и знаменовало собой начало новой эры в молекулярной генетике — эры геномики.

В этой же работе показано, что некоторые цистроны локализованы внутри других цистронов или в их интронах, или частично перекрываются с другими цистронами.

1978 — определена полная нуклеотидная последовательность генома вируса SV-40 размером 5243 п. н.

1980 — В геноме человека обнаружены гипервариабельные области ГЦ-(ГВО), содержащие короткие, обогащенные и тандемно повторяющиеся единицы. Их используют в качестве маркеров-зондов при картировании генов.

1980 — Дж. Шелл с соавт. доказали, что почвенная бактерия *Agrobacterium tumefaciens*, вызывающая у растений болезнь «корончатые галлы», способна переносить и встраивать часть своей Ti-плазмиды (Т-ДНК) в ядерный геном растений. Эта работа легла в основу генетической инженерии растений.

1981 — Определена полная нуклеотидная последовательность митохондриального генома человека (мт-ДНК).

1982 — Определена полная нуклеотидная последовательность генома бактериофага лямбда размером 48 502 п. н.

1984 — В. Макгинис, М. Скотт и А. Вейнер открыли ДНК-гомеобокс-содержащую ДНК (Нох-гены) или гомейотические гены. Гомеобокс ДНК найден у всех эукариот от дрожжей до человека (у дрозофилы и других размером 180 п. н.). Соответствующая последовательность из 60 аминокислот в кодируемых этими генами белках, обогащенная аргинином и лизином, была обозначена термином *гомеодомен*. Нох-гены млекопитающих составили 38 генов, собранных в 4 кластера, область по 120 кб каждый, расположены у человека в хромосомах II, VII, XII, XVII. Они играют формообразовательную роль в развитии и ответственны за построение общего плана тела животных.

1985–1988 — К. Муллис и Р. Сайки разработали метод амплификации (умножения числа копий) фрагментов нуклеиновых кислот *in vitro*, с помощью которого можно достаточно быстро (в течение нескольких часов) получить миллионы копий определенных нуклеотидных последовательностей-генов — метод полимеразной цепной реакции (ПЦР). Этот метод был назван изобретением века, удостоен Нобелевской премии (1993), ус-

корил реализацию программы «Геном человека», а также способствовал внедрению в практику клинической диагностики многих заболеваний высокоэффективных диагностических наборов нового поколения.

**1987** — Клонирование и определение нуклеотидной последовательности ДНК, выделенной из древней египетской мумии.

Созданы первые дрожжевые искусственные хромосомы — YAC (Yeast Artificial Chromosomes). Они сыграют большую роль как векторы для клонирования больших фрагментов геномов.

**1990** — Создан Международный консорциум по секвенированию генома человека (IHGSC). Определена полная нуклеотидная последовательность генома вируса осповакцины размером 192 т. п. н. Осуществлено В. Андерсоном первое успешное применение генной терапии для лечения больной с наследственным иммунодефицитом. К 1997 г. в США и Европе одобрено 100 проектов генотерапии.

**1992** — В Китае впервые начали выращивать в промышленных масштабах трансгенный табак, устойчивый к насекомым.

**1994** — В США зарегистрировано первое трансгенное растение, предназначенное для употребления в пищу — помидор «Флавр-Савр» с длительным сроком хранения.

**1995** — Закончено секвенирование геномов у бактерий *Mycoplasma genitalium* (0,6 млн п. н.) и первого самостоятельно существующего организма бактерии *Haemophilus influenzae* (1.830.137 п. н.).

### Становление геномики как самостоятельного раздела молекулярной генетики

**1996** — Определены полные нуклеотидные последовательности геномов:

- архебактерии *Methanococcus jannaschii* (1,7 млн п. н.),
- микоплазмы *Mycoplasma pneumoniae* (0,8 млн п. н.),
- бактерии *Synechocystis PCC6803* (3,6 млн п. н.),
- эукариот — дрожжей *Saccharomyces cerevisiae* (12,0 млн п. н.).

**1996** — Группой Я. Вилмута (Англия) получен клон овец (5 идентичных животных), донором ядер которых была культура эмбриональных клеток.

**1996** — Открытие С. Прузинером существования *прионов* — инфекционных агентов, вызывающих трансмиссивные спонгиозные (от «спонги» — губка) энцефалопатии человека и животных (скрейпи овец, губчатая энцефалопатия крупного рогатого скота, трансмиссивная энцефалопатия норки, куру человека). Возбудители этих наследственных болезней не содержат генов. Прион — особая изоформа нормального клеточного белка,

отличающаяся от него плохой растворимостью и склонностью к агрегации, которая и является причиной образования амилоидов в тканях мозга. Присвоена Нобелевская премия 1997 г.

**1996** — Секвенированные последовательности генов накапливаются в компьютерных *банках данных*, обрабатываются, сравниваются и классифицируются. Это основной источник информации о структуре генов (М. Дайхоф, У. Фитч и др.).

**1993–1996** — Обычно гены наследуются и передаются по «вертикальному» пути — от родителей к потомкам, но не исключены явления «горизонтального переноса», особенно у бактерий, когда переносчики генов — вирусы, плазмиды, мобильные генетические элементы — преодолевают барьеры видовой изоляции (М. Кидуэлл, 1993).

**1997** — Определены полные нуклеотидные последовательности геномов:

- архебактерий *Archaeoglobus fulgidis* (2,3 млн п. н.),
- *Methanobacterium thermoautotrophicum* (1,8 млн п. н.),
- бактерий *B. subtilis* (4,2 млн п. н.),
- *B. burgdorferi* (1,3 млн п. н.),
- *Helicobacter pylori* (1,7 млн п. н.),
- *Escherichia coli* (4,6 млн п. н.).

**1998** — Разработана технология синтеза моноклональных антител для терапии определенного типа метастазирующего грудного рака.

Определена полная нуклеотидная последовательность генома первого высшего организма — *Caenorhabditis elegans* (нематода) размером  $8 \times 10^7$  п. н.

Определена полная нуклеотидная последовательность геномов прокариот:

- архебактерии *Pyrrococcus horikoshii* (1,7 млн п. н.),
- бактерий *Aquilex aecolicus* (1,5 млн п. н.),
- туберкулезной палочки *Mycobacterium tuberculosis* (4,4 млн п. н.),
- *Rickettsia prowasekii* (1,1 млн п. н.),
- бледная спирохета *Treponema pallidum* (1,1 млн п. н.),
- *Chlamydia trachomatis* (1,0 млн п. н.).

**1999** — Расшифрован геном арабидопсиса размером  $8 \times 10^7$  п. н.

Трансгенные растения выращивают в 11 странах в промышленных масштабах на общей площади 39,9 млн гектаров (исключая Китай).

**2000** — Полностью секвенированы геномы дрозофилы (размер  $1,75 \times 10^8$  п. н.) и риса.

Американские и японские исследователи получили клон телят путем трансплантации ядер фибробластов уха 17-летнего быка.

2001 — Полное секвенирование генома человека двумя коллективами исследователей:

- 1) IHGSC (International Human Genome Sequencing Consortium — Международный консорциум по секвенированию генома человека.) — *Nature*, 2001. Vol. 409. P. 850–921;
- 2) Частная фирма Celera Genomics — *Venter J. C. et al. Science*, 2001. Vol. 291. P. 1304–1351.

Размер генома человека  $3,2 \times 10^9$  п. н.

2002 — Полностью секвенирован геном мыши размером  $2,2 \times 10^9$  п. н.

2006 — Секвенирован геном неандертальца.

## 2. Объекты и методы молекулярной генетики

Молекулярная генетика — это отрасль современной генетики, предметом исследований которой является молекулярный уровень организации наследственных структур и процессов. Она изучает принципы молекулярной организации и эволюции вирусов, прокариот и эукариот, а также молекулярные механизмы генетических процессов на разных уровнях организации живого. Перед молекулярной генетикой стоят следующие научно-практические задачи:

- расшифровка структуры геномов, пути их эволюции;
- создание библиотек генов: хромосомспецифических библиотек генов из ДНК отдельных хромосом, библиотек к-ДНК, геномных библиотек;
- изучение трансгеноза и конструкция новых геномов животных и растений;
- клонирование ДНК и получение продуктов клонирования эукариотических генов в клетках прокариот;
- клонирование соматических клеток животных;
- картирование хромосом и геномов;
- создание методов диагностики и лечения генетических болезней;
- изучение молекулярно-генетических основ онтогенеза, биоразнообразия, эволюции.

Исторически важной концепцией молекулярной генетики явилось соотношение генетической роли нуклеиновых кислот на весь мир живых организмов. Это краеугольное биологическое обобщение основано на многочисленных опытных данных:

- нуклеиновые кислоты ДНК и РНК, кодирующие наследственную информацию и способные к конвариантной редупликации локализованы в таких структурах, как плазмиды, вирусы, пластиды и митохондрии, центриоли, хромосомы;
- мутации возникают в результате спонтанных или индуцированных изменений нуклеиновых кислот;
- универсальность генетического кода для всех живых организмов и коллинеарность (полное соответствие) полипептидов кодирующим их генам;
- молекулы ДНК бактерий, вирусов, плазмид, митохондрий и пластид имеют кольцевидную форму, что отражается на их кольцевых генетических картах;
- фрагменты ДНК прокариот, внесенные посредством различных векторов-переносчиков в клетки эукариот, показали адекватное функционирование.

Природно ДНК является достаточно инертной молекулой (Заварзин, 2004). Заложенная в ней генетическая информация реализуется в жизненные функции под действием ферментов, а с наибольшей полнотой — в размножающемся организме. Нуклеиновые кислоты проявляют заложенные в них генетические функции только в живых клетках или в специально приготовленной бесклеточной системе, состоящей из клеточного экстракта с определенным набором ферментов и других белков (Уотсон, 1967).

Важным фактором, обусловившим эпохальные достижения молекулярной генетики, явилось внедрение и использование в экспериментах простых, технически удобных, одноклеточных объектов — вирусов, бактерий и протистов (низших грибов). По сравнению с многоклеточными и высшими организмами, они обладают следующими преимуществами:

- гаплоидное состояние, короткое время генерации и быстрая смена поколений — у бактерий около 30 мин. (для сравнения — у дрозофилы около 12 дней, у высших организмов длительность поколений — месяцы и годы);
- простота строения геномов и их небольшие размеры, методически несложное выделение нуклеиновых кислот в химически чистом виде и анализ их геномов;
- более легкая доступность при выделении и изучении просто детерминированных биохимических признаков у микроорганизмов по сравнению с морфологическими признаками высших организмов;
- подопытный материал при исследованиях на дрозофиле может достигать численность  $10^3$ – $10^4$  особей, тогда как популяции микроорганизмов насчитывают  $10^8$ – $10^{12}$  особей. Разрешающая способность в та-



ких экспериментах позволяет устанавливать события с вероятностью  $10^{-7}$ – $10^{-11}$  (Ратнер, 1983).

Основные методы молекулярной генетики включают генетические приемы скрещиваний, а также современные геномные технологии. К последним относятся методы:

- получения и обработки ДНК;
- поиска, выделения и идентификации определенных участков ДНК;
- выявления мутаций, в том числе методика полимеразной цепной реакции;
- картирования и скрининга генов.

### 3. Молекулярно-генетические аспекты организации жизни на Земле

Проблема жизни изучается представителями разных естественных наук с разных позиций — биологами, химиками, физиками, математиками, геологами и географами. В настоящее время доминируют две точки зрения на возникновение и развитие жизни:

- химическая точка зрения, основанная на достижениях молекулярной биологии и генетики, вскрывших механизмы наследственности, обмена веществ, размножения, роста и развития организмов;
- геохимическое направление, рассматривающее взаимодействие организма с внешней средой на атомно-молекулярном уровне, ведущее начало от Ж.-Б. Ламарка и сформулированное В. И. Вернадским на уровне достижений естествознания XX века.

Углубленное понимание элементарных процессов путем редукционизма и сведения их на суборганизменный, молекулярный уровень, что проявляется в области генетики, биохимии и биофизики, обозначило два аспекта в понимании явлений жизни:

- жизнь как особое состояние материи с присущими ей отличительными признаками и «носителями жизни»;
- с другой стороны, жизнь как специфическое природное явление с пространственно-временными характеристиками, необходимыми и достаточными для существования этого особого состояния материи.

При сравнении свойств живых организмов Ж.-Б. Ламарк выделил десять присущих им отличий от свойств неорганических тел. Современной попыткой охарактеризовать основные присущие явлениям жизни признаки является формулирование Б. М. Медниковым (1982) так называемых *аксиом теоретической биологии*.

### 1. Аксиома А. Вейсмана

У всех живых организмов проявляется единство фенотипа и программы его построения (генотипа), которое передается в ряду поколений.

### 2. Аксиома Н. К. Кольцова

Генетическая программа формируется по матричному принципу, при этом ген предшествующего поколения представляет собой матрицу для построения гена будущего поколения.

### 3. Первая аксиома Ч. Дарвина

В ходе передачи в ряду поколений наследственных программ по различным причинам происходят их случайные и направленные изменения, которые могут случайно оказаться удачными в условиях данной среды.

### 4. Аксиома Н. В. Тимофеева-Ресовского

При формировании фенотипа проявляется принцип усиления, то есть случайные изменения генетических программ многократно усиливаются.

### 5. Вторая аксиома Ч. Дарвина

Многократно усиленные изменения генетических программ подвергаются отбору факторов внешней среды.

Перечисление аксиом Б. М. Медникова можно продолжить, но это не убедит нас в охвате данной проблемы.

В. И. Вернадский (1980) с позиций современного естествознания выделил шестнадцать отличий живых организмов («живого вещества» по Вернадскому, т. е. биоты) от косных естественных (неорганических) тел. Эти признаки подразделяются на три группы: различия по энергии, различия по химическим проявлениям, а также различия по пространству-времени. Особое значение В. И. Вернадский придавал открытому Л. Пастером явлению диссимметрии разных молекулярных изомеров у живых организмов. Проблема сводится к тому, что количество правых и левых изомеров одного и того же химического соединения, одновременно образующихся в природных условиях, одинаково. Правизна и левизна в них химически и геометрически тождественны, изомеры присутствуют в одинаковом количестве и химически одинаковы.

Такая способность молекул существовать в двух зеркально-противоположных формах называется *хиральностью*. Оказалось, что живой природе присуща абсолютная *хиральная чистота*: белки содержат только ле-

606498

С.Торайғыров

МУ-дің  
Многократно усиленные  
атындағы ғылыми

КІТАПХАНАСЬ

вые аминокислоты, а нуклеиновые кислоты — только «правые» сахара. По Вернадскому, состояние пространства жизни обладает диссимметрией, т. е. особой геометрией, которая не является обычной геометрией Эвклида (*принцип Кюри*). Поддержание длительности жизни в течение всего геологического времени делением, почкованием или рождением является основным проявлением особого пространства-времени живых тел, его особой не-Эвклидовой геометрией (*принцип Реди*).

В последние годы проблема хиральной чистоты живых организмов приобрела новое теоретическое значение в области молекулярной эволюции, использующей аппарат молекулярной генетики. Физики-теоретики В. И. Гольданский, Л. Л. Морозов, В. В. Кузьмин (1985, 1986) и др. развили представление о том, что жизнь характеризуется двумя основными свойствами — способностью к репликации и хиральной чистотой. Первый атрибут жизни — способность к репликации — или более точно, принцип конвариантной редупликации дискретно построенных кодов наследственной информации (Тимофеев-Ресовский, 1969, 1983) — означает самовоспроизведение ДНК и живых частиц с изменениями и наследственными вариациями, осуществляемые на основе матричного принципа. Этот принцип будет нами подробно рассмотрен в главе «Структура и функционирование генетического материала». Здесь мы остановимся на втором атрибуте жизни — феномене оптической изомерии и хиральности биологических молекул.

*Положение гидроксильной группы* при асимметрическом атоме углерода обозначается буквами D (когда OH-группа расположена справа) и L (когда OH-группа расположена слева). Эти буквы определяют конфигурацию вещества. С другой стороны, направление вращения поляризованного луча обозначается (+) — вращение вправо, или (–) — вращение влево.

*Энантиоморфные типы оптических изомеров* (например, изомеры D — и L — винной кислоты) относятся друг к другу, как предмет к своему зеркальному отражению. Они обладают одинаковыми физическими свойствами и вращают плоскость поляризации света на равную величину, но в противоположных направлениях. Оптическая изомерия представляет собой один из видов *стереоизомерии*, а другим ее типом является *геометрическая*, или *цис-транс-изомерия*. Четырехвалентные атомы углерода могут присоединять к себе четыре других атома или группы, которые в пространстве имеют вид *тетраэдра* (рис. 1).

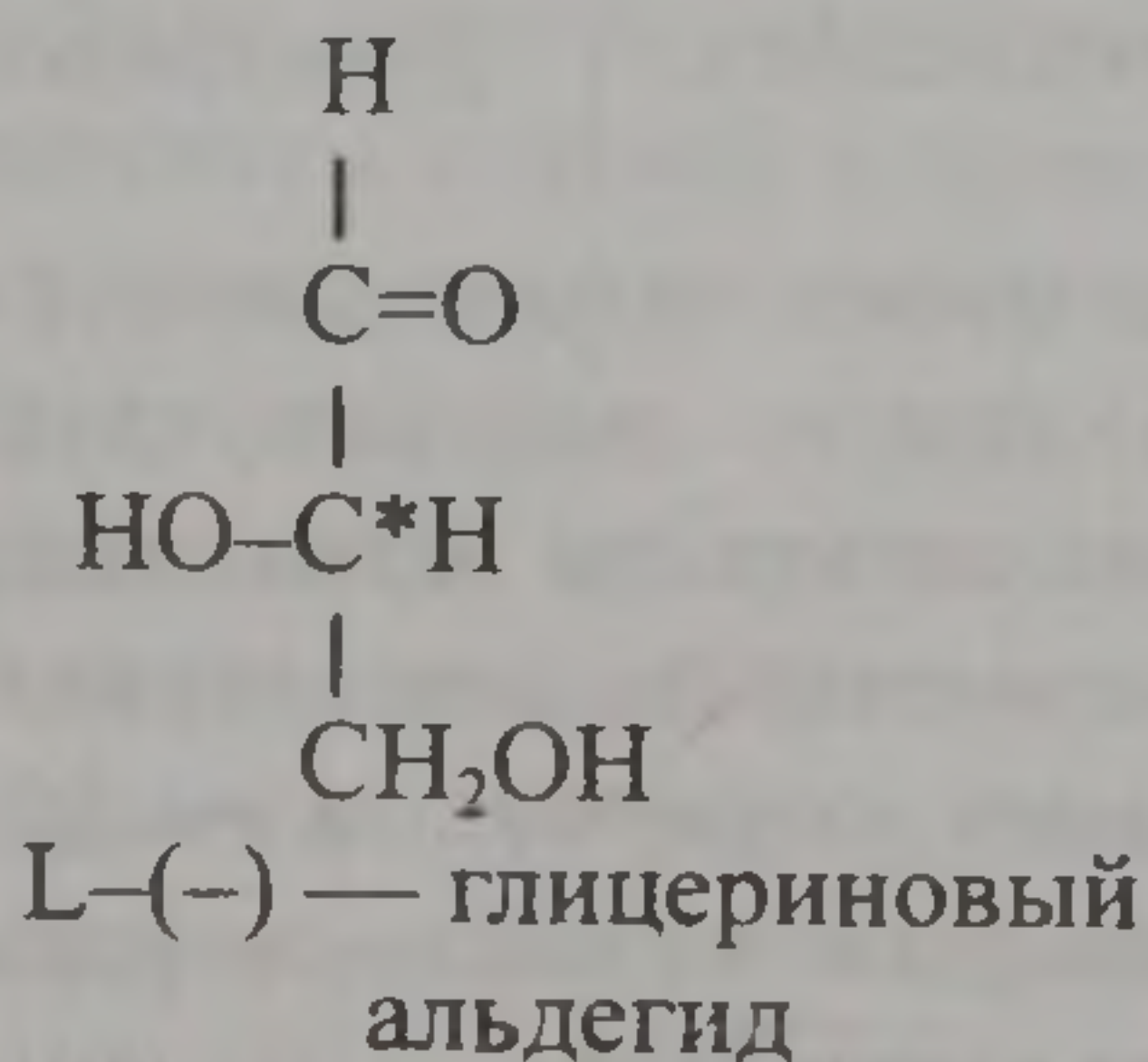
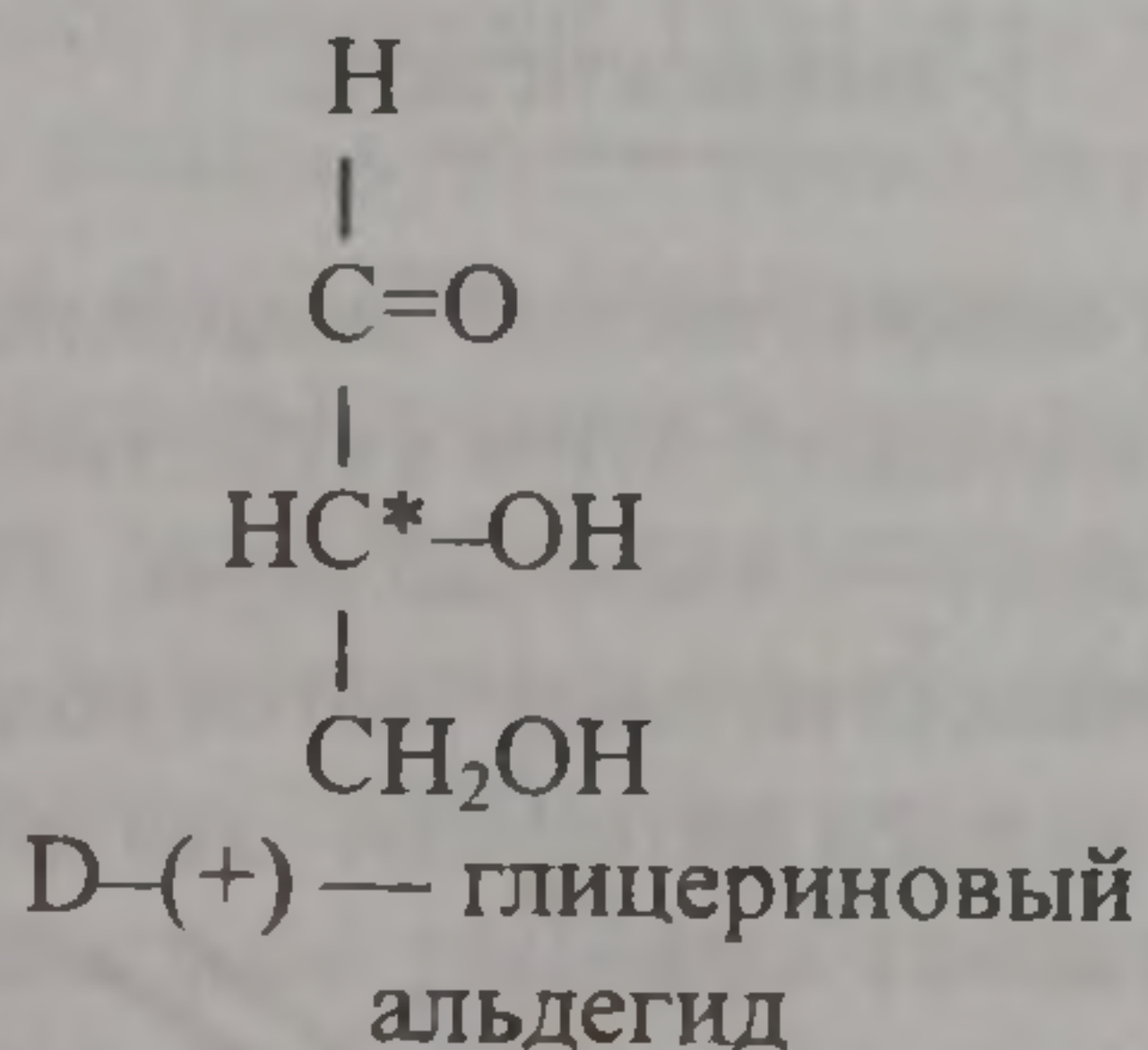
### Хиральность биологических молекул

Способность кристаллов кварца ( $\text{SiO}_2$ ) и растворов природных органических веществ (сахара, камфары) вращать *плоскость поляризации света* была установлена в XIX веке. Позднее Л. Пастер обнаружил два типа кристаллов винной кислоты, разделил их и показал, что их растворы вра-



Рис. 1. Хиральность асимметрического атома углерода

щают в разных направлениях плоскость поляризации света: вправо — «правые», а влево — «левые». *Рацемат*, т. е. эквимольный раствор тех и других, плоскость поляризации света не вращает. В 1874 г. Вант-Хофф и Ле-Бель пришли к выводу, что причиной оптической активности изомеров-антиподов является асимметрия молекулы с четырехвалентным углеродом. Такой атом углерода, связанный с четырьмя различными радикалами, называется *асимметрическим* и обозначается в формулах  $-C^*$ . Впоследствии *тетраэдричность связей углерода* была подтверждена точными физическими измерениями. Вращение плоскости поляризации света наблюдают и измеряют с помощью специального прибора — *поляриметра* (призма Николя — «николь»). В качестве стереохимического стандарта принят глицериновый альдегид:

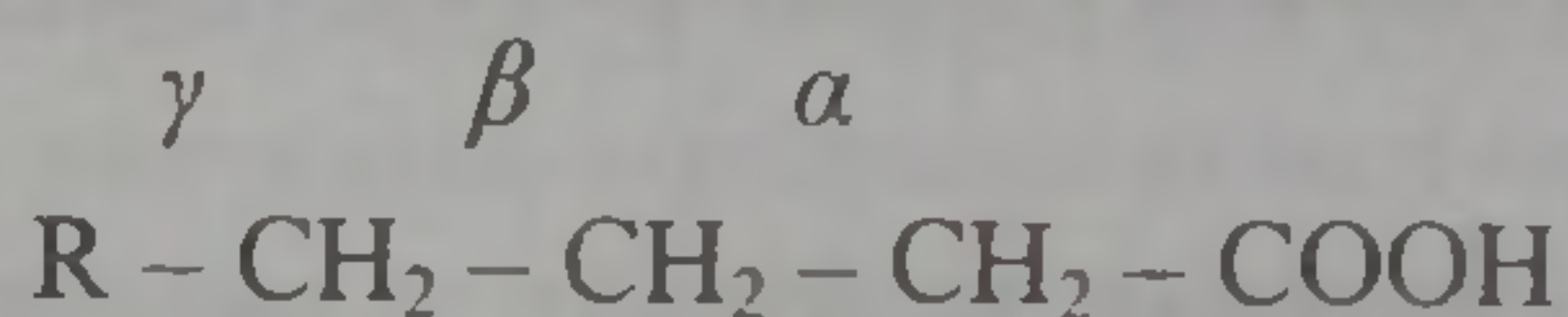


На рис. 1 две связи направлены в сторону от нас (заканчиваются острием), а две в нашу сторону (острием начинаются от атома углерода). После присоединения к атому углерода четырех разных атомов или групп окажется, что возможны два варианта последовательности их присоединения. В варианте L-конфигурации группы COOH, NH<sub>2</sub>, R и H расположены по часовой стрелке. Эта же последовательность в варианте D-конфигурации расположена против часовой стрелки. Обе структуры имеют практически одинаковые химические свойства, но их пространственные конфигурации представляют зеркальное отражение друг друга и напоминают

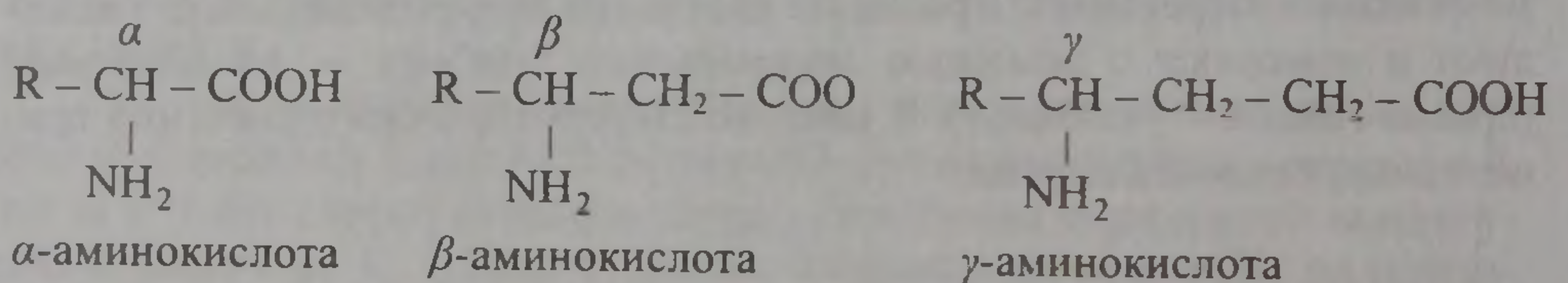
правую и левую руку. Отсюда происходит термин «хиральность» от греческого слова «хира» — рука (по аналогии «хиромантия» — гадание по руке). Во всех многочисленных белках у всех живых организмов наблюдается явление *хиральной чистоты аминокислот*: белки состоят из  $\alpha$ -аминокислот L-конфигурации.

Аминокислоты, а именно 20 канонических пептидогенных аминокислот, являются отдельными звеньями белков с пептидными связями между ними:  $-\text{COOH} + \text{NH}_2 - \rightarrow -\text{CONH}_2 -$ . Они имеют единый тип строения — это органические кислоты, содержащие аминогруппы.

Общая формула аминокислоты



В этой формуле в качестве R могут быть боковые цепи, COOH-карбоксильная (кислотная) группа, а группы CH<sub>2</sub> — фрагменты углеводородной цепи. На основе Женевской номенклатуры нумерации химических систем, идущие после карбоксильной CH<sub>2</sub>-группы, обозначаются как  $\alpha$ ,  $\beta$ ,  $\gamma$ . Поэтому аминогруппа NH<sub>2</sub> в аминокислотах может находиться в положении  $\alpha$ ,  $\beta$ ,  $\gamma$ :



Установлена хиральность девятнадцати канонических аминокислот — только одна из двадцати, глицин (гли), не обладает этим свойством. Зеркальные антиподы девятнадцати аминокислот — *энантиомеры*. При химическом синтезе симметричных молекул вещества получается *рацемическая смесь*, содержащая по 50 % правого и левого антипода. Это объясняется *вторым началом термодинамики*: рацемат отвечает максимальной энергии смешения.

В клетках живых организмов хиральность присуща нуклеиновым кислотам, белкам и ряду низкомолекулярных соединений. Углеводы дезоксирибоза и рибоза в составе ДНК и РНК фигурируют в D-форме, аминокислоты белков — в L-форме, тогда как азотистые основания аденин (А), тимин (Т), гуанин (Г) и цитозин (Ц) имеют плоскую конфигурацию и лишены хиральности. В процессах метаболизма клетка усваивает лишь те антиподы, которые соответствуют структуре ее биологических молекул. Организм различает D- и L-антиподы, а вещества в одной форме могут быть безвредными, а в другой — ядовитыми.

В лабораторных условиях *асимметрический синтез*, т. е. выделение *in vitro* чистых антиподов из рацемических смесей, проводят с помощью хиральных веществ биологического происхождения, обычно — алкалоидов. Например, для разделения рацемической смеси (D, L) необходимо воздействовать изомером L'. Получают:  $(D, L) + L' \rightarrow DL' + LL'$ . Полученные смеси DL' и LL' уже не являются зеркальными антиподами — ими были бы DL' и D'L. Физико-химические свойства смесей DL' и LL' различаются, но их в свою очередь также можно разделить, например, путем кристаллизации.

Хиральность в живой природе определяется хиральной регуляцией ферментативных процессов и существует на молекулярном и более высоких уровнях организации жизни (Волькенштейн, 1988). Для разделения антиподов необходимо асимметрическое воздействие химического соединения или существа, различающего правое и левое. У моллюска корненожки *Neogloboquadrina pachyderma* раковины, закрученные по часовой стрелке, образуются при температуре ниже 7°C, а закрученные против часовой стрелки — при температуре выше 7°C. Л. Пастер в 1848 г. отсортировал чистые зеркальные антиподы винной кислоты, сам при этом играя роль асимметрического фактора: *человек сам хирален* и знает разницу между левым и правым.

Наличие в белках только левых аминокислот объясняется удобством для матричного синтеза. Некоторые организмы могут синтезировать сложные пептидные цепочки специального назначения *нематричным путем*, например, антибиотики типа грамицидина, а также пептиды, слагающие оболочки бактерий. При этом снимаются жесткие запреты матричного синтеза и используются неканонические аминокислоты в левой и правой форме. Так, в состав кольцевой молекулы грамицидина входят два остатка неканонической аминокислоты *орнитина*, также *правый фенилаланин (D-фен)*. Другим примером является наличие в стенке капсул сибиреязвенной бактерии полипептида, построенного из правой глутаминовой кислоты (*D-глу*) (Медников, 2004). Имеется свыше 150 аминокислот, которые никогда не встречаются в составе белков, но могут находиться в различных клетках и тканях в свободном или связанном состоянии ( $\beta$  — аланин,  $\gamma$  — аминокислотная кислота, цитруллин, орнитин, 3-метилгистидин и др.) (Кунижев, Денисова, Андрусенко, 2005).

В настоящее время возникновение и фиксация хиральности рассматривается в свете общей теории добиологической эволюции и теории информации. Так, в начале возникновения порядка из хаоса (т. е. информации) выбор зеркального антипода означал создание информации, равной *1 бит на молекулу мономера*. Раз возникнув, хиральность как флуктуационное отклонение от равномерного рацемического распределения, неограниченно нарастала в самовоспроизводящейся системе и закрепились, так как хи-

ральные системы по сравнению с рацемическими более специфично взаимодействуют с окружающей средой (Волькенштейн, 1988; Романовский и др., 1984). Многочисленные данные из области физики, химии и молекулярной генетики подтверждают вывод о том, что важнейшее *атрибутивное свойство живой материи* — процесс конвариантной редупликации — оказывается возможным лишь в хирально чистой среде.

### Ключевые слова и термины

Молекулярная генетика — область изучения, основные понятия, этапы развития; аксиомы теоретической биологии по Б. М. Медникову; «живое вещество» по В. И. Вернадскому; хиральность, хиральная чистота, рацемическая смесь, энантиомеры; конвариантная редупликация.

Часть первая

---

---

# Структура и функционирование генетического материала

---

---

Нет другой молекулы в истории науки, которая достигла бы иконного статуса ДНК... История породила всего несколько суперимиджей, которые настолько пронизали наше... сознание, что они далеко превзошли свое первоначальное содержание. Это символизировано Монай Лизой, изображенной Леонардо да Винчи примерно в 1503 году. Двойная спираль ДНК не имеет равных в качестве символа биологических наук. Оба символа говорят публике много больше, чем об их узкоспециализированном мире и оба несут громадный багаж ассоциаций.

*Kemp M. Mona Lisa of modern science // Nature. 2003. V. 421. P. 416–420.*



## Структура нуклеиновых кислот ДНК и РНК

Принято считать, что молекулярная генетика и молекулярная биология в целом начались со статьи американского биолога Джеймса Уотсона и английского физика Френсиса Крика «Молекулярная структура нуклеиновых кислот: структура дезоксирибонуклеиновой кислоты» в журнале *Nature* в марте 1953 г. Она занимала объем менее двух страниц и состояла из 900 слов. В статье они описали созданную ими на основе рентгенограмм, полученных Р. Франклин и М. Уилкинсом, теоретическую модель строения молекулы ДНК. Согласно их «научной фантазии», основанной на определенных твердо установленных фактах (Тарантул, 2003), молекула ДНК должна состоять из двух генетических полимерных цепочек (двойная спираль). Позднее Д. Уотсон (1967) так оценил эту модель: «Она легко объясняла, каким образом генетическая информация может быть записана в молекулах ДНК, и в то же время позволяла выказать предположение о химическом механизме самовоспроизведения этих молекул. С этого момента начался совершенно новый этап в изучении гена. Генетики, строя свои гипотезы, могли теперь исходить уже не только из результатов генетических скрещиваний. Они получили возможность истолковывать эти результаты, сопоставляя их со структурой ДНК» (с. 250).

Звенья каждой цепи ДНК состоят из нуклеотидов, каждый из которых содержит по одному из четырех азотистых оснований — аденин (А), гуанин (G), тимин (Т) и цитозин (С). Последовательность нуклеотидов может быть любой, но она строго связана с последовательностью нуклеотидов в другой (парной) полимерной цепи на основе правила *взаимной комплементарности* — аденин связан с тиминем (А–Т), а гуанин с цитозином (G–С). В природе имеется много примеров комплементарности, например, рука и перчатка, ключ и замок, гнезда розетки и штырьки вилки. Далее будет пока-

зано, что именно комплементарность оснований определяет механизм наследственности, а через него — все основные законы биологии. Модель Уотсона—Крика решала проблемы кодирования генетической информации и удвоения (репликации) гена. В 1962 г. Д. Уотсон, Ф. Крик и М. Уилкинс награждены Нобелевской премией (Р. Франклин к этому времени умерла).

## Структура ДНК

Модель строения ДНК Уотсона и Крика учитывала ряд предшествующих открытий. Одним из них были *правила Э. Чаргаффа* (1951) о сумме пуриновых и пиримидиновых оснований в составе ДНК разных видов. На основе этих правил, количество аденина в молекуле ДНК всегда равно количеству тимина ( $A = T$ ), а количество гуанина равно количеству цитозина ( $G = C$ ). При этом сумма пуриновых оснований равна сумме пиримидиновых оснований ( $A+G = T+C$ ). Любая молекула ДНК по ее составу может быть охарактеризована на основе отношения  $(G+C)/(A+T)$ , которое постоянно для каждого вида варьирует в пределах 26–74 мольных %.

С другой стороны, данные Р. Франклин и М. Уилкинса (1951) по рентгеноструктурному анализу показали, что ДНК имеет форму регулярной спирали диаметром 20 ангстрем, делающей один полный оборот каждые 34 ангстрема (3,4 нм). Расстояние между соседними нуклеотидами в цепи составляло 3,4 ангстрема (0,34 нм), следовательно, в одном витке должно быть 10 нуклеотидов. *Размерность*: 1 нм =  $10^{-9}$  м. В цепи ДНК сахарофосфатный остов расположен дальше от оси спирали, чем азотистые основания, поэтому в двойной спирали имеются желобки (бороздки) — большой и малый (рис. 2).

Нуклеиновые кислоты являются биологическими полимерами, мономерами которых служат *нуклеотиды*. Каждый нуклеотид состоит из трех частей — моносахарида, остатка фосфорной кислоты и азотистого основания (рис. 3). Сахар, входящий в состав нуклеотида — пентоза — может присутствовать в одной из двух форм: D-рибоза в РНК и D-2-дезоксирибоза в ДНК.

Азотистые основания являются производными одного из двух соединений — пурина и пиримидина. В нуклеиновых кислотах преобладают два пуриновых основания — *аденин (А)* и *гуанин (G)* и три пиримидиновых — *цитозин (С)*, *тимин (Т)* и *урацил (U)*. В дезоксирибонуклеотидах ДНК присутствуют А, G, C и T, а в рибонуклеотидах — А, G, C и U.

Длинные полинуклеотидные цепочки ДНК и РНК образуются при соединении нуклеотидов между собой с помощью *фосфодиэфирных связей* за счет 3'-ОН одного нуклеотидного остатка и 5'-ОН другого. Цепи ДНК и РНК обладают определенной полярностью, или направлением, благодаря которым каждая полинуклеотидная цепь имеет 5'-конец и 3'-конец. Благодаря такой полярности появляется возможность определить направление

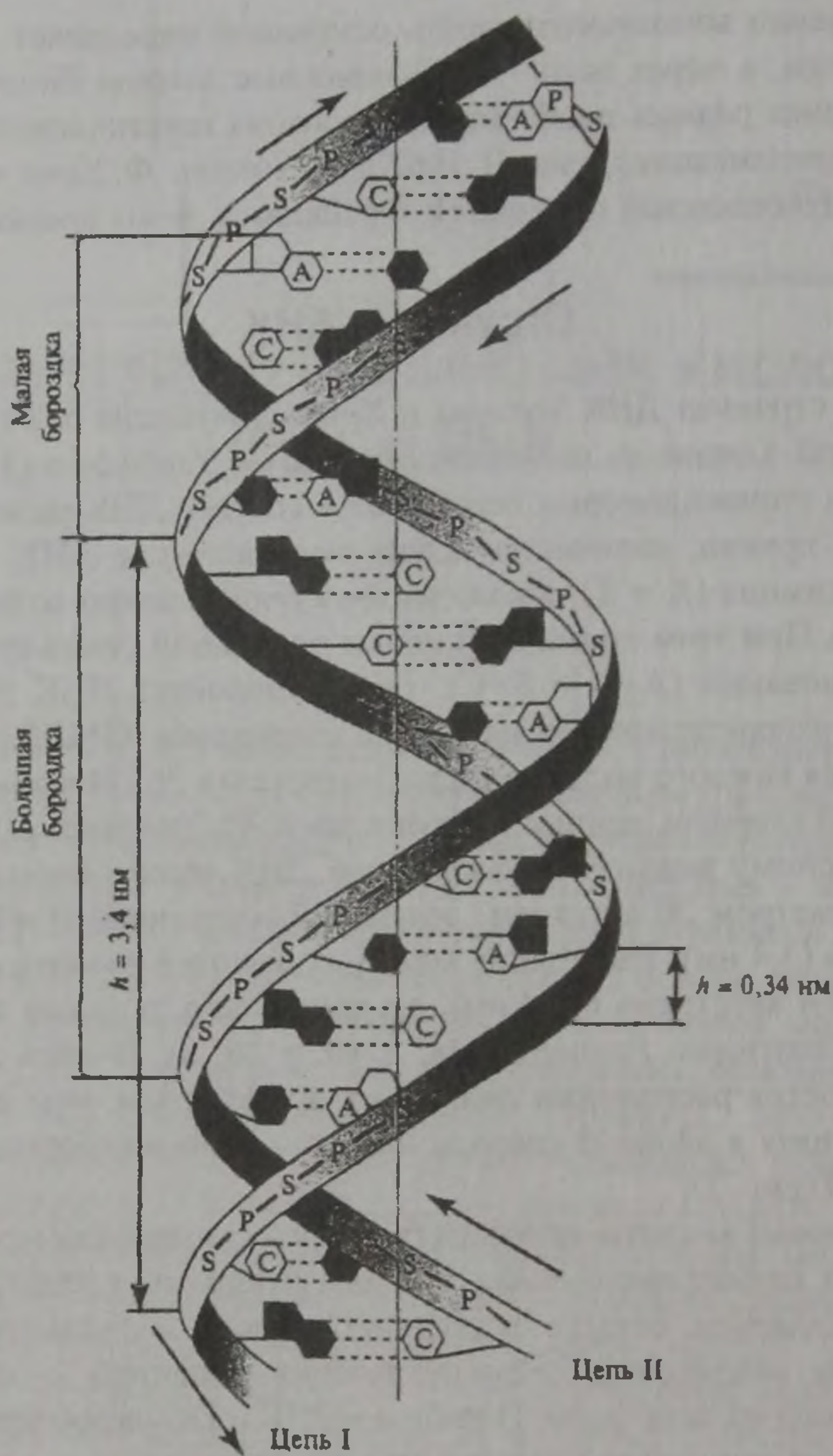


Рис. 2. В-форма двойной спирали ДНК: P — фосфат, S — дезоксирибоза

чтения последовательности нуклеотидов. Большинство ферментов, участвующих в синтезе и гидролизе нуклеиновых кислот, работают в направлении от 5'-к 3'-концу (5'→3') цепочки нуклеиновой кислоты. Согласно принятому соглашению, последовательность нуклеотидов в цепочках нуклеиновых кислот тоже читается в направлении 5'→3' (рис. 4). С точки зрения биохимии, такое противоположное направление цепей в двойной спирали ДНК является антипараллельным (рис. 4).

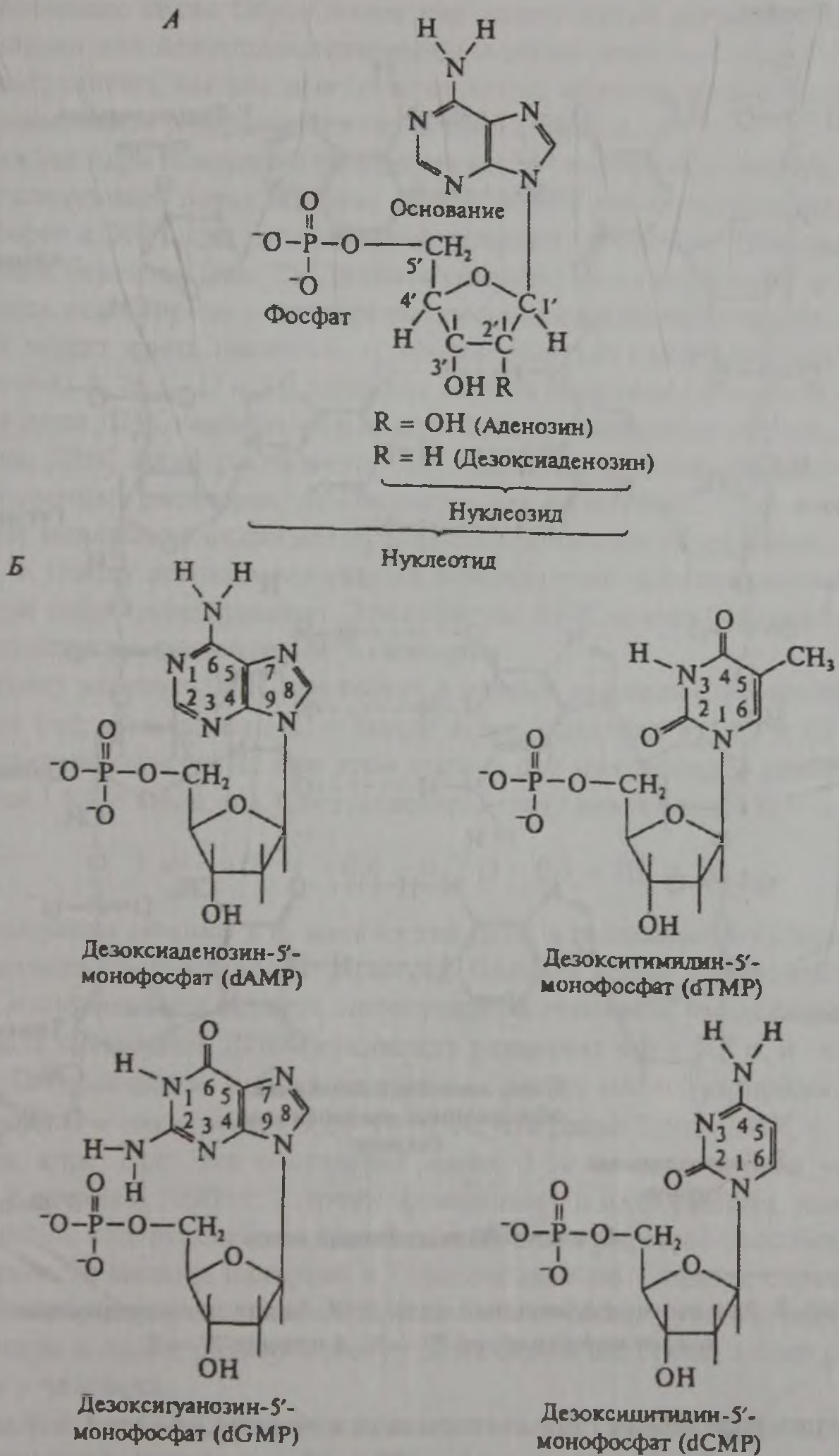


Рис. 3. Структурные формулы нуклеозидов и нуклеотидов: А — химическое строение нуклеотида; Б — дезоксирибонуклеотиды ДНК; цифрами обозначена нумерация атомов углерода

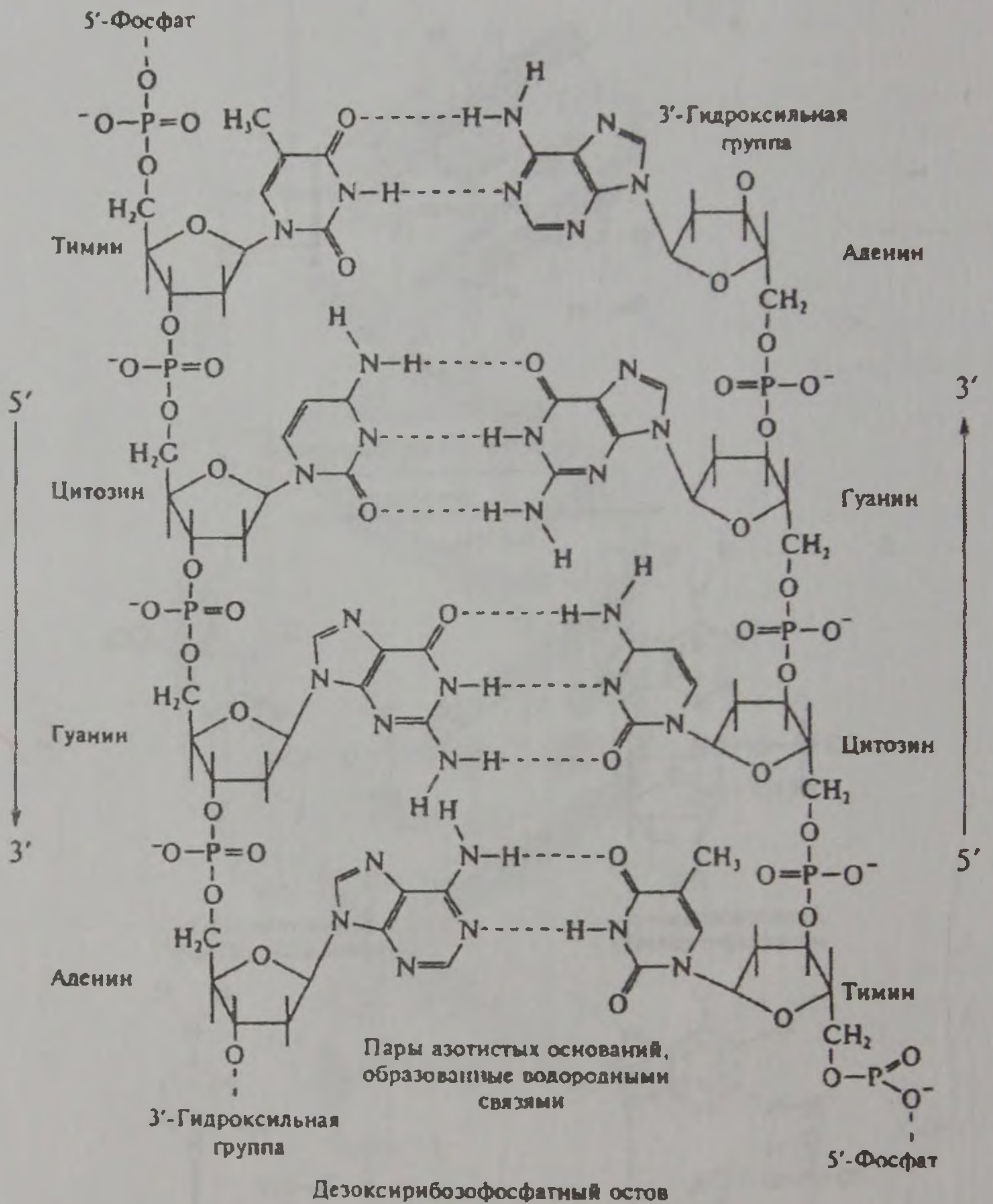


Рис. 4. Две антипаралельные цепи ДНК. Левая цепь сверху вниз имеет направление  $5' \rightarrow 3'$ , а правая  $3' \rightarrow 5'$

Реакция взаимодействия азотистых оснований G с C и A с T получили название *спаривание оснований* с образованием комплементарных пар. Между основаниями, расположенными друг напротив друга в противоположных цепях, формируются специфические водородные связи: в АТ-паре основания соединены двумя водородными связями, а в GC-паре имеются

три водородные связи. Образование пар между двумя пуринами, двумя пиримидинами или некоплементарными основаниями А–С или G–T стерически затруднено, так как при этом не могут образовываться подходящие водородные связи и нарушается геометрия спирали.

Каждая пара оснований повернута на  $36^\circ$  вокруг оси спирали относительно следующей пары, поэтому 10 пар оснований составляют один полный оборот в  $360^\circ$ . Две цепи двойной спирали уложены с наличием малой и большой бороздок (рис. 2). Двойная спираль *правосторонняя*: если смотреть вдоль оси спирали — повороты следуют по часовой стрелке. Молекула ДНК может иметь несколько конформационных (полиморфных) состояний (формы А, В, С, D и Z), тогда как данное описание относится к В-форме. Две цепи ДНК связаны между собой нековалентными связями, поэтому молекула ДНК легко распадается на отдельные цепочки при нагревании или в щелочных растворах (*денатурация или плавление*).

При медленном охлаждении (*отжиг*) цепи способны вновь ассоциировать, и между комплементарными основаниями восстанавливаются водородные связи (*ренатурация*). Эти свойства ДНК имеют большое значение для методологии *генетической инженерии*.

Размер молекул ДНК выражают в разных единицах измерения: *пикограммах* (пг), *дальтонах* ( $1D = \text{массе атома водорода} = 1,67 \times 10^{-24} \text{ г}$ ) или *парах нуклеотидов* (п. н., при этом тысяча пар нуклеотидов равна 1 килобазе, или  $1 \text{ Кб} = 1 \text{ т. п. н.}$ ). Соотношения между ними следующие:

$$1 \text{ пг} = 10^{-12} \text{ г} = 0,6 \times 10^{12} D = 0,9 \times 10^9 \text{ п. н.}$$

Измерения *генома*, т. е. количества ДНК в гаплоидном наборе хромосом, выявили следующие его размеры. Полный геном кишечной палочки *E. coli*, излюбленного объекта молекулярной генетики, представлен одной кольцевой молекулой ДНК (нуклеоид) размером  $4,6 \times 10^6$  п. н. и длиной 1,4 мм. В гаплоидном наборе человека — а это «венец природы», «мера всех вещей»! — содержится  $3,2 \times 10^9$  п. н., что равно 3,5 пг ДНК, т. е. в диплоидном ядре 7 пг. Это составляет менее 1% от веса клетки человека, равной в среднем 1000 пг. С точки зрения теории информации, количество информации, содержащейся в одной клетке, типографским способом можно воспроизвести мелким шрифтом в 1 тысяче книг по 1 тысяче страниц в каждой. В живой природе это не наибольший геном по размеру: например, у саламандры и лилии размер молекул ДНК одной клетки во много раз больше, чем у человека.

Важной стадией изучения генома является определение последовательности нуклеотидов в ДНК и РНК, т. е. их первичной структуры (*секвенирование*). Эта последовательность азотистых оснований уникальна для каждой нуклеиновой кислоты, так как в ней записана генетическая программа организмов. Разработаны современные методы секвенирования нук-

леиновых кислот и белков (последовательность расположения аминокислотных остатков). Проблему последовательности нуклеотидов и аминокислот решают также с помощью высокоэффективных компьютерных программ, пригодных для поиска структурных и биологических особенностей этих последовательностей. Созданы международные банки о последовательностях нуклеотидов в ДНК разных организмов (Gen Bank, EMBL/DDBJ) и о последовательностях аминокислот в белках (PIR/Swiss Pot, на основе нуклеотидной последовательности соответствующих генов).

## Хроматин. Компактизация молекулы ДНК

В неделящихся клетках каждая хромосома состоит из одной молекулы ДНК. Она связана с основными (гистоновыми) и кислыми (негистоновыми) белками, а также с РНК. До 60–80 % хромосомных белков представлены *гистонами*, которые обеспечивают компактную упаковку ДНК в клетках. Они подразделяются на 5 разновидностей — H1, H2A, H2B, H3 и H4. Гистоны обогащены аминокислотами с основными (аргинин, лизин) и гидрофобными (валин и др.) радикалами, имеют небольшую молекулярную массу.

В интерфазных ядрах клеток эукариот хромосомы не различимы, и их выделяют все вместе как *хроматин*. При этом различают *гетерохроматин* и *эухроматин*.

Гетерохроматин представляет собой сильно конденсированные и поэтому генетически неактивные участки хромосом. Они имеют вид плотных глыбок, многие из которых находятся на периферии ядра и прилежат к ядерной оболочке.

Эухроматин — это функционально активные, деконденсированные, светлые участки хромосом, расположенные между глыбками гетерохроматина, доступные для ферментов и белковых факторов транскрипции. В тех локусах эухроматина, на которых в данный момент протекает репликация, транскрипция или репарация, ДНК высвобождается из взаимодействия с гистонами.

Негистоновые (кислые) белки хромосом, насчитывающие несколько сотен — это многочисленные ферменты, обеспечивающие процессы репликации, модификации, транскрипции и репарации ДНК. Самой разнообразной среди них по составу является группа *регуляторных белков*, которые контролируют активность ферментов и доступность участков хромосом (ДНК) для этих ферментов.

ДНК хромосом клеток эукариот свойственна высококомпактная упаковка. Так, общая длина 46 хромосом человека на стадии интерфазы, т. е. в деконденсированном состоянии, составляет около 190 см. Длины разных хромосом сильно различаются в зависимости от их размера (средний размер равен  $190/46 \approx 4$  см). Длина самой маленькой хромосомы № 22 равна

1,4 см в интерфазе, тогда как на стадии метафазы она укорачивается в 7000 раз до длины 2 мкм. От интерфазы до метафазы в митозе и мейозе хроматин проявляет различные уровни своей упаковки, организации, сохраняя эффективность основных генетических процессов. Компактизация (спирализация) молекулы ДНК в составе хромосомы проходит пять этапов, или уровней, в ходе клеточного цикла (рис. 5).

Низшим уровнем иерархической организации хромосом является *нуклеосомный*. Когда хроматин имеет вид «бусин», нанизанных на нить ДНК. Хроматин проходит уровни структурной организации в виде фибрилл и

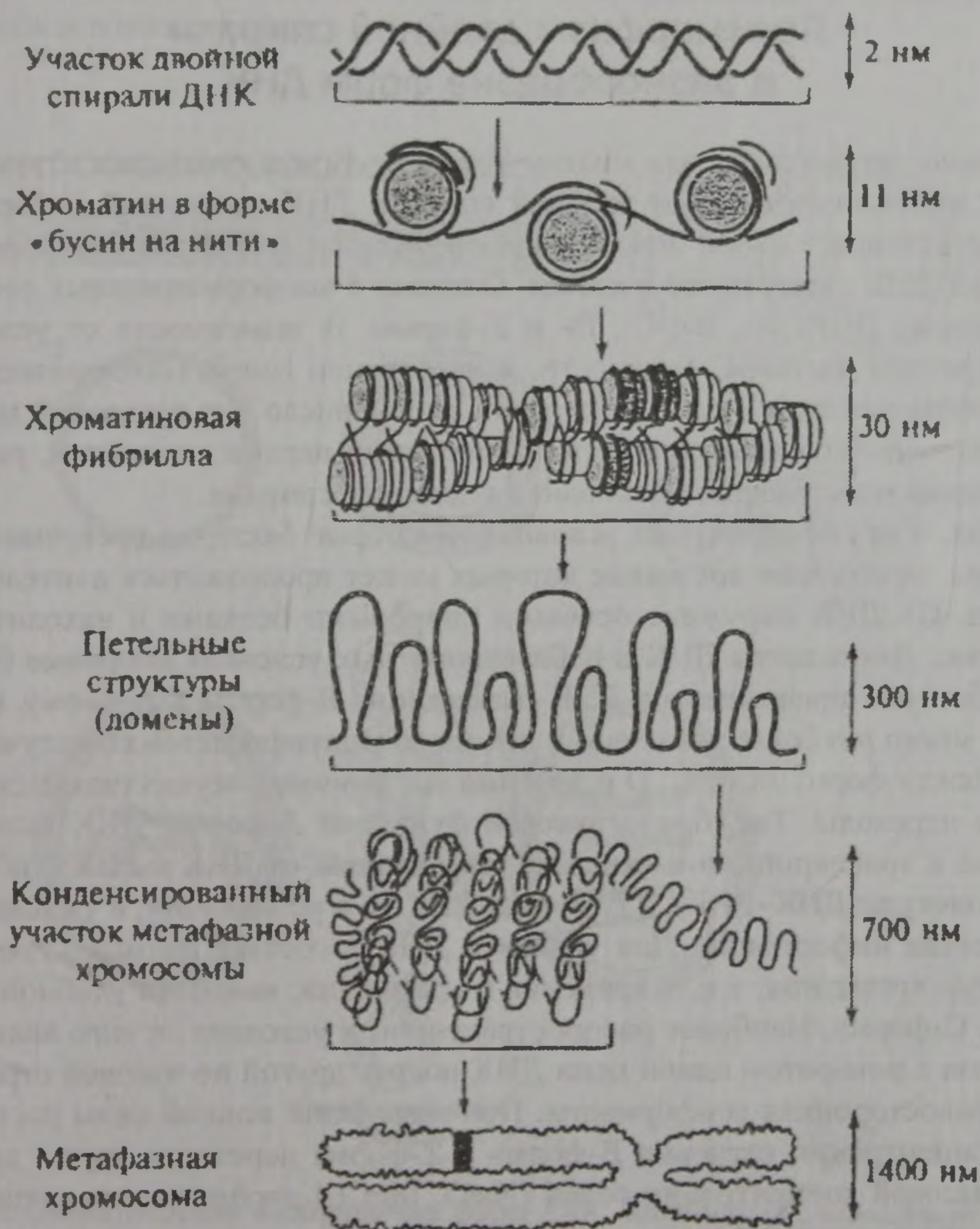


Рис. 5. Компактизация молекулы ДНК в составе хроматина



и 30 нанометров, петельные домены. Высшим уровнем компактизации (в 7000 раз) являются *метафазные хромосомы* с диаметром 1400 нм. Общая длина 46 метафазных хромосом человека в 100 000 раз меньше общей длины молекул ДНК на стадии интерфазы. В цикле компактизации молекулы ДНК важную роль играют *ДНК-топоизомеразы* — ферменты, изменяющие топологию ДНК. Выяснена роль одного из этой группы ферментов — ДНК-топоизомеразы II — в формировании высших уровней структуры хроматина. Этот фермент участвует в образовании петель хроматина во время конденсации хромосом.

### Полиморфизм двойной спирали и разнообразие форм ДНК

Было установлено, что модель Уотсона—Крика описывает структуру одной из нескольких форм двойной спирали ДНК, названной В-формой. Она представляет собой основную конформацию, в которой большинство молекул ДНК существуют в клетке. Описано 5 конформационных состояний (форм) ДНК: А-, В-, С-, D- и Z-формы. В зависимости от условий (температуры раствора, влажности, концентрации солей, соотношения ионов и катионов и др.), они могут иметь разное число пар оснований на виток спирали, угол вращения между соседними парами оснований, разное расстояние между парами оснований и диаметр спирали.

Так, в неблагоприятных условиях некоторые бактерии превращаются в споры, неактивное состояние которых может продолжаться длительный период. Их ДНК окружена особыми споровыми белками и находится в А-форме. Для защиты ДНК в неблагоприятных условиях споровые белки способствуют превращению ДНК бактерий из В-формы в А-форму, которая во много раз более устойчива к действию ультрафиолетового излучения.

Между формами В, С, D и другими состояниями осуществляются взаимные переходы. Так, биологической функцией А-формы ДНК является участие в транскрипции и передаче информации от ДНК к РНК (гибридные молекулы ДНК–РНК), а В-формы ДНК — в репликации, в умножении количества информации. Для упаковки ДНК в составе надмолекулярных структур хроматина, т. е. в хранении информации, наиболее удобной оказалась С-форма. Наиболее распространенной в условиях *in vitro* является В-форма с поворотом одной цепи ДНК вокруг другой по часовой стрелке, т. е. правосторонняя конформация. При изменении ионной силы раствора или концентрации катионов В-форма и Z-форма переходят друг в друга. При высокой концентрации солей (NaCl, MgCl<sub>2</sub>) двойная правая спираль переходит в левую Z-форму.

Наряду с линейно двуцепочечной ДНК у многих эукариот, геномы некоторых вирусов, бактерий, растений и животных имеют форму ковалентно

замкнутых колец, состоящих из одной цепи. Для репликации таких вирусных ДНК происходит их превращение в двуцепочечную форму с последующим образованием одноцепочечных кольцевых ДНК вирусного потомства (рис. 6).

Двуцепочечная ДНК является субстратом для транскрипции последовательностей ДНК в РНК, поэтому экспрессия генетической информации в таких геномах всегда осуществляется в фазе двуцепочечной ДНК. В форме единственной ковалентно замкнутой кольцевой дуплексной молекулы ДНК представлены бактериальные плазмиды, хромосомы некоторых бактерий, большинство ДНК митохондрий и хлоропластов, а также геномы вирусов млекопитающих.

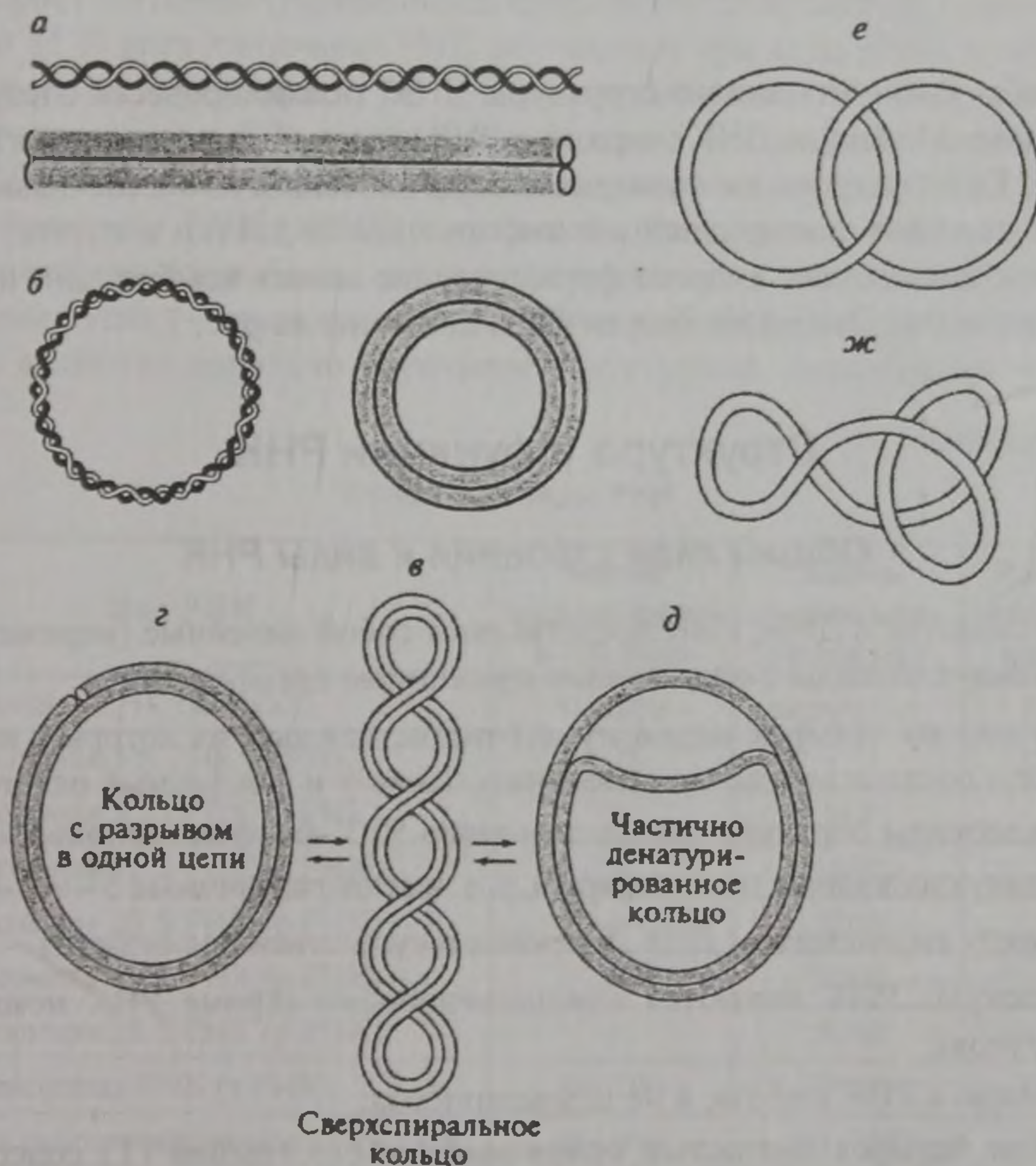


Рис. 6. Схематическое изображение форм ДНК: линейная (а); кольцевые — двуцепочечная (б) и сверхспиральная (в); релаксированные кольцевые формы — разрыв одной из двух цепей (г) или локальное расхождение двух цепей (д); катенан (е); узел (ж)



Рис. 7. Генетические особенности молекулы ДНК

Чтобы закончить анализ структуры ДНК, можно провести следующее обобщение. Молекула ДНК (наряду с РНК) является уникальной в живой природе. Ее структурные и функциональные особенности обеспечивают хранение и передачу наследственной информации от клетки к клетке, от поколения к поколению, а также формирование новых комбинаций признаков у потомства. Эти особенности ДНК показаны на рис. 7.

## Структура и функции РНК

### Общий план строения и виды РНК

По сходству с ДНК, РНК представляют собой линейные (неразветвленные) полинуклеотиды с одинаковым принципом организации:

- 1) состоят из четырех видов нуклеотидов, каждый из которых имеет в своем составе азотистое основание, пентозу и фосфатный остаток;
- 2) нуклеотиды образуют цепь с помощью 5', 3'-фосфодиэфирных связей;
- 3) полинуклеотидные цепи полярны, т. е. имеют различимые 5'- и 3'-концы;

Наряду со сходством ДНК, имеются и существенные отличия:

- 1) молекулы РНК являются одноцепочечными (кроме РНК некоторых вирусов);
- 2) пентоза в РНК рибоза, а не дезоксирибоза;
- 3) среди четырех азотистых оснований вместо тимина (Т) содержится урацил (U), который отличается от тимина отсутствием в 5-м положении метильной группы.

Указанные отличия состава РНК от ДНК имеют важное биологическое значение — свою функцию молекулы РНК могут выполнять только в

одноцепочечном состоянии. Это особенно проявляется в процессе трансляции мРНК на рибосомах. Несмотря на одноцепочечность, цепь РНК в некоторых участках может образовывать *петли* или «шпильки» с двухцепочечной структурой. Содержание нуклеотидов в молекулах РНК варьирует от 75 до многих тысяч. Общее содержание РНК в любых клетках превышает содержание ДНК в 5–10 раз. Ее основные функции заключаются в трансляции генетической информации с образованием белков, а также в осуществлении специализированных эндонуклеазных функций, регулирующих этапы экспрессии генов.

Молекулы РНК во всех клетках представлены следующими типами (видами): мРНК (матричная или информационная), тРНК (транспортная), рРНК (рибосомная), мцРНК (малая цитоплазматическая) и только в клетках эукариот — гяРНК (гетерогенная ядерная), а также мяРНК (малая ядерная). До 85 % всех клеточных РНК составляют три вида рРНК у прокариот и четыре вида у эукариот, 10 % — около 100 видов тРНК. Другие виды РНК представлены: 5 % приходится на долю нескольких тысяч мРНК, менее 2 % — на долю малых ядерной и цитоплазматической РНК (табл. 1).

**Рибосомные РНК (рРНК) и рибосомы.** Рибосомные РНК лежат в основе формирования *рибосом*, в которых происходят центральные события проявления гена — превращение генетической информации в белок. Рибосомы являются довольно крупными структурами, состоящими из боль-

Таблица 1

## Основные виды РНК

Вид РНК	Число разных видов в клетках	Длина (число нуклеотидов)	Распространенность *
1. Матричная РНК (м РНК)	тысячи	варьирует	П, Э
2. Рибосомная 5 S РНК (р РНК)	1–2	120	П, Э
3. Рибосомная 5,8S РНК (р РНК)	1	158	Э
4. Рибосомная 16 S РНК (р РНК)	1	1600	П
5. Рибосомная 23 S РНК (р РНК)	1	3200	П
6. Рибосомная 18 S РНК (р РНК)	1	1900	Э
7. Рибосомная 28 S РНК (р РНК)	1	5000	Э
8. Транспортная РНК (т РНК)	80–100	75–90	П, Э
9. Малая цитоплазматическая (мц РНК)	десятки	90–330	П, Э
10. Гетерогенная ядерная (гя РНК)	тысячи	варьирует	Э
11. Малая ядерная (мя РНК)	десятки	58–220	Э

\* П — прокариоты, Э — эукариоты, S — единица Сведберга (константа седиментации).

шой и малой субъединицы, напоминающих по форме цифру «8». Рибосомы хлоропластов и митохондрий сходны с рибосомами сине-зеленых водорослей (цианобактерий) и бактерий. Рибосомы, локализованные в цитоплазме эукариотической клетки, и рибосомы прокариот имеют различия, при характеристике которых используется *показатель S-коэффициент седиментации*, характеризующий скорость осаждения в центрифуге, и показатель молекулярной массы MDa.

Свойства рибосом	Прокариоты	Эукариоты (в цитоплазме)
Общие размеры	70 S	80 S
Малая субъединица	30 S	40 S
Размер РНК (число нуклеотидов)	5 S (120)	5 S (120)
Большая субъединица	16 S (1500)	18 S (2000)
Большая субъединица	50 S	60 S
Размер РНК (число нуклеотидов)	23 S (3000)	28 S (5000)

Выявлено особое значение 16 S РНК, ген которой 16 S рДНК был положен Карлом Везе в основу современной *филогенетической классификации бактерий* на основе последовательности нуклеотидов по длине ДНК этого гена у разных видов. Картина сходства рибосом явилась критерием для построения универсального филогенетического дерева жизни, включая Eucaryota (Woese, 1977, 1981, 1987).

Основу каждой субъединицы рибосомы составляют рРНК, вокруг которых группируются белки. Элементы структуры рибосом играют определенную роль в процессах биосинтеза белков (закрепления мРНК и тРНК, выведения синтезированного полипептида). Среди азотистых оснований в рРНК имеет место более высокое, чем обычно, содержание нуклеотидов гуанина и цитозина, а гены рРНК являются ГЦ-богатыми. Во вторичной структуре рРНК много двухцепочечных участков и *петель*. Митохондриальные рибосомы грибов имеют коэффициент седиментации 75S, а «мини-рибосомы» митохондрий млекопитающих — 55S. Установлено большое число генов, контролирующих синтез рибосомных РНК: сотни генов 18S и 28S у млекопитающих и человека, а также около 2000 генов 5S-рРНК у человека и 24 000 у лягушки *Xenopus*. На рис. 8 представлено сравнение рибосом прокариот и эукариот.

**Матричные (информационные) РНК (мРНК).** В клетке имеется большое количество разных мРНК, так как каждая из них содержит информацию о синтезе непосредственно на ней той или иной полипептидной цепи белка, в том числе о времени этого синтеза, его количестве, месте и

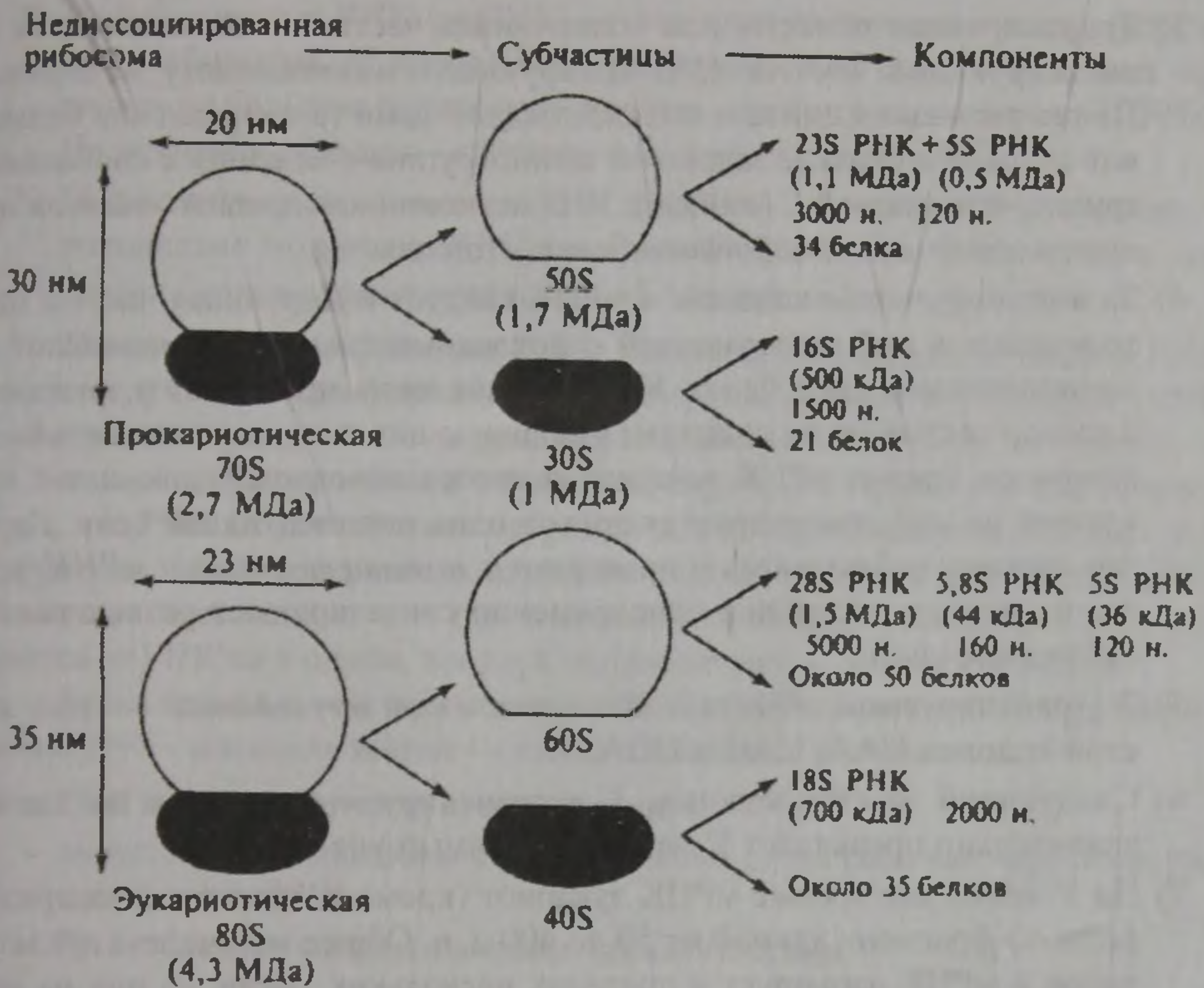


Рис. 8. Состав рибосом прокариот и эукариот

условиях. Несмотря на их разнообразие, все разновидности мРНК составляют всего 5 % от общей массы РНК в клетке. Кроме того, зрелые мРНК имеют сходный тип строения своей линейной цепи и расположения на ней участков с различной функциональной ролью (рис. 9).

Эти участки имеют следующую последовательность расположения:

- 1) На 5'-конце находится «колпачок», или *кэп* — специфическая структура, необходимая для прикрепления мРНК к рибосоме. Кэпирование происходит вскоре после начала синтеза мРНК, а структура кэпа обеспечивает защиту 5'-конца молекулы от действия экзонуклеаз.
- 2) За «колпачком» следует 5'-нетранслируемый участок (5'-НТО) из нескольких десятков нуклеотидов. Этот участок необходим для первичного связывания мРНК с рибосомой, но сам не транслируется.

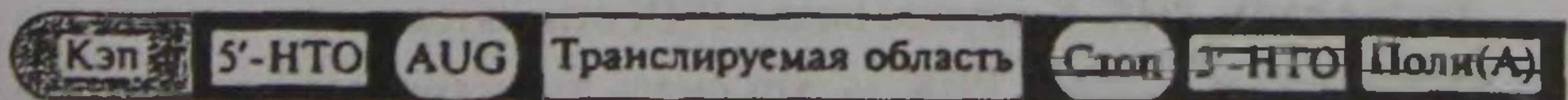


Рис. 9. Составные части зрелой мРНК эукариот

- 3) Транслируемая область, или кодирующая часть, всегда начинается с *иницирующего кодона AUG*, кодирующего аминокислоту метионин. После окончания синтеза полипептидной цепи (а направление белковой цепи при синтезе задают от аминогруппы с ее конца к свободной группе, называемой С-концом), этот метионин как правило, если он не нужен для функционирования белка, отщепляется.
- 4) За иницирующим кодоном в мРНК следует кодирующая часть с находящейся в ней информацией о последовательности аминокислот в полипептидной цепи белка. Кодирующая часть прокариот и эукариот в зрелом состоянии не содержит некодирующих последовательностей — *интронов*. Зрелые мРНК эукариот являются *моноцистронными*, т. е. на каждой из них синтезируется только одна полипептидная цепь. Другая картина наблюдается у прокариот с *полицистронными мРНК*, когда на разных цистронах одновременно синтезируются разные полипептидные цепи.
- 5) За транслируемой областью находится *кодон терминации* — один из стоп-кодонов UAA, UAG и UGA.
- 6) Следующий за стоп-кодомом 3'-нетранслируемый участок по длине значительно превышает 5'-нетранслируемый участок.
- 7) На 3'-конце все зрелые мРНК эукариот (кроме гистоновых) содержат поли(А)-фрагмент длиной от 50 до 400 п. н. Общее количество нуклеотидов в мРНК варьирует в пределах нескольких тысяч, из них на кодирующую часть приходится 60–70 % нуклеотидов.

**Транспортные РНК (тРНК).** Транспортные РНК составляют 10 % суммарной клеточной РНК. Они представляют собой низкомолекулярные соединения (17–35 кДа), а их главная функция заключается в акцептировании аминокислот и переносе их в рибосому для синтеза белка. У живых организмов разных таксонов (микроорганизмов, растений, животных) выделено более 2000 тРНК. В связи с эволюционной консервативностью, у всех разновидностей тРНК один способ двумерной укладки в форме «клеверного листа». В этой структуре — 5 одноцепочечных стеблей и 3 петли. Акцепторный стебель имеет участок из 4 нуклеотидов на 3'-конце, к крайнему нуклеотиду ковалентно присоединяется аминокислота. На антикодонном стебле находится *антикодон* — последовательность из трех нуклеотидов на молекуле тРНК, комплементарно взаимодействующая в процессе репликации с кодоном мРНК. Число генов тРНК и их повторов измеряется сотнями и тысячами — у человека имеется 1300 генов.

**Другие виды РНК.** К другим видам РНК относятся:

- 1) Гетерогенная ядерная РНК — это смесь транскриптов многих ядерных генов, гетерогенных по размеру и обладающих низкой стабильностью.

- 2) Малые ядерные РНК (мяРНК) — это присутствующие в ядре короткие стабильные молекулы РНК размером от 63 до 220 н. В них содержится большое количество урацила, поэтому их называют U-РНК. Их основная функция — участие в сплайсинге про-мРНК.
- 3) Малые цитоплазматические РНК (мцРНК) — это присутствующие в цитоплазме молекулы РНК размером 90–330 н. Они участвуют в переносе белков через липидный слой эндоплазматической сети.

**Неканонические функции РНК.** Каноническое представление о РНК как инструменте трансформации генотипа, заложенном в структуре ДНК, в конкретный фенотип, т. е. в разнообразие белков по пути ДНК → РНК → белок, было поколеблено открытием *обратной транскрипции*. Оказалось, что РНК может служить матрицей для воспроизведения собственной структуры в РНК-содержащих вирусных геномах, а также для синтеза ДНК у высших организмов, т. е. поток генетической информации распространяется от РНК не в одном, а в двух направлениях — к белку и к ДНК.

Точка зрения о том, что молекула РНК представляет собой первичную молекулу — носитель жизни — имеет следующее обоснование:

- РНК несет генетическую информацию;
- вирусные РНК способны к рекомбинации с участием как вирусных, так и клеточных РНК;
- РНК обладает каталитическими способностями.

В настоящее время ведется поиск *фермента РНК-репликазы*, способного к аутокатализу репликации РНК.

## Ключевые слова и термины

Нуклеиновые кислоты ДНК и РНК; нуклеотиды; фосфат, дезоксирибоза, рибоза; азотистые основания; фтородиэфирные связи; антипараллельные цепи ДНК; пикограммы, дальтоны, пары нуклеотидов, килобаза, мегабаза; хроматин, эухроматин, гетерохроматин; типы РНК, рибосомы, кодоны, иницирующие кодоны, кодоны терминации.



## Репликация ДНК

Для реализации своего основного свойства — наследственности — молекула ДНК должна точно копировать саму себя, сохраняя информацию исходной молекулы. Это обеспечивается за счет *репликации ДНК* — процесса, предшествующего делению любой живой клетки. Репликация ДНК представляет собой, наряду со способностью кодирования генетической информации, одно из *двух проявлений наследственности*. Параллельно происходит удвоение количества хромосомных белков, репликация ДНК и является составной частью репликации хромосом.

Репликация ДНК связана с делениями клеток двух типов: митоза и мейоза. В одинарном наборе хромосом количество ДНК выражается символом  $n$ , тогда как в ходе митоза и мейоза происходят следующие его изменения (рис. 10).

Ядро исходной клетки содержит двойной набор хромосом  $n$ , соответственно, диплоидное количество ДНК ( $2n$ ). В процессе репликации на стадии интерфазы количество ДНК возрастает до тетраплоидного ( $4n$ ), а в ходе последующего деления поровну распределяется между дочерними клетками. В отличие от митоза, в мейозе две дочерние диплоидные клетки (сперматоциты и ооциты) делятся еще раз без предшествующей репликации ДНК.

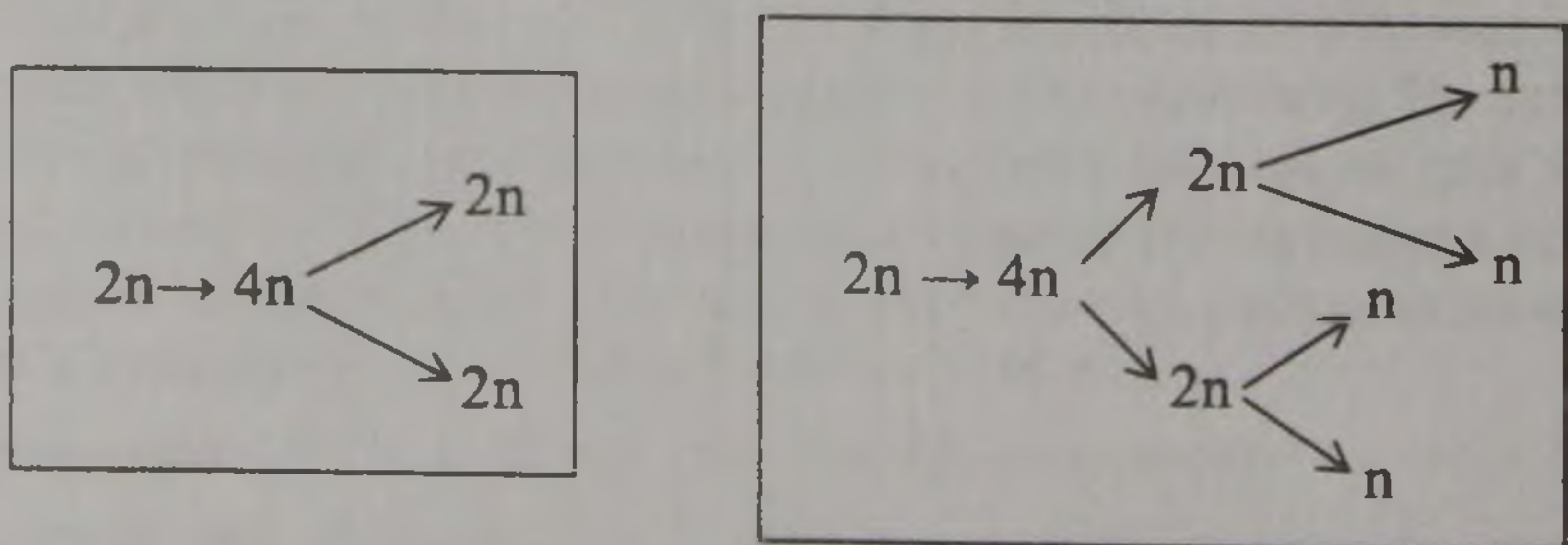


Рис. 10. Схемы митоза (слева) и мейоза (справа)



Рис. 11. Клеточный цикл у эукариот

Все события в ходе деления клетки выделены в особый *клеточный цикл* (рис. 11). В этом цикле выделяют следующие периоды:

- *постмитотический (пресинтетический) период  $G_1$* : интервал от окончания митоза до начала синтеза ДНК и ядерных белков в дочерней клетке. Одновременно восстанавливается содержание цитоплазматических белков и происходит рост дочерней клетки до размера материнской. Содержание ДНК в клетке —  $2n$ ;
- *синтетические  $S$ -период* — это период репликации ДНК и удвоения хромосом, количество ДНК в ядре возрастает от  $2n$  до  $4n$ . Удваивается также количество хромосомных белков в ядре;
- *постсинтетический (премитотический)  $G_2$ -период* — это период подготовки клетки к делению;
- *$M$ -период* — период деления клетки и хромосом, приводящий к образованию двух диплоидных клеток.

$G_1$ -,  $S$ - $G_2$ - периоды вместе составляют интерфазу и занимают 90 % времени. Продолжительность клеточного цикла сильно варьирует в клетках разных типов, занимая от минут до суток.

## Общая характеристика репликации ДНК

Процесс репликации ДНК имеет ряд принципиальных особенностей. Среди них обращают внимание следующие:

1. Репликация представляет собой матричный процесс на основе принципа комплементарности. Каждая синтезируемая нить строится на матрице исходной цепи ДНК, включая в свой состав один из четырех нуклеотидов, который соответствует нуклеотиду в родительской цепи (рис. 12).
2. В отличие от синтеза РНК, матрицами служат обе цепи родительской ДНК, т. е. процесс является симметричным. В то же время синтез *полу-консервативный*, так как в каждой из дочерних молекул одна цепь — родительская, а вторая — новосинтезированная.
3. Направление роста и полярности цепей ДНК отличается тем, что удлинение цепи ДНК всегда происходит в направлении *от 5'-конца к 3'-концу*, а очередной нуклеотид присоединяется к *3'-концу растущей цепи*.

Комплементарные цепи в любой молекуле ДНК антипараллельны, следовательно, растущая цепь антипараллельна матричной цепи, которая считывается в направлении  $3' \rightarrow 5'$ .

4. В процессе репликации участвует *сложный ферментный комплекс*, в состав которого входят 15–20 различных белков.

Детальное изучение процесса репликации ДНК было связано с исследованием поведения и активности ферментов, обеспечивающих работу ДНК-реплицирующего комплекса. Наиболее полно изучен механизм репликации бактериальной ДНК, особенно *E. coli* и размножающихся в ней бактериофагов. Описаны ферменты репликации дрожжей и дрозофилы, но в проблеме репликации в эукариотических клетках остается еще много неясных вопросов (Бокуть, Герасимович, Милютин, 2005).

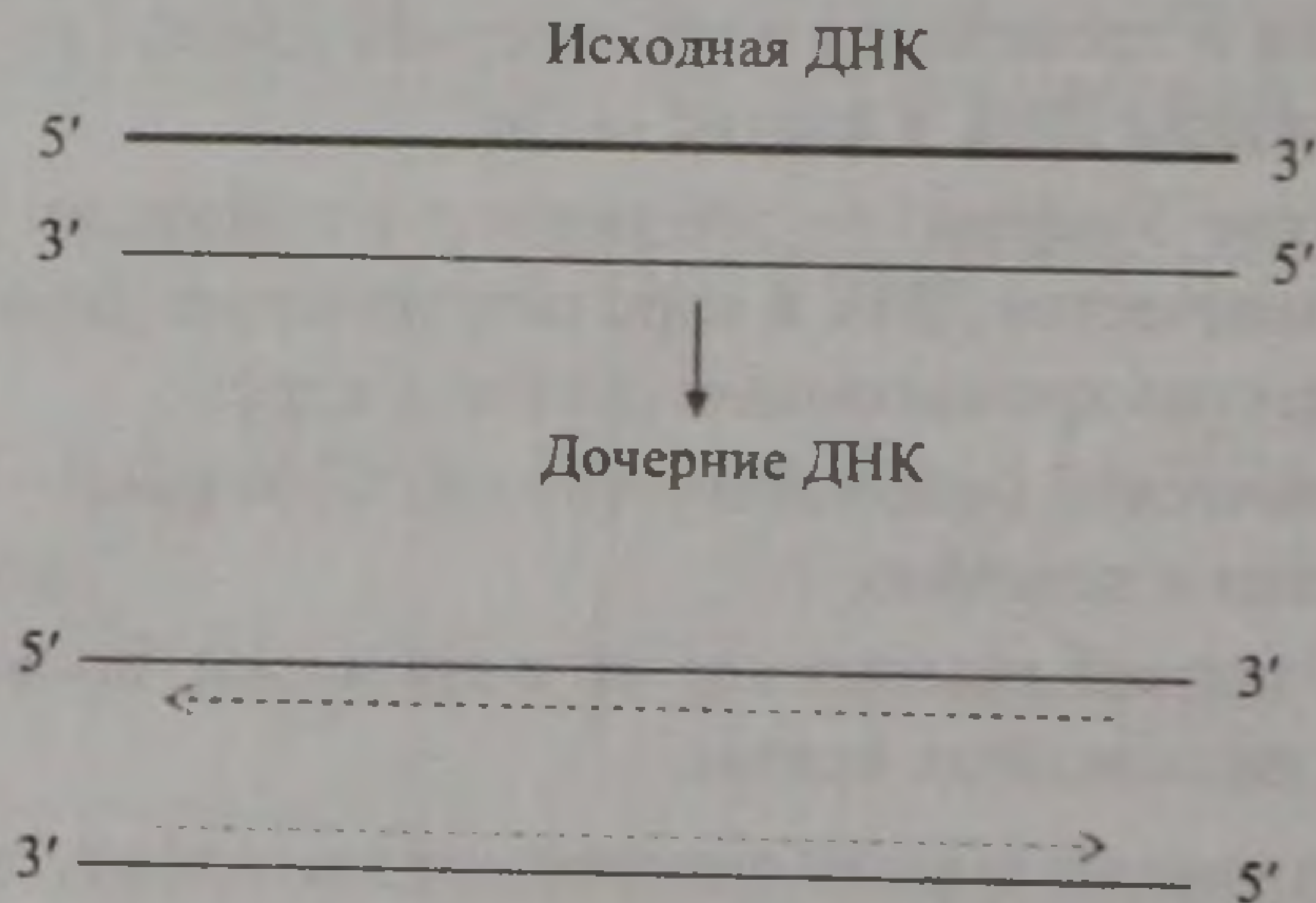
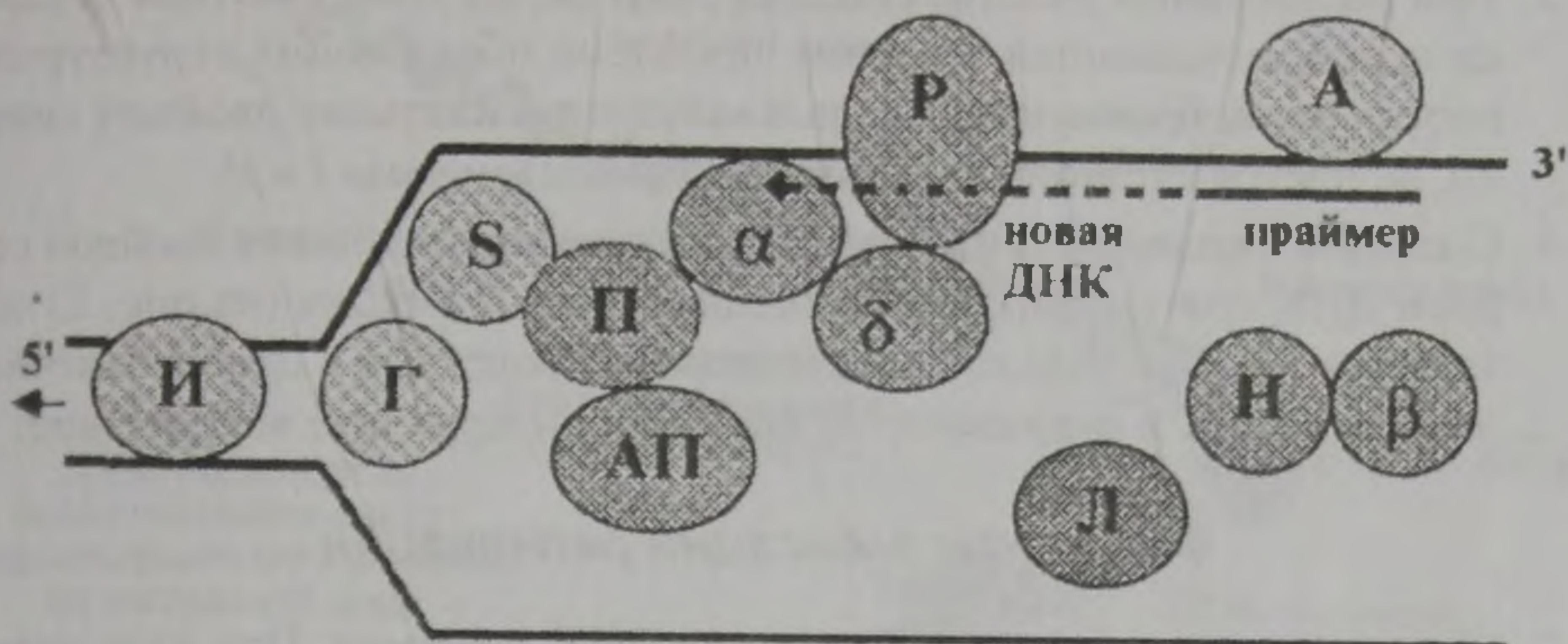


Рис. 12. Схема репликации ДНК



**Рис. 13.** Белки репликативного комплекса. А — узнающие места инициации белки, Г — хеликаза, И — топоизомераза, S-SSB — белки, П — праймаза, АП — активатор праймазы;  $\alpha$ ,  $\beta$ ,  $\delta$  — альфа, бета, дельта — ДНК — полимеразы. Р — белок (ядерный антиген) PCNA, Н — нуклеаза, Л — ДНК — лигаза

Все белки репликативного комплекса можно подразделить по их роли в обеспечении и *инициации, элонгации и терминации* этого процесса (рис. 13).

### Белки инициации репликации

1. АТ — богатые последовательности характерны для точек начала репликации на молекуле ДНК, с которыми связываются узнающие белки, которые у бактерий называются *Dna A* (на рис. 13 обозначены А). Они контролируют начало репликации и обеспечивают связывание ДНК-реплицирующего комплекса. Хромосома *E. coli* содержит единственную область начала репликации — *ori C* (*ориджин репликация*). Репликация ДНК эукариот протекает с образованием большого количества *репликативных вилок*. Размеры репликонов, т. е. расстояния от одной точки начала репликации до соседней, у разных авторов различаются, но средние размеры у эукариот могут достигать 100–200 т. п. н. Следовательно, в гаплоидном геноме млекопитающих должно содержаться 20 000–30 000 репликонов. У дрозофилы и дрожжей репликоны меньше — 40 000 п. н. Это соответствует количеству 3500 репликонов на гаплоидный набор дрозофилы или 500 репликонов дрожжей. Оценена также скорость движения репликативных вилок во времени и составила от 1000 до 3000 п. н. в 1 минуту у млекопитающих, около 1000 п. н. в 1 минуту у растений и до 50 000 п. н. в 1 минуту у прокариот.
2. Фермент *хеликаза* (на рис. 13 обозначен Г) обеспечивает расплетание двойной спирали родительской ДНК в районе репликативной вилки и разъединяет ее на одноцепочечные участки.

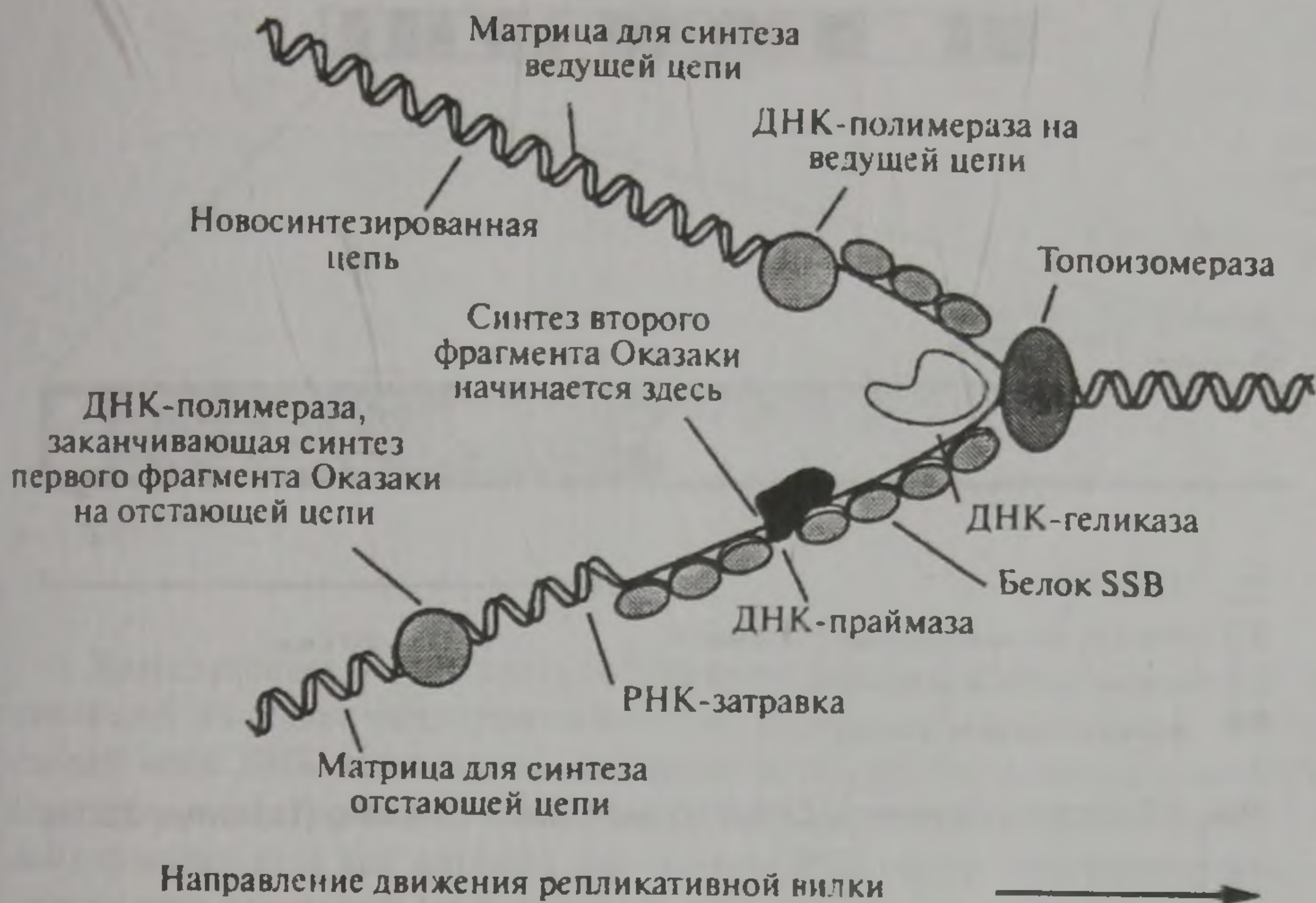
3. При расплетании участка спирали ДНК перед этим участком создается *суперспирализация*, поэтому проблема образования структурного напряжения, блокирующего дальнейшее расплетание двойной спирали, решается с помощью ферментов *топоизомераза I и II*.
4. С каждой отдельной нитью после локального расплетания двойной спирали ДНК сразу связываются специальные *SSB-белки* (на рис. 13 обозначены S). Они обладают повышенным сродством к одноцепочечным участкам ДНК и оказывают на них стабилизирующее воздействие.

### Ферменты элонгации репликации

1. Синтезируется *праймер* — короткая РНК-затравка. При этом специальный белок *активатор праймазы* (на рис. 13 обозначен АП) индуцирует праймазу на матрице одноцепочечного участка ДНК к образованию праймера (праймаза обозначена П на рис. 13). В дальнейшем короткие РНК праймеры, служащие началом фрагментов Оказаки, замещаются участками ДНК под действием ДНК-полимеразы I.
2. В репликации ядерной ДНК из 5 разных ДНК-полимераз у эукариот (*альфа-, бета-, гамма-, дельта-, эпсилон-, дзета-*) участвуют только альфа- и дельта-полимеразы (на рис. 13 обозначены  $\delta$ ). В свою очередь установлена связь альфа-полимеразы с праймазой, а дельта-полимеразы с белком PCNA (на рис. 13 обозначен P). ДНК-полимеразы проводят последовательное включение нуклеотидов в строящуюся цепь ДНК на основе комплементарности. Активность ДНК-полимераз лучше изучена у прокариот, по сравнению с эукариотами. Оказалось, что все прокариотические и эукариотические ДНК-полимеразы способны синтезировать ДНК путем присоединения каждого последующего дезоксирибонуклеозидтрифосфата (нуклеотида) только к 3'-ОН-концу уже имеющейся цепи — *праймеру*. Таким образом, цепи синтезируются в направлении  $5' \rightarrow 3'$ , и наращивание полинуклеотидов в обратном направлении не происходит (рис. 14).
3. У прокариот ДНК — полимеразы I присоединяется к 3 — концу растущего фрагмента, нуклеотиды включаются к 3 — концу (ДНК — полимеразная активность). Растущий фрагмент доходит до нуклеотидов предыдущего фрагмента Оказаки, и функция ДНК — полимеразы I прекращаются.

### Ферменты терминации репликации

В результате процесса репликации и действия всех описанных ферментов получаются состоящие из фрагментов новосинтезированные цепи. Происходит «сшивание» соседних фрагментов с помощью лигазы (на рис. 11



**Рис. 14.** Схема синтеза ДНК в репликативной вилке у прокариот (Филиппович, 1999)

обозначена Л), т. е. образование межнуклеотидных фосфодиэфирных связей. Для полного окончания репликации молекулы ДНК осуществляется специальный процесс репликации ее концов — теломерных участков, в котором главная роль принадлежит ферменту теломеразе.

### Репликация теломер

Теломеры представляют собой концевые участки хромосом, которые не несут генетической информации, а служат для защиты хромосом от потерь при репликации и от атак эндонуклеаз. Теломерная ДНК выделена, а ее структура расшифрована. Выявлено сходное строение теломер у разных видов из многократно повторяющихся фрагментов — у человека это повтор TTAGGG (рис. 15). За короткими повторами располагаются более длинные повторяющиеся последовательности — так называемые повторы, ассоциированные с теломерой.

Установлено, что для линейных ДНК характерна их концевая недорепликация. Это объясняется тем, что описанная ДНК — полимеразная система оставляет 3'-концы материнских цепей ДНК недореплицированными и, соответственно, что новые цепи являются укороченными с 5'-концов.

Известно, что все ДНК-полимеразы осуществляют синтез с использованием РНК-праймеров, которые потом удаляются, поэтому 5-конец ново-

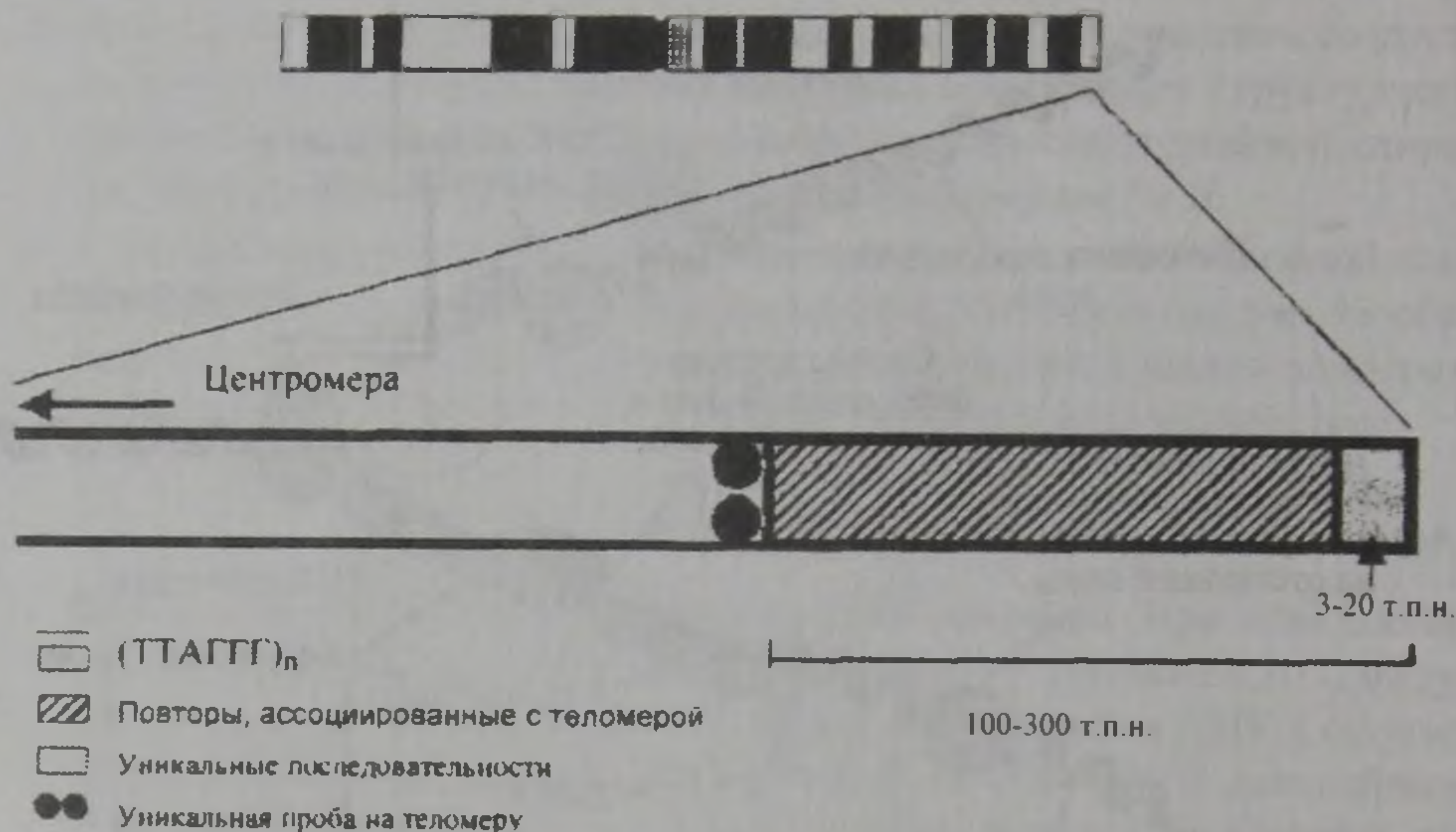


Рис. 15. Структура концов ДНК в хромосомах — теломер (Тарантул, 2003)

синтезированной цепи должен укорачиваться на длину праймера, т. е. на 10–30 нуклеотидов. Такое исчезновение концевых участков должно приводить к потере генетического материала.

Для предохранения от потерь генетического материала при делении клеток у эмбрионов конмолекул ДНК перед каждой репликацией наращиваются короткими повторяющимися последовательностями, с участием фермента теломераз.

Этот фермент перестает работать в клетках взрослого организма, поэтому в каждом цикле репликации ДНК хромосомы постоянно укорачиваются. Укорочение до определенной критической длины является своего рода «сигналом» к прекращению делений. Уставлено, что у соматических клеток существует определенный лимит на число клеточных делений. Ограничение на число клеточных делений названо порогом Хейфлика. Клетки обычно не преодолевают барьер из 20–90 делений.

Фермент теломеразы в раковых клетках и в клетках зародышевого пути добавляет перед каждым циклом репликации повторы TTAGGG к теломеразе, поэтому число делений у них не ограничено.

### Ключевые слова и термины

Репликация, клеточный цикл, полуконсервативный синтез; инициация, элонгация и терминация процесса; репликативная вилка, репликон, ориджин репликации, праймер; репликация теломер, порог Хейфлика.

## Транскрипция ДНК

Транскрипция — это процесс перевода генетической информации, записанной на языке последовательности дезоксирибонуклеотидов в смысловой цепи ДНК на языке последовательности рибонуклеотидов в мРНК. При этом определенный участок одной из двух цепей ДНК (*антисмысловой*) используется как матрица для синтеза РНК путем комплементарного спаривания оснований (рис. 16).

Как и при репликации ДНК, строящаяся цепь РНК имеет направление —  $5' \rightarrow 3'$ , т. е. очередные нуклеотиды у этой цепи присоединяются к  $3'$ -концу. По отношению к матричной цепи ДНК строящаяся цепь РНК антипараллельна, поэтому она транскрибируется ферментом в направлении  $3' \rightarrow 5'$ .

Ранее отмечалось, что репликация ДНК строго связана с делением клеток, предшествуя ему. В ходе реализации генетической информации транскрипция ДНК происходит многократно постоянно в делящихся и неделящихся клетках.

Смысловая цепь ДНК ( $5'$ ) — ТТЦ-АГТ-ЦАГ-ГАЦ-ГАТ-АЦГ — ( $3'$ )

Матричная цепь ДНК ( $3'$ ) — ААГ-ТЦА-ГТЦ-ЦТГ-ЦТА-ТГЦ — ( $5'$ )

↓  
ТРАНСКРИПЦИЯ

Матричная РНК (мРНК) ( $5'$ ) — УУЦ-АГУ-ЦАГ-ГАЦ-ГАУ-АЦГ — ( $3'$ )

↓  
ТРАНСЛЯЦИЯ

Пептидная цепь белка ( $\text{NH}_2$ ) — Фен—Сер—Гли—Асп—Асп—Тре— ( $\text{COOH}$ )

Рис. 16. Принцип записи и реализации генетической информации



В отличие от полуконсервативного синтеза ДНК синтез РНК носит консервативный характер, так как молекула ДНК после транскрипции восстанавливает исходное состояние. При этом в качестве матрицы используется только одна цепь ДНК, а процесс синтеза РНК для инициации не нуждается в затравке, необходимой при репликации ДНК.

Транскрипция ДНК происходит не на всей молекуле ДНК, а в определенных участках молекулы — *транскриптонах*. Транскриптон ограничен с одной стороны последовательностью ДНК — зоной начала транскрипции, так называемым *промотором*, а с другой стороны — зоной остановки транскрипции — *терминатором*.

Процесс транскрипции обеспечивает фермент *ДНК — зависимая РНК-полимераза*. У бактерий синтез мРНК, рРНК и тРНК осуществляется одной и той же РНК-полимеразой. Общее количество молекул этого фермента в клетках *E. coli* может достигать 7000. Наиболее полно изучена РНК-полимераза *E. coli*, структура которой аналогична структуре этого фермента у других бактерий. Процесс транскрипции и его ферментативного обеспечения более полно изучены у прокариот, РНК-полимеразы которых представляют собой сложные белки, состоящие из нескольких субъединиц. Хорошо изучен *полный фермент*, или *холофермент* РНК-полимеразы *E. coli*. Его структуры составляют пять полипептидных субъединиц: две альфа-цепи, одну бета- и одну бета-штрих-цепи, а также сигма-цепи. Холофермент РНК-полимеразы способна узнавать промоторную область в оперонах бактерий и инициировать процесс транскрипции.

У эукариот процесс транскрипции осуществляется с участием трех видов РНК-полимераз, с именно:

- РНК-полимераза I — для синтеза пре-рРНК,
- РНК-полимераза II — для синтеза пре-мРНК,
- РНК-полимераза III — для синтеза пре-тРНК.

## Инициация транскрипции

Транскрипция протекает в три различных этапа: *инициации*, *элонгации* и *терминации*. На первом и последнем этапах с ДНК взаимодействует фермент и инициирует синтез РНК, а по его завершению он отделяется от РНК. Основным процессом синтеза РНК — транскрипция — происходит между этими стадиями с участием главного действующего фермента РНК-полимеразы.

Продукты транскрипции представляют собой отдельные молекулы РНК, имеющие 5'- и 3'-концы. Следовательно, инициация и терминация транскрипции осуществляется в определенных сайтах ДНК с характерными специфическими последовательностями оснований. Для обеспечения транскрипции РНК-полимераза взаимодействует с последовательностью

ДНК промотором. У бактерий фермент имеет характерный компонент промотора — бокс (последовательность) Прибнова:



Общая протяженность промотора составляет несколько десятков пар нуклеотидов. У эукариот промотор является более сложной структурой, так как с ним связывается совокупность белков — *общих факторов транскрипции*, кратко обозначаемых TFIID. Вместе с полимеразой II эта совокупность называется *основным инициаторным комплексом*. В промоторе эукариот различают сходный с боксом Прибнова бактерий *TATA-бокс*, реже — *GC-*, *CAAT-* и другие боксы. Транскрипцию генов у эукариот усиливают регулирующие участки ДНК посредством специфических белков — *энхансеров*.

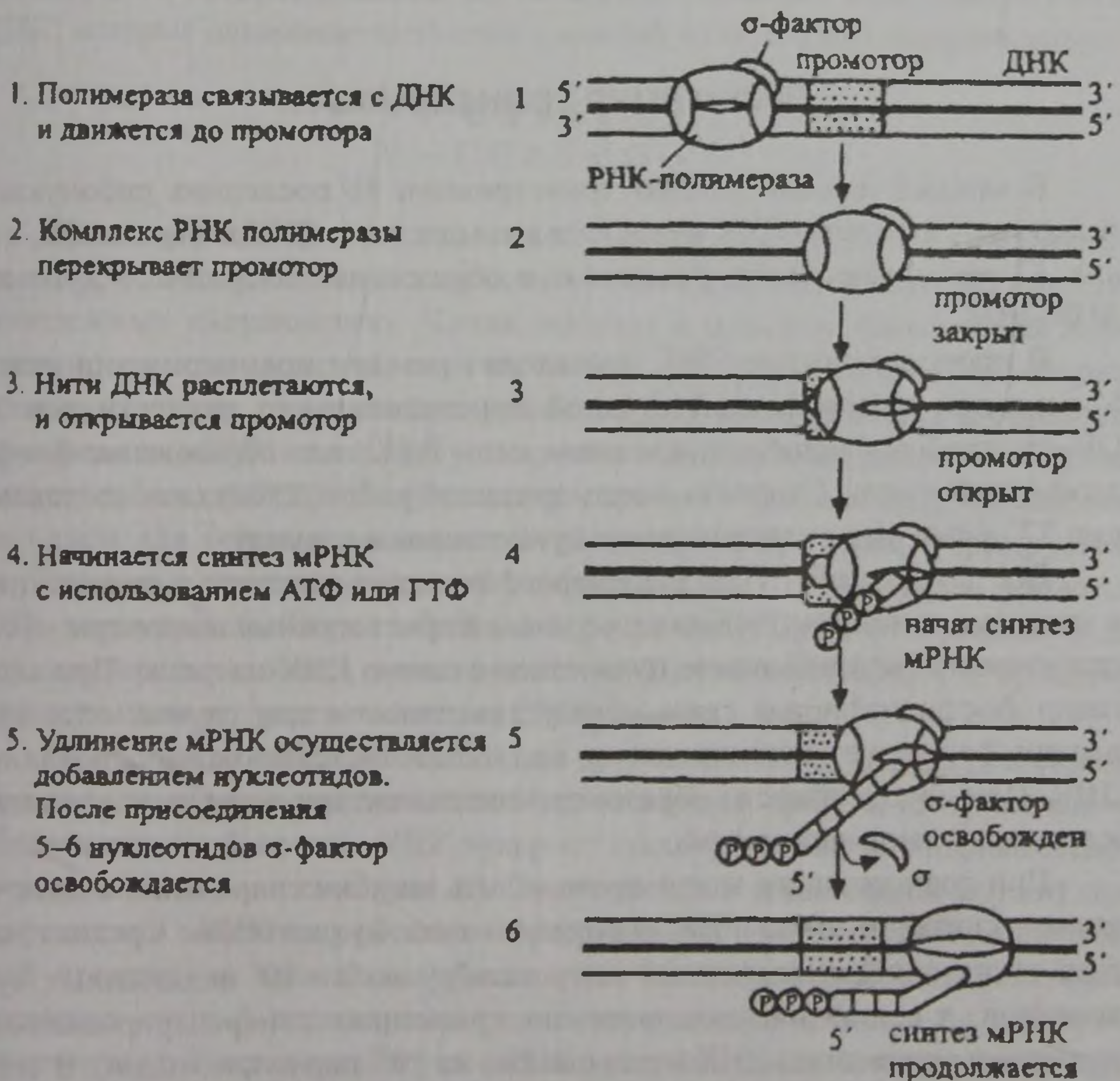


Рис. 17. Инициация и элонгация транскрипции (по Perry, Staley, 1997)

РНК-полимераза после связи с промотором вызывает разделение цепей ДНК на протяжении 15 пар нуклеотидов с образованием транскрипционного «глазка» (рис. 17).

Двойная нить (дуплекс) ДНК в этом месте расплетается и считывается лишь одна нить. Первым в строящуюся цепь РНК всегда включается пуриновый нуклеотид — А или G. Затем образуется первая 5', 3'-фосфатная связь со вторым нуклеотидом. После этого у бактерий сигмафактор теряет связь с ферментом. Промотор может инициатировать отдельный ген или же комбинацию генов, объединенной общей функцией — *оперон*. РНК-полимераза движется вдоль цепи ДНК, раскрывая двойную спираль подобно «молнии». На открывающихся коротких участках единичной матричной цепи происходит удлинение синтезируемой цепи РНК по принципу комплементарности (А–U, G–C), в направлении 5'→3' как и при репликации.

## Элонгация транскрипции

В каждый данный момент транскрипции 50 последних рибонуклеотидов растущей цепи РНК находятся в комплексе с ДНК и ферментом, а из них 12 рибонуклеотидов участвуют в образовании гибридного дуплекса ДНК–РНК.

В процессе синтеза РНК происходит реакция полимеризации между 5'-концевой трифосфатной группой присоединяемого нуклеотида и 3'-ОН-группой последнего нуклеотида цепи РНК, т. е. образование фосфодиэфирной связи. Скорость полимеризации рибонуклеотидов составляет при 37° около 40 присоединенных нуклеотидов в секунду.

Все четыре рибонуклеозидтрифосфата могут потенциально поступать в участок элонгации. РНК-полимераза отбирает нужные мономеры путем спаривания новоприбывшего нуклеотида с цепью ДНК-матрицы. При элонгации фосфодиэфирная связь образуется только в том случае, если нуклеотид формирует каноническую пару с соответствующим основанием ДНК. При неспособности образовать комплементарную пару он удаляется и его место занимает другой.

При транскрипции могут происходить ошибки спаривания с включением в синтезируемую РНК «неправильных» нуклеотидов. Средняя частота таких ошибок составляет одну ошибку на  $2 \times 10^4$  включенных нуклеотидов, т. е. значительно чаще по сравнению с нерепарированными ошибками репликации ДНК (одна ошибка на  $10^{10}$  пар нуклеотидов). В связи с явлением вырожденности генетического кода 67 % ошибок, связанных с положением третьего кодона, не изменяют смысл триплета.

## Терминация транскрипции

На стадии элонгации РНК-полимераза продвигается вдоль матрицы и синтезирует РНК до тех пор, пока фермент не подойдет к терминирующей последовательности. В этом участке происходит остановка включения нуклеотидов в растущую цепь РНК, освобождение образовавшегося транскрипта и отделения РНК-полимеразы от ДНК-матрицы. На стадии терминации разрываются водородные связи, участвующие в формировании временного гибрида ДНК–РНК, а затем происходит восстановление двухцепочечной структуры ДНК.

Последовательность ДНК, необходимая для остановки транскрипции, называется *терминатором*. Он содержит особую последовательность нуклеотидов — *стоп-сигнал*. У прокариот в каждом терминаторе обнаружены *палиндромы*, представляющие собой двухцепочечные последовательности ДНК, которые одинаково читаются в каждой цепи в обоих направлениях:



Палиндром имеет ось симметрии — центральную точку, относительно которой последовательность остается одинаковой при чтении в противоположных направлениях. Копии повтора в одной и той же цепи ДНК взаимно комплементарны, если их читать в противоположенных направлениях. Наличие таких инвертированных повторов приводит к образованию *шпилек в РНК* или структуры *креста в ДНК*. Такие шпильки в терминаторах в большой степени обогащены парами G–C. Полагают, что важным сигналом для терминации является образование шпильки, формируемой на 3'-конце в синтезированной молекуле РНК. Эта шпилька узнается РНК-полимеразой. Таким образом, отрезок цепи ДНК от промотора до терминатора называется *транскриптоном*. Важно отметить, что процессу транскрипции подвергается не вся ДНК генома.

При синтезе мРНК, в отличие от других ее типов, в ходе транскрипции происходят еще два процесса. Это, во-первых, процесс *кэпирования* — образование на 5'-конце мРНК эукариот специфической последовательности нуклеотидов — *кэпа* (рис. 8). Этот отрезок в дальнейшем служит для закрепления мРНК на рибосоме перед началом трансляции. Вторым важным процессом — *полиаденилирование* — присоединение под действием фермента поли-А-полимеразы остатков адениловой кислоты (поли-А-фрагмента) к 3'-концу молекулы мРНК эукариот после завершения ее синтеза. Этот участок мРНК выполняет ряд важных функций и принимает участие в трансляции.



только одну молекулу мРНК, несущую информацию о синтезе *одной полипептидной цепи*. Исключением из этого правила является гистоновая пре-мРНК.

У эукариот кластер гистоновых генов транскрибируется как единое целое в виде одной длинной пре-мРНК, содержащей информацию о пяти гистоновых белках. При созревании она разрезается на пять отдельных гистоновых мРНК. Таким образом, сохраняется общее правило: все зрелые мРНК эукариот является моноцистронными, т. е. каждая из них несет информацию о структуре только одной полипептидной цепи.

2. Пре-рРНК 45S-РНК содержит последовательности трех зрелых рРНК — 18S-, 5,8S- и 28S-рРНК. Кластер этих трех членов рРНК транскрибируется как единое целое (рис. 18, Б). Последовательности разделены спейсерами, но не содержат интронов.

3. Все пре-тРНК (рис. 18, В) содержат последовательности лишь одной тРНК. В первичном транскрипте образуется типичная *структура «кленового листа»*, но в нем не сформирована последовательность акцепторной петли (ЦЦА), а *антикодон* не занимает своего положения.

Полученные продукты транскрипции — предшественники РНК — для превращения в функционально активные цепи РНК должны пройти процесс созревания.

## Эпигенетические факторы подавления и активации транскрипции

В последние десятилетия интенсивно развивается *эпигенетика*, раздел генетики, который изучает формирование и наследственную передачу специфического функционального состояния генома (Корочкин, 2005). Использование молекулярно-генетических методов позволило выявить *эпигенетические события*, влияющие на процессы транскрипции: метилирование, ацетилирование, фосфорилирование, аллельные исключения, дифференциальный сплайсинг, лайонизация, убиквитинилирование и др.

Российский биохимик Борис Ванюшин впервые установил явление *метилирования генетического материала*. Метилирование осуществляется с помощью специального фермента — метилазы — и заключается в присоединении метильной группы (радикал метана) к цитозину ДНК с образованием 5-метилцитозина (5 – Me – Cyt) (рис. 36). При метилировании зоны ДНК с определенным геном включается механизм генного сайленсинга, или *подавления транскрипции*, вызванного изменением конформации ДНК. Наряду с метилазой предполагается наличие другого фермента — *деметилазы*. Ярким примером является слабое метилирование ДНК первичных половых клеток, тогда как гены яйцеклетки и спер-

мия метилированы интенсивно. Установлены различия в структуре метилирования конкретных генов в мужских и женских гаметах — так называемый *импринтинг*. Он заключается в том, что аллели, унаследованные от отца и матери, экспрессируют по-разному вследствие их различного метилирования при образовании половых клеток. Состояние метилирования может передаваться от клетки ее потомкам в ходе клеточного деления.

Другим эпигенетическим фактором регуляции транскрипции генов эукариот является *ацетилирование гистонов*. Процесс катализируется ферментом *гистонацетилацетилтрансферазой (HATS)*, который переносит ацетильные остатки с молекулы *ацетил-коэнзима А* на остатки лизина в молекулах гистонов, снижая их положительный заряд. Ацетилирование необходимо для начала транскрипции. Американский биохимик В. Олфри (1964) установил, что неацетилированные после трансляции гистоны тормозят транскрипцию ДНК, тогда как их ацетилирование снимает это торможение (Корочкин, 2002). Предполагается, что ацетилирование «разрыхляет» связь ДНК с гистонами, делая генный локус доступным действию регуляторных белков транскрипции. Напротив, *деацетилирование* уплотняет структуру хроматина, делая его недоступным действию активирующих генов. Ацетилирование обратимо и осуществляется по лизиновым остаткам гистонов, которые локализованы в  $\text{NH}_2$ -терминальном домене гистонов.

*Фосфорилирование* — еще один путь активации транскрипции. Он представляет собой АТФ-зависимый механизм фосфорилирования радикалов серина (Ser) гистонов, с образованием отрицательно заряженного остатка *фосфосерина*. При этом специфический белковый комплекс ослабляет связи гистонов с ДНК. В результате менее плотная упаковка ДНК в составе нуклеосом создает условия для инициации и элонгации транскрипции.

## Созревания РНК: процессинг и сплайсинг

Процессинг — это модификация изначально синтезированных в ходе транскрипции молекул РНК, называемых первичными транскриптами или РНК — предшественниками (пре-РНК), приводящая к образованию зрелых, функционально активных молекул РНК. Все происходящие при этом процессы подразделяют на удаление одних нуклеотидов, присоединение других и модификация тех же или третьих нуклеотидов. Для удаления «лишних» нуклеотидов используются *нуклеазы*: *срединых экзонуклеазы* отщепляют с 3'- или -5'-концов по одному нуклеотиду, а *эндонуклеазы* разрезают цепь в средних участках на фрагменты. Для отщепления спейсерных последовательностей нуклеотидов используются *эндонуклеазы*, так же как для разрезания на индивидуальные цепи 45S-пре-рРНК и гистоновой пре-мРНК. В ходе *сплайсинга* из всех средних участков, кроме гистоновых, пре-мРНК и пре-тРНК вырезаются некодирующие интрон-

ные последовательности, а экзонные последовательности сшиваются ферментом лигазой в непрерывную цепь.

В 1977 г. в нескольких лабораториях было показано, что эукариотические гены имеют прерывистое строение. Например, гены  $\beta$ -цепей гемоглобина млекопитающих прерываются в области, кодирующей аминокислотную последовательность, одной длинной не кодирующей вставочной последовательностью (включающей от 574 до 904 п. н.) и одной короткой последовательностью (116–130 п. н.). Это показало, что геномная ДНК действительно содержит нетранслируемые последовательности между кодирующими последовательностями. Вставочные последовательности — интроны — транскрибируются вместе с кодирующими последовательностями — экзонами, а затем в процессе образования зрелой мРНК интроны удаляются. Число интронов в прерывистых генах эукариот колеблется от 2 (в гене  $\beta$ -глобина) до 50 (в гене миозина). Размеры интронов также многообразны. Например, два фрагмента иммуноглобулинового гена разделены интронами длиной 1250 п. н. В некоторых генах общая длина интронов существенно превышает длину экзонов. Так, ген овальбумина куриного яйца содержит около 7700 п. н., а его РНК имеет длину всего 1859 нуклеотидов. Как правило, вставочные последовательности прерывистых генов длиннее, чем экспрессирующиеся последовательности. Установлено, что эволюционные изменения в последовательностях интронов происходят быстрее, чем в экзонах.

При сплайсинге мРНК происходит разрыв двух фосфодиэфирных связей и образование одной новой связи. Предполагается, что эндонуклеазная и лигазная активности содержатся в одном ферментном комплексе, и последовательные реакции сплайсинга согласованы между собой. Нуклеотидные последовательности в местах стыковок экзонов и интронов в первичных транскриптах мРНК характеризуются высокой степенью консервативности: интроны начинаются с ГУ, а заканчиваются АГ.

Существуют по меньшей мере две ферментные системы сплайсинга: одна для образования мРНК, а другая для тРНК. Процесс сплайсинга тРНК сходен у эволюционно далеких видов. Для выявления вопроса о том, может ли ген одноклеточного представителя эукариот транскрибироваться, а его РНК-продукт — подвергаться правильному процессингу в клетках амфибий, гены одной из тРНК дрожжей вводили в ооциты шпорцевой лягушки *Xenopus laevis*. Опыт показал, что транскрипция гена тРНК дрожжей в клетках *Xenopus* осуществляется, а 14-нуклеотидная вставочная последовательность первичного тРНК-транскрипта удалялась. Другие экспериментальные исследования также показали, что специфичность ферментов сплайсинга сохранилась на протяжении большого временного периода эволюции. В заключение можно отметить, что сплайсинг РНК включает разрыв фосфодиэфирных связей на границах экзон — интрон и



образование связей между концами экзонов. Разрыв фосфодиэфирных связей происходит и при других реакциях процессинга, но их образование в молекулах РНК является уникальной особенностью сплайсинга.

Результатом транскрипции ДНК является образование в ядре зрелых молекул РНК:

- 1) тысяч различных мРНК — копий генов, функционирующих в клетках разных типов;
- 2) 4 вида рРНК: 28S-, 18S-, 5,8S- и 5S-РНК;
- 3) несколько десятков видов тРНК — по несколько РНК для каждой из 20 канонических аминокислот.

В ядре мРНК связываются с белками, выполняющими защитную и транспортную функции, после чего перемещаются в цитоплазму. В свою очередь рРНК формируют с рибосомными белками большие и малые субъединицы рибосом, которые в цитоплазме участвуют в конвейерной сборке полипептидных цепей.

## Регуляторные последовательности в ДНК эукариот

Регуляция активности у эукариот проявляется в активации или инактивации транскрипции с отдельных генов с участием регуляторных последовательностей. К ним относятся ТАТА-боксы, энхансеры (усилители), сайленсоры (глушители), а также адапторные элементы (рис. 19).

Элементы регуляции транскрипции могут находиться в тысячах пар нуклеотидов от промотора эукариотического гена, а их дистанционное влияние на транскрипцию осуществляется посредством регуляторных белков.

Энхансеры обычно располагаются достаточно далеко от регулируемого гена, при этом один энхансер может активировать различные гены. Действие их может быть ткане- и видоспецифичными. Они могут находиться как в 5', так и 3'-участках гена, а также в интронах. Энхансеры могут быть удалены от промоторов на расстояние до 20 тыс. п. н.

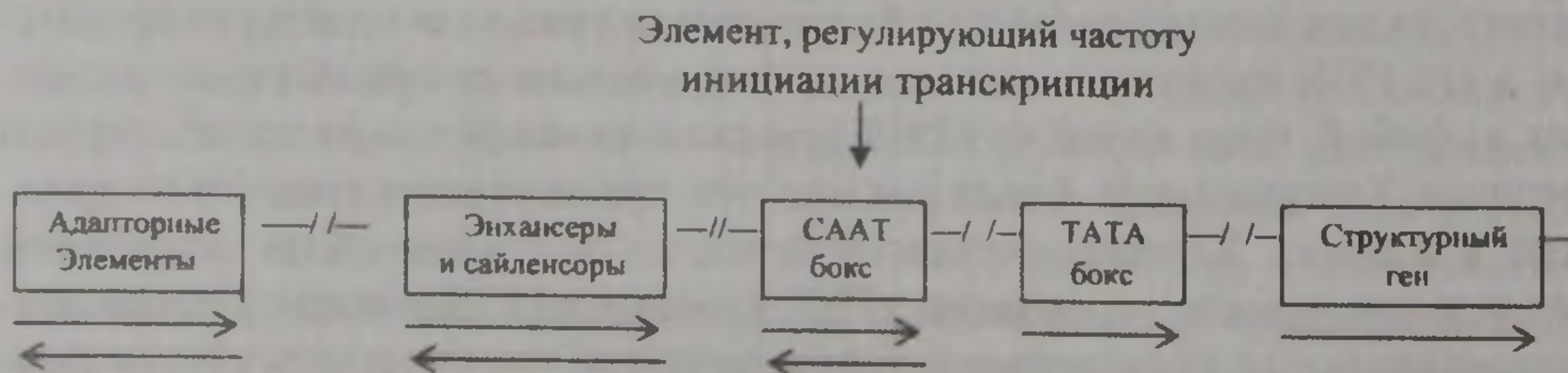


Рис. 19. Модульное строение регуляторных последовательностей в ДНК эукариот (Коничев, Севастьянова, 2003 г.). Стрелки указывают на возможность различной ориентации регуляторных последовательностей

*Сайленсоры* — это ослабители транскрипции, являющиеся негативными элементами по отношению к транскрипции. Они блокируют транскрипцию генов, могут оккупировать большие области ДНК и закрывать доступ белкам, которые инициируют транскрипцию.

*Адапторные элементы* генома эукариот обладают избирательной чувствительностью к стероидным гормонам, глюкокортикоидам, специфически регулируют клеточный ответ на тепловой шок, действие металлов (кадмий, цинк) и некоторых химических соединений.

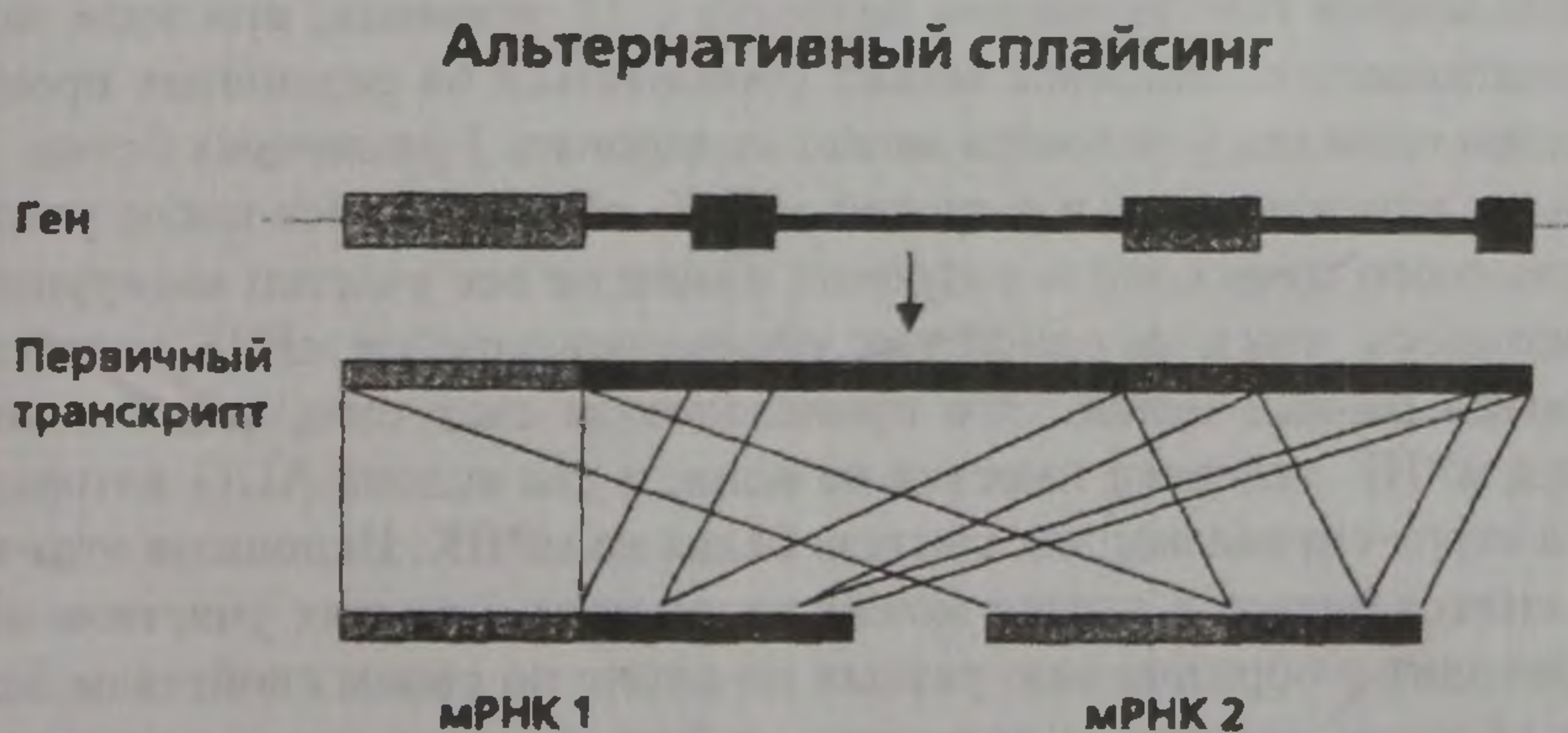
*CAAT-бокc* выявлен в области промотора эукариотических транскриптов как элемент, регулирующий частоту инициации транскрипции.

*TATA-бокc* — небольшая область инициации в промоторе эукариот, сходная с боксом Прибнова у бактерий.

## Альтернативный сплайсинг. Редактирование РНК

Для эукариот большое значение имеют посттранскрипционные события, связанные с пре-мРНК. Это обусловлено мозаичным строением генов эукариот, содержащих в некоторых случаях до нескольких десятков не кодирующих последовательностей — интронов. В ходе последующего сплайсинга из первичного транскрипта пре-мРНК удаются интроны, соседние кодирующие структуру белка экзоны соединяются и образуется зрелая молекула мРНК.

В настоящее время установлено, что содержащиеся в мРНК интроны могут сшиваться в разных комбинациях и образовываться в разные матричные последовательности. Соединение может происходить не только



**Рис. 20.** Схематическое изображение альтернативного сплайсинга (Тарантул, 2003). Прямоугольниками разного оттенка обозначены экзоны, черными линиями — интроны. Показаны лишь две из множества возможных мРНК, образующихся на одном гене

между соседними экзонами, но и между экзонами, отстоящими в гене на значительном расстоянии. Это создает возможность синтеза многих белков на базе одной кодирующей последовательности гена. Оказалось, что мозаичное устройство большинства генов представляет чрезвычайно важное эволюционное приобретение высших организмов, так как создает почву для осуществления альтернативного сплайсинга (рис. 20).

Альтернативный сплайсинг дает возможность организму синтезировать изоформы белков в различных клетках, тканях и органах многочисленных организмов, а также разные по структуре и свойствам белки на базе одного гена.

Выделены три механизма альтернативного сплайсинга:

1. При образовании различных мРНК могут использоваться разные промоторы. Это приводит к образованию транскриптов с разными по длине 5'-концами и разным количеством экзонов, что показано для пре-мРНК легкой цепи миозина позвоночных животных.
2. Другим механизмом является изменение сайта полиаденилирования первичного транскрипта и последующее изменение размера и структуры 3'-конца пре-мРНК. Этот способ установлен при образовании двух видов мРНК тяжелой цепи иммуноглобулинов.
3. При выборе различных экзонов из одинаковых пре-мРНК для формирования зрелых мРНК используются различные экзоны, часть которых не включается в сплайсинг. Это показано на примере сплайсинга пре-мРНК тропонина Т-скелетных мышц млекопитающих.

В геноме человека альтернативный сплайсинг характерен для более трети генов. Примером образования более чем одного белка при экспрессии гена явился ген тропонина человека с 18 экзонами, при этом за счет альтернативного сплайсинга может образоваться 64 различных продукта. В среднем один ген у человека может копировать 3 различных белка.

Было отмечено, что и в зрелой мРНК, образующейся после удаления из первичного транскрипта интронов, также не все участки кодируют белки. Оказалось, что и на одной уже сформировавшейся мРНК могут образовываться разные белки. Это происходит за счет того, что в половине молекул мРНК человека имеется не один, а два кодона AUG, которые являются старт-сигналами для синтеза белка на мРНК. Используя этот механизм, синтез белков в клетке может начинаться с разных участков мРНК, что приводит к образованию разных по длине по своим свойствам белков. Таким образом, многие, а возможно, и большинство генов кодируют целые семейства родственных, но существенно различающихся белков. За счет компактной записи в геноме осуществляется принцип экономии генетического материала и образуется большой набор генных продуктов (белков) на относительно небольшом числе генов.

Наблюдается значительная вариабельность разных генов, а также их экзонов и интронов. В основе лежит изменчивость числа и длины интронов. Максимальный по размерам ген дистрофина в геноме человека состоит из 24 млн п. н., а его подавляющую часть составляют интроны со средним размером несколько десятков тысяч п. н. Наибольшим по кодирующей последовательности в геноме человека является ген белка *титина*: его размер 81 тыс. п. н., в том числе длина единичного экзона равна 17 тыс. п. н., а число интронов составило 178.

Редактирование РНК — это ряд модификаций ее первичной структуры в процессе созревания. Она включает модификации азотистых оснований ведущие к образованию необычных оснований (инозин, тимин, дигидроурацил и др.). К другим изменениям относятся вставки уридилловых нуклеотидов внутрь цепи РНК. Все эти модификации могут приводить к синтезу белков с измененной последовательностью аминокислот, не соответствующей нуклеотидной последовательности гена.

### Ключевые слова и термины

Транскрипция ДНК, транскриптон; ДНК-зависимая РНК-полимераза; бокс (последовательность) Прибнова; кэпирование и полидеанилирование; спейсеры; процессинг и сплайсинг; экзоны и интроны; энхансеры и сайленсоры; альтернативный сплайсинг, редактирование РНК.

## Структура гена у прокариот и эукариот

Ген — самое фундаментальное понятие генетики. В своем историческом развитии оно постепенно наполняется конкретным содержанием. По Г. Менделю, дискретные признаки контролируются наследственными факторами, которые при расщеплении неделимы и не смешиваются при слиянии гамет. Дальнейшим развитием классической генетики является хромосомная теория наследственности (Морган, 1927). В рамках классической генетики накапливались противоречия в представлении о генах и других фундаментальных понятиях этой науки. Все эти противоречия концепции генов нашли достаточное разрешение в рамках молекулярной генетики.

*Гены представляют собой протяженные участки (локусы) генома — ДНК, а у некоторых вирусов РНК, кодирующие структуру одной молекулы: полипептида, мРНК, тРНК или рРНК.* По строению гены вирусов, прокариот и эукариот имеют много общего, но есть и существенные различия. С другой стороны, открытие явления альтернативного сплайсинга существенно изменило представление о гене, как единицы наследственности, кодирующей только одну полипептидную цепь. Все это привело к тому, что в современной генетике нет единого общепринятого определения понятия «ген». У одних авторов в основу определения положена структурная организация гена, у других — функция в организме, у третьих ген рассматривается как единица транскрипции и, наконец, к этим функциям добавляется возможность транскрипции с одного гена нескольких вариантов мРНК. С одного участка генома за счет разных рамок считывания, альтернативного сплайсинга и различных промоторов с одного гена могут транскрибироваться несколько мРНК со сходными или различающимися функциями.



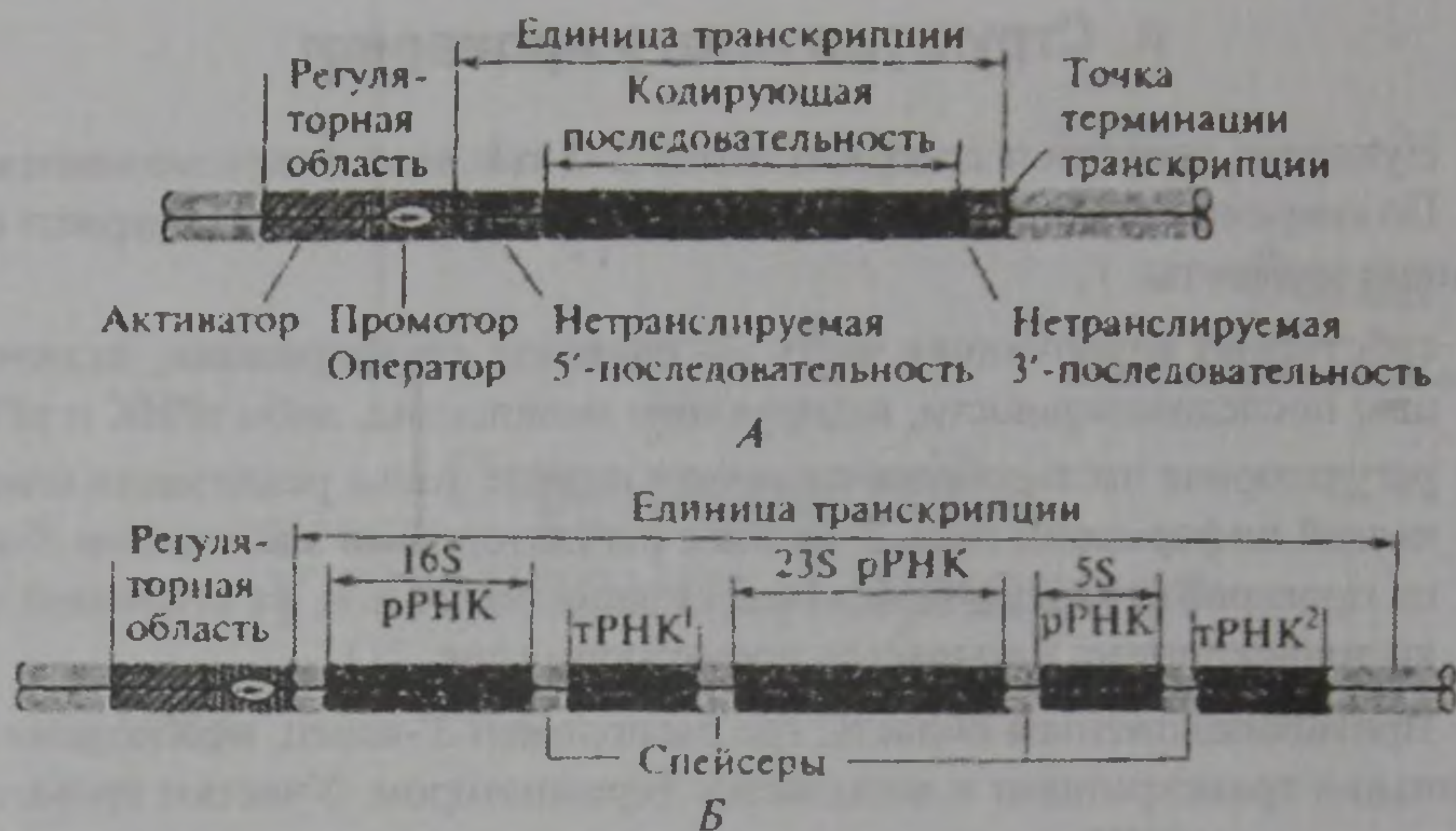


Рис. 22. Структура прокариотических генов (Сингер, Берг, 1998):  
 А — ген, кодирующий один белок; Б — гены, кодирующие рРНК и тРНК

Расстояние от стартовой точки транскрипции в направлении 5'-конца выражается в числе пар нуклеотидов со знаком минус. Так, 35-последовательность на расстоянии  $-35$  оснований участвуют в связывании РНК-полимеразы, а в боксе Прибнова ( $-10$  оснований) начинается локальное раскручивания РНК-полимеразы для создания условий для старта транскрипции РНК. Последовательности 3'-конца формируют в РНК структуру шпильки с образованием Г—Ц пар. Это облегчает терминацию транскрипции и приводит к ее остановке.

У прокариот находят *моноцистронные гены*, кодирующие один белок. Цистрон — это участок ДНК, кодирующий одну полипептидную цепь (рис. 22). Такие гены имеют регуляторные области, а их транскрипция подвержена регуляции. В большинстве случаев единицы транскрипции у *E. coli* и у прокариот являются *полицистронными* и содержат последовательности, кодирующие белки одного и того же метаболического пути или определяющие близкородственные функции. Гены таких последовательностей часто бывают сцепленными (рис. 22Б).

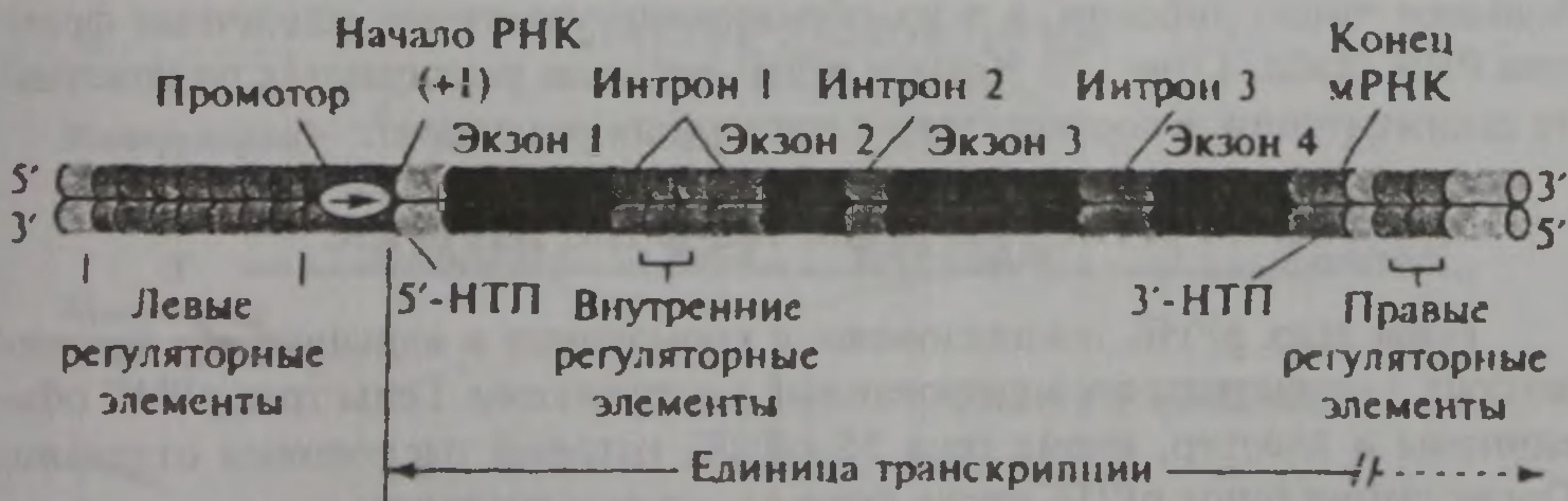
Как правило, транскрипция кодирующих последовательностей в полицистронной единице осуществляется согласованно, с участием общих 5'- и 3'- регуляторных элементов. Такая группа координировано экспрессирующихся генов называется *опероном*. При этом последовательности, кодирующие несколько полипептидов, транскрибируются с образованием зрелой мРНК, которая не претерпевает модификации перед трансляцией. Последовательности, кодирующие разные типы РНК, специфически расщепляются в ходе процессинга с образованием зрелых стабильных РНК-

продуктов. Французские микробиологи Ф. Жакоб и Ж. Моно (1961) показали, что отдельные опероны, как правило не работают изолированно. Они могут быть связаны крест-накрест, образуя так называемый *генетический триггер*. В этом случае продукт, вырабатываемый под контролем одного оперона, управляет состоянием другого, и наоборот.

## Б. Структура гена у эукариот

Число генов в геномах эукариот на порядок больше, чем у прокариот. Сама организация их генов сложнее и характеризуется многими отличительными особенностями. В составе эукариотических генов имеются мозаичные единицы транскрипции, в которых чередуются кодирующие (экзоны) и не кодирующие (интроны) последовательности. Другая особенность состоит в том, что наряду с уникальными генами встречаются многократно повторяющиеся гены. Ген эукариот состоит из следующих элементов (рис. 23):

1. Единица транскрипции с кодирующей областью и расположенные по ее флангам нетранслируемые последовательности. Кодирующие транскрипты большинства эукариотических генов имеют экзон-интронную структуру. Как исключение не имеют интронов гены пяти гистонов и  $\alpha$ - и  $\beta$ -интерферонов. Экзоны чередуются с интронами. Размеры и число экзонов и интронов индивидуальны для каждой мРНК. Обычно число интронов на ген возрастает пропорционально длине кодирующей части, а средний размер экзонов составляет 300 п. н. Считается, что общая длина интронов превышает суммарную длину экзонов от 2 до 10 раз и больше. Так, в геноме человека среднее число экзонов на белок-кодирующий ген составил 8,8, а размер интронов — 3365 п. н.



**Рис. 23.** Структура эукариотического гена, кодирующего белок (Сингер, Берг, 1998) +1 — точка инициации транскрипции; 5'-НТП и 3'-НТП — 5' и 3'-нетранслируемые последовательности



в гене размером 27 тыс. п. н. В то же время ген белка *титина* имеет размер 81 тыс. п. н., при этом длина единичного экзона 17,1 тыс. п. н., а число интронов достигает 178 штук!

Представление об интронах как вариантах некодирующей, «эгоистической» ДНК подвергнуто сомнению после открытия альтернативного сплайсинга, в процессе которого с одного гена транскрибируется несколько вариантов мРНК за счет различных комбинаций экзонов и интронов. Изменилось представление о роли и функции интронов, которые в некоторых случаях проявляют себя как экзоны, а экзоны — как интроны. Кроме того, показано, что в интронах могут находиться промоторы транскрипции, а в одном из интронов гена фактора VIII расположен другой ген («ген в гене»).

2. Кроме единицы транскрипции, в состав генов эукариот имеются основные регуляторные последовательности, выполняющие функции инициации транскрипции (промотор) и ее конца (терминатор). Промоторные элементы у высших эукариот имеют размеры 100–200 п. н. и состоят из коротких нуклеотидных последовательностей. На расстоянии –30 п. н. от точки +1 расположен ТАТА-мотив, который определяет 5'-конец транскрипта. В промоторной зоне тканеспецифических генов имеется последовательность ЦЦААТ на расстоянии –80–(–50) от точки +1.
3. Сегменты гена, повышающие и понижающие уровень транскрипции (энхансеры и сайленсеры), могут быть расположены внутри гена или на значительном расстоянии от него.

## В. Гены рРНК, гены тРНК и гистоновые гены

Типы этих генов имеют свои структурные и функциональные особенности.

Для обеспечения синтеза многочисленных белков в клетке образуется большое число рибосом, а в их образовании участвуют различные фракции РНК (табл. 1, рис. 7). Четыре рРНК рибосом различаются по константе седиментации, в соответствии с которой их различают:

5S-рРНК, 5,8S-рРНК, 18S-рРНК, 28S-рРНК.

Гены всех рРНК локализованы в хромосомах в ядрышковых организаторах — участках, ассоциированных с ядрышками. Гены трех рРНК объединены в кластер, кроме гена 5S-рРНК, который расположен отдельно. Организация генов рРНК очень сходна с генами гистонов:

- эти гены представлены большим числом копий: от 100 у мыши до 280 у человека, а в созревающих ооцитах шпорцевой лягушки за счет амплификации число генов рРНК увеличивается до 2 млн копий на ядро;

- в генах рРНК отсутствуют интроны, при этом относительное содержание пар ГЦ более высокое, чем в среднем по ДНК;
- длина кластера трех генов рРНК достигает 8 тыс. п. н., а гены кластера разделены спейсерами, так же как соседние кластеры между собой;
- транскрипция каждого кластера осуществляется как единое целое с образованием одной молекулы пре-рРНК, которая под действием эндонуклеазы расщепляется на три зрелые молекулы рРНК: 5.8S, 18S и 28S.

В геноме эукариот содержится большое количество генов тРНК: 850 — у дрозофилы, 1300 — у человека. Они также сгруппированы в кластеры, локализованные в разных хромосомах. Всего в эукариотических клетках насчитываются 40–60 основных типов тРНК, по 10–20 копий для каждого типа. Промотор у эукариот расчленен и находится внутри гена, один блок которого находится на расстоянии 8–30 п. н. левее иницирующего кодона, а второй — правее на + 51 п. н. В промоторе гена тРНК имеются А- и В-боксы. В митохондриях млекопитающих гены тРНК чередуются со структурными генами.

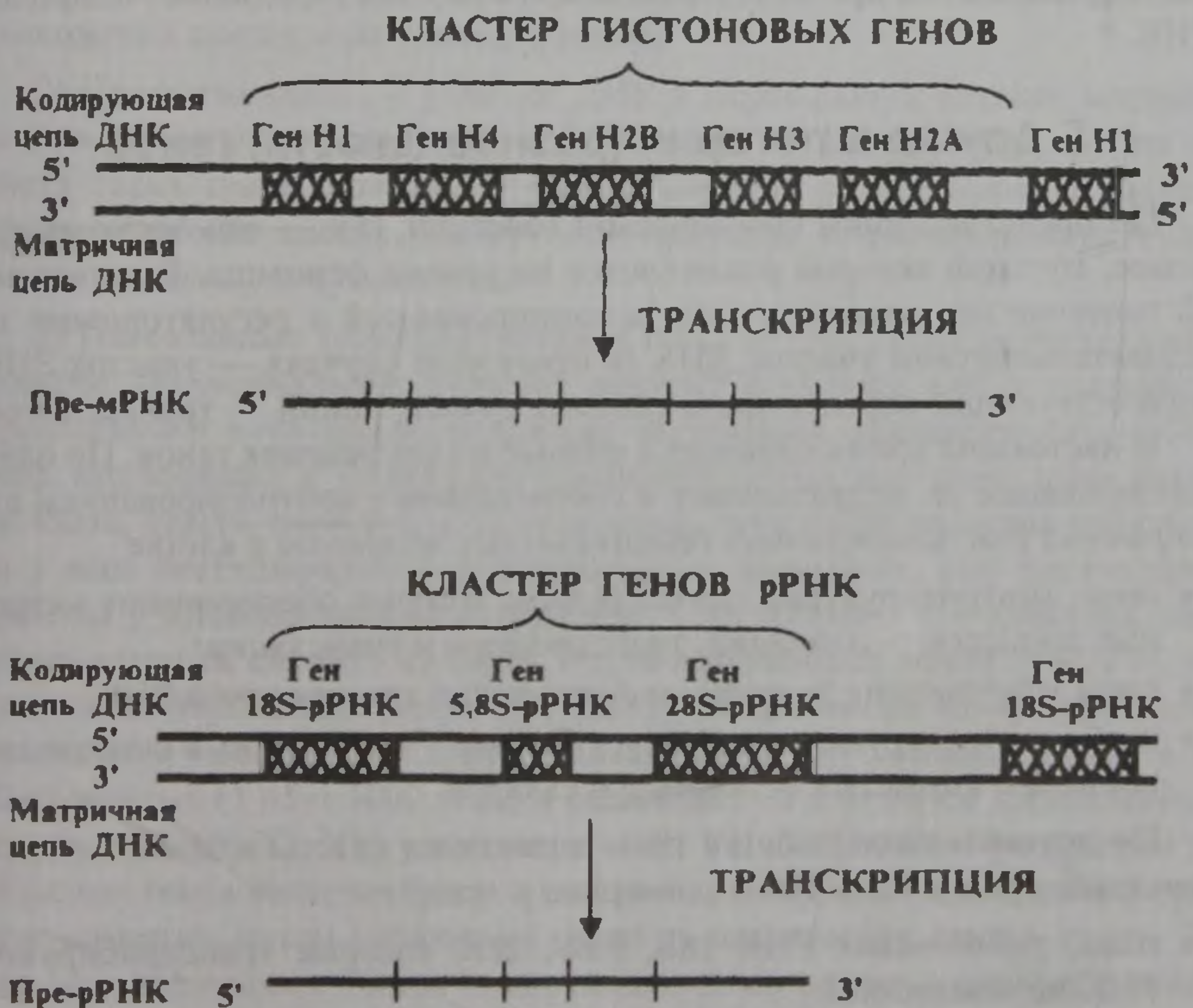


Рис. 24. Кластеры гистоновых генов и генов рРНК:  
А — кластер гистоновых генов, Б — кластер генов рРНК

*Гены гистонов* по своей организации очень сходны с генами рРНК (рис. 24). По кислотно-щелочным свойствам гистоны относятся к основным белкам, глобулы которых в месте с ДНК формируют нуклеосомную структуру хроматина. По количеству они составляют 60–80 % хромосомных белков, обогащенных аминокислотами аргининот и лизином. К ним относятся пять видов белков — Н 1, Н 2А, Н 2В, Н3 и Н4, каждый из которых кодируется отдельным геном. Все пять гистоновых генов соединены в кластеры размером 6900 п. н., повторяющиеся многократно в геномах разных видов, а в хромосомах тандемно следующих друг за другом. Гены, разделены между собой спейсерами, располагаются в кластерах в одной и той же последовательности. К особенностям гистоновых генов относится отсутствие в них интронов и высокое содержание пар GC — например, в геноме человека их относительное содержание 0,39.

У большинства видов эукариот и человека для всех пяти гистонов одна и так же цепь ДНК является кодирующей. И, наконец, транскрипция кластера гистоновых генов протекает как единое целое с образованием одной целой молекулы пре-мРНК с информацией о пяти гистоновых белках. При созревании эта пре-мРНК разделяется на пять отдельных гистоновых мРНК.

## Г. Другие категории (разновидности) генов

По представлениям классической генетики, ген — это локус на хромосоме, мутации которой реализуются на уровне фенотипа. В молекулярной генетике ген трактуется как ассоциированный с регуляторными последовательностями участок ДНК (в некоторых случаях — участок РНК), соответствующий определенной единице транскрипции — транскриптону.

В настоящее время сложились разные подразделения генов. По одной классификации их подразделяют в соответствии с контролем ими *матричных* или *нематричных биохимических процессов* в клетке:

- гены, контролирующие синтез белков, которые обеспечивают матричные процессы репликации, транскрипции и трансляции;
- гены, отвечающие за синтез рибосомной и транспортной РНК;
- гены, контролирующие структуру белков, участвующих в синтезе аминокислот, азотистых оснований и сахаров.

По другой классификации гены делятся на классы в зависимости от транскрибирующей их РНК-полимеразы у эукариот:

- гены, рибосомных РНК 18S, 5.8S, 28S, которые транскрибируются РНК-полимеразой I;
- гены, кодирующие белки, транскрибируемые РНК-полимеразой II;
- гены, 5S-рРНК, тРНК и малых ядерных (мя) РНК, транскрибируемые РНК-полимеразой III.

В основу разделения генов на две группы может быть положено их функциональное назначение: одна группа представлена генами, кодирующими собственные белки, а гены другой группы контролируют синтез разных РНК. По характеру экспрессии гены подразделяют на две группы:

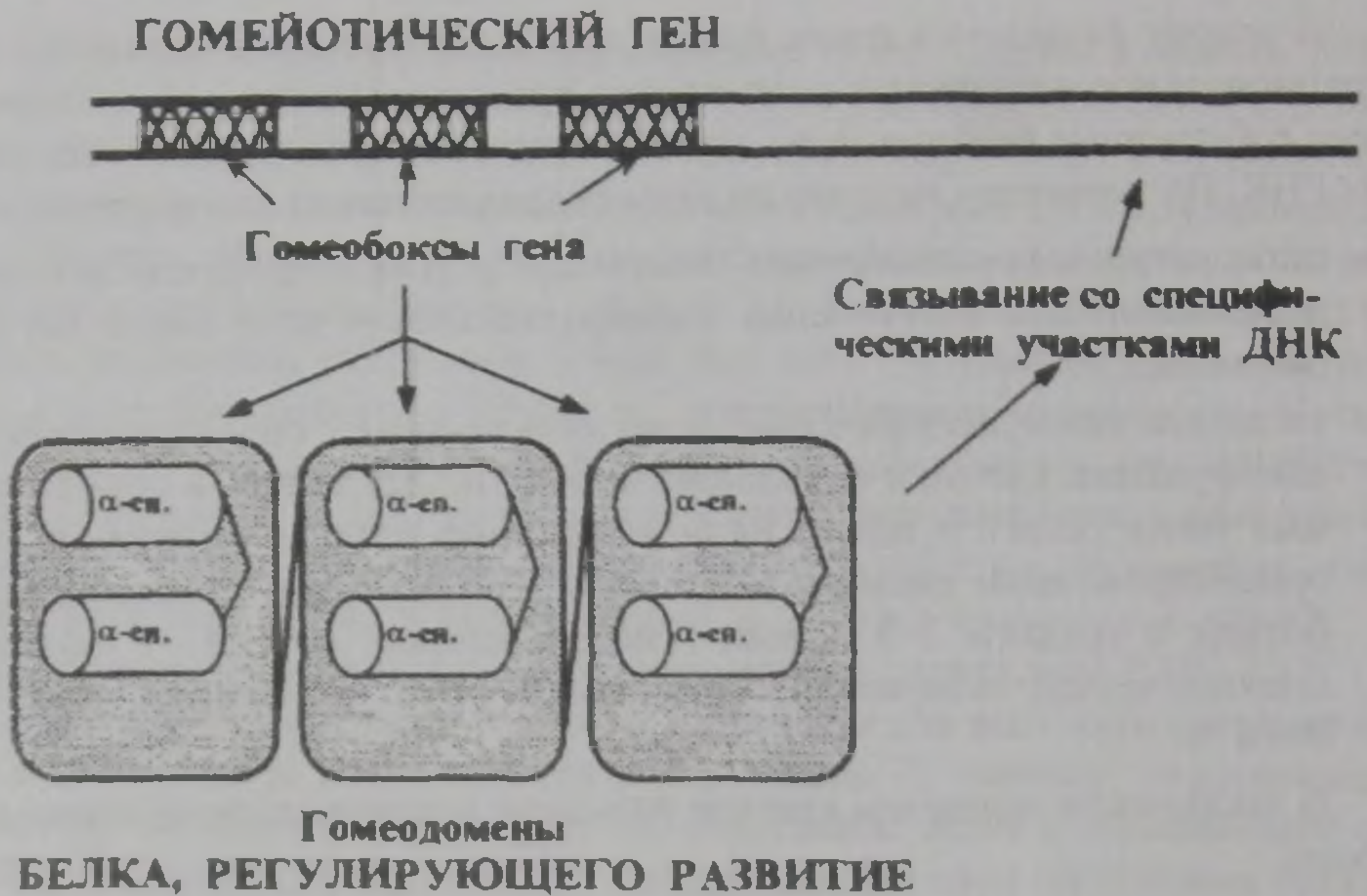
- гены, «домашнего хозяйства» (*housekeeping genes*), продукты которых необходимы для обеспечения жизнедеятельности всех типов клеток организма,
- тогда как тканеспецифические гены обеспечивают специализированные функции клеток и функционально активны только в определенных типах тканей и только на определенных стадиях онтогенеза. По ориентировочным оценкам в тканях млекопитающих и человека работают в среднем 2–3 % всех генов, в клетках печени — основой биохимической лаборатории организма — 5 %, а в клетках мозга — 9–10 %.

В заключении приведем краткое описание некоторых групп генов эукариот:

1) *Онкогены* — их экспрессия подавляет механизмы контролируемого размножения клеток и их трансформации;

2) *Протоонкогены* — участки ДНК в нормальных клетках млекопитающих, сходные по строению с РНК вирусов. Обнаружена группа семейств таких генов, которые играют ключевую роль в пролиферации и дифференцировке клеток. Их мутации приводят к трансформации в *клеточные (целлюлярные) онкогены*;

3) *Гомеозисные (гомейотические) гены* — это гены, регулирующие развитие. Их нормальная функция состоит в выборе или поддержании определенного пути развития клеток, экспрессии определенного набора генов, приводящих к появлению конечного результата этого пути развития: глаза, крыла, ноги и т. д. Гомеозисные гены были найдены практически у всех беспозвоночных и позвоночных животных, они богато представлены у млекопитающих и человека. Они собраны в 4 кластера, каждый из которых состоит из серии тесно сцепленных генов (рис. 25), так называемых *НОХ-генов* (Корочкин, 2002), общее число которых у млекопитающих и человека 39 генов. Каждый кластер занимает область по 120 кб каждый (1 кб = 1000 п. н.) и расположены в четырех хромосомах: у мышей — II, VI, XI, XV, у человека — II, VII, XII, XVII. Продуктами гомеозисных генов является белки *гомеодомены*, которые осуществляют управление развитием путем включения одних и выключения других генов. Экзоны гомейотических генов, кодирующие гомеодомен соответствующего белка, называются *гомеобоксами*. Гомеодомен образован 60 аминокислотами полипептидной цепи. При этом гомеодомены разных гомеозисных генов гомологичны на 80–90 %. Была установлена *коллинеарность*, т. е. кор-



**Рис. 25.** Гомеозисный (гомейотический) ген и гомеодоменный белок

реляция в расположении генов и контролируемых ими признаков, или органов вдоль оси тела: например у млекопитающих активность НОХ-генов соответствует расположению головы зародыша, его грудной области, тазовища и задних отделов.

4) *Митохондриальные гены* — это гены в составе генома митохондрий. Митохондрии представляют собой цитоплазматические органеллы в клетках животных и растений. Они являются энергетическими фабриками клетки, а их функция заключается в синтезе АТФ. Наибольший митохондриальный геном растений в 150 раз превышает геном митохондрий животных. В мтДНК растений имеются интроны, отсутствующие у животных.

У человека каждая клетка содержит несколько сотен митохондрий, а в каждой из них находится по 10 и более небольших молекул ДНК. МтДНК полностью секвенирована в лаборатории молекулярной биологии Медицинского исследовательского центра в Кембриже (Anderson et al., 1981) — ее двух нитевая кольцевая молекула состоит из 16 569 нуклеотидов. Генный состав мт ДНК следующий: 2 гена рРНК, 22 гена тРНК и 13 генов полипептидов (всего 37 генов). В кодировании 1100 митохондриальных белков участвуют также ядерные гены (более 1000): они транскрибируются в ядре, транслируются в цитоплазме и импортируются митохондриями. Установлено, что причиной многих митохондриальных патологий у человека является нарушения взаимодействия между ядерным и митохондриальным геномами. Митохондрии наследуются по материнской линии, то

есть в цитоплазме зиготы у клеток развивающегося организма все митохондрии получены от матери.

Общее количество молекул мтДНК в клетке может достигать десятков тысяч: например, в одном *кардиомиоците* может содержаться до 50 тыс. молекул мтДНК. Во всех клетках и тканях организма имеется нормальная разновидность мтДНК, и это состояние обозначается как *гомоплазмия*. При возникновении мутации и последующей амплификации мутантного клона возникает *гетероплазмия*, при которой в клетках ткани находится совокупность нормальной и мутантной популяции мтДНК.

5) *Гены ДНК пластид (хлоропластов)* представляют собой гены в составе хлоропластной ДНК (хл ДНК) растений. Хлоропласты содержат полный набор ферментов для сложного процесса фотосинтеза, тогда как генетическая информация о структуре этих компонентов распределена между ядерным геномом ядра и геномом самого хлоропласта. Хлоропласты находятся не во всех клетках растений, а только в фотосинтезирующих листьях и стеблях высших растений. Их число на клетку сильно варьирует — от нескольких сотен в крупной клетке водорослей до одного хлоропласта у хламидомонады (одноклеточная водоросль). В клетках мезофилла высших растений содержится только 30 хлоропластов.

ДНК хлоропластов имеет кольцевую структуру. Проведено секвенирование геномов хлоропластов у некоторых видов, в том числе у табака и риса среди покрытосеменных. Средний размер геномов хлоропластов составляет 120–160 т. п. н. Проведена идентификация большинства генов, установлен порядок их расположения. По проведенным расчетам, для функционирования пластид необходимо наличие 120 генов. В опытах установлено, что в хлДНК локализовано 28 генов, кодирующий аппарат фотосинтеза. Гены хлоропластной ДНК кодируют 4 субъединицы РНК-полимеразы, все 4 типа рРНК, 32 типа тРНК, а также третью часть из 60 белков хлоропластных рибосом. Кроме фотосинтеза, пластиды выполняют много других функций, но гены этих функций локализованы в ядре.

После краткого рассмотрения молекулярно-генетических представлений о структуре гена у прокариот и эукариот возникает потребность еще раз окинуть взором процесс интенсивных полувековых исканий в этой области. Одним из первых нужно отметить изучение Н. В. Тимофеевым-Ресовским и его сотрудниками в середине 30-х гг. закономерностей возникновения мутаций под действием рентгеновских лучей с попыткой оценить *физические размер гена*. В этих исследованиях установлено, что ген — неделимая элементарная единица, а его мутации — изменение внутреннего состояния. *Ген* — это макромолекула, поэтому, исходя из законов квантовой химии, сложные молекулы могут иметь изомерные формы, т. е. при том же составе иметь различные структурные формулы. Большая молекула имеет много дискретных устойчивых состояний — *аллелей*. Позднее, в

50-е гг., была подтверждена идея об изомерных переходах между квантовыми состояниями генов. Азотистые основания нуклеотидов ДНК могут иметь *таутомерные варианты*, которые, нарушая канонические правила комплементарности А–Т и G–C, могут спариваться с другими нуклеотидами. Такие нарушения вносят свой вклад в спонтанный мутагенез.

В 1935 г. Н. В. Тимофеев-Ресовский, К. Циммер и М. Дельбрюк сделали доклад о структуре гена и механизме мутации в Геттингенской академии естествознания в Германии «О природе генных мутаций и структура гена» (Timofeeff-Ressovsky, Zimmer, Delbruck, 1935; Тимофеев-Ресовский, Циммер, Дельбрюк, 1996). Наряду с квантовой моделью гена-молекулы, Н. В. Тимофеев-Ресовский сформулировал принцип *конвариантной редупликации дискретных кодов наследственной информации*, т. е. редупликации живых частиц, включающих наследственные вариации: «Всюду, где какие-то элементарные существа размножаются, строят себе подобные рядом, — всюду имеются репликации молекулы. В отличие от процесса роста кристаллов, где тоже есть репликация молекулы, мы называли эти присущие живому процессы репликации *редупликацией*. У всех живых организмов существует спонтанный мутационный процесс, мутации наследственны и они посредством редупликации передаются следующим поколениям живых существ» (Тимофеев-Ресовский, 1983). По современным представлениям, способность к конвариантной редупликации и хиральная чистота является важнейшими атрибутивными свойствами живой материи.

Исследования Тимофеева-Ресовского, Циммера и Дельбрюка оказали огромное влияние на многих выдающихся физиков (Э. Шредингер, Г. А. Гамов, Ф. Крик, М. В. Волькештейн и др.) и математиков (А. А. Ляпунов, В. А. Ратнер, Ю. М. Свирежев и др.). Эти ученые внесли большой вклад в выяснение молекулярной природы гена, зарождение новых научных направлений молекулярной биологии и генетики, создание в генетике информационно-кибернетической концепции. Большое значение для развития молекулярной генетики имело использование в качестве объектов исследований микроорганизмов, их направленность в сторону поиска носителей генов и анализа молекулярной организации фундаментальных генетических процессов. В 90-е гг. оказалось возможным сформулировать фундаментальное понятие генетики — *ген* — на уровне современных достижений (Ратнер, 1983):

Гены представляют собой протяженные участки генома (локусы), сохраняющие свою непрерывность в генетических процессах и осуществляющие в них свою матричную или опознавательную функцию. Они обладают следующими свойствами:

- 1) способность к *конвариантной редупликации* в клетке или эквивалентной ей системе;
- 2) способность к *дискретным случайным изменениям* порядка мономеров, не затрачивающим других генов, в клетке или эквивалентной ей системе;

- 3) способность *обмениваться* любыми частями с гомологичными генами в процессе общей рекомбинации в клетке или эквивалентной ей системе;
- 4) определенный *дискретной локализацией* в геноме относительно других генов;
- 5) способностью *кодировать дискретную порцию информации* (молекулярную функцию или свойство) при помощи порядка нуклеотидов, которая может быть реализована в клетке или эквивалентной ей системе через процессы транскрипции, трансляции или контактного узнавания;
- 6) способность *вступать в синапсис* с гомологичными генами в клетке или эквивалентной ей системе;
- 7) способность *длительно сохранять генетическую информацию* в клетке или эквивалентной ей системе.

После формулирования фундаментальной молекулярной концепции гена за последние 20 лет получены новые экспериментальные факты, расширяющие концепцию. Назовем некоторые из них:

- в составе генов эукариот имеются кодирующие сегменты (экзоны) и некодирующие (интроны);
- в составе генома прокариот и эукариот имеются *мобильные генетические элементы* — подвижные фрагменты;
- некоторые цистроны локализованы *внутри других цистронов* или в их интронах;
- открытие *прионов* (Prusiner, 1996) — инфекционных агентов, которые не содержат генов, но вызывают трансмиссивные спонгиозформные (от «спонги» — губка) энцефалопатии человека и животных;
- явление *«горизонтального переноса» генов*, когда их переносчики (вирусы, плазмиды, мобильные элементы) преодолевают барьеры видовой изоляции.

Эти и другие факты получены экспериментальными и теоретическими методами *генной инженерии* (секвенирование, перенос генов векторами, создание геномных библиотек и др.) и *биоинформатики* (создание компьютерной геномики и протеомики, сравнение генетических текстов, построение деревьев сходства и др.).

## Ключевые слова и термины

Моноцистронные и полицистронные гены, единица транскрипции, кодирующая последовательность, оперон; левые, правые и внутренние регуляторные элементы; кластеры генов; гены «домашнего хозяйства», гомеозисные (гомейотические) гены; гомеодомены, гомеобоксы; мобильные генетические элементы; прионы.



## Трансляция. Биосинтез белка

*Трансляция (биосинтез белка)* — это перевод информации о белке, записанной в виде последовательности нуклеотидных триплетов в первичной структуре нуклеиновых кислот, в полипептидную цепь, формирующую затем белковую структуру. Биосинтез белка является центральным механизмом, в котором сочетаются все элементы отношений кодирования: матричная РНК, транспортные РНК, аминокислоты и готовый белок. Все рассмотренные ранее матричные процессы (репликация, транскрипция) сводились к различным вариантам синтеза *полинуклеотидных* цепей на комплементарных *полинуклеотидных* же матрицах, т. е. новые тексты переписывались на одном языке. В ходе трансляции, т. е. синтеза *полипептидной* цепи на *полинуклеотидной* матрице, осуществляется перевод с четырехбуквенного «языка» нуклеиновых кислот (4 нуклеотида) на двадцати буквенный «язык» белков (20 аминокислот) в соответствии с правилами *генетического кода*. Процесс трансляции осуществляется на рибосомах, с участием основных типов РНК (мРНК, рРНК, тРНК), а также большого количества вспомогательных белков — факторов инициации, элонгации и терминации трансляции.

### Генетический код

Реальные предпосылки для возникновения идей кодирования в генетике явились результатом ряда открытий после 1950 г. Первым было определение в 1953 г. Ф. Сенджером последовательности аминокислот в *инсулине*, белковом гормоне, регулирующем содержание сахара в крови. Впервые было доказано, что белки имеют строго определенную структуру, обусловленную последовательностью аминокислот. Они являются индивидуальными соединениями, у которых изменение положения и характера аминокислоты в полипептидной цепи приводит к возникновению другого белка, с иными физико-химическими и биохимическими свойствами.

В этом же году Дж. Уотсон и Ф. Крик предложили модель строения молекулы ДНК в виде двойной спирали и доказали правильность этой модели на основе рентгенограмм, полученных ранее Р. Франклин и М. Уилкинсом. В 1950 г. английские исследователи П. Колдуэлл и Ч. Хиншелвуд высказали предположение о возможности синтеза белка на матрице молекулы ДНК, а А. Л. Даунс в 1952 г. выдвинул гипотезу, согласно которой синтез белка осуществляется на матрице РНК (Ичас, 1971). Однако окончательное решение проблемы генетического кода осуществил выдающийся русский физик-теоретик Георгий Антонович Гамов, работавший после 1933 г. в Германии и Дании и умерший в 1968 г. в США.

После публикации в мае 1953 г. статьи Дж. Уотсона и Ф. Крика, в которой они объясняли, как наследственная информация хранится в молекулах ДНК в виде последовательности четырех азотистых оснований, Гамов первым сформулировал проблему генетического кода. Генетическая информация в полинуклеотидах ДНК записана в виде последовательности четырех букв, символов нуклеотидов А, Т, G и С. Она затем перекодируется в последовательность букв 20 символов (аминокислот). Гамов и его коллеги постепенно пришли к выводу, что единицей кодирования является *триплет, т. е. кодон*, состоящий из трех нуклеотидов. Правила соответствия триплетов нуклеотидов (кодонов) и аминокислот образуют *генетический код*. Предельное число триплетов составляет  $4^3 = 64$ , т. е. больше чем нужно для кодирования 20 аминокислот. Это указывало на то, что генетический код должен быть *вырожденным*, т. е. содержать для каждой аминокислоты несколько сходных по значению *синонимические* триплетов. Анализ экспериментального материала подсказывал, что код является *неперекрывающимся*, т. е. триплеты должны быть расположены один за другим. Важность работы Гамова, по оценке Ф. Крика, состояла в том, что это была действительно абстрактная теория кодирования. В. А. Ратнер (2001) считает, что это был информационно-кибернетический подход, который позже полностью себя оправдал при разработке теории молекулярно-генетических систем управления и генетического языка.

Первые успехи в экспериментальном решении этой проблемы получили М. Ниренберг и Х. Маттеи (1961) в бесклеточной системе на основе кишечной палочки. При введении в систему полиуридиловой кислоты в продукт белкового синтеза включалась радиоактивная аминокислота фенилаланин. В этом опыте было установлено первое соответствие генетического кода *триплет-аминокислота*: UUU-Phe (УУУ-Фен). Оказалось, что полицитидиловая кислота (поли-С) стимулировала включение в синтезируемый продукт меченого пролина: CCC-Pro (ЦЦЦ-про), а полиадениловая кислота (поли-А) — лизина (AAA-Lys). Вскоре к исследованиям М. Ниренберга подключились лаборатории биохимика С. Очоа и химика-органика Г. Кораны. К 1964 г. была установлена большая часть соответствий триплет — аминокислота, а результатом этой работы является генети-

		Вторая буква								
		U		C		A		G		
Первая буква (5'-конец)	U	UUU	Phe	UCU	Ser	UAU	Tyr	UGU	Cys	U
		UUC		UCC		UAC		UGC		C
		UUA	Leu	UCA		UAA	стоп	UGA	стоп	A
		UUG		UCG		UAG	стоп	UGG	Trp	G
	C	CUU	Leu	CCU	Pro	CAU	His	CGU	Arg	U
		CUC		CCC		CAC		CGC		C
		CUA		CCA		CAA	Gln	CGA		A
		CUG		CCG		CAG		CGG		G
	A	AUU	Ile	ACU	Thr	AAU	Asn	AGU	Ser	U
		AUC		ACC		AAC		AGC		C
		AUA		ACA		AAA	Lys	AGA	Arg	A
		AUG	Met	ACG		AAG		AGG		G
	G	GUU	Val	GCU	Ala	GAU	Asp	GGU	Gly	U
		GUC		GCC		GAC		GGC		C
		GUA		GCA		GAA	Glu	GGA		A
		GUG		GCG		GAG		GGG		G

Рис. 26. Генетический код

ческий код (рис. 26). К указанным его свойствам (триплетность, вырожденность, неперекрываемость) следует добавить *универсальность*, т. е. он един для всех живых организмов. В генетическом коде 61 триплет кодирует одну из протеиногенных аминокислот, а три кодона (UAA, UGA, UAG) являются сигналами остановками трансляции и называются стоп-кодонами или терминирующими кодонами.

Кодон UGA может использоваться для включения в некоторые белки — фермент необычной протеиногенной аминокислоты — *селеноцистеина*.

Триплет, кодирующий метионин (AUG), в начале синтеза кодируется *формилметионин* (fMet) — аминокислоту, которая инициирует процесс синтеза белка.

И, наконец, в ходе генетических экспериментов по расшифровке генетического кода доказано, что кодоны в нуклеиновой кислоте и соответствующие им аминокислоты в белке расположены в одном и том же линейном порядке. Эта закономерность называется *колинеарностью* кода. После расшифровки генетического кода были проведены расчеты «генетического разнообразия жизни»: сколько различных «слов-генов» можно составить из четырех «букв» генетического алфавита? (Дромашко С. Е., 2006). Если размер «алфавита» обозначить через  $n$  ( $n=4$ ), а длину гена —

через  $k$ , тогда из  $n$  элементов можно составить  $n^k$  полимеров. При этом различных полимеров будет в два раза меньше: молекулы, превращающиеся одна в другую поворотом на  $180^\circ$ , например А1, А2, А3, А4 и А4, А3, А2, А1 мы различать не будем. Тогда число различных генов из  $k$  нуклеотидов составит  $n^k/2$ . В среднем ген состоит из 1000 нуклеотидов, т. е.  $n = 1000$ , тогда число разнообразных генов, которые можно построить на основе «четырёхбуквенного алфавита», равно  $n^k/2 = 4^{1000}/2 = 2^{1999}$ . Это число превосходит количество атомов в солнечной системе.

## Биосинтез белка

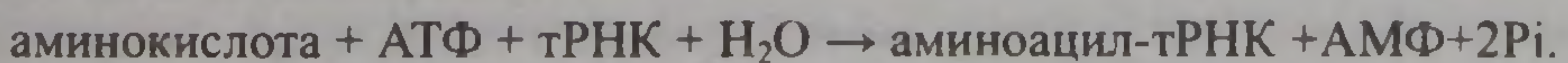
В процессе трансляции различают стадии активации аминокислот, аминоацилирование тРНК и, соответственно, трансляции.

На стадии активации протеиногенные аминокислоты должны присоединиться к тРНК, которая осуществляет их доставку к рибосомам. Важными участками тРНК являются *антикодоны*, служащие для взаимодействия с комплементарным кодоном мРНК, и *акцептирующий* конец для присоединения аминокислоты. Они обеспечивают перевод последовательности нуклеотидов в последовательности аминокислот синтезируемого белка в соответствии с правилами генетического кода.

Присоединение аминокислот к тРНК, с одной стороны, активирует ее карбоксильную группу для последующего образования пептидной связи, а с другой стороны, обеспечивает узнавание соответствующего кодона мРНК. Образование пептидной связи между аминокислотой одной аминокислоты и карбоксильной группой другой аминокислоты *термодинамически невыгодно*. Этот термодинамический барьер преодолевается путем активации карбоксильной группы аминокислоты-предшественника с помощью активированных промежуточных продуктов синтеза белка, названных *аминоацил* — *тРНК* (рис. 27).

П. Замечник и М. Хогленд (1957) установили, что активация аминокислот и их последующее присоединение к тРНК катализируется специфическими ферментами — *аминоацил-тРНК-синтетазами* (рис. 27А). В катализируемых этими ферментами реакция на первой стадии из аминокислот и АТФ образуется *аминоациладенилат* с высвобождением неорганического пирофосфата. Аминоациладенилаты обладают очень высокой реакционной способностью, а их стабилизация осуществляется благодаря прочному связыванию с ферментом. Вторая стадия реакции состоит в переносе *аминоацильной* группы от связанного в активном центре фермента *аминоациладенилата* на 2'- или 3'-ОН-группу концевой рибозы тРНК с образованием *аминоацил-тРНК* (рис. 27Б).

Суммарная реакция превращений, приводящих к образованию *аминоацил-тРНК*, может быть записана в виде:



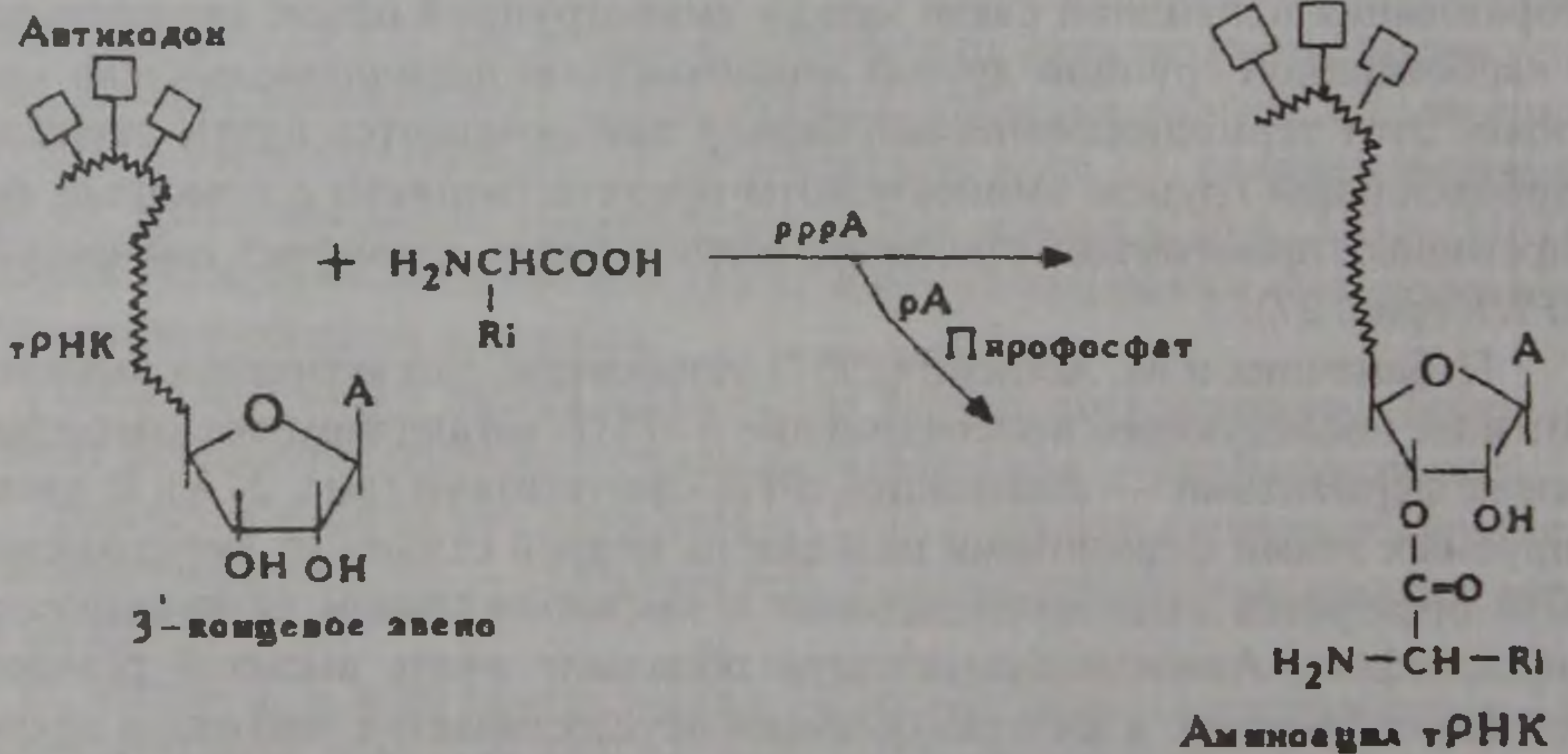
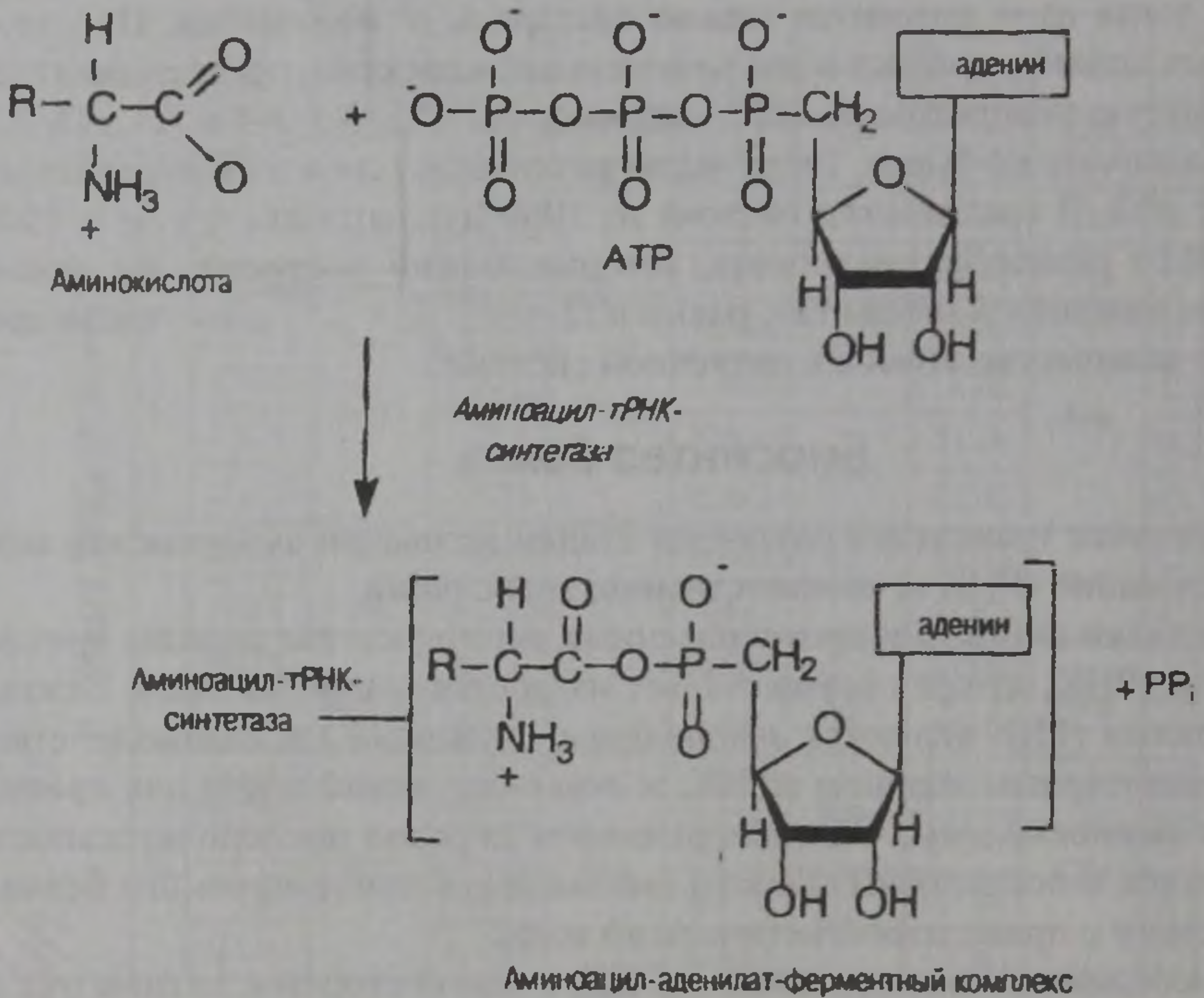


Рис. 27. Активация аминокислот при трансляции: А — образование аминоксил-аденилат-ферментного комплекса; Б — образование аминоксил-тРНК

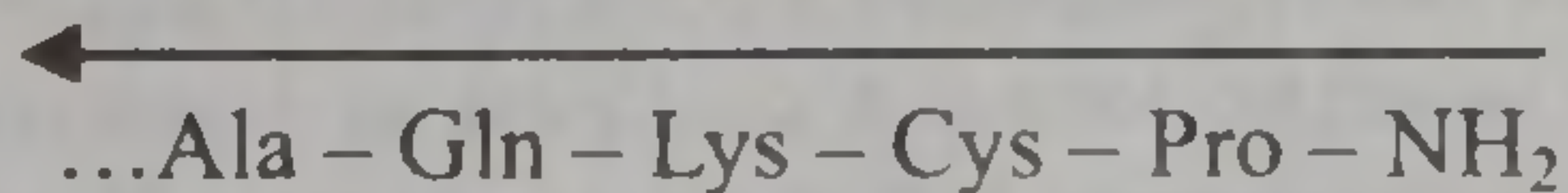
Из реакции следует, что на синтез аминоксил-тРНК затрачивается две богатые энергией фосфатные связи, одна из которых расходуется на образование эфирной связи между аминокислотой и тРНК, а другая при гид-

ролизе пиррофосфата сдвигает равновесие всей реакции в сторону образования продукта.

Ферменты, участвующие в реакции аминоацилирования тРНК обладают высокой субстратной специфичностью. Присоединение к тРНК каждой из 20 кононической аминокислот катализируется определенной аминоацил-тРНК-синтетазой. Фермент должен отличать аминокислоту от 19 других и переносить ее на одну или несколько тРНК, которые называются *изоакцепторными* и обозначаются соответственно  $\text{тРНК}_1^{\text{Gly}}$ ,  $\text{тРНК}_2^{\text{Gly}}$  и т. д. Это связано с вырожденностью генетического кода. После образования аминоацил-тРНК эти молекулы направляются к рибосомам — рибонуклеопротеиновым частицам, специально приспособленным к биосинтезу полипептидных цепей белков. Цитоплазматические рибосомы эукариот и рибосомы прокариот сходны по структуре (рис. 8). Каждая из них состоит из большой и малой субъединиц. Молекулярная масса рибосомы у прокариот составляет 2,5 млн (70S), а у экариот — 4,2 млн (80S). В малой субъединице происходит связывание мРНК и тРНК, а большая участвует в образовании пептидной связи.

Суть процесса трансляции состоит в последовательном декодировании мРНК в направлении  $5' \rightarrow 3'$  с помощью аминоацилированных тРНК, в ходе которого происходит последовательная конденсация аминокислотных остатков, начиная с амино-(N)-конца полипептидной цепи, в направлении к карбоксильному (C)-концу. Матричный принцип процесса соблюдается при узнавании комплементарных нуклеотидов в составе очередного кодона мРНК и антикодона тРНК. Последовательность процесса биосинтеза белка на рибосомах подразделяют на этапы *инициации, элонгации и терминации*.

Для инициации трансляции необходим иницирующий (стартовый) триплет мРНК, инициаторная аминоацил-тРНК и белковые факторы инициации, например, имеются ферменты цепи белка:



В этом фрагменте N-конец, с которого начинается синтез белка, находится справа. Цепь удлиняется справа налево. Главную роль в преобразовании информации с 4-буквенного алфавита в 20-буквенный играют молекулы адапторов — тРНК (на рисунке заштрихованы) (рис. 28). В нижней части этих молекул находится антикодон — тройка оснований, которая на основе принципа комплементарности узнает триплеты (кодоны) на мРНК.

В верхней части к тРНК присоединены с помощью ферментов аминокислоты, соответствующие узнаваемым триплетам.

Процесс движения рибосомы (выделена серым цветом) вдоль мРНК происходит в направлении от 5'-конца к 3'-концу (показано стрелкой). На

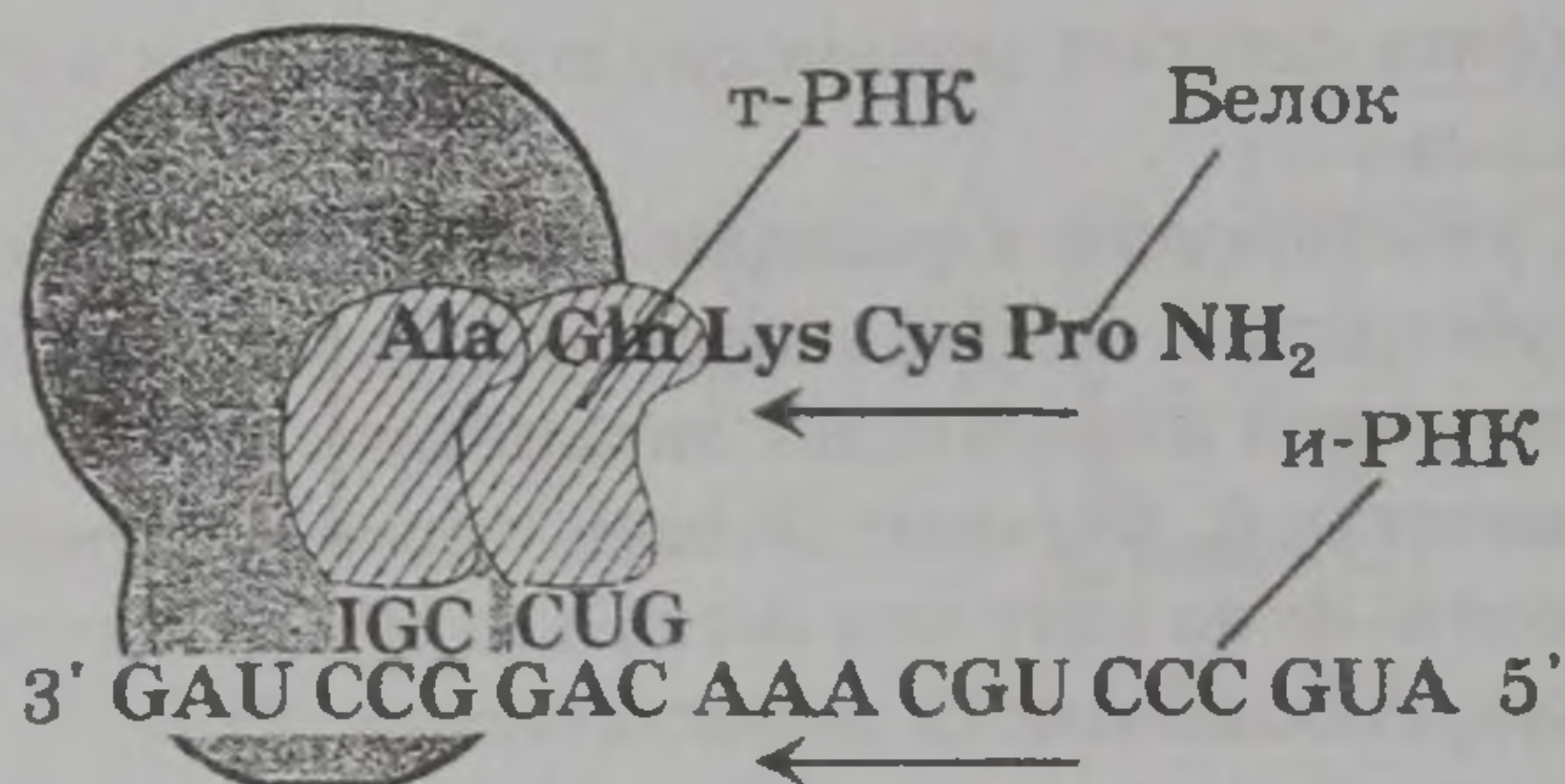


Рис. 28. Принципиальная схема биосинтеза белка

каком-то этапе при участии рибосом на мРНК садятся молекулы тРНК, которые своими антикодонами узнают комплементарные кодоны.

Аминокислоты, которые переносят молекулы тРНК, с помощью рибосомы соединяются в белковую цепь, растущую от N-конца. Начало синтеза полипептидной цепи инициируется триплетом AUG, который кодирует формилметионин (fMet). На мРНК он расположен первым со стороны 5'-конца, читается справа налево. Сигналом конца биосинтеза является триплет UAG (на мРНК он последний, рядом с 3'-концом) и ряд других триплетов, называемых *стопкодонами*. Для них тРНК отсутствует. В процессе синтеза цепь белка и мРНК имеют одинаковое направление, а последовательность кодонов мРНК соответствует последовательности аминокислот (*принцип коллинеарности*). Коллинеарный характер этих цепей обеспечивает возможность синтеза белка.

Скорость синтеза белков у различных представителей прокариот и эукариот варьирует в широких пределах. У бактерий при 37° C за 1 секунду может включаться в растущую полипептидную цепь от 10 до 20 аминокислот. Подсчитано, что белковая молекула, состоящая из 300 аминокислот, синтезируется бактериальной клеткой за 20 секунд. Скорость синтеза белка у эукариот заметно ниже: в среднем за 1 секунду к растущей цепи полипептида присоединяется 1–2 аминокислоты. У эукариот мРНК имеет моноцистронную структуру, а ее 5' – лидерная нетранслируемая часть занимает около 100 нуклеотидных остатков. Далее следуют кодирующая и 3' – нетранслируемая часть, размер последней — около 1000 нуклеотидных остатков. Обычно мРНК транслируется одновременно на кластере рибосом называемом *полисомами* (*полирибосомами*). Каждая стадия трансляции (инициация, элонгация, терминация) осуществляется каждой рибосомой. Преимущество такого процесса заключается в возможности синтеза нескольких копий полипептида, прежде чем мРНК будет расщеплена. Размеры полисом сильно различаются, но обычно они зависят от размера молекулы мРНК. Очень большие молекулы мРНК, состоящие из несколь-

ко тысяч нуклеотидов, могут образовывать полисомные комплексы из 50–100 рибосом, но обычно кластеры содержат по 3–20 рибосом. В отличие от эукариот, процессы транскрипции и трансляции у прокариот тесно связаны и совершаются одновременно. По мере роста цепи мРНК в ходе транскрипции с ней связываются все большее число рибосом (рис. 29А). У эукариот эти процессы разобщены во времени и пространстве: транскрипция проходит в ядре, а трансляция (биосинтез белка) — в цитоплазме. Между транскрипцией ДНК и трансляцией РНК происходит процессинг и транспорт РНК (рис. 29Б).

По мере выхода из рибосомы полипептидная цепочка сворачивается, приобретая структуру, необходимую для проявления ферментативной активности. Последовательность аминокислот в белке служит связующим звеном между генетической информацией, заложенной в ДНК, и трехмерной структурой белка, лежащей в основе его биологической активности. Особенность белков состоит в том, что каждый из них имеет четко определенную трехмерную структуру. Полипептидные цепи, будучи развернутыми или уложенными случайным образом, лишены биологической активности.

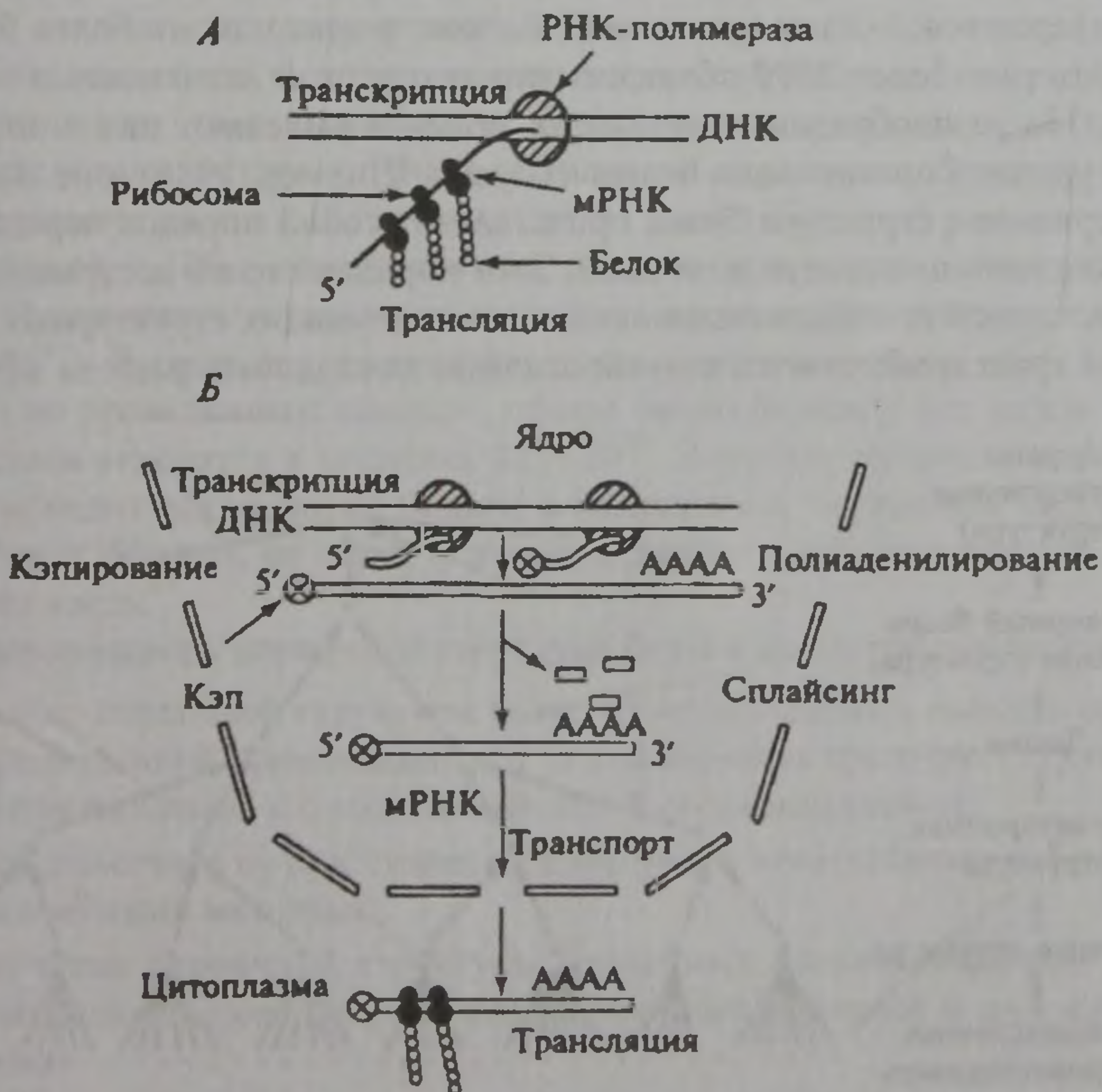


Рис. 29. Различия в транскрипции и трансляции у прокариот (А) и эукариот (Б)



Изменения в последовательности аминокислот могут привести к нарушению нормальной функции белка и, следовательно к соответствующей болезни. Например, серповидноклеточная анемия, тяжелая болезнь, нередко приводящая к летальному исходу, возникает в результате одной замены аминокислоты валина на глутамин в 6-м положении полипептидной цепи  $\beta$ -глобина в состав гемоглобина. В последнее время стало известно, что процесс формирования структуры белка проходит чаще всего уже в момент синтеза. Это показал академик А. С. Спирин на примере присоединения гема к молекуле синтезируемого гемоглобина — присоединение происходит только до определенного момента, когда молекула белка синтезирована не более чем на две трети. Этот и другие примеры показали, что самосборка белка происходит почти одновременно — *ко-трансляционно* — с его синтезом.

## Иерархические уровни структуры белковых молекул

Огромное разнообразие структур и функций белков обусловило их участие во всех молекулярно-биологических процессах. В клетках прокариот содержится 2–3 тыс. различных белков, у эукариот их более 50 тыс., среди которых более 2500 обладают каталитической активностью — *ферменты*. По разнообразию структур их молекул выделяют шесть иерархических уровней организации белков (Шульц, Ширмер, 1986) (рис. 30).

Первичная структура белка представляет собой порядок чередования аминокислот в полипептидной цепи. Этот порядок строго детерминирован в нуклеотидной последовательности соответствующих структурных генов, которые транскрибируются в нуклеотидные последовательности мРНК по

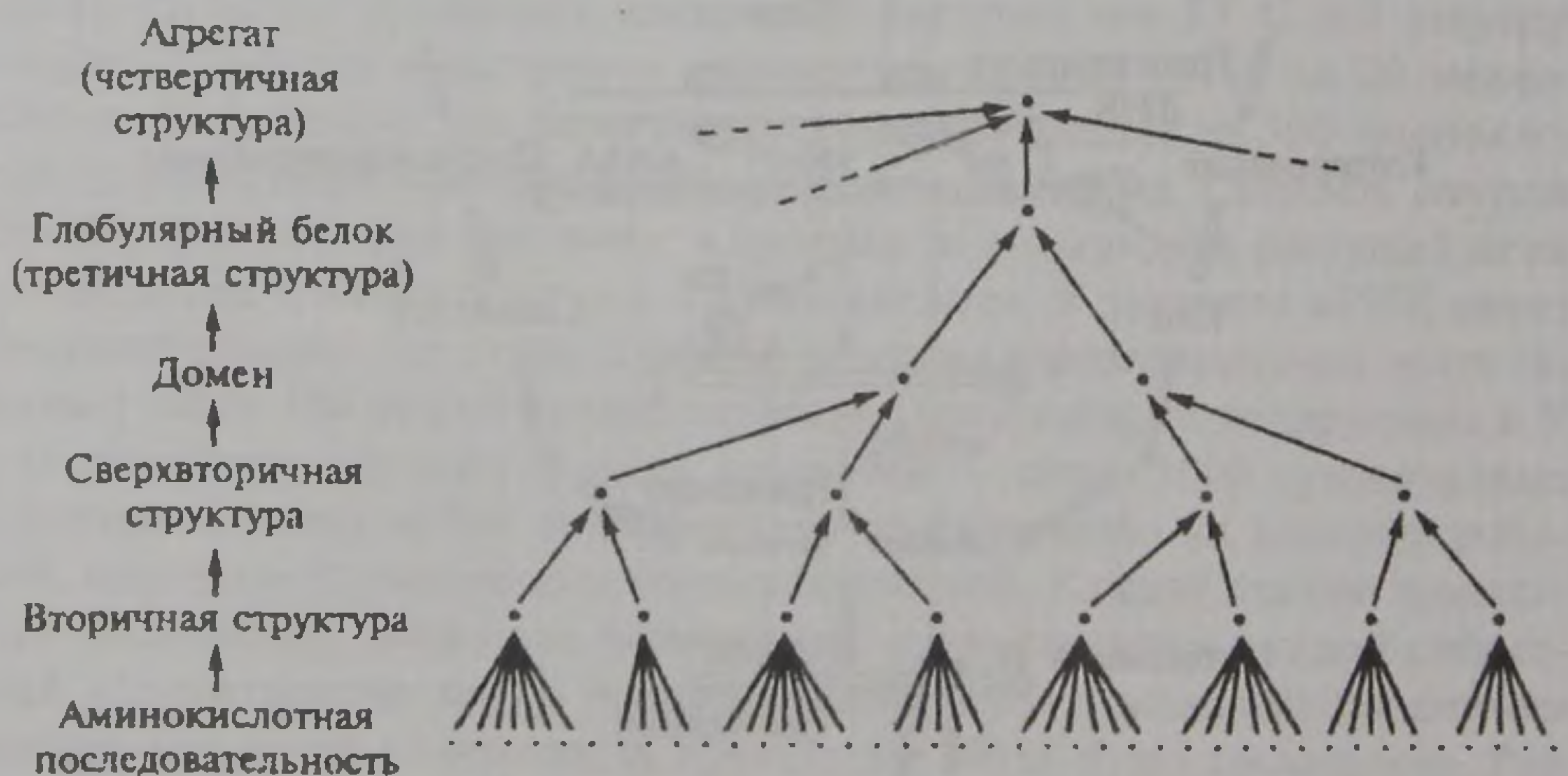


Рис. 30. Иерархические уровни структурной организации белковых молекул

принципу комплементарности. В процессе биосинтеза полипептидных цепей генетическая информация декодируется (транслируется) на рибосомах в цитоплазме. Мономерные единицы белков, природные аминокислоты, являются  $\alpha$ -(L)-аминокислотами. В их молекулах при атоме  $C_\alpha$  имеются четыре различных заместителя. Все  $\alpha$ -аминокислоты, кроме глицина, имеют асимметрический (хиральный)  $\alpha$ -углеродный атом и существуют в виде двух энантиомеров — L- и D-аминокислот (рис. 1).

Некоторые белки включают одну полипептидную цепь, другие — две или более. В состав белков входят до 20 боковых групп (радикалов)  $R_1, R_2, \dots, R_n$ . Для каждого белка имеется свой специфический порядок расположения боковых групп вдоль полипептидной цепи. Боковые группы (радикалы) имеют самое различное строение. Именно это химико-физическое свойство белков — сочетание в одной макромолекуле разнообразных белковых групп (алифатических, ароматических, гетероциклических, полярных, неполярных, заряженных, способных образовывать водородные связи) обеспечивает неограниченные возможности пространственной организации макромолекул, их богатства и разнообразия химических свойств и биохимических функций. Содержание аминокислотных звеньев в белках может заметно отличаться: полипептидные цепи могут включать несколько тысяч аминокислотных остатков — цепь тиреоглобулина состоит из 2750, а цепи отдельных трансмембранных белков-рецепторов содержат до 5000 аминокислотных остатков. Число возможных первичных структур белков неограниченно велико. Например, для не самой большой полипептидной цепи длиной в 100 аминокислотных звеньев (при числе мономеров, равному 20) возможное количество вариантов первичной структуры составит  $20^{100}$  ( $10^{130}$ ). В природе реализована лишь ничтожная часть этих возможностей: по проведенным оценкам, общее число белков у всех видов живых организмов находятся в пределах  $10^{10}$ – $10^{12}$ . У живых организмов постоянно происходит образование белков с измененной первичной структурой (мутантных белков), но через сито естественного отбора проходит только их малая часть.

Информация о первичной структуре белков имеет огромное значение:

- а) Знание первичной структуры белка позволяет сделать вывод о его пространственной организации, т. е. о вторичной и третичной структурах, а следовательно, и о механизмах его функционирования;
- б) Она облегчает путь к синтезу белка чисто химическими и биотехнологическими методами;
- в) Изучение первичной структуры мутантных белков открывает путь к выяснению характера ряда наследственных болезней и их генной терапии.

Анализ структуры белков в настоящее время ускорено путем создания секвенатора — специального прибора для автоматического определе-

ния последовательности аминокислот. С помощью этого прибора можно определить аминокислотную последовательность полипептида или белка, содержащего до ста аминокислотных остатков.

До недавнего времени считалось, что последовательность нуклеотидов мРНК всегда транслируется непрерывно от иницирующего до терминирующего кодона. Открытие явления *динамического репрограммирования* трансляции показало, что первичная структура синтезируемого белка может изменяться по ходу трансляции, а мРНК может служить матрицей для синтеза различных по структуре белков. Кроме механизма альтернативного сплайсинга (рис. 20), это может происходить за счет перепрограммирования (перекодирования) мРНК.

Перепрограммирование трансляции обычно происходит на один–два нуклеотида в разных направлениях: в сторону 5'-конца мРНК (-1 нуклеотид) или в сторону 3'-конца (+1, +2 нуклеотида) (рис. 21). Для изменения рамки считывания необходимо наличие определенных сигналов в транслируемой области:

- а) У ретровирусов рамка считывания сдвигается на (-1) в области гептануклеотидной последовательности перед шпилечной структурой.
- б) Рибосомы в ходе трансляции могут совершать «прыжки», обходя терминирующий кодон и пропуская часть кодирующей последовательности мРНК, в которой имеется особая петельная структура.
- в) Включение в синтезируемые полипептидные цепи ряда белков редкой, функционально важной 21-й белковой аминокислоты *селеноцистеина* с помощью уникальной по структуре аминацил-тРНК (тРНК<sup>Sec</sup>). Для присоединения селеноцистеинила-тРНК<sup>Sec</sup> служит кодон UGA, а в некоторых условиях также два других стоп-кодона (UAA и UAG) могут служить для включения селеноцистеина в синтезированный пептид.
- г) К перепрограммированию относятся также процесс *транс-трансляции* у зубактерий с участием необычного вида РНК-тмРНК (10S РНК), совмещающей в себе свойства транспортной и матричной РНК. При этом синтез полипептида, начатый рибосомой на одной мРНК, переключается на трансляцию другой РНК.

*Вторичная структура белков.* Пространственные структуры белков обладают значительной сложностью, поэтому различают несколько уровней этих структур. *Молекулярный уровень* организации отражают *вторичная и третичная структуры*, а их *надмолекулярную организацию* — *четвертичная структура*. В качестве вторичной структуры рассматривается любая пространственная организация участков полипептидной цепи, а особое внимание уделяется регулярным пространственным структурам. Наиболее часто встречаются *α-спирали (α-структуры)* и *β-складчатые слои (β-структуры)*. В этих структурах атом кислорода группы СО каждого звена

образует внутримолекулярную водородную связь с атомом водорода группы NH. По представительству  $\alpha$ -спиралей и  $\beta$ -структур различают белки типа  $\alpha$  (содержат только  $\alpha$ -спирали), типа  $\beta$  (имеют только  $\beta$ -слои), типа  $\alpha + \beta$  (содержат  $\alpha$ -спирали и  $\beta$ -слои, отчетливо разделенные вдоль полипептидной цепи), типа  $\alpha / \beta$  (имеют перемежающиеся участки  $\alpha$ -спирали и  $\beta$ -структур).

*Сверхвторичные структуры* — это ансамбли, агрегированные из регулярных вторичных структур. Этот уровень представлен у некоторых фибриллярных и многих глобулярных белков.

*Домены* — это две или более пространственно разделенные области молекулы белка, которые слабо взаимодействуют между собой и могут считаться самостоятельными до 150 звеньев. Пространственные структуры доменов могут быть сходными для нескольких и для многих белков. Известны случаи, когда разные белки с совершенно непохожими биологическими функциями включают сходные по структуре домены. Можно предположить, что в природе были отобраны некие основные структурные каркасы, используемые во многих случаях. Возможно, что эти каркасы (домены) играют роль «блоков конструктора», из которых разными способами складываются разные «конструкции» (третичные структуры).

*Третичная структура белков.* Под ней понимают расположение в пространстве всех атомов полипептидной цепи белковой молекулы. В большинстве случаев полипептидные цепи белков сворачиваются в глобулы, образуя *глобулярные белки*. Внутри глобулы образуется гидрофобное ядро из аминокислотных остатков Ala, Val, Ile, Leu, Met, Phe. Глобулярная структура хорошо приспособлена для выполнения ферментативных функций. Оказалось, что не любая произвольная созданная полипептидная цепь способна к образованию устойчивой пространственной (третичной и др.) структуры. Полипептидные цепи природных белков, т. е. их первичные структуры, претерпели отбор в процессе эволюции, в том числе и по способности к образованию единственной устойчивой конформации. Особые белковые комплексы — *шапероны* — способствуют процессу формирования пространственной структуры (*фолдингу*) синтезированных полипептидов.

*Четвертичная структура белков.* Эта структура представляет собой способ взаимного расположения субъединицы (полипептидных цепей) в пространстве, связанных между собой нековалентными взаимодействиями. Отдельные цепи (субъединицы) названы *мономерами*, а белки носят соответствующие название — *димеры* (состоящие из двух цепей) *тримеры* (состоят из трех цепей), а молекула в целом — *мультимер*. Мультимерный блок может состоять из двух субъединиц, например, алькогольдегид-рогеназа печени. Более сложными примерами могут служить *тетрамер гемоглобин* (содержит по две пары  $\alpha$ -цепей и  $\beta$ -цепей), *пируватдегидрогеназа* (содержит 24 цепи). Белковые оболочки вирусов — *капсиды* — содержат тысячи полипептидных цепей, а белковая оболочка вируса табачной мозаики со-

держит 2130 идентичных цепей. В капсидах субъединицы объединяются в спиральных структуры. Четвертичная структура, как надмолекулярный уровень организации, закодирована в первичной структуре белка.

Четвертичная структура лежит в основе возникновения изоформ многих ферментов (*изоферментов*). Независимая регуляция экспрессии генов в клетках разных тканей одного организма приводит к образованию изоферментов как тканеспецифических белков. Так, фермент *лактатдегидрогеназа* у человека представляет собой тетрамер, протомеры которого кодируется генами, локализованными в разных хромосомах. Эти субъединицы А и В могут формировать 5 изоферментов различного состава, функционирующих в разных органах и тканях (в скелетных мышцах, в сердечной мышце и др.). Изоферменты широко используются в молекулярных генетике как маркеры генов, в популяционной генетике, а также для диагностики заболеваний.

## Ингибиторы трансляции

Многие антибиотики являются ингибиторами трансляции у микроорганизмов. Непосредственными объектами их воздействие служат функциональные центры рибосом в составе малой или большой субъединицы. В области малой (30S-) субъединицы установлено ингибирующее действие стрептомицина и тетрациклина; первый ингибирует инициацию белков синтеза, а второй — подавляет процесс связывания очередной тРНК. Среди других антибиотиков исследовано подавляющее действие левомицетина (хлорамфеникола) и эритромицина на функции большой (70S-) субъединицы. Отмечается, что эти эффекты имеют избирательный характер, и в отношении эукариотических рибосом они не проявляются.

На трансляцию у эукариот, в том числе животных и человека, способны влиять некоторые антибиотики, дифтерийный и холерный токсины, а также интерфероны. Объектом воздействия циклогексимида является большая субъединица (80S-) рибосом. Другой антибиотик, пурамицин, *обрывает элонгацию* пептидной цепи.

Выделяемый бактериями дифтерийный токсин поражает ткани, заселяемые при инфекции дифтерийными палочками (слизистые оболочки верхних дыхательных путей), а также с током крови внутренние органы. В клетках он *инактивирует* один из факторов элонгации — *транслоказу*, при этом ингибируется синтез белка в целом. Сходный механизм действия холерного токсина в эпителиальных клетках кишечника.

Вырабатываемые в клетках животных и человека, инфицированных вирусами, *интерфероны* являются белками-гликопротеинами. Их образование индуцируют вирусы с двухцепочной РНК (реовирусы и др.), при этом они не проникают внутрь клеток, а остаются на рецепторах клеточной

мембраны. Связываясь с вирусной РНК, интерфероны стимулируют ее распад в клетках и тормозят процесс трансляции. Это вызывает резкое ограничения синтеза как вирусных, там и собственных белков клетки. Трансляция блокируется из-за дефицита основных компонентов инициаторного комплекса — иницирующей тРНК и мРНК.

## Организация генетического материала и передача генетической информации в клетке (заключение)

После расшифровки генома человека оказалось, что собственно кодирующее последовательности ДНК составляют всего около 1 % всего генома. Близкая картина оценивается и у других видов млекопитающих. В любой клетке организма содержится полный набор генов данного вида, но функционально активной, экспрессирующейся, является лишь небольшая их часть в каждой ткани. *Экспрессия гена* — это способность к транскрипции в данном типе клеток на пути к синтезу первичный молекулярных продуктов — РНК и белка (рис. 31).

Различная временная и тканевая избирательность экспрессии генов определяет процессы онтогенеза, дифференцировки и функционирования различных органов, тканей и клеток. Согласно *центральной догме молекулярной биологии*, процесс передачи наследственной информации происходит в направлении ДНК → РНК → белок. На первом этапе происходит *транскрипция*, «переписывание» нуклеотидной последовательности гена путем синтеза комплементарной молекулы мРНК с участием фермента РНК-полимеразы. Первичный РНК-транскрипт (премРНК) проходит стадию созревания — *процессинг*. На этой стадии происходит модификация концевых участков для стабилизации всей молекулы: *полиаденилирование* (присоединение многих адениловых остатков — *поли-А* — к 3' концу), *«кэпирование»* (образование на 5'-конце мРНК эукариот особой структуры — *«кепа»*, необходимой для прикрепления мРНК к рибосоме, с помощью 7-метил-гуанозина). Важнейшим событием процессинга пре-мРНК является «удаление» некодирующих звеньев — интронов. Этот процесс удаления — *сплайсинг* — приводит к тому, что в составе РНК остаются лишь «сшитые» друг с другом смысловые участки гена — *экзоны*.

Следующий этап передачи генетической информации — трансляция — происходит в цитоплазме и заключается в сборке на рибосомах молекул белка на матрице мРНК. Иницирующим триплетом является AUG, кодирующий аминокислоту метионин и открывающий *рамку считывания* генетической информации.

Сигналом окончания трансляции служит один из трех стоп-кодонов (UAA, UAG, UGA), распознавание которых на рибосоме прекращает синтез полипептидной цепи. По окончании трансляции первичная полипеп-

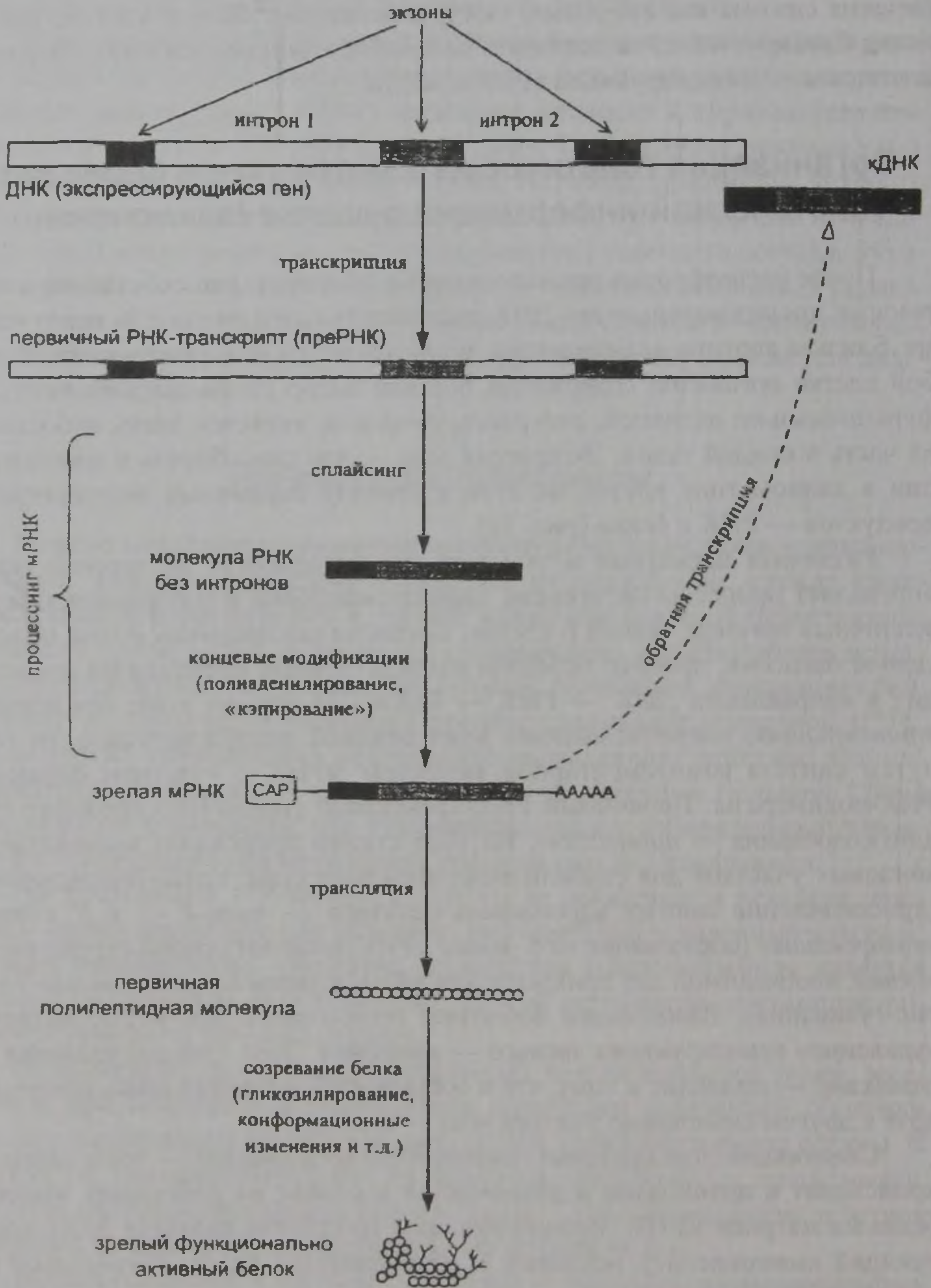


Рис. 31. Передача генетической информации в клетке

тидная молекула претерпевает «дозревание» белка, т. е. определенные посттрансляционные модификации, превращаясь в функционально зрелый продукт.

В природе существует возможность другого пути передачи генетической информации — от молекулы РНК к молекуле ДНК, т. е. в направлении, обратном положению центральной догмы. В естественных условиях такой процесс осуществляется у РНК-содержащих вирусов, обладающих специальным ферментом — *ревертазой (обратной транскриптазой)*. У высших организмов обратная транскрипция проводится генно-инженерными методами *in vitro* с использованием ревертазы. В результате получают искусственные молекулы ДНК, не содержащие интронов и полностью комплементарные матрице мРНК (пунктирная стрелка на рис. 31). Они называются *кДНК, т. е. комплементарная ДНК*. Таким способом получают библиотеки (банки) кДНК разных видов или отдельных хромосом для проведения различных молекулярно-генетических исследований, в том числе для практических целей генодиагностики и генотерапии.

Объектами изучения молекулярной генетики являются информационно-генетические системы, построенные по *иерархическому принципу*. Так, в процессе реализации наследственной информации согласно центральной догме молекулярной биологии в направлении ДНК → РНК → белок, на каждом этапе процесса возрастает количество молекул — носителей управляющих сигналов. Последствием реализации этого принципа является высокий «*коэффициент усиления*» молекулярно-генетических систем управления. Он достигает усиления порядка в  $10^{16}$ – $10^{23}$  раз (Ратнер, 1975; Дромашко, 2006). Это явление Н. В. Тимофеев-Ресовский назвал «*принципом усилителя в биологии*».

### Ключевые слова и термины

Трансляция, биосинтез белка, полипептидная связь; генетический код, триплет, универсальность и вырожденность генетического кода; принцип коллинеарности; иерархические уровни белковых молекул (первичная, вторичная, третичная, четвертичная структуры); ревертаза (обратная транскриптаза).



## Мутации и репарация повреждений ДНК

Мутации представляют собой дискретные изменения разных единиц общего кода наследственной информации живых организмов (Тимофеев-Ресовский, Воронцов, Яблоков, 1969). Эти скачкообразные изменения в генетических структурах возникают спонтанно и передаются последующим поколениям потомков по законам Менделя. В зависимости от затрагиваемой единицы изменения, выделяют следующие типы мутаций:

*Генные мутации* — единицей изменения являются отдельные гены;

*Хромосомные мутации* — единицами изменения являются структурные перестройки отдельных хромосом;

*Геномные мутации* — единицами изменения являются числа отдельных хромосом или числа наборов хромосом у данного вида;

*Цитоплазматические мутации* — дискретные изменения внеядерных автономных и константных генетических структур митохондрий и пластид у растений.

Впервые теорию непрерывно протекающего в природе мутационного процесса создавал директор ботанического сада Петербургского университета С. И. Коржинский. В 1899 г. он издал монографию «Гетерогенез и эволюция», которая в 1901 г. была переведена на немецкий язык и стала известна голландскому ботанику Гуго де Фризу.

Де Фриз проводил исследования на растении *Oenothera lamarckiana* (ослиник) в чистых культурах и наблюдал появление многих мутаций. В своем классическом труде «Мутационная теория» (Де Фриз, 1904) он сформулировал ее основные положения, которые до наших дней не утратили своего значения:

- 1) мутации возникают внезапно, без всяких переходов;

- 2) новые формы вполне константны в своем проявлении;
- 3) мутации в отличие от ненаследственных изменений (флуктуаций) не образуют непрерывных рядов, не группируются вокруг среднего типа (моды);
- 4) они идут в разных направлениях — могут быть полезными, вредными или нейтральными;
- 5) выявление мутаций зависит от количества проанализированных особей;
- 6) одни и те же мутации могут возникать повторно.

Мутации генов возникают у всех живых организмов-вирусов, прокариот и эукариот. Мутационная изменчивость как процесс качественных скачкообразных изменений является всеобщей для всех органических форм. В то время как Г. де Фриз оперировал результатами мутаций — их фенотипическим проявлением, современная теория мутационного процесса концентрирует влияние на природе мутаций и процессе их возникновения. Сам термин «мутация» употребляется для обозначения не только процесса возникновения мутационной формы, но также и носителя конкретного мутационного признака, а иногда — и всех групп мутационных индивидов в популяции.

В настоящее время созданы различные классификации мутаций: по происхождению, по проявлению в гетерозиготном состоянии, по силе проявления аллелей и т. д. Мы рассмотрим только один тип мутации — по уровню организации изменяемого генетического материала, а более подробно-генные мутации.

## Классификация мутаций

Классификация мутаций по уровню организации изменяемого материала представлена на рис. 32.

В результате мутаций возникают новые аллели соответствующих генов, на чем основана генетическая изменчивость живых организмов.

Аллели различаются между собой по составу нуклеотидов в ограниченных участках ДНК. Количество аллелей одного гена может составлять от двух-трех до нескольких десятков, образуя *серии множественных аллелей*, например, серии аллелей по системам групп крови у человека и домашних животных. Во многих случаях мутации не вызывают заметных фенотипических проявлений, как это отмечено при ее локализации в функционально не значимых сайтах молекулы ДНК, а также при возникновении нового кодона с прежним информационным смыслом в связи с вырожденностью генетического кода.

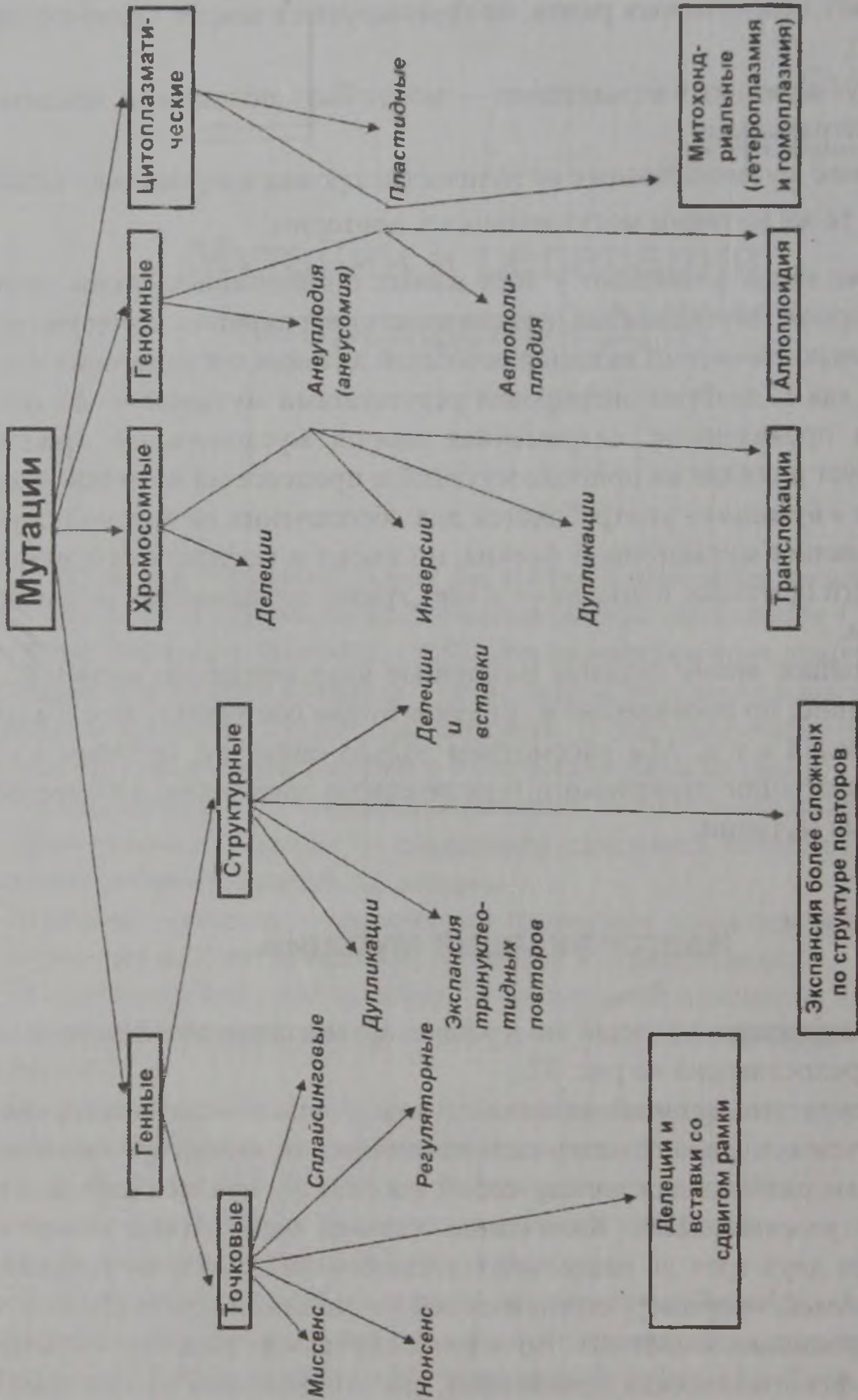


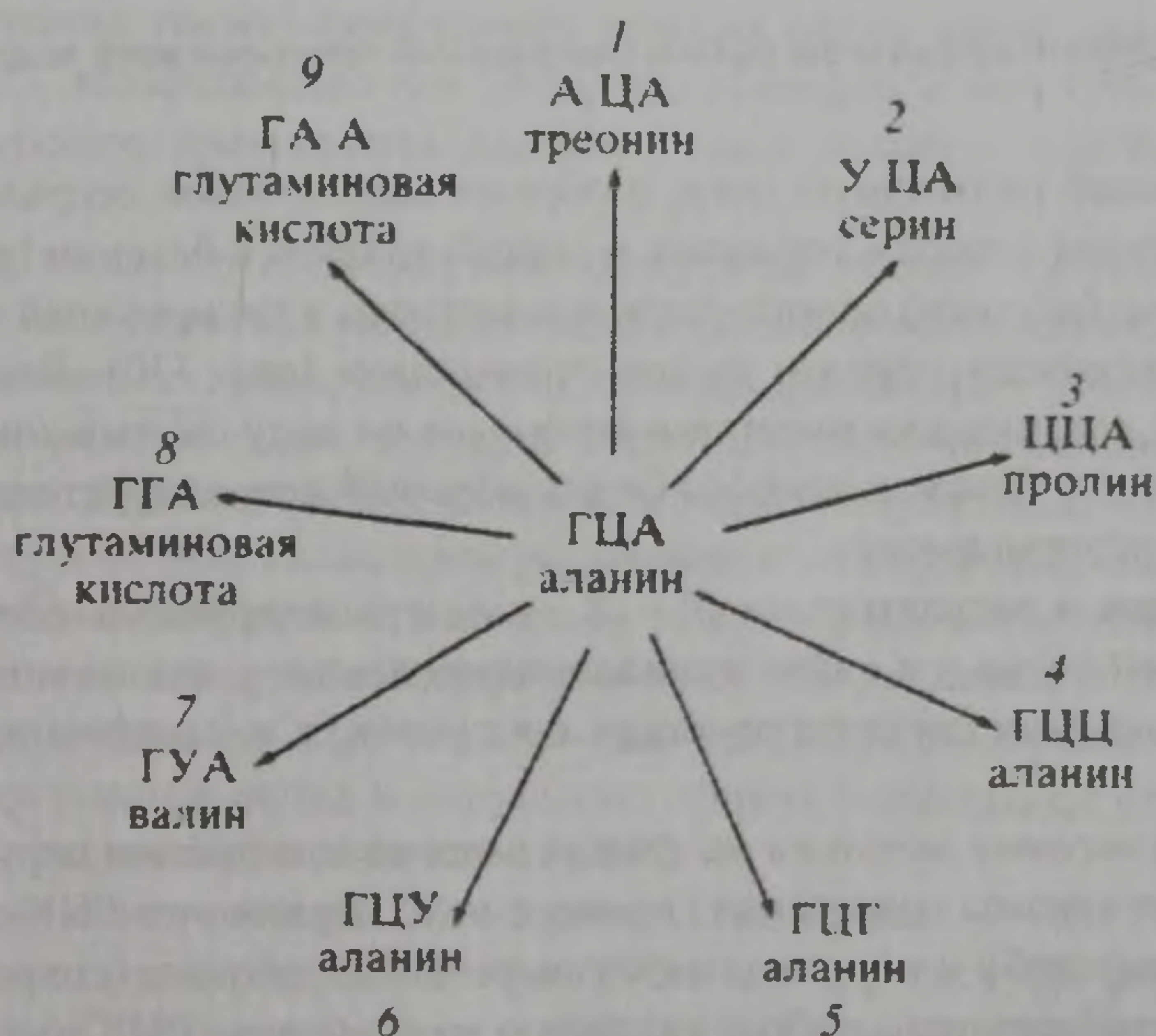
Рис. 32. Классификация мутаций

## Генные мутации

*Генные мутации* представляют собой изменения числа или последовательности нуклеотидов в структуре ДНК (вставки, выпадения, перемещения, замещения нуклеотидов) в пределах отдельного гена, приводящие к изменению количества или качества белковых продуктов. Условно генные мутации можно подразделить на точковые и структурные.

*Точковые мутации генов* — это нуклеотидные замены в кодирующей части гена, то есть в экзонах. В результате могут образовываться мутантные кодоны в виде *миссенс-мутации* (изменяется смысл), *нейтральные мутации* (не измененный смысл) или *нонсенс-мутации* (затрагивает терминирующий кодон). Миссенс-мутации приводят к замене одной аминокислоты в полипептиде на другую, а ее фенотипическое проявление зависит от функциональной значимости затронутого домена. Многие моногенные болезни человека вызываются появлением таких мутаций.

При нонсенс-мутации кодон вместо определения аминокислоты мутирует в один из стоп-кодонов (UAA, UAG, UGA) и не транслируется на рибосомах. Стоп-кодон внутри структурного гена останавливает трансляцию.



**Рис. 33а.** Возможный спектр мутаций одного кодона ГЦА, кодирующего синтез аминокислоты аланина. Благодаря вырожденности кода (молекулярно-генетической основе нейтральности части возникающих мутаций) часть нуклеотидных замещений не приводит к информативно-существенным итогам. В результате нейтральности мутации 4–6 оказываются фенотипически невыраженными

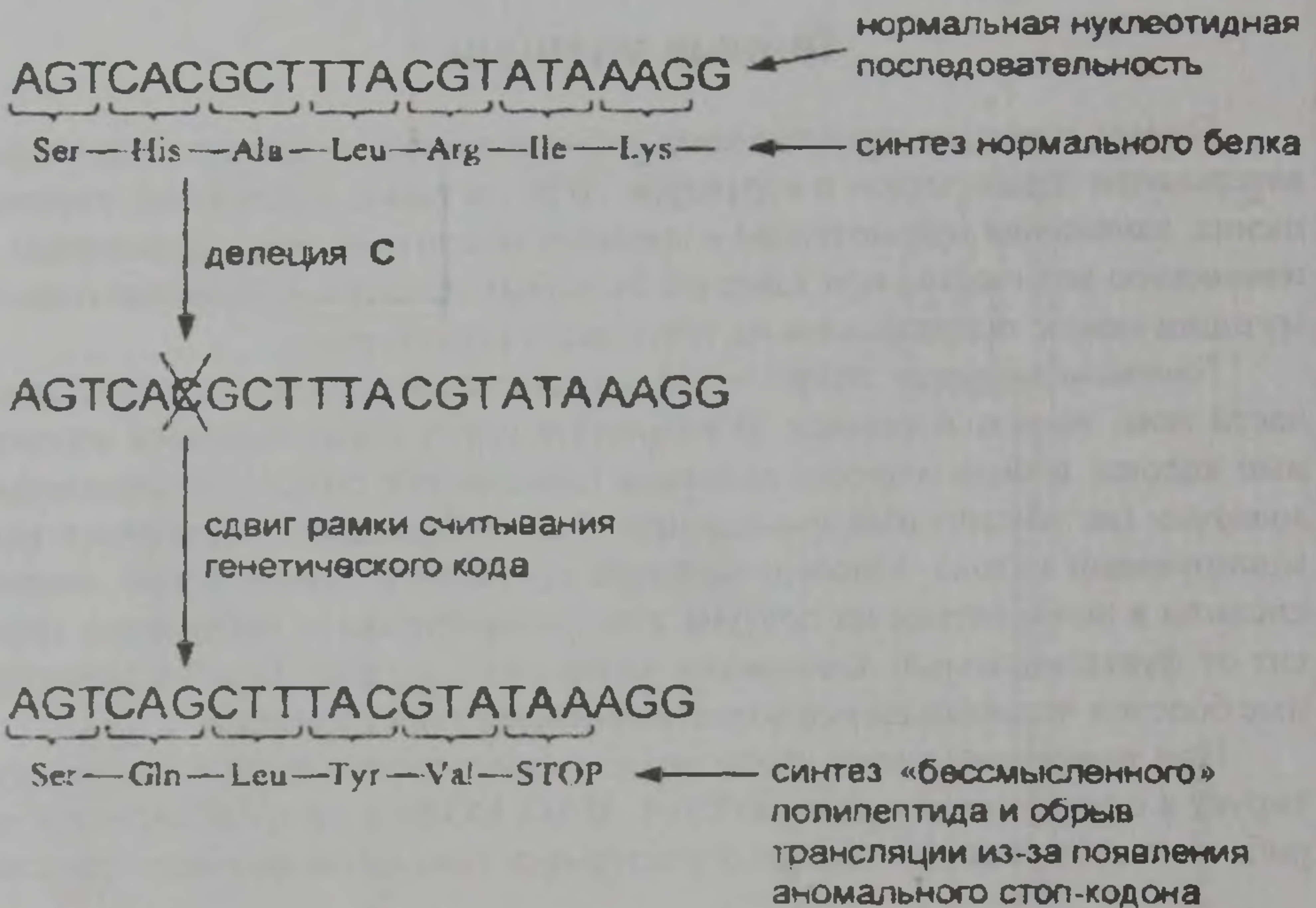


Рис. 336. Мутация со сдвигом рамки считывания генетического кода.

цию и обрывает пептидную цепь, а образующийся белок деградирует. Распространенным классом точковых мутаций являются *делеции* (выпадения) или *инсерции* (вставки) одного—двух нуклеотидов в кодирующей части гена, сопровождающиеся *сдвигом рамки считывания* (рис. 336). Все триплеты ниже сайта делеции или инсерции меняются по ходу считывания, что приводит к синтезу бессмысленного белка, который деградирует под действием внутриклеточных ротеаз.

Мутации в регуляторных 5' — 3' — не транслируемых областях — в промоторной части и в сайте полиаденилирования — вызывают количественные изменения соответствующих продуктов, т. е. содержание белков в клетке.

*Сплайсинговые мутации* на стыках экзонов и интронов нарушают нормальные механизмы созревания (процессинга) первичного РНК-транскрипта и приводят либо к неправильному вырезанию интрона (сохранению его в составе транскрипта), либо к удалению из молекулы РНК информационно значимой экзонной последовательности. Все это приводит к значительному нарушению структуры белка, особенно при сдвиге рамки считывания из-за нарушения сшивки экзон-интронных соединений. Такая мутация в акцепторном участке сплайсинга гена ХРА приводит к полной инактивации белка и к развитию тяжелой формы пигментной ксеродермы у человека.

## Структурные мутации гена

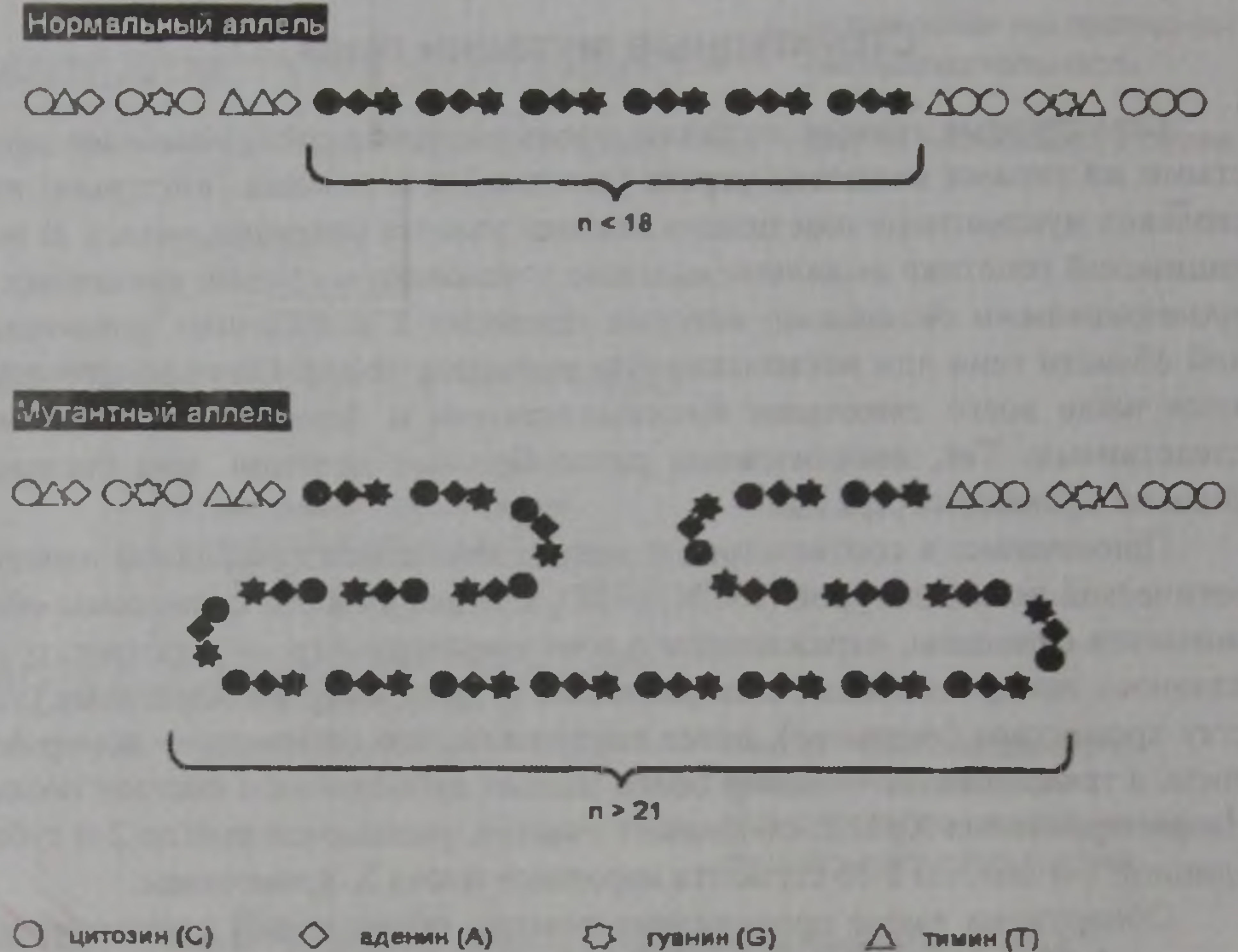
Структурные генные мутации очень разнообразны. Наиболее простыми их типами являются утрата (делеция) или вставка (инсерция) нескольких нуклеотидов или целого генного участка (описано выше). В медицинской генетике выявлено наличие у человека мутаций, связанных с *протяженными делециями*, которые приводят к выпадению значительной области гена или нескольких близлежащих генов. Они характеризуются чаще всего тяжелыми биохимическими и фенотипическими последствиями. Так, локализованы разнообразные делеции гена *дистрофина* на хромосоме Xp21.2.

Примечание: в соответствии с действующей международной цитогенетической номенклатурой (ISCN, 1978), каждый участок хромосомы обозначается символом, отражающим плечо хромосомы (p — короткое, q — длинное), номер сегмента по направлению от центромеры к концевому участку хромосомы (теломере), далее внутри каждого сегмента — номер полосы, а также иногда — номер более мелких субъединиц в составе полос. Например, символ Xp21.2. обозначает участок, расположенный во 2-й субъединице 1-й полосы 2-го сегмента короткого плеча X-хромосомы.

Обнаружена также протяженная делеция обеих копий хромосомного участка 5q13, включающего ген SMN, что приводит к отсутствию у больных нормального транскрипта данного гена и лежит в основе развития болезни.

Другим типом структурных генных мутаций являются *дупликации* части гена, целого гена или определенного хромосомного сегмента, включающего несколько генов, например, полная дупликация участка длиной 1,5 млн п. н. на хромосоме человека 17p11.2. Отмечена склонность митохондриальной ДНК (мт ДНК) ко множественным делециям, что объясняется несовершенством ее системы репарации и высокой повреждаемостью при постоянном контакте с высокими концентрациями активных форм кислорода в митохондриях. Вследствие того, что репликация мт ДНК находится под контролем ядерного генома, то его разнообразные мутации вносят значительный вклад в нарушение тонких механизмов структурной организации кольцевой молекулы мт ДНК.

*Динамические мутации*, или *мутации экспансии* — это совершенно необычный тип мутаций, открытый в 1990-е гг. и зарегистрированный только у человека. Они представляют собой превышение числа повторяющихся элементов, нуклеотидных последовательностей, локализованных в кодирующих или регуляторных областях гена. Например, в определенном участке гена CACNL1A4 нуклеотидная последовательность представлена цепочкой повторяющихся тандемных (следующих друг за другом) триплетов «цитозин — аденин — гуанин», или (CAG)<sub>n</sub> (рис. 34).



**Рис. 34.** Мутация экспансии tandemных тринуклеотидных повторов

Число таких повторов у здоровых индивидов варьирует в строго определенных пределах, а нормальные аллели генов отличаются между собой количеством tandemных тринуклеотидных повторов. Мутация проявляется при увеличении числа копий повторов, превышающих некоторое пороговое значение. На рис. 34 нормальные аллели гена содержат до 18 копий CAG-повторов, тогда как у больных число копий превышает 21 CAG-повторов. Наследование таких мутаций отличается от классического менделеевского типа. Для них характерны:

- различная *пенетрантность* (частота в процентах проявления признака у носителей определенного гена) в сочетании с неполным доминированием;
- *геномный импринтинг*, то есть различие в фенотипическом проявлении в зависимости от того, получена мутация от матери или отца;
- *феномен антиципации* — нарастание тяжести проявления заболевания в последующих поколениях.

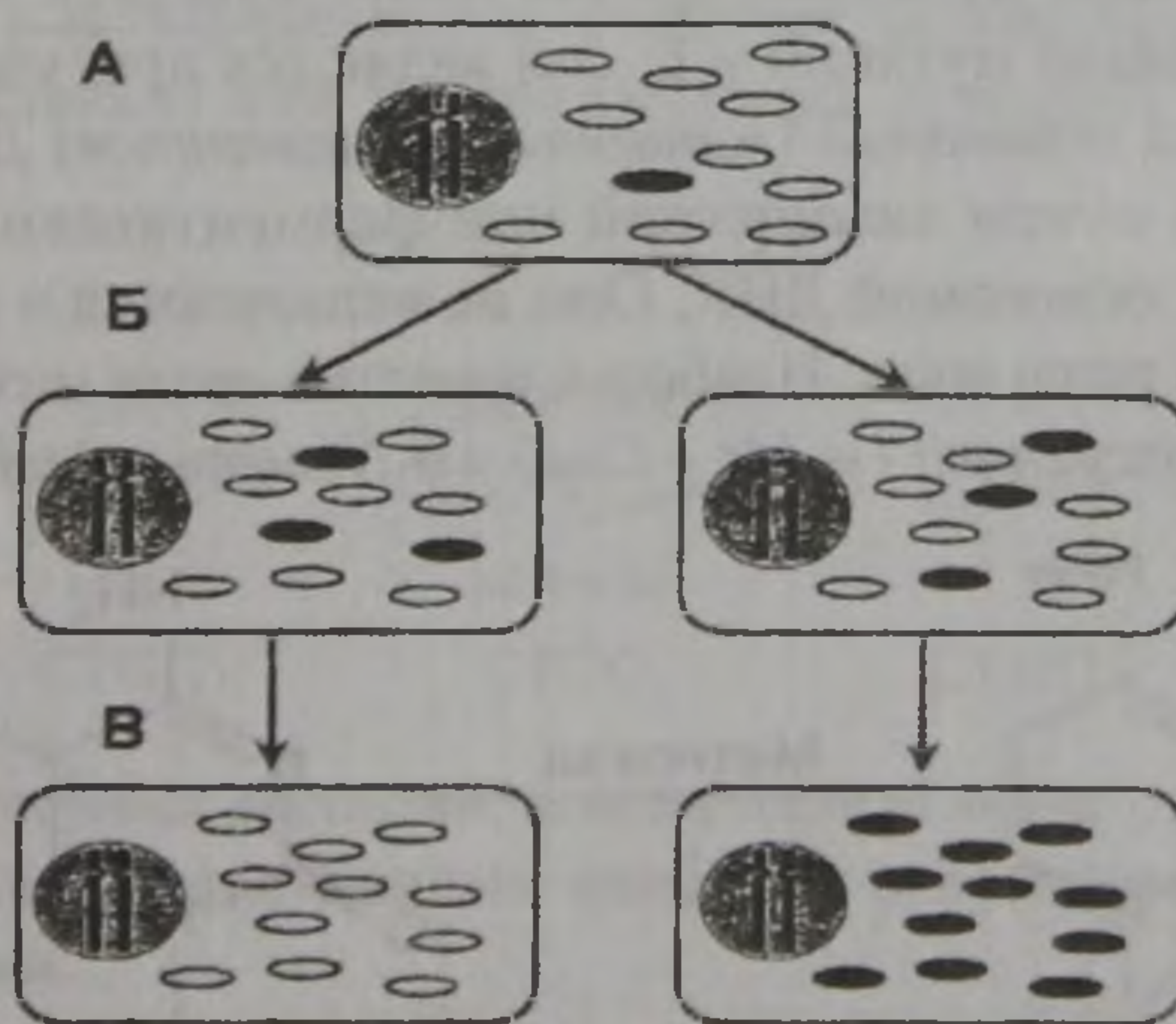
К настоящему времени известно около 20 наследственных заболеваний, обусловленных динамическими мутациями. Этиология этих заболеваний связана с экспансией триплетных повторов  $(CGG)_n$ ,  $(CTG)_n$ ,  $(CAG)_n$ .

Причиной повреждающего действия одних динамических мутаций является утрата функции вследствие блокирования генной экспрессии, других — появление белковых продуктов с аномальной структурной, оказывающих повреждающее действие на жизнеспособность клеток.

Близкой к описанному типу мутаций являются другие мутации экспансии: экспансия *пентануоклеотидных повторов* — тандемные повторы АТТСТ в некодирующей области гена (в норме около 20 копий повторов, в мутантном аллеле — от 800 до 4500). Установлена также экспансия повторяющихся участков ДНК, состоящих из десятков нуклеотидов, например, экспансия 12-нуклеотидного повтора в 5'-области гена CST6 на хромосоме 21q22.

### Цитоплазматические мутации

К этой категории относятся мутации хлоропластной ДНК растений, а также митохондриальной ДНК (мт ДНК) растений, животных и человека. Наиболее изученным является мутагенез мт ДНК человека. В мт ДНК человека возникают в основном точковые мутации, делеции и дупликации, в том числе в сочетании с делециями. Они приводят к изменению белков в комплексе дыхательной цепи митохондрий, что вызывает снижение энергетического обмена клеток. Недостаток синтеза молекулы АТФ приводит к заболеванию тех тканей и органов, в которых происходят энергоемкие процессы. Патогенез этих болезней достаточно изучен благодаря знаниям биохимических процессов, протекающих в митохондриях. В 1958 г. обнаружена крупная делеция в мт ДНК протяженностью от 2 до 10 т. п. н.



**Рис. 35.** Распределение митохондриальной ДНК по дочерним клеткам: А — исходная родительская клетка (мутантная молекула мтДНК обозначена темным цветом); Б — феномен гетероплазмии в дочерних клетках; В — феномен гомоплазмии по нормальной (слева) и мутантной (справа) митохондриальной ДНК





Они появляются в составе ДНК уже после репликации. В составе ДНК прокариот и эукариот наиболее часто встречаются 5-метилцитозин, которой модифицируется под воздействием фермента *метилазы*, добавляющей в определенных участках ДНК к цитозину метильную группу (рис. 36).

Такая модификация затрагивает лишь небольшую часть остатков цитозина. При изучении гена *lac I* в составе *lac*-оперона у *E. coli* все горячие точки спонтанных и индуцированных точечных мутаций были локализованы в сайтах, содержащих 5-метилцитозин. Результатом таких мутаций были транзиции (мутации замены одного пурина на другой или одного пиримидина на другой, что может приводить к замене пары С–G на Т–А (рис. 37).

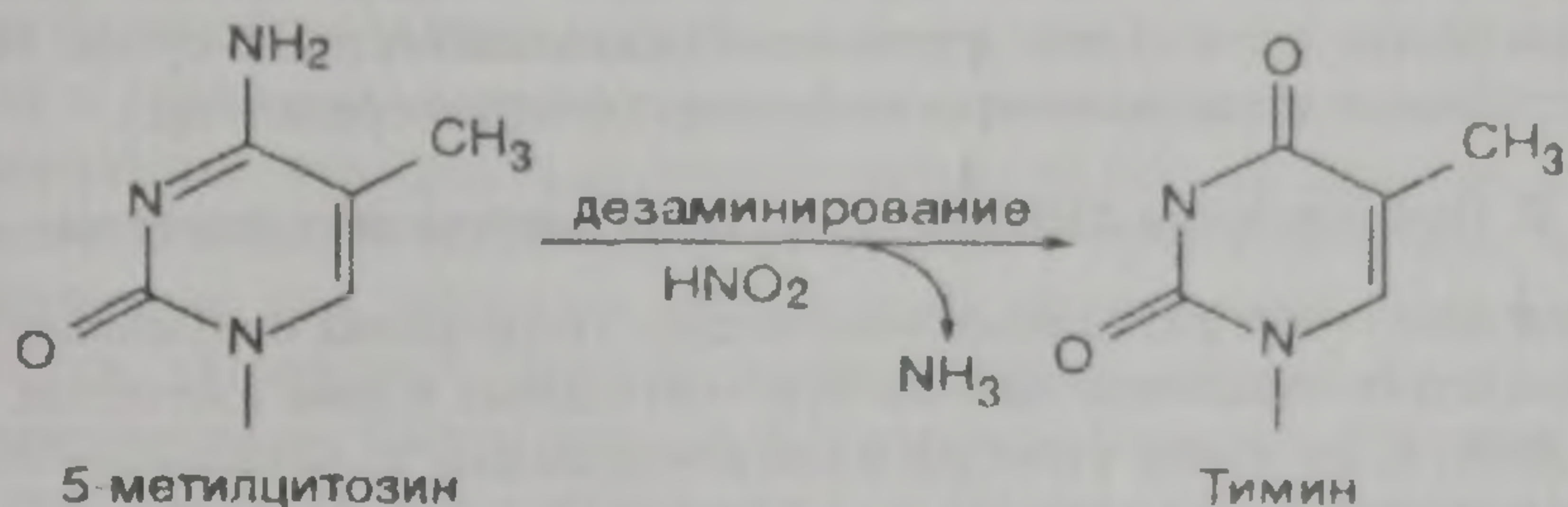


Рис. 37. Деаминирование 5-метилцитозина и мутация транзиции G–C → A–T

К транзиции приводит спонтанное или индуцированное деаминирование 5–Me–Cyt, образование тимина (Thy) и ошибочное спаривание G с T. При последующей репликации может появиться замена G–C → A–T.

Еще одной горячей точкой является трижды повторяющаяся последовательность CTGG. В определенных условиях может происходить добавление (мутация вставки) или потеря (мутация делеции) четырех буквенных единиц.

Мутант	CTGG	CTGG		
	↑	делеция		
Дикий тип	CTGG	CTGG	CTGG	
	↓	вставка		
Мутант	CTGG	CTGG	CTGG	CTGG

Появление точечных мутаций замены одной пары оснований на другую пару обусловлено следующими наиболее известными и хорошо изученными факторами.

### 1. Не исправленные ДНК-полимеразой ошибки репликации

Корректирующая 3'→5'-экзонуклеазная активность ДНК-полимеразы обеспечивает высокую точность репликации. Причиной ошибок в процессе

репликации является способность всех азотистых оснований образовывать термодинамически не выгодные таутомерные формы за счет миграции атома водорода, а они, в свою очередь, могут образовывать неправильные пары с другими основаниями. Так, таутомерная *имино-форма цитозина* образует пару не с гуанином, а с аденином, а в последующей репликации происходит замена А–Т на G–C. Сходным образом аденин может формировать имино-форму, способную к комплементарному спариванию с цитозином.

## 2. Феномен апуринизации ДНК

Вследствие термальных разрывов N-гликозидных связей между пуриновым основанием и дезоксирибозой ДНК каждой клетки организма человека теряет около 5 тыс. пуриновых оснований А и G в сутки. На этот процесс влияют также химические факторы (кислая среда и др.).

## 3. Повреждение ДНК под действием химических факторов

Под действием различных химических соединений окружающей среды происходит изменение весьма чувствительных к ним азотистых оснований ДНК. К их числу относятся многочисленные *ксенобиотики* — яды, лекарства, канцерогены, пестициды, инсектициды, гербициды. Наиболее полно изучено воздействие азотистой кислоты  $\text{HNO}_2$ , гидроксилamina  $\text{NH}_2\text{OH}$ , а также алкилирующих агентов (диметилсульфата, диметилнитрозамина, азотистого иприта). Наиболее часто наблюдаются *транзиции* G–C → A–T и обратные транзиции A–T → G–C.

## 4. Ошибки ДНК-полимеразы и включение аналогов природных нуклеотидов

Ферменты ДНК-полимеразы обладают высокой точностью функционирования при катализе репликации. Тем не менее, они могут не различить нормальные нуклеотиды от других нуклеотидов с той же структурой. Одним из таких примеров является природный нуклеотид с азотистым основанием — 2-аминопурин. Его встраивание в ДНК вместо аденина (А) и последующее спаривание с цитозином (С) способствует транзиции A–T → G–C.

Другой структурный аналог природных нуклеотидов — 5-бром-дезоксиуридинтрифосфат (5 – Br – dUTP), являющийся аналогом тимидина, а также галопроизводное урацила 5 – Br – Ura в *енольной и кето-форме*.

Часто образуемая енольная форма спаривается с Gua и приводит к транзиции A–T → G–C. Кето-форма является аналогом тимина и спаривается с Ade.

## 5. Интеркаляция акридинов и этидий бромида

Это явление сводится к способности ряда органических соединений встраиваться (интеркалировать) между парами оснований ДНК. Это при-

водит к появлению вставок или делеций одной или более пар оснований. Если они не кратны трем парам, то происходит сдвиг рамки считывания. К таким соединениям относятся акридиновые красители и этидий бромид.

## Теория мутаций

Теория мутации начала интенсивно развиваться после открытия законов Менделя и экспериментального обоснования Морганом и его школой закономерностей сцепления генов и их рекомбинации в процессе кроссинговера, иными словами после установления наследственной дискретности хромосом. Мутации, возникающие под влиянием обычных природных факторов внешней среды или в процессе биохимических и физических изменений в организме, получили название «спонтанных мутаций». Первые обнаруженные мутации были спонтанными, то есть не вызванными какими-либо заметными внешними причинами. С другой стороны, мутации, возникающие под влиянием специальных воздействий (ионизирующей радиации, химических веществ, экстремальных условий и т. д.), названы *индуцированными*. Индуцированный мутагенез впервые наблюдали у дрожжей под действием радиоактивности Г. А. Надсон и Г. С. Филипов в 1925 г. В 1927 г. Г. Меллер сообщил о действии рентгеновых лучей на мутационный процесс у дрозофилы и разработал количественный метод учета рецессивных летальных мутаций в X-хромосоме у плодовой мушки. В 30-х гг. открыт химический мутагенез у дрозофилы В. В. Сахаровым (1932), а также М. Е. Лобашевым и Ф. А. Смирновым (1934).

Развитие количественных аспектов классической радиационной генетики связано с именами многих ученых. В 1936 г. Н. В. Тимофеев-Ресовский, К. Циммер и М. Дельбрюк обосновали радиобиологический *принцип попадания* и сформулировали так называемую *теорию мишени*. Она объясняла наблюдаемый мутагенный эффект элементарными попаданиями квантов излучения в чувствительные генетические мишени — хромосомы (Циммер, 1962; Тимофеев-Ресовский, Иванов, Корогодин, 1968.).

## Репарация повреждений ДНК

Под репарацией *понимается процесс исправления ошибок в последовательности нуклеотидов ДНК и восстановления ее исходной структуры.*

Механизм исправления ошибок включает большое число ферментативных реакций. Уникальность молекулы ДНК, наряду с ее другими свойствами, проявляются и в том, что она является единственной среди всех макромолекул, которая не деградирует, а исправляет возникающие повреждения. Репарация обеспечивает чрезвычайно высокую стабильность ДНК, сохранение ее первичной структуры и генетической информации, долговременное существование видов и их выживание. В природных условиях *всегда*

присутствуют и действуют различные факторы, воздействующие на структуру любых биополимеров. Важнейшими из них является тепловые соударения молекул ДНК с беспорядочно движущимися молекулами. Столь же постоянно влияние и других факторов на структуру ДНК-ультрафиолетового излучения солнечных лучей, многочисленных химических соединений.

Репарация обеспечивает поддержание целостности клеточного генома. В результате из 1000 случайных изменений ДНК в мутацию может реализоваться статистически менее одного случая. Расчеты показали, что белок со средней молекулярной массой, кодируемый геном размером 1 тыс. п. н., изменяется мутационно лишь один раз в ходе  $10^6$  клеточных делений. Установлено, что в ДНК генома млекопитающего размером  $3 \times 10^9$  п. н. за один год происходит всего лишь 15 замен нуклеотидов, затрагивающих зародышевый путь. При такой частоте замен средняя полипептидная цепь из 400 аминокислотных звеньев заменяет одно звено а среднем один раз за 200 тыс. лет (Алексеев, Каминский, 2004).

Различают несколько видов репарации:

- 1) *Фотореактивация* — восстановление исходной структуры молекулы ДНК, поврежденных УФ-излучением, в результате последующего воздействия видимого света;
- 2) *эксцизионная репарация* — удаление неправильно спаренных или поврежденных оснований из ДНК и синтез новых последовательностей, взамен удаленных поврежденных;
- 3) *рекомбинационная (пострепликативная) репликация* — происходит с участием рекомбинации;
- 4) *SOS-репарация* — медленная пострепликативная репликация с участием системы ферментов, которую индуцирует облучение.

### Фотореактивация

Под влиянием УФ-лучей в одной цепи ДНК образуются димеры двух соседних пиримидинов (тиминовые димеры) циклобутанового типа. Каждый из димеров задерживает репликацию на 10 сек. Процесс регенерации оснований-мономеров активируется под действием видимого света с длиной волны 300–600 нм и осуществляется под действием специфических фотореактивирующих ферментов *дезоксирибодипиримидин-фотолиаза*. Этот фермент является продуктом гена *phr*. Фотолиаза узнает димер и расщепляет ковалентные связи между двумя пиримидинами.

### Эксцизионная репарация

Этот вид репарации не является строго специфичным в отношении повреждений ДНК, как фотореактивация, так как с его помощью могут

исправляться не только пиримидиновые димеры, но также и многие другие повреждения структуры ДНК. Механизм эксцизионной репарации состоит из нескольких стадий.

1. Измененный участок, несущий повреждение в цепи ДНК, узнается и удаляется с помощью специфических ДНК-эндоуклеаз, которые осуществляют гидролиз фосфоэфирных связей с 5'-конца от повреждения, а 5'→3'-экзонуклеаза затем удаляет поврежденный участок.
2. Ресинтез ДНК — заполнение образовавшейся брешы, при этом одноцепочечный участок ДНК используется в качестве матрицы.
3. ДНК-лигаза ковалентно сшивает 3'-конец нового материала с 5'-концом старого материала.

## Рекомбинационная (пострепликативная) репарация

Этот тип репарации основан на процессах рекомбинации и репликации поврежденной ДНК и является наименее специфичным из всех рассмотренных типов репарации, так как в них отсутствует этап узнавания повреждения. У *E. coli* эти обмены осуществляются с помощью *продуктов гес-генов*, а в заключительном этапе ресинтеза и сшивания участвуют ДНК-полимераза I и лигаза.

## SOS-репарация

Многие воздействия, повреждающие ДНК или ингибирующие ее репликацию у *E. coli*, вызывают целую серию фенотипических изменений — «SOS-ответ». Например, тепловой шок активирует целый набор генов, продукты которых могут защитить клетки. Индуцируется синтез особых «шоковых» белков, часть которых помогает репарировать другие белки, частично денатурированные тепловым шоком.

Непосредственным стимулом к запуску механизмов SOS-репарации является накопление одноцепочечных разрывов ДНК, которые индуцируют протеазную активность белка Rec A, специфически взаимодействующего с белком Lex A — репрессором для генов *гес* (B, C, E, F, J). В результате работы репарационных систем происходит уменьшение количества одноцепочечных разрывов ДНК, снижается инициирующий SOS-репарацию сигнал и механизмы SOS-репарации выключаются.

## Ключевые слова и термины

Мутации — генные, хромосомные, геномные, сплайсинговые, динамические (мутации экспансии); транзиции; репарации повреждений ДНК, фотореактивация.

## Молекулярно-филогенетическая систематика

Одним из важнейших результатов биологической эволюции явилось наблюдаемое разнообразие живых организмов, представляющее определенную иерархию филогенетических групп и их свойств. Исторически такая иерархия со времен Э. Геккеля (1866) явилась плодом филогенетического мышления и филогенетической концепции исторического развития биоты и реконструкцией на основе комплексного анализа морфологических структур, сравнительного исследования закономерностей филогенеза в разных группах организмов. В основе *филогенетики* лежит изучение процесса появления новых и исчезновения существующих групп организмов и их свойств. С ней связана *систематика* — раздел биологии, изучающий свойства и закономерности таксономического разнообразия как одного из аспектов биологического разнообразия путем отражения этих свойств в таксономической системе. В системе таких представлений филогенетике отводится роль *одной из ключевых биологических наук*, а ее идеалом является *построение всеобщего «Древа жизни»* (Павлинов, 2005).

Филогенетическому мышлению противостоит *популяционный подход (мышление)*, базисом которой служит *модель минимальной эволюции (микроразволюции)*. Она сводит эволюционные изменения к изменениям генных частот в популяциях и методологически ориентирована на *фенетическое мышление*, которое редуцирует организм до совокупности равноценных «элементарных признаков» и структуры биополимеров.

При рассмотрении высших уровней иерархического деления живых организмов каждый исторический этап изучения находил выражение в новой системе. Так, в прошлом было принято традиционное для того времени (до 30-х гг. прошлого столетия) деление организмов на растения и животных, при этом прокариоты тогда относили к растениям. Стениер и ван Ниль в своем сообщении «The concept of a bacteria» (1962) обосновали предложенное Шаттоном в 1937 г. противопоставление прокариотной и эукариотной клетки. Изучение клетки позволило разграничить прокариот и

эукариот. Такая классификация на основе фундаментальности прокариот по их клеточному строению не сразу была принята специалистами.

До конца 70-х гг. XX в. вирусы рассматривали в качестве самостоятельного таксона, противопоставляемого другим группам организмов (Воронцов, 1987, 2005). Мир клеточных организмов делили на две части:

- организмы с оформленным клеточным ядром, окруженным двойной ядерной оболочкой и имеющим хромосомы, в которых ДНК связана с белками гистонами; общее название этой группы — *эукариоты* (Eukaryota);
- другую группу организмов, лишенную отмеченных эукариотических свойств составляли бактерии, синезеленые водоросли (цианофиты), лучистые грибки, миксобактерии с общим названием — *прокариоты* (Prokaryota).

Бурное развитие молекулярно-генетических и геномных реконструкций при изучении нуклеиновых кислот и белков внесло огромные изменения в классификационную практику. В настоящий момент анализ последовательностей рибосомных РНК у многих тысяч видов привел к выделению особой группы *архебактерий* в рамках *трехдоменной* (от англ. domain — *домен, доминион*) *системы организмов*. Архебактерии в этой классификации оцениваются в качестве третьей формы жизни наряду с *эубактериями* и *эукариотами*. Монолитность прокариот как таксономической группы по молекулярным признакам оказалась под вопросом (Woese et al., 1990; Woese, 2000). Открытие архебактерий стало крупнейшим достижением систематики последней четверти прошедшего столетия. В своих объемных работах крупный систематик Кавалье-Смит (Cavalier-Smith, 2001; 2002a, в) сделал попытку наметить диапазон и границы новой мегасистематики, которая распределяет все организмы по трем доменам.

В результате на сегодняшний день существует несколько альтернативных классификаций, отвечающих самым высшим уровням иерархического деления (Шаталкин, 2004а, в; 2005). По этой причине сообщество систематиков разделилось на сторонников новой классификации и тех, кто считает ее неприемлемой по разным причинам.

Задачей нашего исследования является молекулярно-генетический и геномный анализ результатов изучения живых организмов разных систематических групп. Поэтому мы считаем необходимым представить краткое рассмотрение альтернативных классификаций:

- во-первых, классическую мегасистематику высших таксонов — систему из двух империй и пяти царств (Воронцов, 1987, 2005);
- во-вторых, молекулярно-генетическое подразделение мира организмов на высшие таксоны — *домены архебактерий, эубактерий и эукариот* (Шаталкин, 2004в, 2005; Cavalier-Smith, 2004).

Это связано с тем, что в числе других читателей данная работа может оказаться полезной аспирантам и студентам вузов, которые используют



учебную литературу, в которой на сегодняшний день отражена только классическая систематика. С другой стороны знакомство с современной методологией молекулярных реконструкций и их применение в мегасистематике углубит их понимание проблем филогенетики и эволюции.

### Классическая мегасистематика высших таксонов: система из двух империй и пяти царств

В 1965 г. Н. Н. Воронцов предложил распространить принципы разделения живых организмов на мегатаксоны и ввел понятие империи (*imperium*) в ранге высшей таксономической единицы (Воронцов, 1965, 1987, 2005). В этой системе вирусы, рассматривавшиеся ранее в ранге самостоятельного царства, составили империю доклеточных (Noncellulata). Все

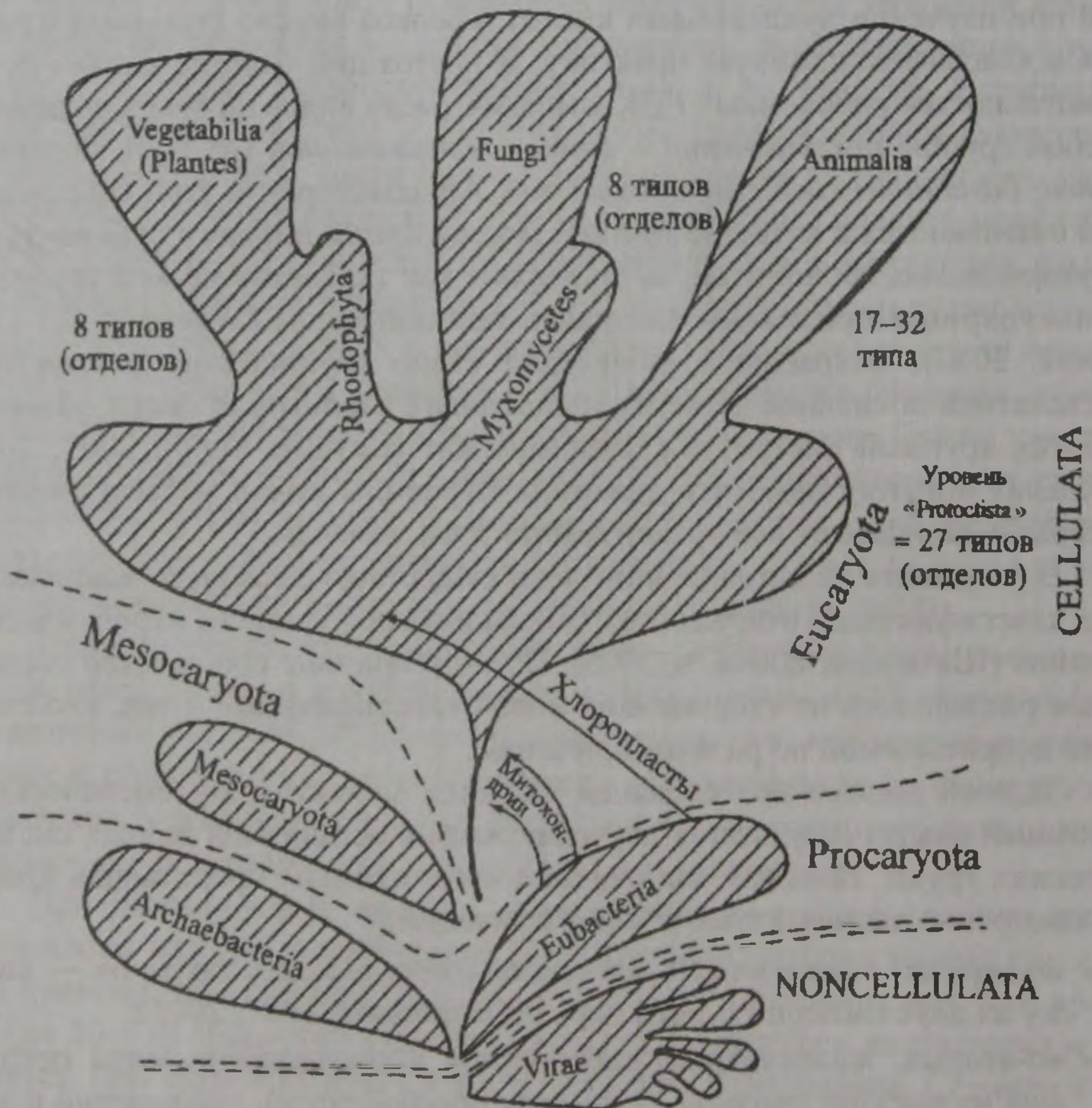


Рис. 38. Схема возможных взаимоотношений между основными группами организмов (Воронцов, 2005)

остальные организмы, отнесенные к империи *клеточных* (Cellulata), были подразделены на *подимперии прокариот* (Prokaryota) и *эукариот* (Eukaryota) (рис. 38). Основные признаки, разделяющие прокариот и эукариот, обобщены в табл. 2 (Шаталкин, 2004а).

Таблица 2

Основные признаки, разделяющие прокариот и эукариот (Шаталкин, 2004а)

Признаки	Прокариоты	Эукариоты
Размер клетки	1 мкм	10–100 мкм
Цитокинез (сократительное кольцо)	Из белков семейства Fts	Из актиновых белков
Плазматическая мембрана		
Наличие стероидов	Известны примеры	Имеются
Ядро		
Ядерная мембрана	Pianctomycetes (гомология?)	Имеется
Ядрышко	Отсутствует	»
ДНК		
Объем (пар оснований)	$2-6 \times 10^6$	От $2-4 \times 10^7$ до $5 \times 10^9$
Строение ДНК	В виде единственной кольцевой молекулы	В виде линейных структур, обычно связанных с гистонами
Компактизация ДНК	Через множественное «перекручивание»	Через разноуровневую спирализацию при участии гистонов и конденсинов
Гистоны	Известны примеры	Есть
Начальная область репликации	Единственная	Многочисленные
Ядерные ДНК полимеразы	Три типа	Четыре типа
Разделение ДНК	Через сопряженный с ростом клеточной стенки процесс	С помощью митотического микротрубочкового аппарата
РНК		
рРНК (число)	Три	Четыре (5.8S уникальна для эукариот)
мРНК (интроны)	Нет	Имеются
мРНК (кэпирование и полиаденилирование)	»	Имеется

Окончание Таблицы 2

Признаки	Прокариоты	Эукариоты
мРНК (последовательность Шайна—Дальгарно)	Есть	Нет
мРНК (пилицистройная)	»	Нет (бывает у предшественника мРНК)
Структура генов		
Опероны	Есть	Теперь найдены
Эукариотическая структура генов	Известны примеры	Есть
Половой процесс	Парасексуальные отношения	Мейоз
Цитоплазма		
Рибосомы (размер)	70S	80S
Органеллы в качестве производных мембраны	Известен пример ( <i>Ignicoccus</i> )	Имеются (эндосомы)
Аппарат Гольджи	Нет	Есть
Митохондрии	»	»
Пластиды	»	»
Цитоскелет		
Актиновые волокна	Близкие к актину белки, определяющие в некоторых группах форму клеток	Актиновый скелет, определяющий форму клеток, типы движения, фагоцитоз и прочее
Микротрубочки	Имеются гомологи тубулина с неизвестной функцией	Имеются
Промежуточные волокна	Отсутствуют	»
Способность вырабатывать целлюлозу	Целлюлозный матрикс вне клеточной стенки ( <i>Acetobaaer</i> )	Целлюлозная стенка (водоросли и высшие растения)

### Империя доклеточные — Noncellulata

Тело не имеет клеточного строения. Имеется лишь один тип нуклеиновых кислот (двунитчатая или однитчатая ДНК; двунитчатая или однитчатая РНК) — либо ДНК, либо РНК. Нет рибосом. Размножение возможно только в живых клетках (облигатные клеточные паразиты). Нет обмена веществ. Используется энергия, получаемая за счет обмена веществ в клетках хозяина.

Один из белковых генов бактериофага Т4 имеет экзон-интронную структуру. Среди ДНК-содержащих вирусов имеются группы с двунитчатой и однонитчатой ДНК кольцевой или линейной формы. РНК-содержащие вирусы обычно содержат однонитчатую несегментированную линейную РНК, но у реовирусов (Reoviridae) двунитчатая РНК представлена 10–12 фрагментами. Вирусы разделены на 20–25 семейств, однако истинный ранг этих таксонов, по-видимому, существенно выше.

## Империя клеточные — Cellulata

Тело имеет клеточное строение. В клетке присутствуют оба типа нуклеиновых кислот ДНК и РНК. Есть рибосомы. Способны к самостоятельному размножению. Самостоятельно синтезируют ферменты. Имеется обмен веществ. Как правило, свободноживущие формы, если паразиты, то межклеточные, реже внутриклеточные.

## Подимперия преядерные или прокариоты — Prokaryota

Ядерный материал (ДНК) концентрируется в центральной части клетки (*нуклеоиде*), но неотделен от периферической части клетки. Нет ядерной оболочки (*кариотеки*). Единичная молекула ДНК замкнута в кольцо. Хромосомных белков нет, настоящих хромосом нет. Нет аппарата митоза, мейоза, ядрышковых организаторов. Репликация ДНК начинается с одной точки. В ДНК нет повторяющихся последовательностей. Нет внеядерной ДНК. Рибосомные ДНК трех типов: 16S, 23S и 5S (первые два типа не встречаются у эукариот). РНК-полимераза одного типа. Рибосомы типа 70S (единиц Сведберга). В состав прокариот входят: 1) грамположительные бактерии; 2) грамотрицательные бактерии; 3) синезеленые бактерии (*цианобиты*). Положение микоплазм в системе не ясно.

## Подимперия эукариот — Eukaryota

Ядро полностью обособлено от цитоплазмы. Есть ядерная оболочка (*кариотека*). Ядерная ДНК не замкнута в кольцо, расположена линейно. ДНК вместе с хромосомными белками образует нуклеопротеидный комплекс, среди хромосомных белков преобладают гистоны. Имеется митотический аппарат. Есть митоз и мейоз. Репликация ДНК во времени приурочена к концу интерфазы, начинается репликация с многих точек. Имеются повторяющиеся последовательности ДНК. Есть внеядерная ДНК типичного для прокариот строения. Рибосомные РНК четырех типов: 18S, 5,8S, 28S сцеплены транскрипционно и не встречаются у прокариот, а 5S рРНК транскрипционно не сцеплена. РНК-полимераза трех типов (в митохондриях — одна). Рибосомы типа 80S (но в пластидах типа 70S).

### **Царство зеленых растений — *Vegetabilia (Plantae)***

Клетки всегда имеют наружную клеточную оболочку из целлюлозы или пектина. Фотосинтезирующие автотрофы. Запасающие вещества — крахмал, иные полисахариды, жирные масла, но не гликоген. Относятся три подцарства (Тахтаджан, 1976): 1) настоящие водоросли (*Phycobionta* с восемью отделами); 2) багрянки или красные водоросли (*Rhodobionta*); 3) высшие растения (*Embriobionta* с восемью отделами).

### **Царство грибов — *Fungi, Micetalia, Mycota***

Клетки с наружной клеточной оболочкой из хитина. Гетеротрофы с пассивным адсорбционным типом питания. Запасающие вещества-гликоген, а также жирные масла. Два подцарства: 1) миксомицеты (*Myxobionta*) с отделом слизистые грибы; 2) настоящие грибы (*Mycobionta*, 1–2 отдела).

### **Царство животных — *Animalia***

Клетки не имеют наружной оболочки. Это, а также запасание энергии в виде растворимого гликогена, создает предпосылки для развития подвижности. Гетеротрофы с активным типом питания.

Подцарство простейших (*Protozoa*) делится не менее чем на пять типов. Подцарство многоклеточных животных (*Metazoa*) подразделяется на:

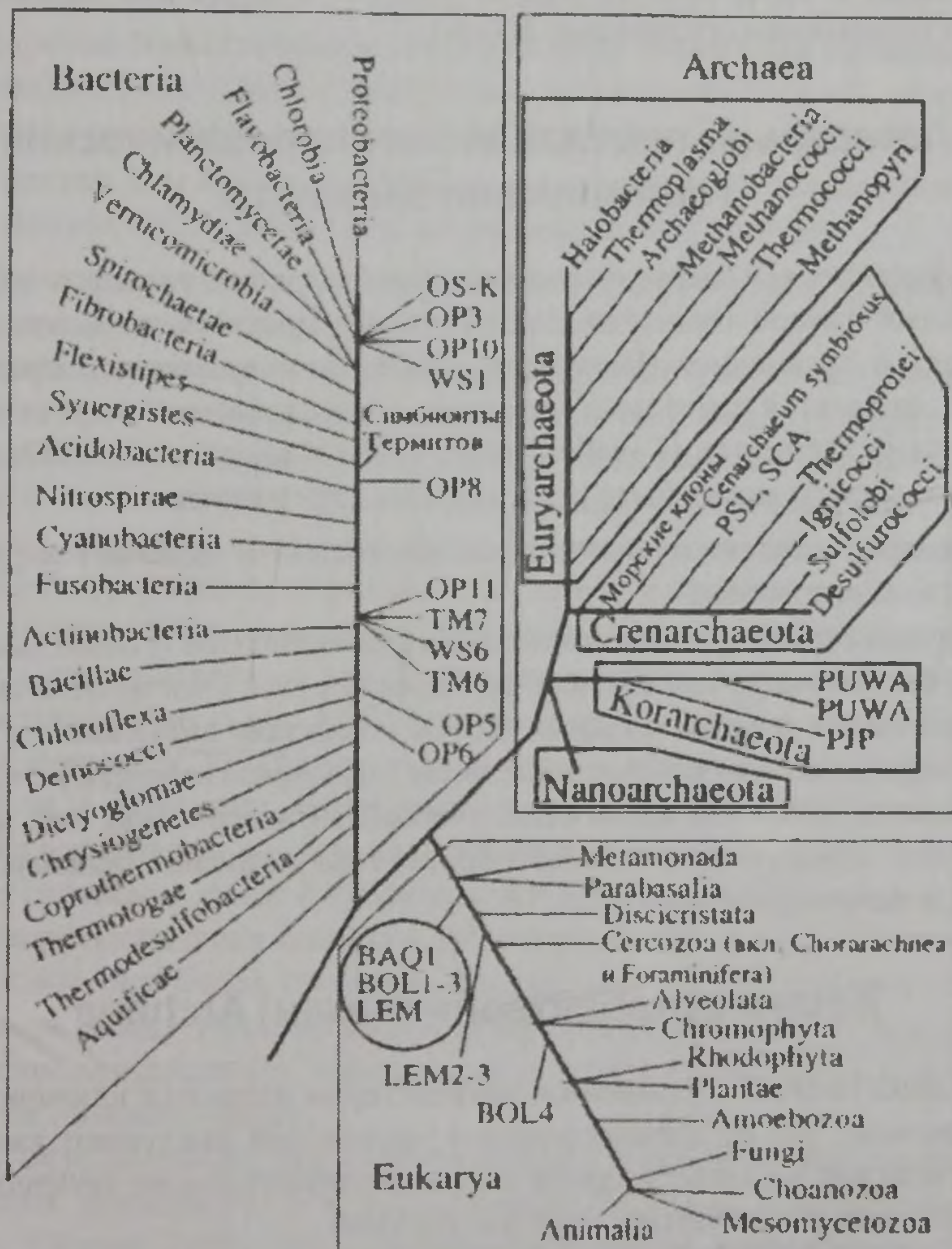
- 1) инфрацарство паразоа (*Parazoa* с единственным типом *Porifera* — губки);
- 2) инфрацарство фагоцителловых (*Phagocitellozoa* с единственным типом *Placozoa*);
- 3) инфрацарство настоящих многоклеточных животных (*Eumetazoa*).

### **Современная мегасистематика на основе молекулярных реконструкций**

В основе разработки современной мегасистематики живых организмов лежат три открытия в 70-х гг. прошлого столетия:

- 1) доказательство Томасом Броком (Thomas Brock) существования жизни в экстремальных условиях выше 70°C на примере термофильных бактерий *Thermus aquaticus* и других видов и выделение термоустойчивого фермента — ДНК-зависимой ДНК-полимеразы (Madigan, Marrs, 1997);
- 2) создание методов прямого определения нуклеотидных последовательностей ДНК и начало эры секвенирования и компьютерной геномики;
- 3) нахождение Карлом Вёзе (Воез, Woese) технически удобной модели анализа последовательностей нуклеотидов в генах рРНК малой субъединицы рибосом как наиболее консервативной органеллы клетки.

В конце 1977 г. появилась работа по рибосомной классификации метаногенных бактерий (Fox et al., 1977), а также предложение по новой таксономической структуре организмов с выделением *трех доменов (надцарств)* архебактерий, эубактерий и эукариот (Woese, Fox, 1977). В последующем изучении архебактерий установлено, что они распространены более широко, чем изначально предполагали. По оценкам специалистов, суммарная биомасса архебактерий составляет  $10^{14}$  тонн, что превышает биомассу всех ранее известных форм жизни на Земле (Madigan, Marts, 1997).



**Рис. 39а.** Универсальное дерево жизни (по: Hugenholtz et al., 1998, Gupta, 2000; Madigan et al., 2000; Dawson, Pace, 2002; Baldauf, 2003, с дополнениями). Буквами обозначены группы некультивируемых бактерий, а также эукариоты, известные лишь по нуклеиновым пробам.

Эубактерии, архебактерии и эукариоты рассматриваются в качестве трех отдельных линий развития (рис. 39а). На этом была построена концептуальная схема филогенетического древа жизни (Woese et al., 1990), состоящего из трех основных доменов, берущих начало от единого гипотетического «корневого» предка.

Далее мы рассмотрим и кратко охарактеризуем все высшие уровни иерархического деления организмов. Три самостоятельные группы развития предложено называть археями (Archaea), бактериями или эубактериями (Bacteria) и эукариями (Eukarya) вместо эукариот для созвучности с другими названиями (Шаталкин, 2004в).

## Латерально передаваемые паразитические генетические элементы

В классической мегасистематике вирусы рассматривались как самостоятельный таксон (империя доклеточных), противопоставляемый другим группам организмов (Воронцов, 1987). По современным представлениям их относят к категории латерально передаваемых паразитических генетических элементов (Cavalier-Smith, 1998; Madigan et al., 2000). Как целое, эти элементы могут быть разделены на две группы:

- одна представлена нуклеиновыми кислотами и включает вирусы, вирионы и плазмиды;
- в другой группе функцию генетических элементов (генов) несут вместо нуклеиновых кислот особые белки *прионы*. Они не отличаются от нормальных функциональных белков по составу аминокислот, но имеют *вырожденную конформационную (пространственную) структуру* (Prusiner, 1998). По своей генетической характеристике они же определяют *конформационное преобразование* незрелых белковых молекул в новые прионы.

## Домен архебактерии (археи) Archaea

В новой системе организмов архебактерии являются ключевой группой. Показано, что на цитологическом уровне они выступают как прокариоты, тогда как на молекулярном они напоминают скорее эукариот. В домене архей выделяют четыре царства или типа.

1. Царство *Crenarchaeota* включает экстремальных термофилов (серные анаэробные архебактерии), встречаются в горячих (55–105°C) кислых (pH 1–3) геотермальных источниках, таких, как, например, гейзеры на Камчатке. Некоторые виды выдерживают рекорд по температуре выносливости — *Picrolobus fumarii* не живет при температуре ниже

97°C, а зарегистрированное верхнее пороговое значение у этого вида составило 113°C.

2. Царство *Euryarchaeota*: метаногенные и галофильные археи, которых известно 25 родов и 50 видов. Строгие анаэробы. Встречаются в кишечнике ряда животных, и иногда человека. Два вида играют роль в разложении растительного материала в рубце коров, где они производят 200–400 л метана в день. Эмиссия метана из прямого кишечника происходит у 37 % мужчин и 63 % женщин, что может свидетельствовать о присутствии метаногенных архей.
3. Царство *Nanoarchaeota*: единственный известный представитель — симбионт или паразит гипертермофильного вида архей, обитающих в гидротермальных системах США, Исландии и на Камчатке.
4. Царство *Korarchaeota*: ДНК видов обнаружены в геотермальных источниках США, Исландии, на рисовых полях Японии.

Молекулярно-генетические признаки архей, а также зубактерий и эукариот представлены в табл. 3.

Генетический код несет кодоны для 20 пептидогенных (белкообразующих) аминокислот. Кроме того, белки включают более ста других аминокислот в ходе посттрансляционных изменений, замещая отдельные атомы в основной аминокислоте. Исключение составляет *селеноцистеин*, отличающийся от *цистеина* наличием атома *селена* вместо *серы*. Кодировать эту аминокислоту у архей, некоторых бактерий, эукариот, включая человека, может триплет UGA, играющий роль стоп-кодона для стандартных аминокислот. Селеноцистеин — это 21-я генетически кодируемая аминокислота.

Недавно (Hao et al., 2002; Srinivasaan et al., 2002) у метанобактерий в составе фермента, катализирующего распад метиламинов с образованием метана, и у некоторых зубактерий открыта 22-я нестандартная аминокислота *пирролизин*, кодируемая стоп-кодоном UAG.

Из сравнения архей с зубактериями и эукариотами вытекает, что хотя археи морфологически близки к зубактериям, все же на макромолекулярном уровне они более сродни с эукариотами. Они занимают срединное положение между зубактериями и эукариотами. Их древность и срединное положение предполагает, что они могут стоять ближе всего к общему для всех трех доменов предку.

К. Вёзе допускал, что существенной стороной жизни древних одноклеточных организмов был *горизонтальный обмен генетическим материалом* (Woese, 2000, 2002). Устойчивые линии вертикально связанных форм на этих этапах развития жизни не прослеживались. Не было и предкового таксона, а лишь конгломерат эволюционно исходных форм, постоянно обменивающихся генами и проявляющих максимально возможные комбинации признаков.



Таблица 3

Основные признаки, разделяющие эубактерий, архебактерий и эукариот  
(Шаталкин, 2004в)

Признаки	Эубактерии	Архебактерии	Эукариоты
Клеточные покровы			
Пептидогликан	Имеется	Отсутствует	Отсутствует
Структура липидов	Сложный эфир глицерина с 2 жирными кислотами	Простой эфир глицерина с 2 молекулами фитанила	Сложный эфир глицерина с 2 жирными кислотами
Хиральность глицерина	<i>d</i> -изомер	<i>l</i> -изомер	<i>d</i> -изомер
Фосфатидилинозит	Актинобактерии	Есть	Есть
Чувствительность к пенициллину	Есть	Нет	Нет
Нуклеосомы	Нет	Есть	Есть
ДНК полимераза	Тип С	Тип В	Тип В
Чувствительность к афидиколину	Нет	Есть	Есть
Топоизомеразы			
Торо VI и гомологи	»	»	Гомолог ТороLV
Транскрипция			
число субъединиц РНК полимеразы	4	6–12	15
Промотор	Бокс Прибнова	TATA-бокс	TATA-бокс
Транскрипционные факторы (TFIIВ, ТВР)	Нет	Имеются	Имеются
Активаторы и репрессоры	Прокариотические	Прокариотически	Эукариотические
Интроны генов тРНК	Нет	Эукариотический тип	Эукариотический тип
Трансляция			
Структура рибосомных белков	Эубактериальная	Эукариотическая	Эукариотическая
тРНК с:	N-формилметионином	Метионином	Метмонином

Окончание Таблицы 3

Признаки	Эубактерии	Архебактерии	Эукариоты
Структура IF-4A	Асимметричная	Симметричная	Симметричная
Гомологи фактора элонгации II	EF-G	EF-2	EF-2
Высвобождающий фактор RF1	Эубактериальный	Эукариотический	Эукариотический
Фактор RRF	Есть	Нет	Нет
Чувствительность			
к дифтерийному токсину	Нет	Есть	Есть
хлорамфениколу	Есть	Нет	Нет
кирромицину	»	»	»
стрептомицину	»	»	»
Протеасомы	Нет (актинобактерии?)	Есть	Есть
Убиквитин	»	Термоплазма	»
Экзосомы	»	Экзосомные гомологи	»
Шапероны			
Низкомолекулярные	Эубактериальные	Эукариотические	Эукариотические
		гомологи	
Шаперонины II и префольдин	Нет	Есть	Есть
Шаперон Hsp90	HtpG (гомолог)	Нет	»
Тип АТФазы	Эубактериальный	Ближе к эукариотическому типу	Эукариотический
Транслокация белков			
Tat-система	Имеется	Имеется	Отсутствует
SRP19	Отсутствует		Имеется
Рецепторы для SRP	FisY	FtsY	SR $\alpha$ и SR $\beta$
Sec A-система	Имеется	Отсутствует	Отсутствует
Котрансляционное N-гликозилирование	Нет	Есть	Есть

## Домен эубактерии (бактерии) Bacteria

В последнее десятилетие резко возросло число групп высокого ранга среди эубактерий. На сегодняшний день речь идет уже о 30 типах (филумах) культивируемых форм (рис. 39а). При учете некультивируемых форм, выделенных на основе структуры ДНК, число типов будет около 50 (Madigan et al., 2000).

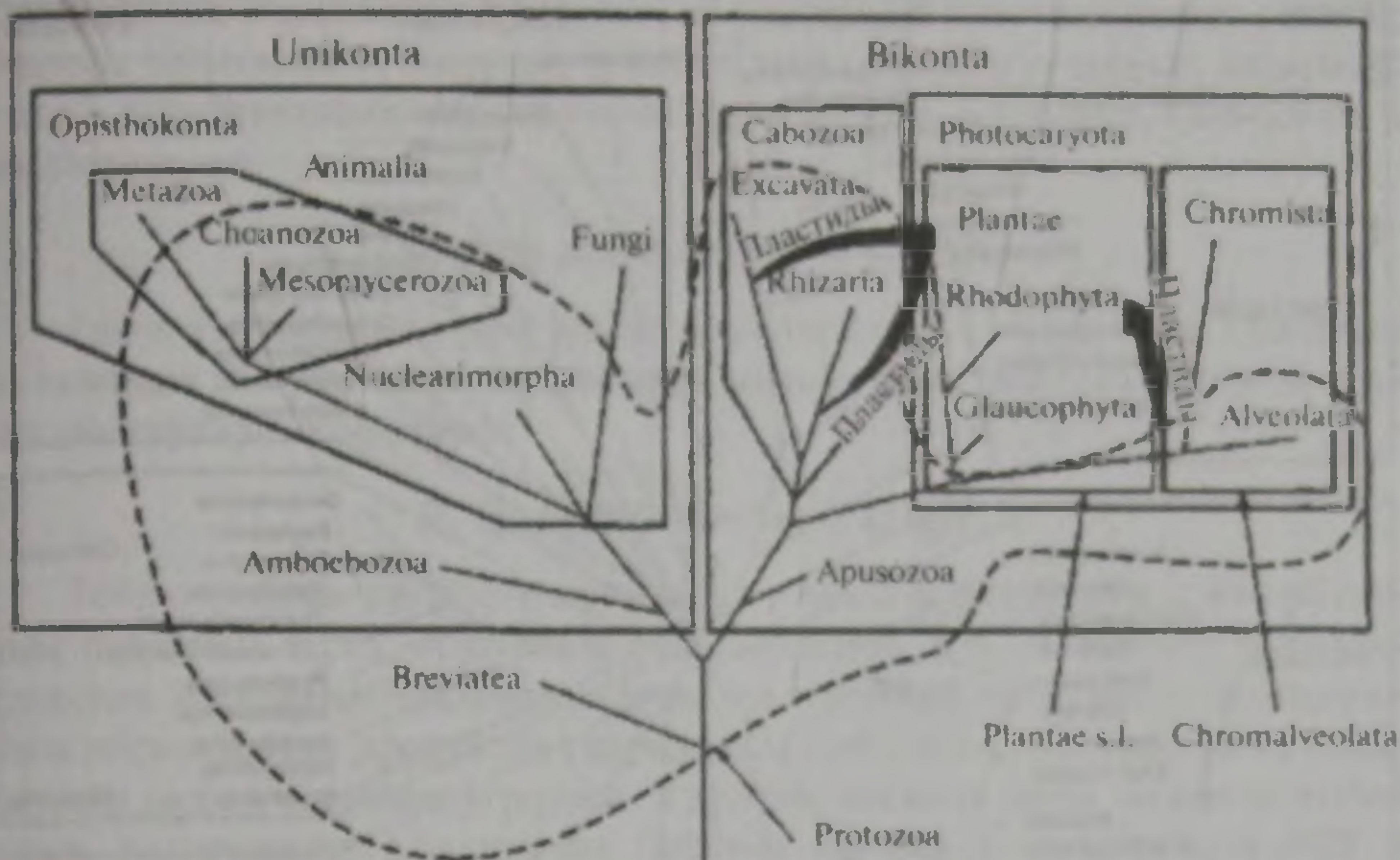
При сравнительном изучении нуклеиновых кислот из природных выборок в бактериальных сообществах выявлены многочисленные формы на видовом уровне, признаки которых показывали на возможность разделения классов, отделов и групп более высокого ранга. В ближайшем будущем возможны новые открытия, способные значительно изменить наши представления о строении эубактерий. Так, в небольшой группе бактерий класса Planctomycetes отмечены небактериальные признаки: отсутствие пептидогликана, наличие одной или двух мембран, окружающих нуклеоид, при этом гены 16S и 23S рРНК разделены.

## Филогенетическая система домена эукариоты Eukaryota

Для устранения некоторых недостатков рибосомных деревьев и кладограмм возникла необходимость поиска медленно меняющихся молекул или их отличий, отличающихся своей уникальностью и редким появлением. Этим требованиям удовлетворяют выпадения (*делеции*), вставки (*инсерции*), участков молекулы белка, слияние и разделение генов и их *дупликации*. Открытие этих и многих новых молекулярных маркеров дало возможность создать новую филогению эукариот (рис. 39б).

Эволюция эукариот шла по двум главным направлениям. Одно было связано с появлением анцестральных двужгутиковых организмов (биконты — *Bikonta*). Сестринской по отношению к биконтам группой являются заднежгутиковые (опистоконты — *Opisthokonta*) и некоторые амебоидные формы (*Amoebozoa*), которые исходно имели один жгутик (*Unikonta*). В настоящее время хорошо обоснован монофилетический статус всех основных групп живых организмов и установлены их связи (Stechmann, Cavalier-Smith, 2002).

При построении дерева эукариот на основе молекулярных реконструкций исследователи сходятся на модели, которая состоит из немногих *гипотетических групп высшего ранга (supergroups)*, каждая из которых включает в свой состав большое число первичных одноклеточных эукариот (протистов и водорослей). В процессе создания этого дерева учитывали многочисленные данные: деревья на основе отдельных генов, мультигенный анализ, другие молекулярные и структурные признаки (Шаталкин, 2004а, б; 2005; Keeling et al., 2005; Dagan, Martin, 2006; Cavalier-Smith, 2006; Parfrei et al., 2006).



**Рис 396.** Кладограмма связей и классификация организмов  
(по: Cavalier-Smith, 2002, 2004 с изменениями)

Далее мы представим краткое описание дерева эукариот из пяти *надцарств (супергрупп)* (рис. 40): экскаваты, ризарии, униканты, хромальвеолы и растения в широком смысле.

## I. Надцарство экскаваты (Excavata)

Экскаваты представляют собой многочисленную группу протистов, многие из которых являются анаэробами и паразитами, наиболее известны среди которых *Trichomonas*, *Giardia* и *Trichomonas*. Ядро группы составляют жгутиконосцы. Группа включает 6 таксонов. Они слабо объединены комбинацией молекулярных и морфологических данных.

## II. Надцарство ризарии (Rhizaria)

Название этого таксона предложил Кавалье-Смит (Cavalier-Smith, 2004) для монофилетического объединения церкозоев, фораминифер и радиолярий. Ядро группы составляют амёбы и амёбофлагелляты. Группа выделена недавно на основании молекулярных данных — структуре РНК-полимеразы II, актина и полиубиквитина. Последний имеет у представителей группы уникальную вставку между убиквитиновыми единицами а и в. Близкие результаты показывают РНК-деревья.

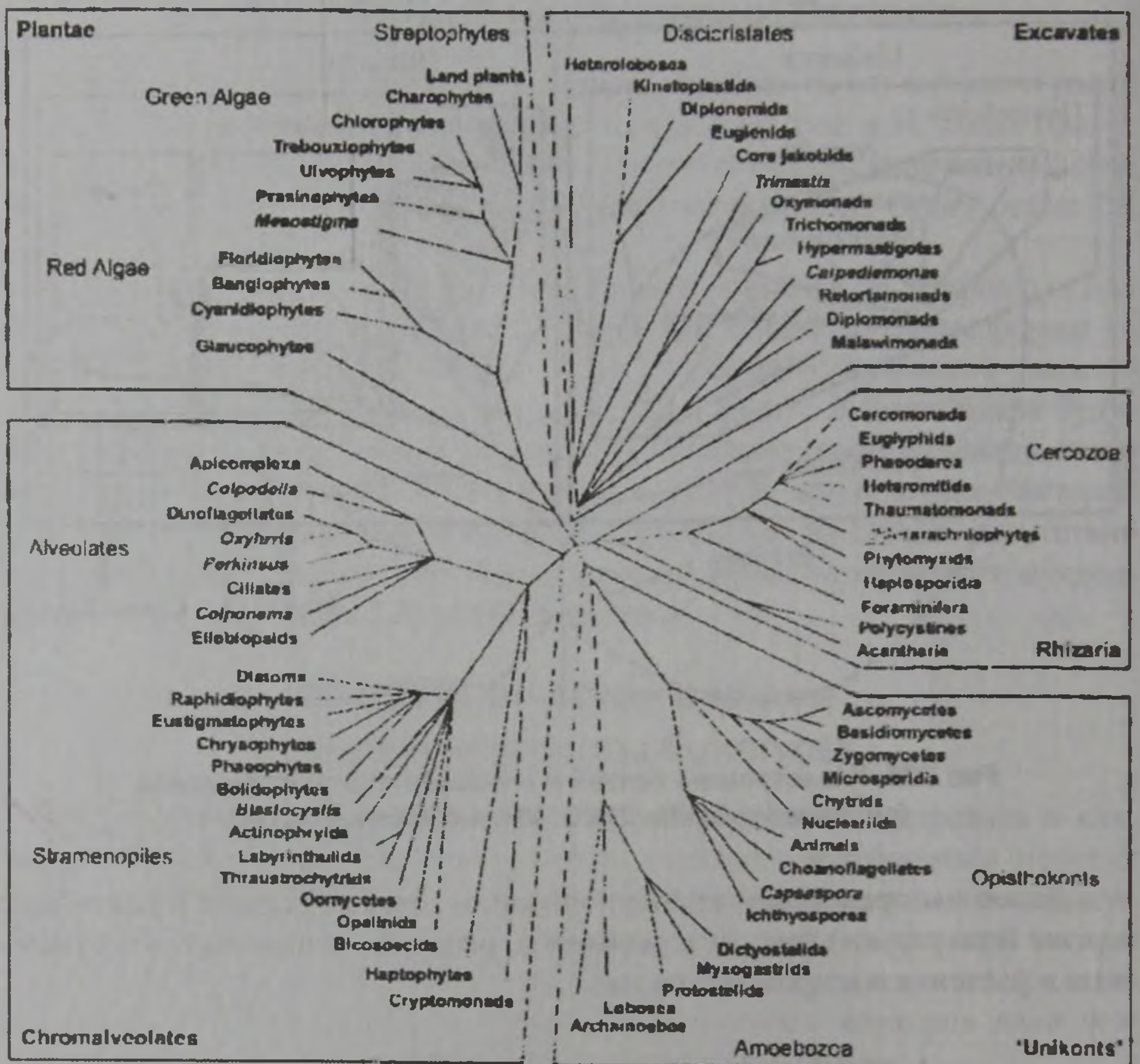


Рис. 40. Дерево эукариот. Гипотетическое дерево составлено на основе данных молекулярной филогении, других молекулярных, морфологических и биохимических признаков. Пять групп высшего ранга (supergroups), каждая из которых включает в свой состав большое число первичных одноклеточных эукариот (протистов и водорослей) (по: Keeling et al., 2005)

### III. Надцарство одножгутиковые (Unikonta)

«Одножгутиковые» является спорным названием для объединения самостоятельных хорошо обоснованных групп: амебозои и заднежгутиковые. В общем и целом, заднежгутиковые включают животных в широком смысле, грибов и некоторых амёб (например Entamoeba), миксомицетов (Dictyostelium), некоторых паразитических протистов и хоанофлагеллят. Эта группа выделена как филогенетическое целое на основе наличия цепочек из четырех 123 и 129 генов, а также внутренней дубликации одного домена фосфофруктокиназы (Streichmann, Cavalier-Smith, 2003). Для линии

Opisthokonta характерно присутствие инсерций при удлинении гена 1a фактора элонгации и энолазы. Многие виды заднежгутиковых используются в исследованиях как биологические модели, например, *Drosophila* и *Saccharomyces*.

### 1. Царство амебозои (*Amoebozoa*)

Ранее амебоидные простейшие распределялись по 6 типам. Эта группа основана на сравнении генов актина, малой субчастицы рРНК и ряда других молекулярных маркеров.

### 2. Царство грибы (*Fungi*)

Гетеротрофные спорообразующие паразиты и сапрофиты с абсорбтивным питанием через «наружное пищеварение», т. е. выделение гидролитических ферментов (экзоферментов) на пищевой субстрат с последующим поглощением продуктов гидролиза. Грибы могут использовать практически любой источник углерода. У грибов найдено лишь по одной копии генов, кодирующих аквапорины (водные каналы в мембранах) AQP и GlpF, тогда как у растений (резуховидка *Arabidopsis*) — 35 генов. Грибные AQP-белки близки к PIP-белкам растений, а кодирующие их гены, возможно, были приобретены независимо от бактерий (Zardoya et al., 2002).

### 3. Царство животные в широком смысле (*Animalia*)

Царство включает следующие таксоны:

- Mesomycetozoa (также Ichthyosporaea): паразиты рыб, головастиков, членистоногих и моллюсков. Возбудитель риноспоридиоза вызывает у человека поражение слизистых носа и глаз;
- Хоанофлагелляты — небольшая группа одиночных или колониальных форм. Они гетеротрофы, однако, есть указание на наличие фототрофных видов;
- Животные в узком смысле (*Animalia*). Из большого числа групп, ранее включавшихся в состав Protozoa, на сегодняшний день филогенетически связанными с животными, осталось два таксона — хоанофлагелляты и мезомицетозои. Многие типы мембранных и связанных с мембраной белков, участвующих в процессах раздражимости животных, отмечены уже у хоанофлагеллят: протокадерины и 7 TM-кадерины, лектины, C-типы тирозинкиназы, рецепторы, сопряженные с G-белками (Шаталкин, 2005).

### IV. Надцарство хромальвеоляты (*Chromalveolata*)

Хромальвеоляты — это таксон, включающий объекты альгологии, микологии и зоологии. Они состоят из двух сестринских групп Chromista

и Alveolata, имеющих вторичные пластиды красноводорослевого происхождения и ряд других уникальных признаков на молекулярном уровне (Cavalier-Smith, 1999). Ключевой фермент гликолиза фруктозобисфосфатальдолаза (FBA) делится на два класса, не связанных структурно и филогенетически. Гомотетрамер FBA<sub>T</sub> имеется в пластидах у большинства эукариот и красных водорослей, тогда как у производных от красных водорослей групп (гетероконтов, криптомонад, динофлагеллят) в пластидах работает гомодимер FBA<sub>II</sub> (Patron et al., 2004).

На основе ряда наблюдений на молекулярном уровне сделано предположение, что хромальвеоляты возникли как фотосинтезирующие виды, но в ходе эволюции способность к фотосинтезу у отдельных видов терялась.

### 1. Царство хромисты (*Chromista*)

Ядро представителей этого царства составляют *гетероконтные грибы*: оомицеты, лабиринтулиды, траутохитридиевые, а также водоросли с хлорофиллами а и с. *Гетероконтность* — это важная характеристика парных жгутиков по длине, направленности и опушению. Обычно один жгутик направлен вперед и по функции является моторным, второй жгутик в этом случае выполняет функцию руля. Гетероконтными являются различные водоросли — бурые, диатомовые, золотистые, золотисто-зеленые, хлоромонадовые. Их общее название — *хромофитовые водоросли*. К хромистам принадлежат также ряд гетеротрофных одноклеточных — бикосоециды, бластоцистиды (кишечные анаэробные паразиты и комменсалы овальной формы), опалиниды (крупные многоядерные паразиты заднего отдела кишечника рыб, амфибий и рептилий). Для гетероконтных водорослей и связанных с ними гетеротрофов предложено название *страменопили* (*Stramenopila* — англ. straw-haired, соломенноволосяные). Самое базальное положение в этой группе занимают опалиниды. К страменопилиям близки гаптофиты и криптомонады.

### 2. Царство альвеоляты (*Alveolata*)

К этому надцарству относятся одноклеточные животные (ранее входили в Protozoa) — инфузории, динофлагелляты, споровики (таксон Apicomplexa). Кавалье-Смит (Cavalier-Smith, Chao, 2004) выдвинул внутри динофлагеллят подтип Protalveolata с примитивными родами Colponema, Ellobiopsis, Colpodella и Perkinsus.

Определяющими особенностями альвеолярных простейших являются: наличие *альвеолярной системы*, из мембранных пузырьков в субпелликулярном слое клеточной стенки; другая особенность — наличие *микропор (пиноцитозных пор)*, через которые осуществляется особый тип пиноцитоза. В последнее время найдены молекулярные маркеры единства

альвеолят. Одним из них явилась кодируемая геном *Cox2* субъединица митохондриальной *цитохром-с-оксидазы*. У апикокомплексных простейших, зеленых водорослей и бобовых растений соответствующая единица кодируется двумя ядерными генами, образованными в результате разделения гена *Cox2*. У всех других эукариот этот ген один и является митохондриальным.

Около половины динофлагеллят являются фотосинтетиками с пластидами, производными от красноводорослевого симбионта, содержат хлорофилл *a* и *c*, а также дополнительный пигмент *перидинин*, способный поглощать свет в зелено-голубой области (470–550 нм).

Паразитические споровики делятся по молекулярным данным на две группы — грегарины и криптоспоридий, а также группа кокцидий и кровяных споровиков (Kuvardina et al., 2002).

## V. Надцарство растения в широком смысле (*Plantae — sensu Cavalier-Smith, 1981*)

Растения представляют собой монофилетическое объединение из трех групп: глаукофит (глаукоцистофит), красных и зеленых водорослей, а также наземных растений. На основе молекулярных данных была обоснована также монофилия красных и зеленых водорослей.

Пластиды этих трех групп таксона произошли от *цианобактериального предшественника*. Группа цианобактерий сформировалась на основе форм с кольцевым *периферическим светособирающим комплексом (ПСК)*. В их дальнейшей эволюции появились *фикобилисомы* с основным ПСК. У зеленых водорослей и наземных растений ПСК представлен кольцевыми структурами. Растения принципиально отличаются от других эукариотических фототрофов (= автотрофов) *первичными пластидами*, имеющими две цианобактериальные мембраны. Эвглени и хромальвеоляты — это формы с *вторичными пластидами*.

Важная особенность растений состоит в том, что они не теряли пластиды даже при переходе на иной тип питания. По признаку *облигатности пластид* растения резко отличаются от всех групп автотрофов (хромальвеолят и эвгленовых), вторичные пластиды которых не образуют жесткого объединения с материнской клеткой и легко исчезает при смене питания. Так, Кавалье-Смит (Cavalier-Smith, 2002) отметил 18 случаев потери пластид у хромальвеолят.

### 1. Царство глаукоцистофиты, или глаукофиты (*Glaucocystophyta, Glaucophyta*)

Глаукофиты представлены небольшой группой одноклеточных организмов, содержащих 2–4 *псевдопластиды (цианеллы)*, производные от циа-



нобактерий. Геном цианелл на порядок величин меньше геномов свободноживущих цианобактерий и составляет 136 кб (килобаз). Глаукофиты встречаются в пресноводных холодных (1–9°C) с низким рН (4,8–5,2) источниках. Известны восемь родов.

Ряд авторов (Шаталкин, 2005) допускают независимое приобретение пластид глаукофитами и предком зеленых и красных водорослей. Они единственные среди эукариот содержат в пластидах цианобактериальную форму *фруктозобисфосфатальдозы (FBA II B)*. В то же время у красных и зеленых водорослей, а также у наземных растений имеется FBA I, причем и в цитозоле, и в пластидах.

## 2. Царство красные водоросли (*Rhodophyta*)

Представители царства — в основном многоклеточные организмы нитчатого, псевдопаренхиматозного и паренхиматозного строения. Известны также одноклеточные формы. Сильная обособленность этой группы обусловлена следующими уникальными особенностями:

- полное отсутствие жгутиковой стадии;
- наличие лишь хлорофилла а (иногда хлорофилл d). Кроме того, присутствуют дополнительные пигменты — красный (*фикоэритрин* — придает красный цвет водорослям) и голубой (*фикоцианин*);
- клеточная стенка состоит из микрофибрилл, а в качестве запасного вещества вырабатывается близкий к гликогену *флоридейский крахмал*.

## 3. Растения в узком смысле (*Viridiplantae*, или *Plantae*)

С наземными растениями непосредственно связана группа зеленых водорослей (*Chlorophyta*), рассматриваемая в качестве исходной для растений (Дьяков, 2000). Между этими группами отмечены многочисленные сходства по многим признакам:

1. Близкий химический состав клеточных стенок на основе полисахаридов и целлюлозы.
2. Наличие в структуре хлоропластов хлорофиллов а и в, а также комплекса каротиноидов и ксантофилов.
3. Накопление крахмала в хлоропластах.
4. Особенности синтеза фермента Рубиско (рибулозо-1,5 бифосфат карбоксилаза/оксигеназа). Этот ключевой фермент фотосинтетического углеродного цикла (цикла Кребса) участвует в образовании фосфоглицериновой кислоты. Он состоит из 8 малых и 8 больших белковых субъединиц. У зеленых водорослей и наземных растений боольшая субъединица кодируется геном *rbcl* ДНК хлоропласта, тогда как малая субъединица — ядерной ДНК. У всех остальных фотосинтезирующих эукариот этот фермент кодируется генами хлоропласта.

Растения выработали много уникальных особенностей, связанных с их способом жизни внутри клеточной стенки: это, прежде всего, наличие плазмодесм и нового способа клеточного деления.

## Скорость молекулярной эволюции.

### Концепция молекулярных эволюционных часов

Развитие методов молекулярной систематики в последние десятилетия привело к революционным изменениям наших представлений о системе животных и их эволюции. Широкое распространение молекулярных методов для описания биологического разнообразия потребовало объединения опыта исследователей, использующих классические методы с бурно развивающимися методами описания биоты по молекулярным признакам.

Скорость эволюции видов теоретически можно определить по темпам их изменений в ходе эволюции. Оставалось неясным, как методически измерить этот темп. Фенотипический уровень изменений регистрируется при изучении ископаемых остатков. По данным Дж. Г. Симпсона (Simpson, 1953), высота одного определенного зуба лошади увеличивалась в ходе ее эволюции в среднем на  $4 \times 10^{-8}$  год, или на 4 % за 1 млн лет. Была предложена *единица скорости эволюции* — *дарвин* как скорость изменения количественного признака, равную  $10^{-6}$  год: один дарвин соответствует увеличению или уменьшению величины признака в  $e$  раз ( $e$  — основание натурального логарифма, равное 2,718) за 1 млн лет. Скорость эволюции лошади в этих единицах составляет 40 миллидарвинов. Длина тела динозавров на протяжении мезозойской эры увеличивалась со скоростью  $2,6 \times 10^{-8}$  или 26 миллидарвинов. Когда эти скорости превышали 1 дарвин, это указывало на фактор вмешательства человека, например, очень быстрые изменения домашних животных и культурных растений в процессе селекции.

Идеальным показателем темпов эволюции Симпсон считал некую «генетическую скорость», хотя было неизвестно, в каких единицах ее измерять. Развитие молекулярной генетики и ее методов давало ключ для оценки скорости включения новых мутантных генов в геномы эволюционирующих видов, и проводить сравнительный анализ аминокислотных последовательностей белков у отдаленных видов. Удобной моделью при анализе скоростей эволюции на молекулярном уровне являются глобины позвоночных — гемоглобин и миоглобин, необходимые для жизни представителей этой группы. Функция гемоглобина — перенос кислорода от легких к тканям и углекислого газа от тканей к легким. У млекопитающих  $\alpha$ -цепи гемоглобина состоят из 141 аминокислотного остатка, а  $\alpha$ -цепи разных млекопитающих *гомологичны*.  $\beta$ -цепи состоят из 146 аминокислотных остатков. У японского макака и человека  $\alpha$ -цепи совершенно идентичны за исключением четырех аминокислотных сайтов, то есть различие

в аминокислотных последовательностях этих цепей равна 2,8 %. С учетом того, что каждый сайт белковой последовательности может быть замещен любой из 20 аминокислот, вероятность случайного совпадения двух последовательностей для сегмента, например, из 10 аминокислотных сайтов будет очень мала и составит  $(1/20)^{10} \approx 10^{-13}$ . Допускается, что гены двух  $\alpha$ -цепей этих двух видов произошли от общего предкового гена в результате примерно пяти замен азотистых оснований. У человека и собаки  $\alpha$ -цепи гемоглобинов различаются по 23 аминокислотным сайтам, то есть на 16,3 % от общего числа сайтов.

При сравнении  $\alpha$ -цепей гемоглобина семи различных видов позвоночных (человека, акулы, карпа, тритона, курицы и др.) было установлено, что скорость эволюционных замен аминокислот на сайт в год составила около  $0,9 \times 10^{-9}$ . Скорость эволюции  $\beta$ -цепей гемоглобина у семи видов приматов также составила примерно  $10^{-9}$  на аминокислотный сайт в год. На основе этих данных был сделан вывод, что дивергенция 6 видов млекопитающих по  $\alpha$ - и  $\beta$ -цепям гемоглобина произошла 80 млн лет назад. Близкую к этим оценкам картину показало сравнительное изучение миоглобинов (состоят из 153 аминокислот) человека, обыкновенного тюленя, барсука, лошади, быка и кашалота, то есть скорость замен на сайт в год составила  $10^{-9}$ . В табл. 4 приведены скорости эволюции белков млекопитающих, выраженные в единицах, равных  $10^{-9}$  аминокислотных замен на сайт в год.

Самая низкая эволюция отмечается у гистона H4 ( $K_{aa} = 0,01 \times 10^{-9}$ ), а самая высокая — у фибринтпептидов ( $K_{aa} = 8,3 \times 10^{-9}$ ). В качестве единицы измерения скорость эволюции белков Мотоо Кимура (Кимура, 1985) предложил «полинг» как величину изменений  $10^{-9}$  на аминокислотный сайт в год. Скорость эволюции гемоглобина близка к 1 полингу, а цитохрома С — 30 сантиполингов ( $K_{aa} = 0,3 \times 10^{-9}$  в год).

Таблица 4

Скорость эволюции белков млекопитающих, выраженные в единицах, равных  $10^{-9}$  аминокислотных замен на сайт в год

Белок	$K_{aa} \cdot 10^{-9}$ в год
Фибринопептиды	8,3
Панкреатическая рибонуклеаза	2,1
Лизоцим	2,0
$\alpha$ -цепь гемоглобина	1,2
Миоглобин	0,89
Инсулин	0,44
Цитохром с	0,3
Гистон H4	0,01

Таблица 5

Число аминокислотных различий, выявленных при сравнении  $\alpha$ - и  $\beta$ -цепей разных гемоглобинов

Число аминокислотных различий *	Число аминокислотных различий		
	$\alpha$ -цепь человека — $\beta$ -цепь человека	$\alpha$ -цепь карпа — $\beta$ -цепь человека	$\alpha$ -цепь утконоса — $\beta$ -цепь утконоса
0	62	61	62
1	55	49	49
2	21	29	28
3	0	0	0
Пробел	9	10	9
Сумма	147	149	148

\* Аминокислотные различия разных типов (обозначенных 0, 1, 2, 3), выявленных в трех сериях сравнений  $\alpha$ - и  $\beta$ -цепей, обусловлены (табл. генетического года) как минимум 0, 1, 2 или 3 нуклеотидными заменами, соответственно. Для каждого случая приводится также число пробелов, выраженное через эквивалентное число аминокислот.

Для изучения вопроса о примерном постоянстве или непостоянстве скоростей молекулярной эволюции при сравнении более далеких таксонов исследовали дивергенцию  $\alpha$ - и  $\beta$ -цепей гемоглобина. Эти две полипептидные цепи явились результатом дупликации в далеком прошлом. В табл. 5 приведен анализ дивергенции  $\alpha$ - и  $\beta$ -цепей гемоглобина человека, утконоса, а также  $\alpha$ -цепи гемоглобина карпа и  $\beta$ -цепи гемоглобина человека.

Для перехода от одной аминокислоты к другой необходимо учитывать минимальное число замен для 20 (горизонтально)  $\times$  20 (вертикально) комбинаций аминокислот. При этом пробелы — это часть нуклеотидных замещений, не приводящих к информативно-существенным итогам в силу вырожденности кода: часть возникающих мутаций молекулярно-генетически нейтральна (рис. 33А).

Ранее при изучении различий между аминокислотными последовательностями  $\alpha$ -цепей гемоглобина позвоночных было установлено, что эти цепи у человека и карпа различаются примерно на 50 % аминокислотных сайтов. Это доказывает независимую и симметричную дивергенцию генов  $\alpha$ - и  $\beta$ -цепей после их появления в результате репликации. Предполагается, что дупликация произошла около полумиллиарда ( $5 \times 10^8$ ) лет, поэтому скорость аминокислотных замен равна примерно  $0,96 \times 10^{-9}$ , что очень близко к величине  $0,9 \times 10^{-9}$ , полученной при сравнении последовательностей  $\alpha$ -цепей гемоглобинов позвоночных. Эти факты свидетельствуют о том, что скорость молекулярной эволюции, выраженная через число мута-

ционных замен, постоянна в расчете на год, при этом она не зависит от продолжительности поколения, размера популяции и условий обитания.

При реконструкции филогений и оценке генетических различий по аминокислотным сайтам белков делаются различия в отношении двух типов гомологических отношений между генами — *ортологическими* и *паралогическими* (Айала, 1984).

Ортологичные гены происходят от предкового гена в генотипе вида, от которого образовались сравниваемые виды. Эволюция таких генов отражает эволюцию видов, в генофонде которых они локализованы. Так, при изучении нуклеотидных замен в ортологичных генах, кодирующих *цитохромы С* у 20 различных видов от дрожжей и нейроспоры до человека, было установлено, что все они произошли от одного предкового гена в генотипе предка всех 20 сравниваемых видов (Fitch, Margoliash, 1967).

Паралогичные гены происходят от предкового гена в процессе дупликации, а эволюционные изменения претерпевают в пределах одного вида, а также параллельно у разных видов (рис. 41).

У человека  $\alpha$ -,  $\beta$ -,  $\gamma$ - и  $\delta$ -цепи гемоглобина кодируют паралогичные гены. На рисунке точками отмечены места дупликаций предковых генов, послужившие началом новых направлений эволюции. В каждой ветви древа указано число нуклеотидных замен, необходимое для возникновения существующих различий между аминокислотными последовательностями белков. Первая дупликация гена произошла примерно 600 млн лет назад: од-

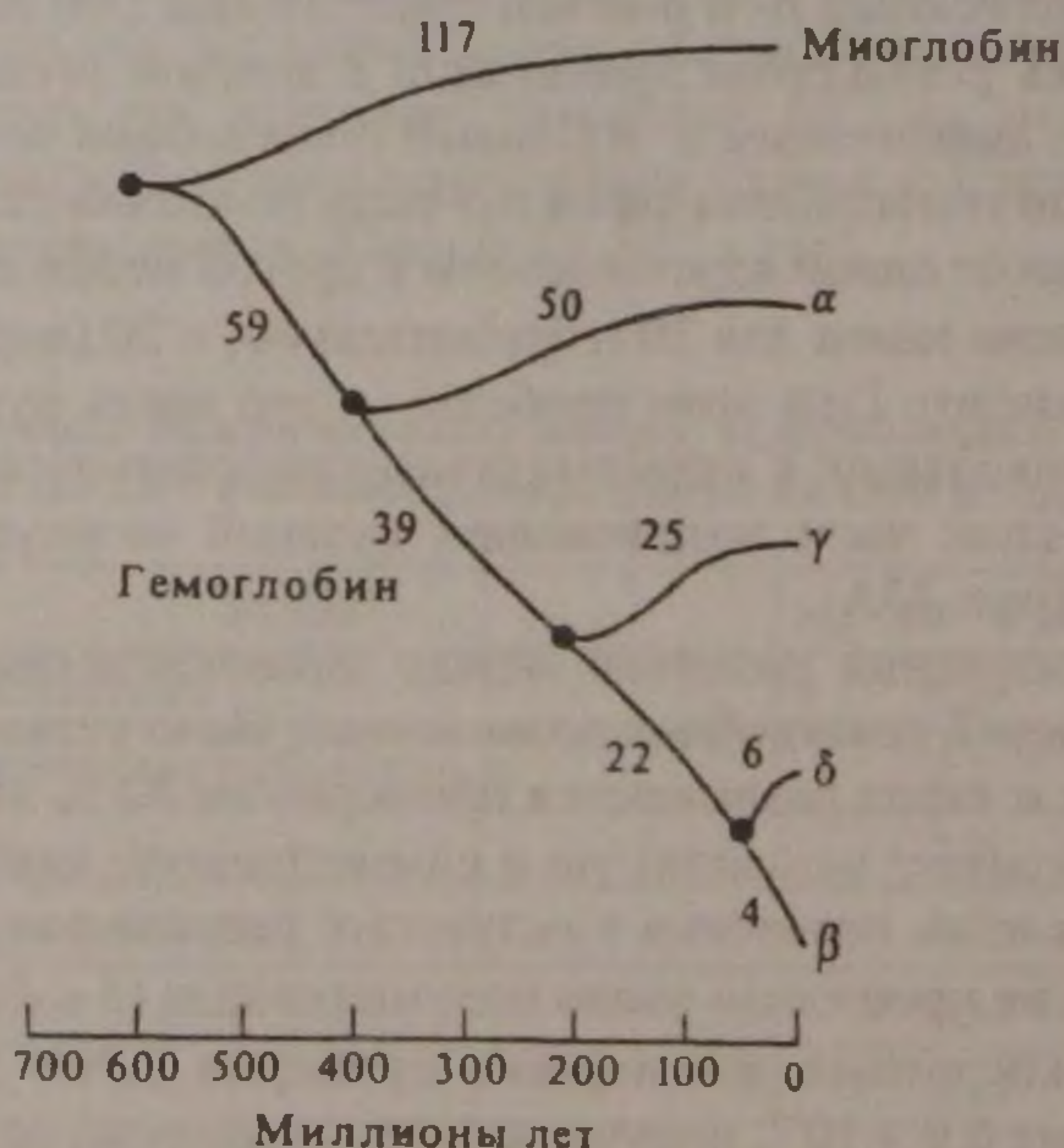


Рис. 41. Эволюционная история генов глобина

на из двух копий исходного гена кодирует *миоглобин*, а другая дала начало всем генам, кодирующим *различные гемоглобины*. Около 400 млн лет тому назад дуплицировался ген гемоглобина, и одна из копий привела к современному гену  $\alpha$ -гемоглобина, а вторая еще раз дуплицировалась 200 млн лет назад, в результате чего возникли гены  $\gamma$ - и  $\beta$ -цепей гемоглобина. В последней ветви, к которой относятся современные высшие приматы, 40 млн лет назад произошла еще одна дупликация, положившая начало эволюции  $\delta$ -цепи гемоглобина. Оценки абсолютного времени дупликаций основаны на палеонтологических и морфологических исследованиях.

Многие генетики отмечают, что проблема постоянства скорости эволюции является одной из наиболее дискуссионных в теории молекулярной эволюции, особенно ее интерпретация с позиции теории нейтральности (Левонтин, 1978). Тем не менее, исходя из фактов примерного постоянства скорости молекулярной эволюции во времени, измеряемом в годах, выдающийся биохимик XX в. Лайнус Полинг сформулировал *концепцию «молекулярных эволюционных часов»* (Zuckerkandle, Pauling, 1965). В этой концепции тезис о постоянстве скорости эволюции, выраженной через число аминокислотных замен, является научной гипотезой, которая приобрела большую популярность. Эти молекулярные часы не имеют ничего общего с точным механизмом измерения времени в обыденной повседневной жизни. Ход этих «часов» носит *стохастический характер, сопоставимый с радиоактивным распадом*. В качестве постоянной принимается лишь вероятность изменений в единицу времени, подчиняющаяся распределению Пуассона. При измерении продолжительных промежутков исторического времени *стохастические часы* показали достаточную точность, с небольшим разбросом скоростей эволюции.

Сформулированы основные положения гипотезы молекулярных часов (Абрамсон и др., 2004):

- 1) скорость молекулярных изменений относительно постоянна в пределах конкретных генов и таксонов;
- 2) данные по молекулярной дивергенции можно использовать для предсказания времени реальной дивергенции;
- 3) создана шкала времени для выведения филогении, часто используемая при отсутствии или недостаточности палеонтологических данных;
- 4) предполагается, что для многих генов и участков генома у многих таксонов существуют «универсальные» молекулярные часы;
- 5) существуют и «локальные» молекулярные часы: у близкородственных видов со сходной биологией, скоростью метаболизма и сменой поколений скорость эволюции конкретного гена стабильна.

Большой заслугой эволюционно-генетического направления явилась формулировка и экспериментальное подтверждение *идеи глобальной клас-*

сификации мира живых организмов. В то же время имеются примеры, когда филогении на основе традиционных и молекулярно-генетических методов оказываются различными. Источниками несовпадения филогений на основе разных методов являются различия в скорости генетических и морфологических изменений в разных ветвях филогенетического древа. Примером является противоречие между скоростями эволюции организма человека в целом и скоростью молекулярной эволюции в ветви, приведшей к возникновению человека (рис. 42). На основе анатомических отличий человека (семейство *Hominidae*) от всех высших обезьян, его и шимпанзе (семейство *Pongidae*) относят к разным таксонам на уровне семейств. Наряду с этим их геномные различия незначительны и сопоставимы с морфологически неразличимыми друг от друга видами дрозофил. В эволюционной линии, приведшей к возникновению человека, изменения организации ( $y$ ) были значительно больше, чем в линии шимпанзе ( $x$ ). Но если судить по скорости нуклеотидных замен в ДНК и аминокислотных замен в белках, то скорости эволюционных изменений в обеих линиях были примерно одинаковы ( $W \sim Z$ ).

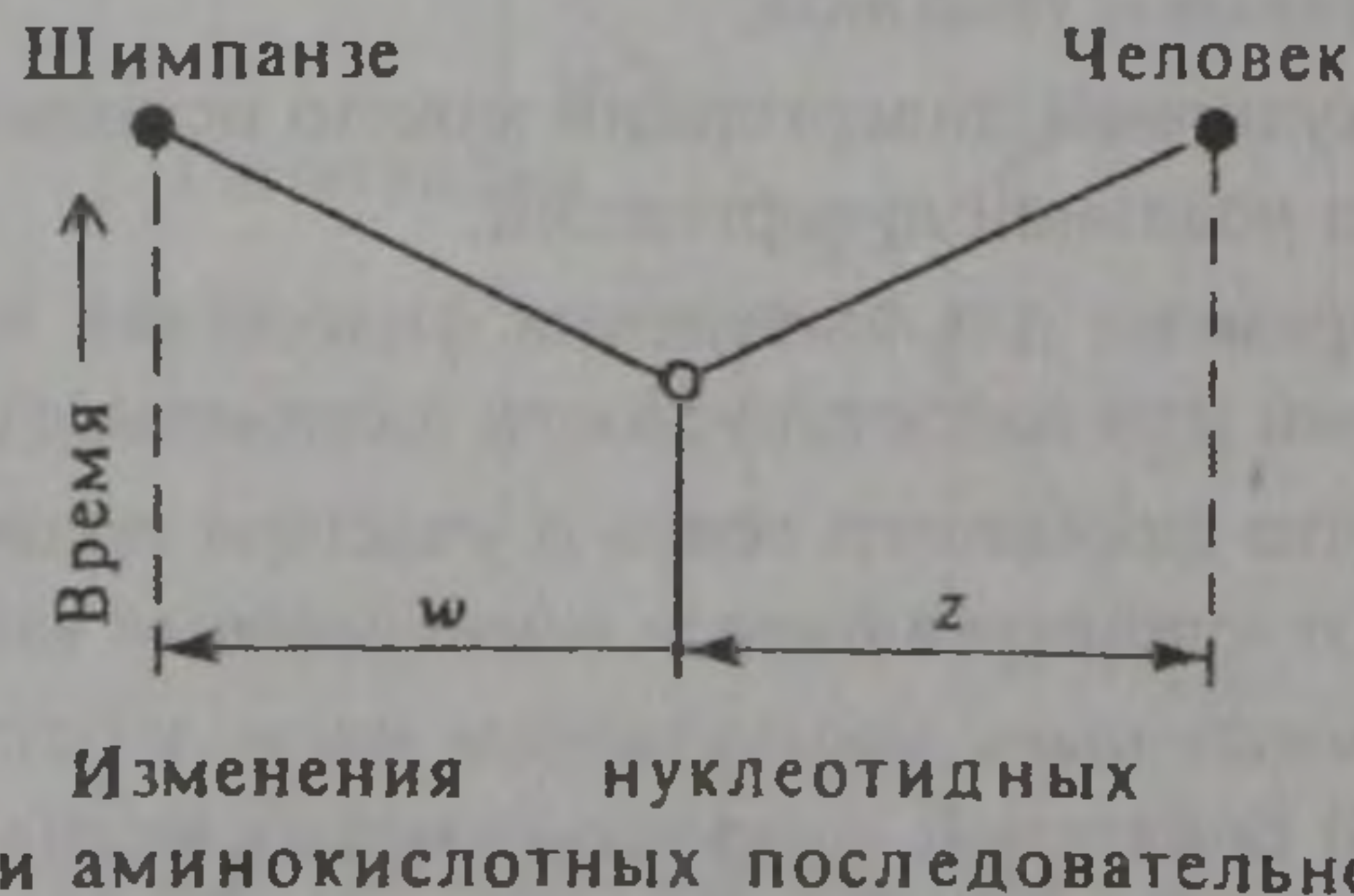
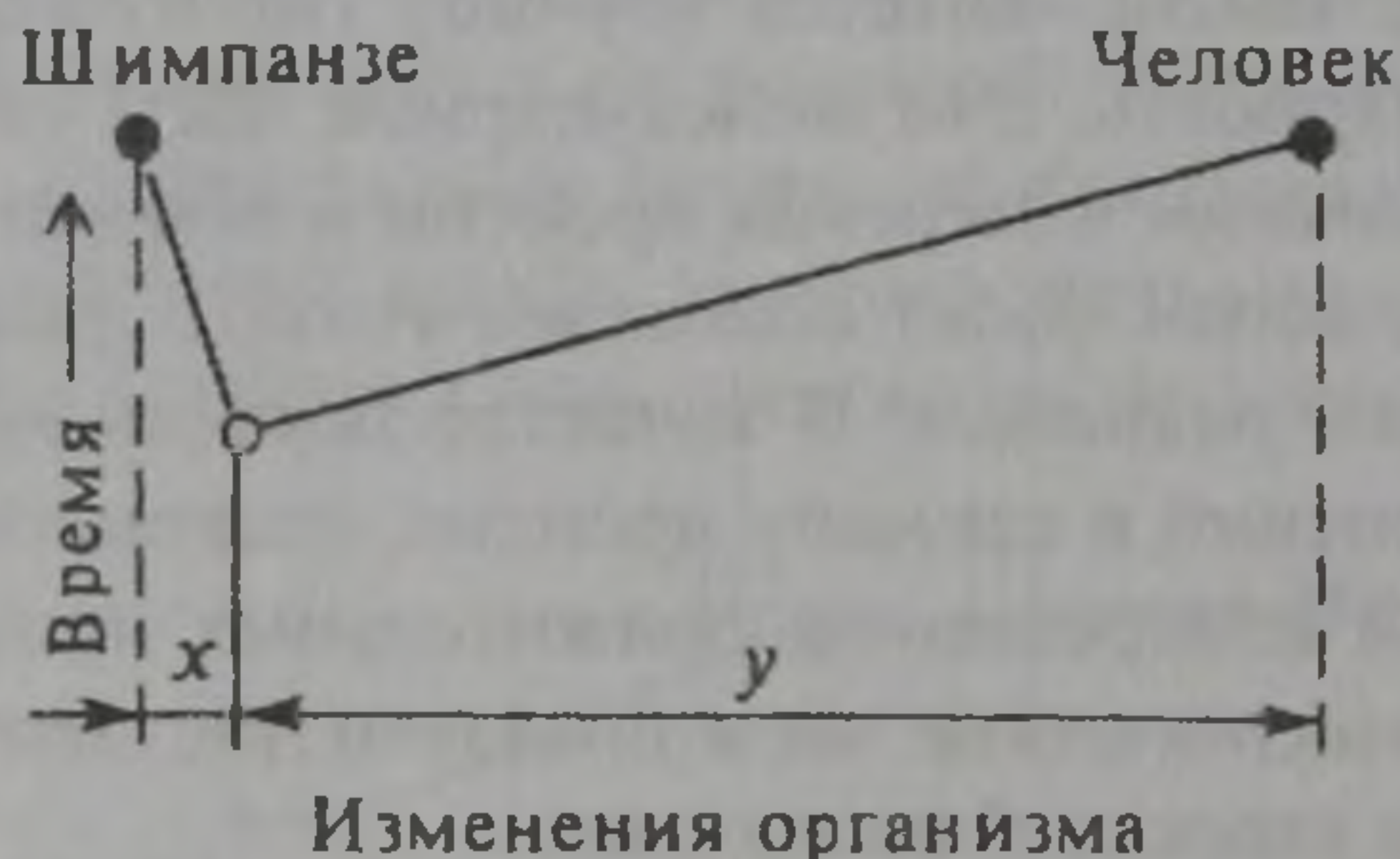


Рис. 42. Несовпадение между морфологической и молекулярной эволюцией на примере дивергенции человека и шимпанзе.

Рядом авторов (Абрамсон и др., 2004) систематизированы наиболее распространенные артефакты («ловушки») молекулярной филогении:

- 1) использование негомологичных (ортологичных) генов у изучаемых видов — филогения видов выводится путем сравнения паралогичных вариантов генов, образованных в результате репликации еще до разделения видов;
- 2) использование в качестве молекулярных маркеров генов, скорость эволюции, которых не адекватна рассматриваемому периоду времени — гены с большим количеством мутаций (несут большое количество филогенетической информации) и гены с медленным эволюционированием и небольшой филогенетической информацией;
- 3) использование коротких цепочек ДНК с недостаточным количеством информации;
- 4) использование ограниченного набора видов и малых выборок.

В связи с тем, что изучение скорости молекулярной эволюции и концепция молекулярных эволюционных часов основаны на теории нейтральности, необходимо кратко рассмотреть ее основные положения. Ее автор, Мотоо Кимура (1985) дает уточненное, более правильное название — «*гипотеза взаимодействия нейтральных мутаций и генетического дрейфа*» — и формулирует эмпирические и теоретические правила «*общей концепции молекулярной эволюции*». Эта теория утверждает, что большая часть мутационных замещений в ходе эволюции обусловлена не положительным дарвиновским отбором, а случайным закреплением нейтральных или почти нейтральных мутаций. Предполагается, что внутривидовая молекулярная генетическая изменчивость проявляется в виде полиморфизма ДНК и белков и селективно нейтральна.

## Правила молекулярной эволюции

**1 правило:** скорость эволюции любого белка (в полингах) постоянна и одинакова в разных филогенетических линиях.

Установлено, что скорость эволюции наиболее быстро изменяющихся белков в почти 400 раз больше, чем у белка с самой низкой скоростью эволюции *гистона*. В одном семействе темп эволюции белков редко различается более чем в 2–3 раза. Постоянство темпов молекулярной эволюции подтверждает тот факт, что филогенетические древа на основе сравнительного изучения аминокислотных и нуклеотидных последовательностей согласуются с обычными филогенетическими схемами.

**2 правило:** функционально менее важные молекулы или их части эволюционируют (накапливают мутационные замены) быстрее, чем более важные.



**3 правило:** мутационные замены аминокислот с меньшими нарушениями структуры и функции белков происходят чаще тех, которые вызывают существенные нарушения структуры и функции этой молекулы.

**4 правило:** появлению нового в функциональном отношении гена всегда должна предшествовать *дупликация гена*.

За последние годы накопилось большое количество данных о широкой распространенности дублированных (повторяющихся) структур в геномах эукариот. Так  $\alpha$ - и  $\beta$ -цепи гемоглобина возникли в результате генной дупликации. Ранее дупликация гена обусловила дивергенцию гемоглобина и миоглобина, а после повторных генных дупликаций у предков человека из  $\beta$ -цепи возникли  $\gamma$ - $\delta$ -цепи:  $\gamma$ -цепь является составной частью фетального гемоглобина плода HbF, а  $\delta$ -цепь входит в состав HbA<sub>2</sub>. Геном человека в норме содержит по одному гену  $\beta$ ,  $\delta$  и других цепей гемоглобина, а гены  $\alpha$ - и  $\gamma$ -цепей дублированы (Фогель, Мотульски, 1990). Важную роль дупликации генов в прогрессивной эволюции от прокариот до млекопитающих показывают данные М. О. Дэйхоф (Dayhoff, 1978), в соответствии с которыми общее количество имеющихся у человека 50 000 белков можно сгруппировать в 500 *надсемейств*, каждое из которых включает около 100 белков.

**5 правило:** случайная фиксация нейтральных или очень слабо вредных мутаций в ходе эволюции происходит чаще благоприятных мутаций.

Примером случайной фиксации слабо вредных, почти нейтральных мутаций является утрата человеком и некоторыми другими видами (обезьяны, морские свинки, питающиеся фруктами рукокрылые, некоторые воробьиные птицы) способности синтезировать аскорбиновую кислоту (витамин С). Способность синтезировать эту кислоту — характерный признак наземных позвоночных, потребляющих пищу с ее богатым содержанием. Можно предположить, что мутации, приводившие к потере способности синтезировать аскорбиновую кислоту, особенно вредными не были и при постоянном мутационном давлении фиксировались в процессе случайного дрейфа генных частот.

В отечественной литературе в поддержку представлений М. Кимуры выступил Ю. П. Алтухов (Алтухов, 1983) и М. В. Волькенштейн (Волькенштейн, 1981, 1984). По современным представлениям, на молекулярном уровне 2/3 замещений нуклеотидных пар вызывает фенотипический эффект в форме изменения полипептидной цепи, а 1/3 замещений не контролируются отбором. Это означает, что в эволюции наблюдается взаимное сочетание как процессов отбора, так и стохастические процессы дрейфа нейтральных мутаций (Воронцов, 1999). В результате «недарвиновская эволюция» за счет стохастических процессов и дарвиновский отбор представляет собой непротиворечивые и взаимодополняющие друг друга концепции.

## Заключение

Концепция «молекулярных часов» допускает стохастическую природу накопления нейтральных мутаций в филетических линиях как результат филетической эволюции и известную скорость их накопления. Метод описания биоразнообразия по молекулярным признакам, как и любой другой, имеет свои недостатки и ограничения. Так, между классическими и молекулярными методами имеются разночтения по точности предсказания времени дивергенции видов, то есть расхождения признаков у родственных организмов в ходе эволюционного процесса. Встает вопрос о том, постоянна ли скорость такой дивергенции у разных таксонов. Проблема калибровки времени молекулярных событий постепенно найдет свое разрешение по мере получения многих независимых оценок времени дивергенции по различным генам и по многим таксонам. Молекулярно-генетические данные, пригодные для исследований филогенеза, накоплены для очень малой доли живых организмов: для нескольких десятков тысяч из почти 2 млн современных видов. По существующим оценкам, в ископаемом состоянии находятся 90 % возникших в процессе эволюции видов. Следовательно, молекулярные данные недоступны для подавляющей доли глобального биологического разнообразия.

Из информационных макромолекул в молекулярно-генетической систематике наиболее широкое применение нашли рРНК, а также митохондриальная и хлоропластная ДНК. Большие перспективы открывает переход в области *генамики и эпигенетики*. Разработка быстрых методов *секвенирования (определения последовательности нуклеотидов ДНК)* сделала возможным определение нуклеотидных последовательностей крупных молекул ДНК и их фрагментов. После опубликования первой полной последовательности геномной одноцепочечной ДНК бактериофага фидесять-174 (Сэнджер, 1977), полного генома самостоятельно существующего организма — грамотрицательной бактерии *Haemophilus influenzae* (1995), первого эукариотического организма и других, началась *эра генамики и биоинформатики* как самых перспективных направлений молекулярной генетики. В настоящее время расшифровкой геномов занимаются многие международные консорциумы. В 2001 г. о завершении секвенирования большей части (более 95 %) генома человека объявили участники американской программы HGP и Международного консорциума по секвенированию генома человека (JHGSC), объединяющего научные организации США, Великобритании, Германии, Японии и Китая. Число секвенированных геномов достигло более тысячи, среди них более 100 видов болезнетворных микроорганизмов и более 600 вирусных геномов. Безусловно, что это направление исследований внесет значительный вклад в развитие геносистематики и филогенетики.

Несомненно, что залогом достоверного описания эволюционных событий, филогении и систематики таксонов является сопоставление результатов, полученных разнообразными классическими (морфологическими) и молекулярными методами.

### Ключевые слова и термины

Филогенетика, систематика, таксономия, таксоны — империя, царство, тип; древо жизни, прокариоты, эукариоты; филогеномика, геносистематика, теория нейтральности биологической эволюции, молекулярные эволюционные часы, дарвин и полинг — единицы скорости эволюции, гены — омологичные, ортологичные и паралогичные; секвенирование геномов, геномика.

## Горизонтальный (латеральный) перенос генов как информационный фактор ЭВОЛЮЦИИ

Современная молекулярная генетика доказала генетическую роль нуклеиновых кислот во всем мире живых организмов. Многочисленные опытные данные показали, что нуклеиновые кислоты ДНК и РНК кодируют наследственную информацию и способны к конвариантной редупликации; они локализованы в разных структурах — хромосомах, митохондриях и пластидах, в плазмидах, вирусах и др. В основе жизненных процессов лежит универсальный для всех организмов генетический код.

В 1972 г. в Стенфордском университете (США) П. Берг с сотрудниками впервые генно-инженерными методами *in vitro* получили *рекомбинантную (гибридную) молекулу, состоящую из ДНК обезьяньего вируса SV-40 и ДНК фага лямбда*. Затем гибридные молекулы ДНК вводили в клетки с целью их размножения, селекции и выделения клонов. Эти опыты положили начало генно-инженерным исследованиям по созданию штаммов бактерий, дрожжей, линий клеток, эффективно продуцирующих биологически активные белки человека и животных. Эти штаммы и линии *содержат и экспрессируют чужеродную генетическую информацию*.

Это направление исследований побудило ученых поставить вопрос: существуют ли аналогичные процессы в природных условиях, т. е. могут ли одновременно существующие в природе близкородственные и отдаленные формы обмениваться генами?

По мере выяснения путей передачи информации используют следующие термины:

- *вертикальный перенос генов* — передачи информации от родителей потомкам;

- *горизонтальный перенос генов* — передача наследственной информации от одних одновременно существующих организмов другим, т. е. особям того же поколения

Многочисленные эксперименты и наблюдения разных авторов показали, что горизонтальный перенос генов является одним из важнейших механизмов эволюции вирусов и прокариот. У прокариот выделяют три способа горизонтального переноса генов — трансдукцию, конъюгацию и трансформацию.

## 1. Горизонтальный перенос генов у прокариот

Содержащаяся в бактериях генетическая информация заключена в *нуклеоиде* — эквиваленте ядра клетки эукариот. Нуклеоид не окружен мембраной и не отделен от цитоплазмы. В его состав входят ДНК, РНК и белки. Каждый нуклеоид содержит двухцепочечную замкнутую в кольцо молекулу ДНК размером  $1-4 \times 10^6$  п. н. Кроме нуклеоида, у бактерий ДНК содержатся также во внехромосомных генетических элементах — *плазмидах*. Они представляют собой двухцепочечные кольцевые молекулы ДНК размером от 2 т. п. н. до более чем 300 т. п. н. Плазмиды могут нести гены,



Рис. 43. Биологические функции, определяемые бактериальными плазмидами

которые обуславливают фенотипическое отличие содержащих их клеток от бесплазмидных клеток. Одни плазмиды в клетке находятся в одной или нескольких копиях, другие могут иметь в клетках 10–200 копий. Плазмиды несут устойчивость к антибиотикам, лекарственным препаратам, тяжелым металлам. Огромно их значение для генетической инженерии: легкость выделения плазмид из клеток, возможность встраивания в них чужеродного генетического материала *in vitro*, хорошо разработанные технологии введения реконструированных плазмид обратно в клетку. Они часто используются в качестве векторных систем, обеспечивающих доставку чужеродного фрагмента ДНК в клетку хозяина и репликацию в ней гибридной молекулы ДНК. (рис. 43)

У бактерий отсутствуют митоз и мейоз, связанные с клеточным делением и гаметогенезом у эукариот. Но у них имеются три возможности для обмена генетическим материалом:

- 1) *трансдукция* — перенос генов из одной бактериальной клетки в другую при помощи бактериофагов;
- 2) *трансформация* — перенос генетической информации в бактериальные клетки при помощи изолированной ДНК с участием или без участия плазмид;
- 3) *конъюгация* — это непосредственный контакт между клетками бактерий, сопровождаемый переносом генетического материала из клеток донора в клетки реципиента.

### Трансдукция

Генетический обмен (трансдукция) был открыт в 1952 г. Н. Циндер и Д. Ледербергом на бактерии мышинного тифа *Salmonella typhimurium*, а генетический перенос осуществлялся с участием фага-переносчика. Бактериофаги, или фаги (их принято обозначать греческой буквой — фи) — это группа бактериальных вирусов двух категорий: вирулентных и умеренных.

Вирулентный бактериофаг проникает в клетку, размножается и лизирует бактерию, т. е. вызывает *литическую реакцию*. Умеренные бактериофаги могут вызывать как литическую, так и *лизогенную реакцию*. В последнем случае фаг переходит в состояние *профага*, а несущие их бактерии называют *лизогенными*, так как они устойчивы к дополнительному заражению тем же фагом. Фаг включает свой генетический материал в хромосому бактерии реципиента механизмом типа кроссинговера.

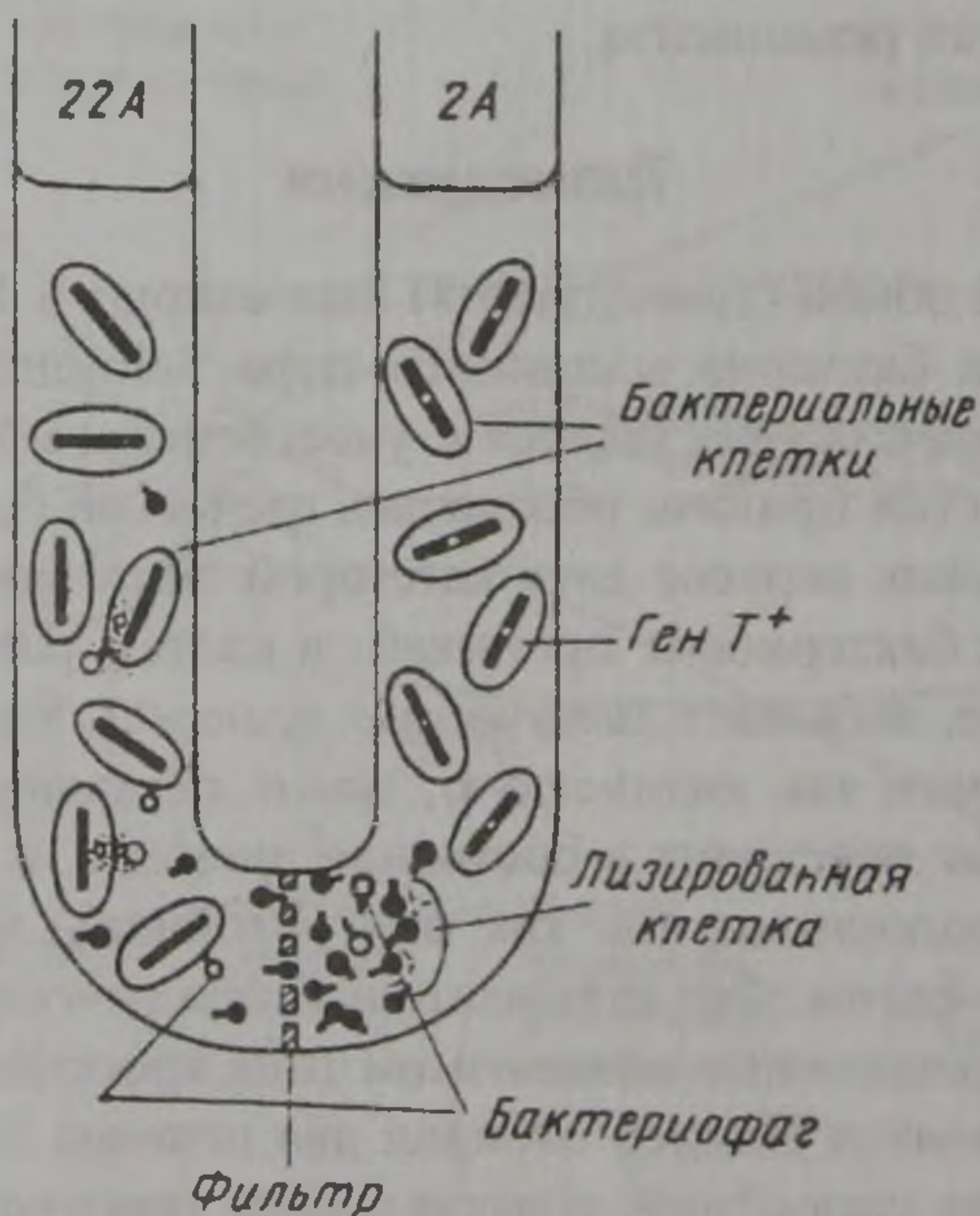
Для экспериментов Циндер отбирал два штамма *S. typhimurium*: ауксотрофный 22А, не способный синтезировать триптофан ( $T^-$ ), и штамма 2А, способный синтезировать триптофан ( $T^+$ ). Штаммы засеивали в U-образную трубку, разделенную внизу бактериальным фильтром. После инкубации штаммов в трубке и рассева их клеток бактерии штамма 22А при посеве на

минимальную питательную среду дали небольшое количество колоний (частота таких трансдуцированных клеток была равна  $1:10^{-5}$ ). Это говорило о том, что некоторые клетки приобрели способность синтезировать триптофан и образовывать колонии на среде без этой аминокислоты (рис. 44).

Анализ показал, что фаг P22 из лизогенного штамма 22А освобождается из культуры, проходит через фильтр и лизировал штамм 2А. Фаг присоединял часть генного материала 2А, возвращался обратно и передавал эти гены штамму 22А, т. е. последний приобретал свойство синтезировать триптофан. Аналогичным образом могут быть трансдуцированы многие другие признаки, например, способность к сбраживанию, устойчивость к антибиотикам.

Величина трансдуцируемого фрагмента определяется размером ДНК донора, способной упаковаться в головку фага. Доказано, что в частицах трансдуцирующего фага практически вся фаговая ДНК может быть заменена на бактериальную (например, фаги P1 и P22 у разных бактерий — эшерихий, шигелл и сальмонелл).

В 1957 г. К. Уоддингтон дал эволюционную оценку открытию трансдукции бактериальных генов от одного вида бактерий другому через инфицированные бактериофаги. Он уже тогда предположил возможность наличия подобного переноса генетической информации не только у прокариот, но и эукариот.



**Рис. 44.** Схема опыта, демонстрирующего явление трансдукции у *Salmonella*. Штаммы бактерий 22А — не способных синтезировать триптофан (Т<sup>-</sup>); 2А — способных синтезировать триптофан (Т<sup>+</sup>) и лизогенных по бактериофагу

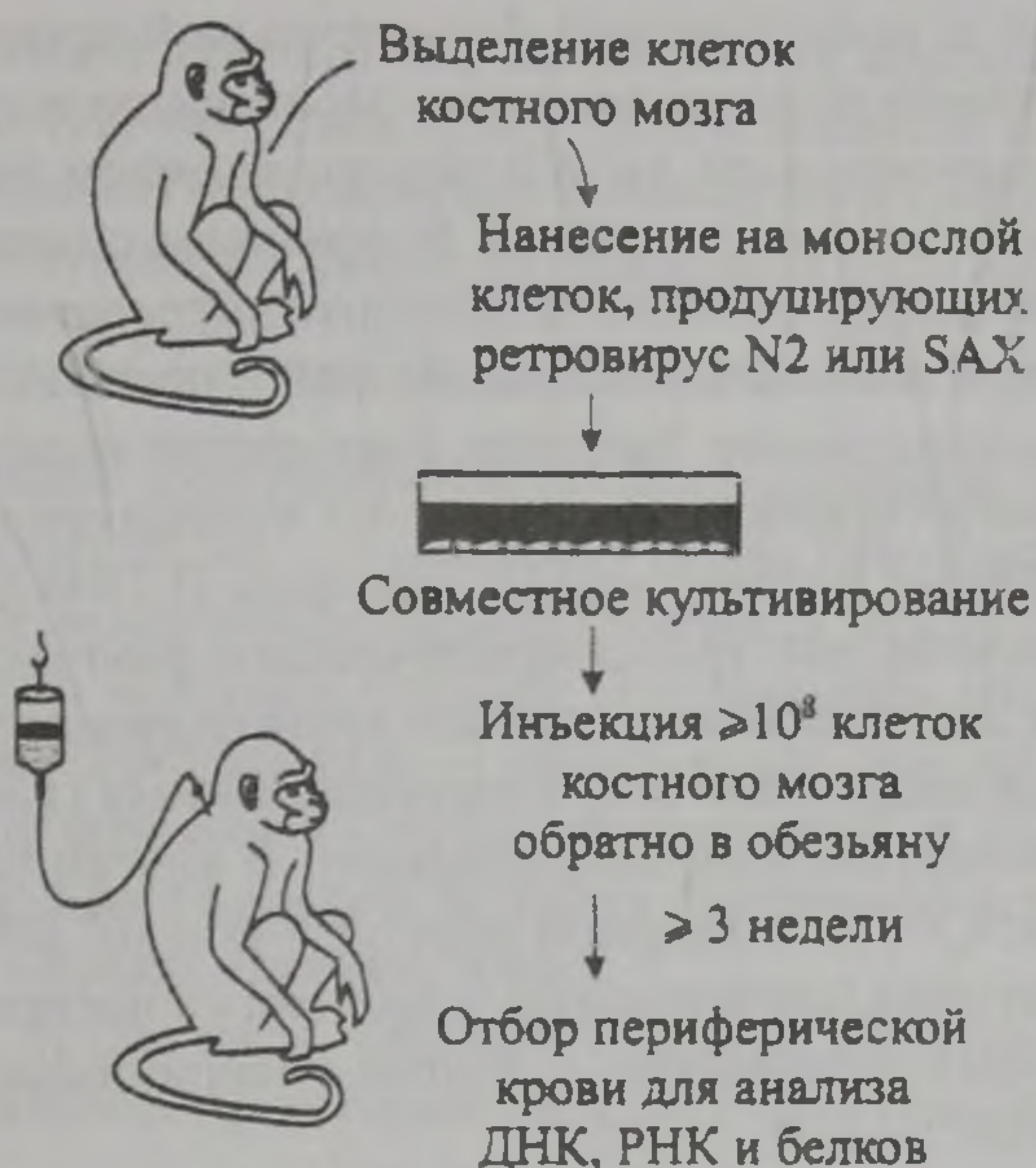


Рис. 45. Схема эксперимента по переносу чужеродных генов в организм макака-резуса

Многие эксперименты по трансдукции генов с помощью ретровирусных векторов проведены на мышах. В 1986 г. С. Андерсон доказал возможность переноса гена аденозиндезаминазы человека в организм макака-резуса с помощью гибридного ретровируса SAX.

Генная терапия с применением ретровирусов основана на их активном использовании в качестве молекулярных векторов (рис. 45).

Внутри капсида вирионов ретровирусов имеется несколько десятков молекул обратной транскриптазы, молекулы тРНК и две идентичные геномные одноцепочечные молекулы РНК длиной от 3,5 до 9 тыс. нуклеотидов (уникальный случай диплоидного вирусного генома). М. Калеко (1991) разработал метод трансдукции генов в клетке печени животного. Гибридный ретровирус, содержащий ген *neo*, инъецировали в митотически активную паренхиму мыши и показали, что при этом происходил перенос гена и его длительная экспрессия *in vitro*. Генная терапия на основе аденовирусов направлена на создание стратегии лечения раковых опухолей. Проводятся клинические испытания гибридных аденовирусов при терапии рака легких, печени, области головы и шеи.

### Трансформация

Трансформация — это поглощение изолированной ДНК бактериидонора клетками бактерии-реципиента. В 1928 г. Ф. Гриффит вводил мы-



шам вирулентный и авирулентный бескапсульный штамм пневмококков. Введение вирулентного штамма вызывало заболевание и гибель мышей, тогда как введение вирулентного, но убитого нагреванием штамма, или авирулентного штамма, не вызывало гибели. В основном опыте вводилась смесь живой культуры авирулентного и убитого нагреванием вирулентного штаммов: мыши при этом заболели пневмонией и погибли, а в их крови имелись вирулентные капсульные бактерии. Был сделан вывод, что живые бактерии авирулентного бескапсульного штамма *трансформировались* — приобрели свойства убитых болезнетворных бактерий. В 1944 г. Эвери, Мак-Леод и Мак-Карти показали, что трансформирующим фактором является ДНК.

Установлены два типа бактериальной трансформации: естественная — у *Bacillus subtilis* (сенная палочка) и индуцированная (клетка должна приобрести компетентность к трансформации). В клетки бактерий проникают фрагменты ДНК с молекулярной массой  $4-5 \times 10^6$  D\* (менее 1 % хромосомы). Эти бактерии (пневмококки и др.) могут неспецифически поглощать ДНК из разных источников, а другие (*Haemophilus*) — только свою гомологичную ДНК.

В ходе трансформации (длительность 10–30 мин) одна из двух цепей ДНК деградирует, а вторая проникает и спаривается с гомологичным участком ДНК реципиента и встраивается в нее посредством кроссинговера, т. е. ДНК становится гибридной (длина интегрированной цепочки 500 п. н. — 200 т. п. н.). Трансформация с низкой частотой может происходить между разными видами бактерий, что помогает установить степень родства между ними. В связи с развитием генетической инженерии широко применяется *плазмидная (векторная) трансформация* — введение в клетки прокариот и эукариот рекомбинантных (гибридных) ДНК.

За последние 50 лет осуществлено более 60 трансформаций микроорганизмов внутри вида и между видами: изменение внешних признаков (появление капсул или жгутиков), биохимических свойств (устойчивость к антибиотикам, лекарственным веществам и др.), способности окислять или сбрасывать те или иные аминокислоты (лизин) или витамины (В12), придание некоторым видам способности осуществлять необычные для них биохимические процессы.

### Конъюгация

*Конъюгация* — это процесс переноса генетического материала от бактерии-донора в бактерию-реципиент при их непосредственном контакте. Установление этого явления связано с открытием в 1946 г. Дж. Ледербергом и Э. Тейтемом процесса рекомбинации у штамма K 12 *E. coli*. В исходном эксперименте они использовали принцип отбора редких *прототроф-*

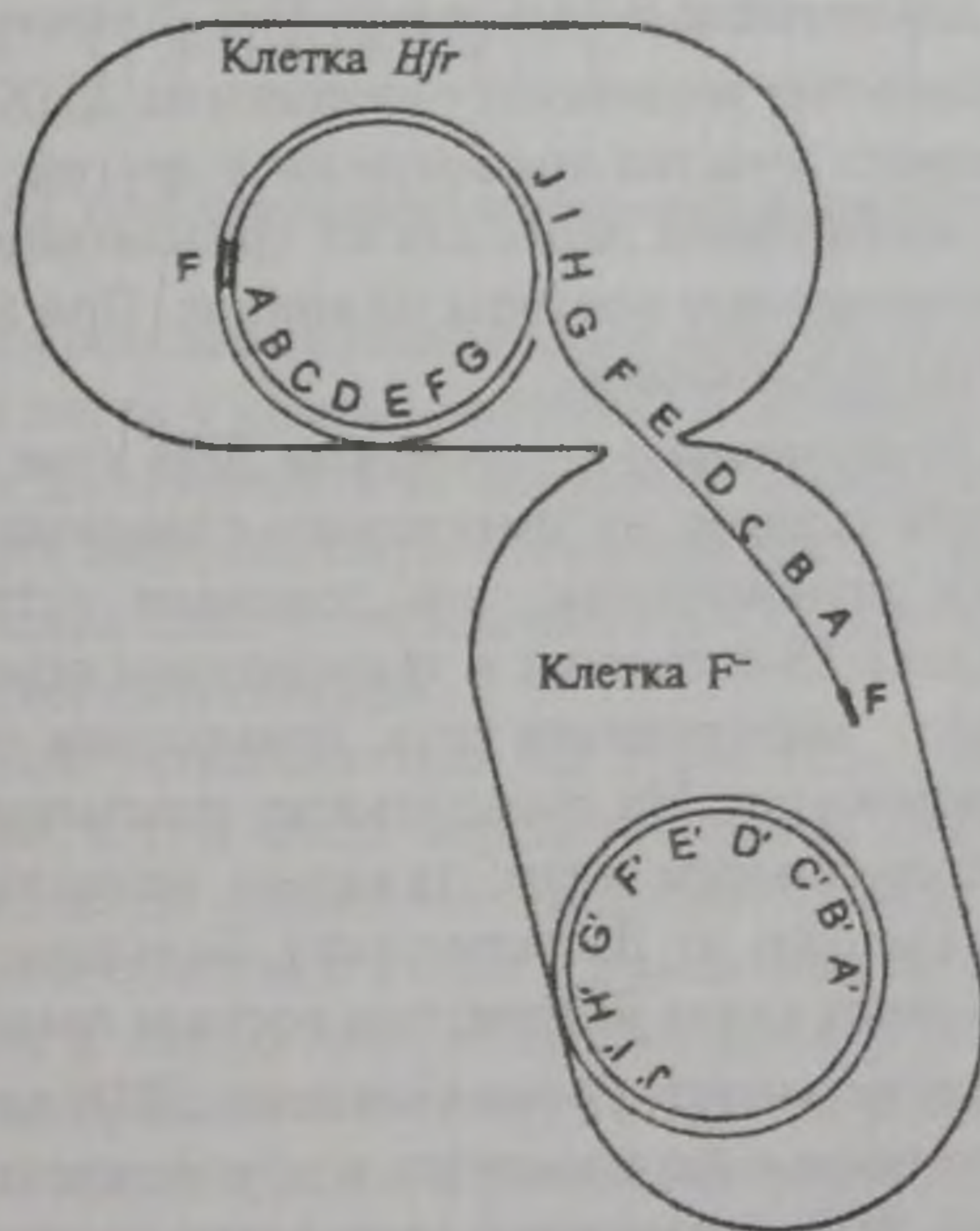
\* D — дальтон, единица измерения молекулярной массы.

ных (способных синтезировать все необходимые компоненты из простых питательных веществ, таких как сахар) рекомбинантов из смеси ауксотрофов (мутанты, не способные расти без добавления в среду некоторых сложных веществ) мутантных штаммов на минимальном агаре без добавок:

*штаммы*  $thr^- leu^- met^+ thi^+$  X *штаммы*  $thr^+ leu^+ met^- thi^-$

При смешанном культивировании этих штаммов выделяли редких прототрофов, растущих на минимальной среде. Это были рекомбинанты, появляющиеся после непосредственного контакта между бактериями (конъюгации) с помощью мостиков F-пилей. Это показало, что у *E. coli* имеется определенный тип скрещивания — конъюгация: это примитивный половой процесс между штаммами (половыми типами)  $F^+$  и  $F^-$ . При конъюгации клеток происходит односторонняя передача генетического материала от клетки  $F^+$  в клетку  $F^-$  (Жакоб, Вильям, Хейс).

Позже был установлен вариант штамма  $F^+$ , названный *Hfr* (*high frequency of recombination*), который передает свои гены клеткам  $F^-$  с более



**Рис. 46.** Перенос ДНК при скрещивании  $Hfr \times F^-$ . Клетка *Hfr* соединяется с клеткой  $F^-$ , и между ними образуется конъюгационная трубка. Фактор F, встроенный в хромосому *Hfr*, начинает реплицироваться. Но так как фактор F встроен в хромосому, он заодно запускает репликацию всей хромосомной ДНК. Поэтому в клетку  $F^-$  переходит копия хромосомы *Hfr* с генами в линейной последовательности. Как только донорская ДНК проникает в клетку  $F^-$ , она может спариваться с хромосомой  $F^-$  и образовывать рекомбинанты с алелями *Hfr* в хромосоме

высокой частотой. Фактор F — это особый класс плазмид — *эписомы*, которые могут существовать в клетках как дискретные элементы, или встраиваться в хромосому хозяев, образуя Hfr-штаммы. Идентифицировано около 30 сайтов интеграции фактора F в хромосому. Когда клетка Hfr соединяется с клеткой F<sup>-</sup>, между ними образуется конъюгационная (протоплазматическая) трубка (рис. 46).

Начало конъюгационного переноса связано с разрезанием одной из нитей ДНК плазмиды ДНК-хеликазой в локусе *ori T*. Свободный конец одной из цепей ДНК передвигается в клетку реципиента F<sup>-</sup> и достраивается до двухцепочечной структуры. В клетке-доноре на цепи ДНК синтезируется вторая цепь. Первым в бактерию-реципиент всегда проникает конец хромосомы, обозначаемый O (ориджин). Метод получил широкое распространение, а расстояние на генетической карте *E. coli* измеряется в минутах. Вся окружность хромосомы составляет 100 мин., на карту нанесено более 1000 генов (около 30 % генетической емкости). При конъюгации половой фактор вместе с фрагментом ДНК иногда переходит в женскую клетку, превращая ее в мужскую.

*К мобильным генетическим элементам (МГЭ) прокариот относятся:*

1. *IS-элементы (insertion sequences)* — сегменты ДНК, способные перемещаться из одного участка локализации в другой; содержат лишь те гены, которые необходимы лишь для их транспозиции (перемещения), а также инвертированные повторы на концах. При встраивании вызывают небольшую дупликацию.
2. *Транспозоны (Tp-элементы)* — сегменты ДНК (так же как IS-элементы), обладающие генами, не имеющими отношения к транспозиции (устойчивости к антибиотикам, гены токсинов — гены, которые имеются в плазмидах). IS-элементы и транспозоны ответственны за генетические явления: инактивации гена, повышения частоты делеций и инверсий, транслокации. На транспозонах перемещаются гены устойчивости к бактериальным ядам, тяжелым металлам, антибиотикам, гены токсинов (делают их патогенными). Большинство полезных для хозяина плазмидных генов находятся в составе транспозонов.
3. *Умеренные фаги* вступают в рекомбинацию ДНК хозяев, захватывают фрагменты хромосом и переносят их в другие клетки или организмы. Копии вирусной ДНК (*провирусы*) становятся постоянными компонентами генома у многих организмов: утрачивает способность образовывать полноценный вирус, но кодирует полезные белки. Как и половой процесс, вирусы способствуют смешению генофондов различных организмов.

Новая роль МГЭ заключается в том, что они создают неравномерность мутационного процесса во времени и в разных популяциях, а появление особо подвижного элемента может вызвать вспышку мутагенеза и

Мобильные генетические элементы (МГЭ) бактерий  
и их участие в ГПГ между видами бактерий

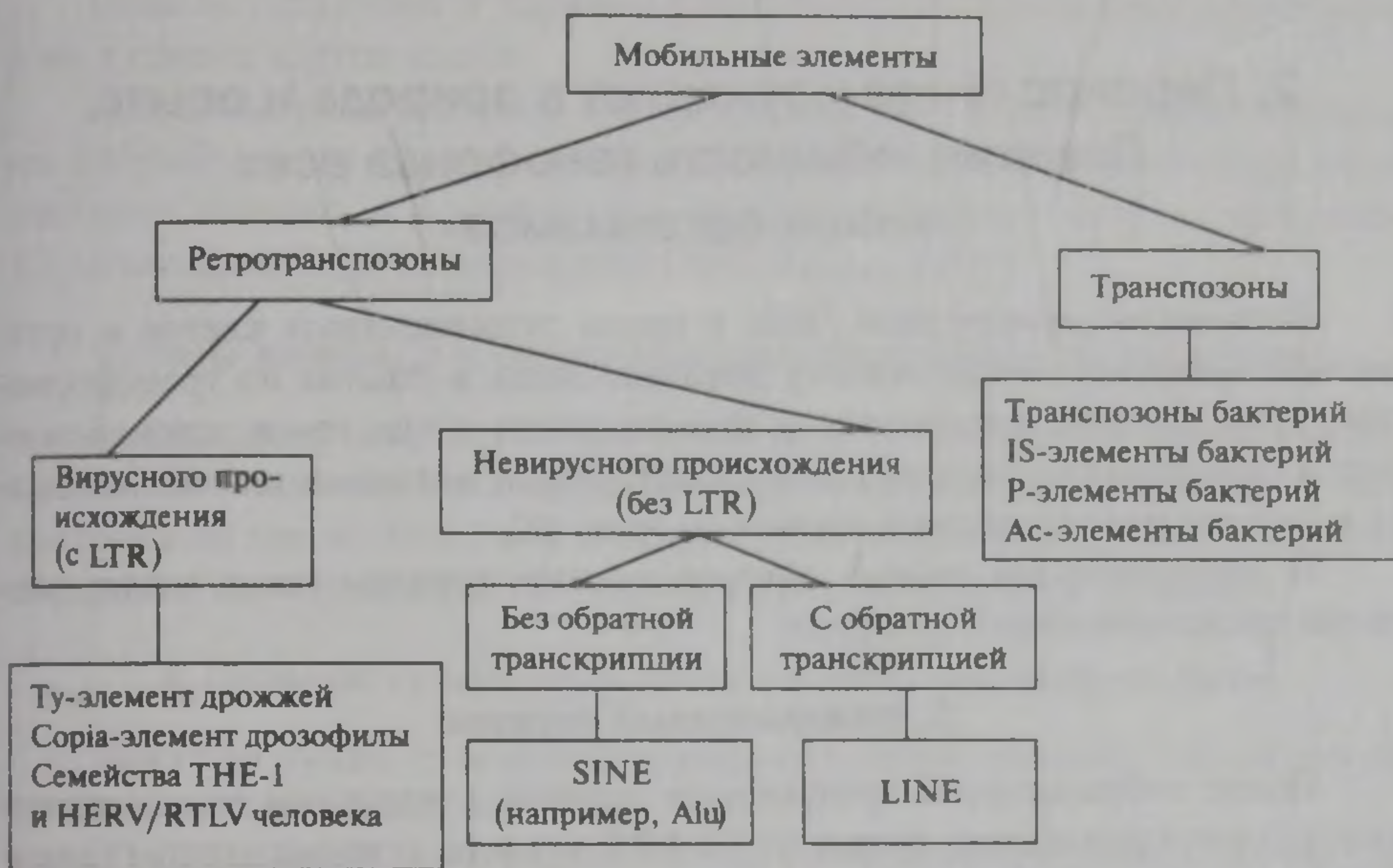


Рис. 47. Классификация мобильных элементов

создавать в популяции особый «мутационный период» («мода» на форму — эвкалиптовая форма листа у разных растений Австралии, «мода на мутации» у дрозофил и т. д., посредством вирусной трансдукции). Направленное изменение разных видов бактерий проявляется в форме *параллельной изменчивости* (свойство сайт-специфичности — включение МГЭ в конкретный определенный сайт генома).

Самой интересной особенностью МГЭ считается активный перенос генов между разными, в том числе неродственными, бактериями. Многие авторы выдвигают идею об *общности генофонда разных видов бактерий*: их объединяют условия обитания. Более аргументированным считается обмен не всеми генами, а только полезными, которых у данного вида ранее не было. Яркой иллюстрацией могут служить гены устойчивости к антибиотикам: это можно назвать «своеобразной коллективной обороной бактерий против «человеческой хитрости» (широкое применение антибиотиков в медицине — выражение Р. Б. Хесина, 1984). Плазмидные детерминанты резистентности к антибиотикам, лекарственным препаратам, тяжелым металлам и др. — эти распространенные у самых разных бактерий «космополиты» — могут быть примером для доказательства общности генетического фонда неродственных бактерий. Транспозоны переносят эти гены от хромосом на плазмиды, а от них — на хромосомы новых видов. Наглядной аналогией распространению крупных инженерных изобре-

ний между народами может представлять в мире бактерий перенос транспозонами и плазмидами генов фиксации азота разными видами.

## 2. Перенос генов у эукариот в природе и опыте. Понятие «общность генофонда всех живых организмов»

Включение чужеродной ДНК в геном реципиентных клеток и организмов эукариот по-настоящему доказано лишь в опытах по трансформации, трансфекции и трансгенезу при введении в них генов, клонированных в большинстве случаев генноинженерными методами или находящих-ся в составе изолированных хромосом (рис. 48).

В экспериментах разных авторов показан перенос генов между разными представителями эукариот:

### 1. Межклассовый перенос

После гибридизации эритроцитов курицы с клетками мыши проявлялся (синтезировался) фермент ГГФРТ курицы (гипоксантин-гуанин-фосфорибозил трансфераза).

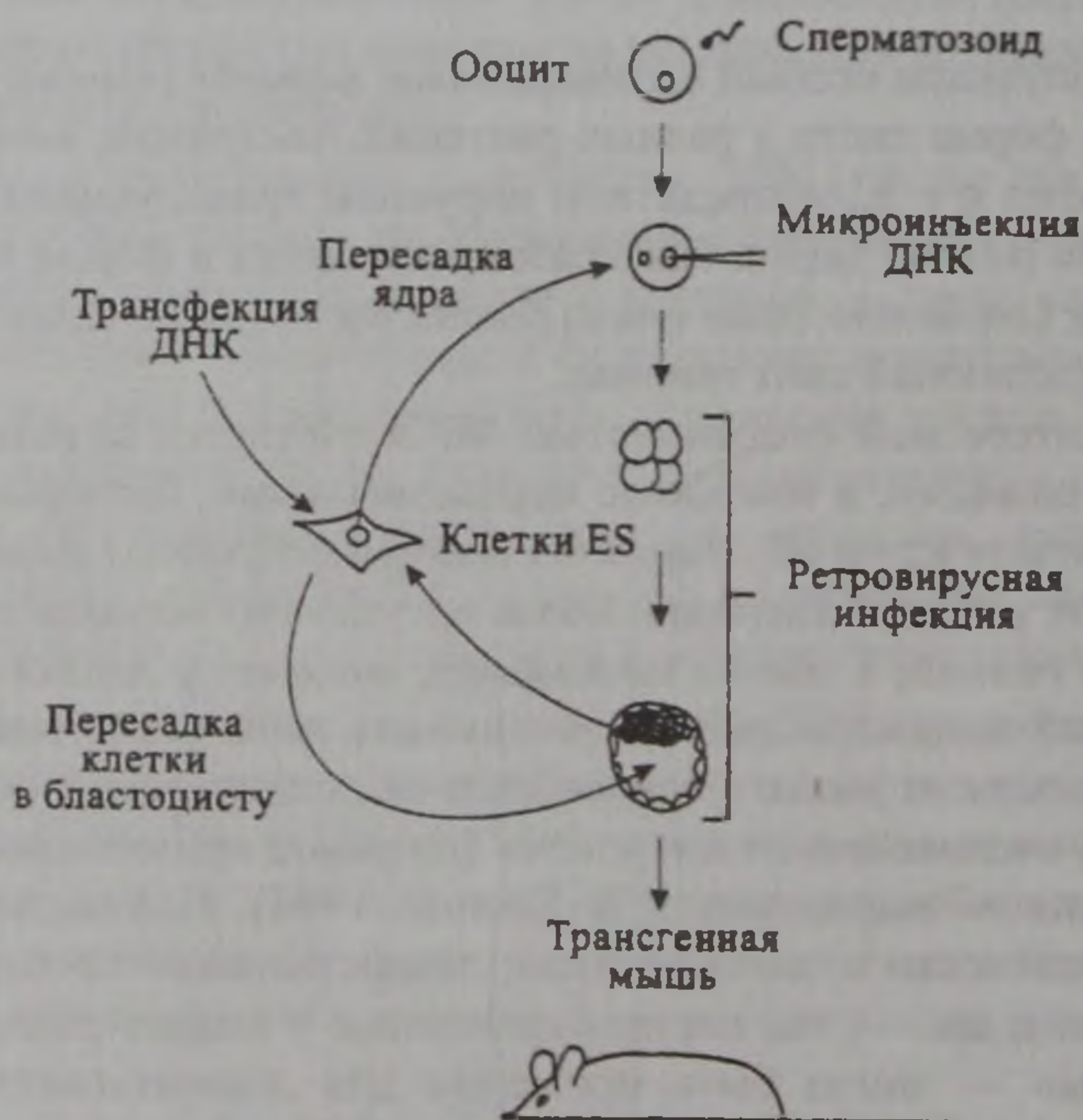


Рис. 48. Пути введения чужеродных генов в клетки млекопитающих на предимплантационной стадии развития

## 2. Перенос человек × мышь, китайский хомячок × мышь

Гены интерферона и гормона роста человека нормально регулировались в геноме клеток мыши.

При слиянии опухолевых клеток китайского хомячка с эмбриональными фибробластами мыши установлен синтез только полного набора полипептидов хомячка, а у мыши, при наличии у гибрида всех ее хромосом, 1/3 ее полипептидов не образуется (Bravo et al., 1982).

## 3. Гены животных и человека в клетках дрожжей (*S. cerevisiae*)

В клетках дрожжей эти гены ведут себя по-разному в зависимости от наличия/отсутствия у них подходящих для реципиента промоторов; от того, построен ли ген из экзонов и интронов, т. к. сплайсинг дрожжей видоспецифичен. Некоторые гены полноценно функционируют в дрожжах.

## 4. Перенос чужеродных генов в клетки зачаткового пути

Могут ли чужие гены интегрироваться в геном половых клеток для передачи их в чреде поколений?

Для этих целей перспективен метод клонирования генов в составе вирусных векторов SV-40 и фага лямбда. Ген  $\beta$ -глобина кролика в составе фага лямбда инъецировали в пронуклеусы яйцеклеток мыши → вводили их в яйцевод ложнобеременных самок → глобиновые последовательности в геноме клеток печени у потомков → самцы передавали ген потомкам.

Данные о переносе генов между видами немногочисленны, тем не менее, можно заключить: *гены, перенесенные между отдаленными видами эукариот, могут успешно функционировать.*

### Другие примеры переноса генов между таксонами

- Из бактерий в клетки животных, от животных в бактерии. В кишечной палочке *E. coli* успешно синтезируют более 30 белков человека и животных — интерферон (его ген не содержит интронов), инсулин, соматостатин и др. Для экспрессии этих генов их кодирующие последовательности присоединяли к участкам с бактериальными промоторами;
- Из бактерий в дрожжи, из дрожжей в бактерии. Показано определенное сходство промоторов в клетках дрожжей и бактерий.

Проблема переноса генов между про- и эукариотами в ходе эволюции в настоящее время поставлена, но не разработана. Наиболее вероятен обмен генами между длительно контактирующими симбионтами и хозяевами:

- обогащение эукариотического генома прокариотическими генами бактерий-симбионтов в ходе эволюционного превращения их в митохондрии;

- передача ДНК между органеллами цитоплазмы, как, например у кукурузы ДНК митохондрий содержит участок из 12 т. п. н., гомологичный участку ДНК хлоропластов.

Перенос генов от почвенной бактерии *Agrobacterium tumefaciens* в клетки растений-хозяев происходит постоянно: в клетки растений переходит бактериальная Ti-плазмида (ее сегмент T-ДНК) и вызывает развитие корончатых галлов и опухолей.

*Сыграл ли такой перенос генов заметную роль в эволюции — точно оценить невозможно.*

Яркий пример природного (произошедшего в естественных условиях) горизонтального переноса генов между представителями разных подимперий органического мира — про- и эукариотами был открыт в 1981 г. Дж. Мартином и И. Фридовичем. В семействе сребробрюшковых рыб (*Leiognathidae*) со светящимся брюшком свечение обеспечивает живущая в специальной железе биолюминесцирующая бактерия *Photobacter leiognathi*. Наряду с типичным для прокариот ферментом — металлопротеидом Fe-супероксидимутазой, этот вид имеет, как и рыбы, эукариотный металлопротеид Cu-Zn-супероксидисмутазу. Его детальное сравнение по аминокислотному составу и физико-химическим свойствам у бактерии и костистой рыбы показало сходство друг с другом. Естественно допускается предположение о горизонтальном переносе этого гена от рыб к симбиотической бактерии.

Другие авторы (Bridges, Salin, 1981) допускают возможность обратного переноса прокариотного гена Fe-супероксидисмутазы в геном голосеменных (сем. *Ginkgoaceae*) и двудольных покрытосеменных (сем. *Nimphaceae* и *Cuscutaceae*) растений. Этот горизонтальный перенос произошел у них независимо.

Палеоботаники (Красилов, 1977) допускают параллельное приобретение посредством вирусной трансдукции эвкалиптовой формы листа у растений Австралии из разных семейств. В других странах растения этих семейств имеют листья иной формы.

Из рассмотренного материала вытекают следующие выводы:

1. Главная эволюционная роль МГЭ про- и эукариотов сводится к переносу чужеродных генов между разными, в том числе отдаленными организмами. Благодаря этому генофонды всех организмов объединяются в общий генофонд всего живого мира.
2. Человек поставил два колоссальных по масштабу природных эксперимента с МГЭ: на плазмидах бактерий с их генами резистентности к антибиотикам, а так же на насекомых с помощью инсектицидов. Популяции насекомых в ответ на инсектициды охвачены генетической экспансией быстро распространяющихся генетических элементов, повышающих устойчивость организма.

3. Основное теоретическое открытие генетической инженерии заключается в доказательстве видоспецифичности белков, их способности синтезироваться в новой клеточной среде. Генетический код универсален, поэтому если клетки присоединяют чужие гены к своим промоторам, то они (гены) синтезируют полноценные белки. Непреодолимое препятствие — наличие интронов у эукариот — разрешается с участием ретровирусов и их способностью к обратной транскрипции на матрице мРНК, прошедшей сплайсинг.
4. Общность генофонда всех организмов означает их потенциальную способность приобретать необходимые гены от неродственных видов, что является редким событием.

### 3. Специальные эксперименты по горизонтальному переносу генов (ГПГ)

В последние годы ГПГ установлен у патогенных микроорганизмов — *Salmonella* (возбудитель сальмонеллеза, в том числе брюшного тифа), *Acinetobacter*, *Streptococcus* (гноеродные, гемолитические и фекальные стрептококки). Перенос генов особенно опасен между стрептококками и кишечной палочкой. ГПГ с участием болезнетворных микробов оказывает влияние на возникновение и развитие различных заболеваний. Этот распространенный в природе механизм привлек внимание эпидемиологов для объяснения неожиданных вспышек опасных заболеваний.

Многие микроорганизмы, в том числе *Streptomyces* spp., *Erwinia carotovora*, *Pseudomonas aureofaciens*, и др., способны вырабатывать антибиотики. Многие авторы считают, что важную роль в эволюции генов антибиотиков у стрептомицетов и других бактерий сыграл процесс ГПГ.

Описанным производителям антибиотиков как бы противостоят бактерии, несущие гены устойчивости к антибиотикам, как, например, устойчивость кишечной палочки к канамицину. У ряда болезнетворных бактерий отмечена устойчивость к нескольким видам антибиотиков, причиной которой является ГПГ, при этом бактерии сохранили «свои» и приобрели «чужие» гены устойчивости.

Так, устойчивость к канамицину и ряду других антибиотиков обеспечивает ген *prt II*. Он часто используется у генетически модифицированных растений (ГМР) в качестве *маркера* — гена, не определяющего изучаемый признак, но позволяющего судить о его передаче. Этот ген — маркер широко распространен в природе — найден в бактериях из стоков, навоза, речной воды, почв, кишечного тракта человека и животных, поэтому вполне возможен его перенос от одних бактерий в другие.



Для сельскохозяйственной биотехнологии актуальность представляет вопрос: возможен ли ГПГ из растений в бактерии? Каков вклад этого процесса в общий диапазон ГПГ? Опасен ли он?

Единственным природным механизмом для этого процесса является *трансформация*. Вероятность реализации ГПГ от растений к бактериям зависит от многих событий и факторов, которые должны произойти и совпасть в естественной экосистеме, чтобы ГПГ произошел. К ним относятся:

1. Выход неповрежденной ДНК во внешнюю среду;
2. Ее абсорбция частицами почв для защиты от разрушений ферментами;
3. Наличие компетентных для трансформации видов бактерий;
4. Эффективное поглощение ДНК на поверхности бактериальных клеток;
5. Эффективный перенос ДНК в эти клетки;
6. Интеграция чужеродной ДНК в геном бактерии-реципиента;
7. Экспрессия генов введенной ДНК в клетке-реципиенте.

Перечисленные факторы подвергнуты изучению, особенно клеточные барьеры ГПГ — абсорбция ДНК на поверхности бактериальных клеток. Подвергнуты также анализу данные о взаимных ГПГ между прокариотами и эукариотами в процессе эволюции. Тем не менее в природных условиях до сих пор не наблюдался ГПГ от ГМР к бактериям.

Для изучения возможности ГПГ от растений к бактериям проведены исследования в специально созданных экспериментальных условиях, способствующих осуществлению событий ГПГ. Результаты представлены в таблице (Конов, 2004) (табл. 6).

Таблица 6

## Эксперименты по ГПГ от растений к бактериям

Растения-бактерии	Вероятность ГПГ
<i>S. tuberosum</i> — <i>E. chrysanthemi</i>	Не обнаружен, не выше $2 \cdot 10^{-17}$ (?) *
<i>Brassica</i> spp. — <i>A. niger</i>	Обнаружен для нескольких колоний
<i>N. tabacum</i> — <i>A. tumefaciens</i>	Не обнаружен
<i>S. tuberosum</i> — <i>Acinetobacter</i> sp.	Не обнаружен, не выше $10^{-18}$ (?)
<i>B. vulgaris</i> — <i>Acinetobacter</i> sp.	Не обнаружен, не выше $10^{-16}$ (?)
<i>B. vulgaris</i> — <i>Acinetobacter</i> sp.	Обнаружен в специальном штамме, частота $3 \cdot 10^{-6}$
<i>B. vulgaris</i> — <i>Acinetobacter</i> sp.	Обнаружен в специальном штамме, частота $3 \cdot 10^{-9}$
<i>S. esculentum</i> — <i>R. solanacearum</i>	Не обнаружен
<i>S. tuberosum</i> — <i>Pseudomonas stutzeri</i>	Обнаружен в специальном штамме, частота $3 \cdot 10^{-4}$
<i>S. tuberosum</i> — <i>Acinetobacter</i> sp.	Не обнаружен, не выше $3 \cdot 10^{-7}$

\* Символ (?) означает, что предсказание не подтверждено в эксперименте.

Опытные данные таблицы показывают, что в условиях опыта можно зарегистрировать ГПГ от растения к бактерии с пренебрежимо малой вероятностью. Ранее (Ратнер, 1983) было установлено, что в генетических исследованиях на микроорганизмах разрешающая способность позволяет устанавливать события с вероятностью  $10^{-7}$ – $10^{-11}$ . Это дает основания считать, что ГПГ мог происходить и в природе и играть определенную роль в эволюции. В то же время его вклад в суммарный ГПГ между организмами пренебрежимо мал.

Способностью к естественной трансформации выявлена у 40 видов прокариотов, среди них несколько случаев относятся к кишечной флоре. Анализ данных по возможной трансформации у бактерий желудочно-кишечного тракта показывает, что вероятность переноса гена *nptII*, вызывающего устойчивость к антибиотику, из пищи микробам тракта оценивается с вероятностью того же порядка, как и вероятность ГПГ от растений к бактериям почвы. Конечно, между этими процессами имеются существенные отличия. Например, в процессе приготовления пищи и ее переваривании, молекулы ДНК подвергаются многим разрушающим воздействиям (механическим, термическим, ферментативным), поэтому возможность «перенесенного» гена уцелеть в желудке снижается.

В заключении необходимо подчеркнуть, что обмен генами, являющийся широко распространенным процессом между бактериями, в том числе болезнетворными и генами устойчивости к антибиотикам — этот процесс пока не установлен в направлении от растений к бактериям в природных условиях. Теоретически можно допустить такие события в процессе эволюции, но они происходили реже, чем в мире прокариот.

#### 4. Информационное давление и информационный фактор эволюции

В 1957 г. выдающийся английский генетик Конрад Хэл Уоддингтон, анализируя открытие вирусной трансдукции бактериальных генов от одного вида к другому, предположил о возможности такого переноса генетической информации также и у эукариот. Позднее, отвечая на аргумент о редкой частоте таких событий, палеонтолог С. В. Мейен показал, что для теории не имеет значение частота события, а важно другое — существует ли данное явление в природе или нет. Новый геном может быть синтезирован из уже существующего путем трансдукции в него участка генома другого вида, рода или отряда.

В современной генетической инженерии разработаны многочисленные методы экспериментального преобразования геномов бактерий и эукариот (Щелкунов, 2004). Возникает законный вопрос: создала ли природа аналогичные механизмы в ходе миллиардов лет органической эволюции?

При всей актуальности этой проблемы до недавнего времени в нашей стране и за рубежом биологи-эволюционисты по молчаливому согласию не подвергали обсуждению и оценке установленные эволюционных явлений при участии чужеродных генов. Только в 1996 г. в Калифорнийском университете (США) состоялась международная конференция по ГПГ и проблемам эволюционного параллелизма.

Задолго до этого в нашей стране киевский генетик В. А. Кордюм (1976, 1982) всесторонне проанализировал новейшие данные о переносе генов между организмами в эксперименте и природных условиях. Чужеродная генетическая информация может достигать генома клетки различными путями — число таких каналов достигает двух десятков. В работе впервые была высказана идея о том, что важнейшим фактором эволюции является отсутствие таксономических преград для взаимного обмена генетической информацией между всеми живыми организмами на Земле. Кордюм выдвинул *целостную всеобъемлющую концепцию биологической эволюции* на основе новых фактических данных. Эта концепция включает следующие основные положения:

1. Информационные каналы биосферы, а также различные переносчики экзогенной (чужеродной) генетической информации (в том числе горизонтальный перенос), создали своеобразный «информационный шквал», воздействующий на живое. Многоуровневая система защиты организма при определенных условиях становится проницаемой для экзогенной информации. Напротив, в основе концепции дарвинизма-селектогенеза лежит утверждение о существовании жесткого информационно-видового барьера.
2. Экзогенная информация, проникая в организмы, создает в них постоянно поддерживаемое депо этой информации, состоящее из ДНК многочисленных симбионтов и комменсалов [форма симбиоза]. На этом основано понятие «*информационный фактор эволюции*», включающий всю систему создания новой информации, ее преобразование и обмен между организациями биосферы. Отсюда автором формулируется *теория как «информационная концепция эволюции»*.
3. По аналогии с мутационным давлением, Кордюм вводит представление об «*информационном давлении*» как непрерывный поток генов из многочисленных информационных каналов. На основе этого давления организмы получают ценную информацию, на приобретение которой по традиционной схеме синтетической теории эволюции (СТЭ) ушли бы миллионы лет. Таким образом, виды при всей их генетической замкнутости являются *генетически открытыми системами*, способными к обмену со всем генофондом биосферы. Положение концепции автора перекликается с выводами биохимика Р. Б. Хесина в его сводке «Непостоянство генома» (1984).

4. Информационный обмен генами происходит в планетарном масштабе, поэтому эволюционный процесс это не эволюция видов как суммы информационно замкнутых групп, а эволюция биосферы как единого центра, в котором каждое ее конкретное проявление передает все всем и черпает все от всех. Выбор единицы эволюции в условиях тотального генетического обмена зависит от уровня рассмотрения (например, ценоз, вид, популяция).
5. Привносимая извне информация у низших и высших организмов проходит доработку на соответствие новому молекулярному окружению и внешним условиям. У прокариот и гаплоидных низших эукариот с их генетическим аппаратом без повторов и дупликаций генов экзогенная информация наследуется и проявляется фенотипически. У высших эукариот поступившая информация какое-то время хранится в молчащем состоянии в неэкспрессивной части генома, без фенотипических изменений. Новая информация должна пройти период «включения».
6. Управление экспрессией молчащего генетического материала экзогенного происхождения может осуществляться с участием транспозонов, «бродячих» промоторов и нуклеотидных последовательностей самой экзогенной информации. Новые признаки могут появляться внезапно и в массовом масштабе. Новые таксоны возникают сразу не как единичные особи, а в виде готовой популяции. Изменчивость приобретает эпидемический характер: *«вирусная эпидемия»* влечет за собой *«генетическую эпидемию»* (Голубовский, 1977).
7. В свете новых представлений об изменчивости получает объяснение феномен молчащей ДНК, представляющий неразрешимую проблему в рамках СТЭ. Ее размер у некоторых таксонов высших эукариот доходит до более 90 %. Для вида это безусловное расточительство, так как на ее репликацию затрачивается много энергии. Ее наличие необходимо для будущего, когда при возможных геологических катастрофах молчащая информация будет востребована, а на ее основе возникнет новая спасительная генетическая информация. С другой стороны, низшие организмы освободились от избыточной ДНК и законсервировались в своем развитии путем отбора. В случае кризиса им грозит вымирание.
8. Вопрос о причинах прогрессивного эволюционного развития, кстати, не разрешимый для СТЭ, объясняется тем, что эволюция идет по вынужденному пути усложнения организации за счет поступления большого массива экзогенной генетической информации. Многие авторитетные специалисты (Бердников, 1981, 1990; Голубовский, 2000 и др.) считают, что усложнение организации детерминируется автогенетическими свойствами самого генетического материала наращивать длину ДНК и увеличивать размер генома.

9. Кордьюм видит в естественном отборе всего лишь один и достаточно второстепенный механизм эволюции среди множества других. Его эволюционным взглядам созвучны номогенетические построения Берга и Вавилова, которым отведена треть книги. В основе общности всего живого лежит универсальность переносчиков информации — нуклеиновых кислот: информация может распространяться по всей биоте. Новый виток эволюции начинается с групп особей, поэтому *полифилия — правило, а монофилия — исключение*. Подобно современным номогенетикам (Мейен, 1984, 1986), Кордьюм считает, что новая генетическая информация создается во влажных тропических лесах. Таким образом, *информационная концепция эволюции с механизмом ГПГ и номогенез взаимно дополняют друг друга*.
10. Формирование нового вида не может быть воспроизведено в эксперименте в связи с медленностью протекания эволюционных преобразований, несоизмеримой с продолжительностью человеческой жизни (точка зрения эволюционистов разных направлений). Но опыты по генетической инженерии, установление существования аналогичного механизма в природе, доказывают принципиальную возможность экспериментов по моделированию начального этапа процесса эволюции, который связан с экспрессией чужеродного генетического материала. «Экспериментальная эволюция» не требует миллионов лет. Теперь уже во многих лабораториях мира постоянно проводятся успешные эксперименты по ее осуществлению: для этого необходим набор генов, методика введения их в геном и соответствующая технология. На *Международной конференции по ГПГ* в Калифорнийском университете в 1996 г. обсуждались доказательства переноса генетической информации разными векторами между про- и эукариотами, вопросы о природе и функциях подвижных элементов генома эукариот, их происхождение. Не была оставлена в стороне также ведущая проблема номогенеза — об общих тенденциях эволюционного параллелизма.

## 5. Латеральная геномика: изучение роли горизонтального (латерального) переноса генов в видообразовании

Исследования в области геномики в последние 15 лет показали, что горизонтальный (латеральный) перенос генов (ГПГ) является одним из главных механизмов видообразования, особенно в мире прокариот. Это позволило Ф. Дулитлу (Doolittle, 1998, 1999) обосновать данное направление как «*латеральную геномику*». Ее основные достижения следующие:

- 1) доказательство феномена ГПГ между близкими и филогенетически отдаленными организмами на разных этапах эволюции;

- 2) создание методов и критериев выявления ГПГ;
- 3) установление путей переноса генов между представителями разных таксонов;
- 4) показаны последствия и выгоды для организмов, приобретающих чужеродные гены путем горизонтального переноса;
- 5) проведен полногеномный анализ большого числа видов бактерий и архей, что позволило обобщить некоторые закономерности ГПГ;
- 6) показаны возможности и пути определения организмов — доноров в ГПГ;
- 7) установление фактов ГПГ от эукариот в клетки прокариот;
- 8) доказательство горизонтального переноса фрагментов генов, содержащих отдельные домены (Попов, 2007а, в).

Горизонтальный перенос генов, наряду с вертикальной эволюцией «вверх» и редуccionной эволюцией (путь «вниз») (Шестаков, 2003а, в), является главным источником обновления, инструментом быстрого приобретения и возникновения новых генов, способных радикально изменить свойства клеток. В ходе вертикальной эволюции наследуется *базовый набор предковых ортологичных генов*, сходных у разных организмов. В результате дупликаций, мутаций и рекомбинаций из них образуются *паралогичные гены*, которые увеличивают набор белков и диапазон фенотипических вариаций. В процессе редуccionной эволюции происходит сокращение числа генов, утрата функций, путей метаболизма, органелл и т. д. Для большинства патогенных бактерий характерна утрата многих генов, так как эти облигатные паразиты используют метаболические системы клетки «хозяина». В результате ГПГ изменчивость организмов реализуется через различные *каналы генетической коммуникации*, к которым относятся процессы переноса генов в составе векторов (плазмид, вирусов, мобильных элементов). Сейчас стало ясно, что на принципах ГПГ базируется *современная генетическая инженерия*, использующая различные векторы; что *генная инженерия широко распространена в природе* и играет важную роль в эволюции. Будучи наиболее высокой на ранних этапах становления биосферы, частота горизонтальных переносов снижалась в ходе эволюции эукариот и развития барьеров, препятствующих ГПГ и «размыванию генов». Среди этих барьеров были ограничение контактов между организмами, действие систем рестрикции, разрушающих чужеродную ДНК, механизмов репарации, обеспечивающих стабильность геномов.

Выявление показателей ГПГ в латеральной геномике проводят по следующим критериям. *Во-первых*, используется такой видоспецифический признак как нуклеотидный состав ДНК (GC — содержание в мол. %). Указанием на присутствие чужеродных генов является отличие по нуклеотидному составу отдельного сегмента, иногда целых кластеров, содержащих профаги, мобильные элементы, «островки» патогенности от остальной

части генома. Вторым удобным показателем явилась частота встречаемости в гене определенных кодонов, так как для генов каждого вида характерен их ограниченный набор. Показано, что у кишечной палочки в гене *gar* содержится 20 кодонов аминокислоты лейцина, при этом из 6 триплетов-синонимов кодон CTG повторяется 19 раз, а кодон TTA только один раз. Для сравнения у *Bacillus subtilis* в гомологичном гене встречается преимущественно кодон TTA, а кодон CTG вообще не используется. В ходе эволюции за счет мутаций и рекомбинации происходит *амелиорация нуклеотидного состава*, т. е. процесс унификации использования кодонов: чужеродные гены становятся неотличимыми от своих собственных. Третьим критерием чужеродности гена может служить существенное отличие этого гена по положению на филогенетическом дереве от большинства других генов. Например, когда типично эукариотический ген обнаруживается у одного вида бактерий при его отсутствии у других бактерий.

Среди путей переноса генов между представителями разных таксонов *инновационную ценность* имеют три основных типа переносов (Koonin et al., 2001):

- *1 путь*: возникновение нового качества за счет привнесения нового гена, для которого нет гомолога в геноме реципиента и в геномах филогенетически родственных видов.
- *2 путь*: привнесение *паралогичного (структурно похожего) гена* от генетически отдаленного донора. Это ведет к увеличению функционального спектра белков в клетке.
- *3 путь*: «*приживание*» гена *ксенолога*, функционально замещающего свой собственный ген. Ксенолог и старый гены имеют структурные различия между собой, но обеспечивают аналогичные физиологические функции.

Приобретение чужеродных генов путем горизонтального переноса ведет к следующим выгодам и последствиям для реципиентов:

1. *Возникновение нового пути биосинтеза или катаболизма*, что обеспечивает реципиенту появление способности утилизировать новый субстрат.
2. *Повышение резистентности к антибиотикам, токсинам, патогенам*, подавляющим рост клеток данного вида. Патогенные микроорганизмы таким путем могут получить гены, ответственные за средства «нападения».
3. При горизонтальном переносе могут быть получены гены, продукты которых повышают эффективность функционирования клеточных систем — повышение термоустойчивости, резистентности к ингибиторам и т. д.

4. Приобретенные гены могут дублировать функции уже имеющихся генов и быть своего рода страховкой для реципиента в случаях повреждения собственного гена мутацией.
5. *Велико значение ГПГ для эволюции видов.* Приобретение чужеродных генов может изменить направление эволюции вида, существенно повлиять на фенотип организма и его способность к адаптации. ГПГ способствует ускорению эволюционного процесса по сравнению с градуальным накоплением мутаций или внутригеномными перестройками. Особо важную эволюционную роль играют мутации в генах, контролирующих стабильность генома (системы репликации, репарации, модификации ДНК).

В табл. 7 приведены данные о ГПГ у бактерий и архей, полученные при полногеномном анализе большого числа видов. Они дают возможность обобщить некоторые закономерности горизонтального (латерального) переноса генов.

Таблица 7

## Горизонтальный перенос генов у архей и бактерий

Вид	Число генов в геноме	Перенесенные гены	
		количество	% в геноме
<b>Архен</b>			
<i>Archaeoglobus fulgidus</i>	2407	179	8,4
<i>Methanococcus jannaschii</i>	1715	77	5,0
<i>Pyrococcus horikoshii</i>	2064	154	7,6
<i>Aeropyrum pernix</i>	2694	370	14,0
<b>Патогенные бактерии</b>			
<i>Mycoplasma pneumoniae</i>	677	39	5,9
<i>Chlamydia trachomatis</i>	894	36	4,3
<i>Rickettsia prowazekii</i>	834	28	3,6
<i>Treponema pallidum</i>	1031	77	8,3
<i>Haemophilus influenzae</i>	1709	96	6,2
<i>Helicobacter pylori</i>	1553	89	6,4
<i>Mycobacterium tuberculosis</i>	3918	187	5,0
<b>Свободноживущие бактерии</b>			
<i>Aquifex aeolicus</i>	1552	72	4,8
<i>Hiemotoga maritima</i>	1846	198	11,6
<i>Escherichia coli</i>	4289	381	9,6
<i>Pseudomonas aeruginosa</i>	4036	411	10,1
<i>Bacillus subtilis</i>	4110	537	14,8
<i>Synechocystis</i> sp.	3169	219*	7,5

Из обзора: Koonin E. V., Makarova K. S., Arvind L. 2001. Annual Rev. Microbiol. V. 55: 709–42



1. У разных видов доля генов, полученных при ГП, широко варьирует и может достигать 10–15 % от общего числа генов в геноме вида.
2. Наибольшее число переносов характерно для свободноживущих бактерий с широкими экологическими ареалами (почвенные бациллы, псевдомонады и др.).
3. Наименьшее число переносов обнаружено у большинства патогенных бактерий, обитающих в узких эконишах. Однако эти переносы имеют важное значение, так как определяют признаки вирулентности и токсичности.
4. Переносы специфичны, поскольку приобретенный ген обнаруживается, как правило, только в клетках определенного вида и даже штамма.
5. Наиболее редко в горизонтальные переносы вовлекаются *гены информационных систем* (репликации, транскрипции, трансляции), составляющие *базовый геном*. Продукты этих генов входят в состав сложных белковых комплексов, куда чужеродные белки не встраиваются или не функционируют.
6. Чаще всего в горизонтальных переносах участвуют *гены операционных систем*, обслуживающих метаболизм, транспортные пути, механизмы сигнальной трансдукции. У многих бактерий латерально перенесенные гены содержат функционально неизученные или гены, не имеющие сходных ортологов у реципиентов.
7. В составе приобретенных сегментов ДНК часто обнаруживаются профаги, плазмиды, участки резистентности и др.

Сравнительная геномика дает информацию, в основном, о степени сходства гомологичных генов в разных геномах, поэтому в большинстве случаев трудно определить, какие конкретно организмы могли быть донорами в горизонтальном переносе. Косвенно о направлении такого переноса можно судить в случаях, когда, например, у многих видов архей какие-то гены встречаются часто, но обнаруживаются в геноме только одного или ограниченного числа видов бактерий. Таким примером может служить возможный перенос генов из термофильных архей в клетки предшественников современных видов термофильных бактерий *Thermotoga maritima* и *Aquifex aeolicus*. У этих бактерий около 15 % генома составляют типично архейные гены, которых нет у других бактерий. ГПГ легко выявляется по критериям гомологии на геномном уровне у эукариот, в ядерном геноме которых есть гены бактериального или архейного происхождения. Это явилось результатом перемещения в ядро генов из митохондрий или хлоропластов, возникших на основе прокариотических геномов, претерпевших редуцированную эволюцию. Обсуждается также другая возможная причина мозаичного строения ядерных геномов эукариот — за счет интеграции чужеродных фрагментов ДНК, поступающих с пищей при поедании бактерий протистами.

Геномный анализ показал случаи ГПГ от эукариот в клетки прокариот в условиях тесных физических контактов в симбиотических или паразитических системах, при наличии челночных векторных переносчиков. Имеется предположение, что донором эукариотического гена термоустойчивой аминоксил-тРНК-синтетазы у архей *Rugosoccus* могли быть полихеты. В геномах патогенных бактерий риккетсии и хламидии обнаружено 20 эукариотических генов, среди которых имеются несвойственные прокариотам гены белков, транспортирующих АТФ и АДФ. Эти транспортные системы позволяют патогенным бактериям «выкачивать» энергию из клетки хозяина. Характерно, что у риккетсии (возбудителя тифа) имеются гены, гомологичные генам животных, а в геноме хламидии (возбудителя трахомы) больше генов, типичных для растений. Это указывает на то, что предки хламидии сперва были паразитами растений, а затем переключились на животных. Возможно, переносы генов патогенности шли между риккетсиями и хламидиями, живущими в одной экологической нише в организме млекопитающих.

По современным представлениям, гены являются сложными структурами, содержащими различные участки, ответственное за различные функции в белковом продукте. ГПГ обеспечивает передачу как целых генов и их блоков, так и фрагментов, содержащих участки — «домены». Такие домены в процессе рекомбинации могут становиться сегментами других генов или принимать участие в образовании других генов. Методы геномного анализа позволяют определять наличие таких доменов в необычных сочетаниях. Встраивание нового домена может менять характеристики, влияющие на физиологическую функцию гена — локализацию белкового продукта в клетке, узнавание и природу сигналов и многие другие. Выявлены более 40 случаев появления «чужих» доменов, не подвергшихся амелиорации в составе генов, кодирующих различные белки. При этом происходит рекомбинационное слияние домена из латерально привнесенного гена с резидентным геном, а партнерами могут быть филогенетически отдаленные организмы. Важность выяснения механизмов перекомбинации доменов и возникновения новых генов в результате ГПГ обусловлена тем, что она связана с задачами генной и белковой инженерии по искусственному созданию белков с новыми свойствами.

Ф. Дулитл и другие (Doolittle, 1998, 1999; Jain et al., 1999; Eisen, 2000; Woese, 2000; Koonin et al., 2001) обосновали последовательность латерально-геномных событий в ходе эволюционного процесса. На ранних этапах эволюции в неустоявшихся предковых клетках существовало общее генное «коммунальное хозяйство», а между клетками происходил активный горизонтальный обмен генами. В ходе последующей автономизации клеток появлялись отдельные таксономические линии. Причиной мозаичности геномов прокариот и эукариот явилось то, что эволюционные связи

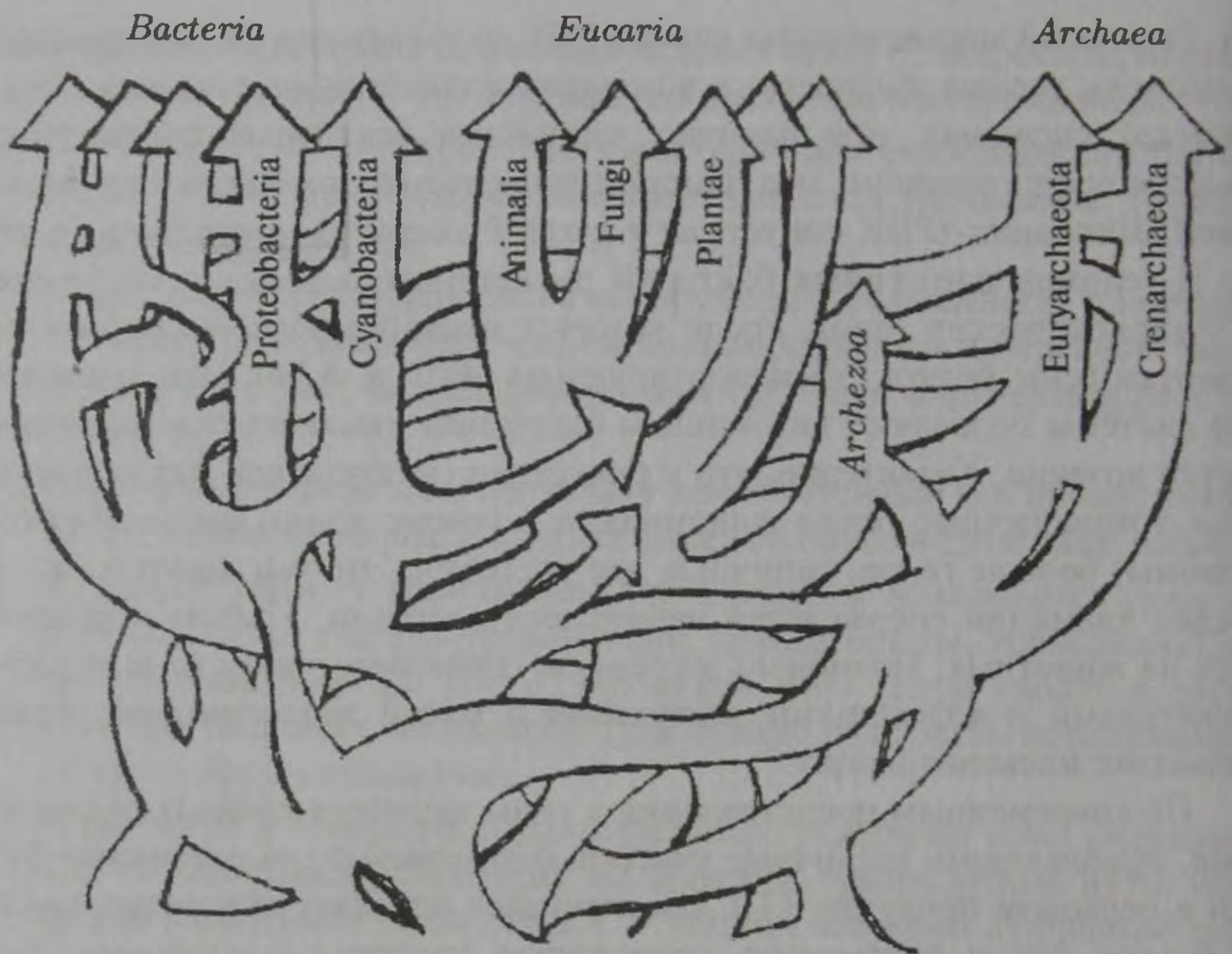


Рис. 49. Мицелий с сетью разнонаправленных горизонтальных переносов у предковых прокариот

предковых прокариот напоминали не ветвящееся дерево, а своего рода мицелий с переплетенной сетью горизонтальных переносов разных направлений (рис. 49). В процессе эволюции массивные генные переносы осуществлялись в пределах мегатаксонов (доменов, надцарств) и между ними, а многочисленные прокариоты в этих процессах играли роль «*проточных емкостей*». Скорости генных потоков, влияющих на темпы эволюции геномов, значительно различаются для разных групп организмов. У модельных видов прокариот и протистов оказалось возможным методами геномного анализа проводить сопоставление времени приобретения и утраты определенных генов с геологическими эпохами и экологическими кризисами биосферы (рис. 49).

### Ключевые слова и термины

Вертикальный перенос генов, горизонтальный перенос генов, плазмиды, трансдукция, трансформация, конъюгация, генная терапия, ретровирусы, вирулентный штамм, авирулентный штамм, мобильные генетические элементы, транспозоны, ретротранспозоны, общность генофонда живых организмов, информационное давление, информационный фактор эволюции.

## Часть вторая

---

---

# Геномика

---

---

Специфический раздел молекулярной генетики — геномика — сегодня занимается анализом структуры и функций геномов как интегрального функционального массива генов, их регуляторных элементов и других последовательностей, необходимых для функционирования генома. Начавшись с исследований генома человека, геномика значительно расширила диапазон своих интересов и включила в них множество модельных организмов — бактерии и дрожжи, нематоду, дрозофилу и мышь, геномы которых исследуются и сравниваются между собой для расшифровки структурных основ их функциональной организации.

Академик *Е. Д. Свердлов*

## Становление геномики как самостоятельного раздела молекулярной генетики

Геномика — раздел генетики, который изучает геномы и отдельные гены на молекулярном уровне, их структуру (*структурная геномика*) и функции (*функциональная геномика*), а также их использование в генной инженерии, биотехнологии и генной терапии (*медицинская геномика, или геномная медицина*). В задачи геномики входит:

- полное секвенирование геномов, сопоставление и оценка их генетической сложности как молекулярных систем;
- выявление ранее неизвестных генов, сравнительный анализ функционального и структурного сходства различных генов и геномов;
- выявление общих принципов организации сложных клеточных молекулярно-генетических систем управления.

По образному выражению новосибирского генетика В. А. Ратнера, геномика является одной из главных точек роста современной молекулярной генетики.

В чреде клеточных поколений происходит передача *элементов генома* как дискретных фрагментов ДНК, различающихся по составу нуклеотидов и по функциональным признакам. Их можно подразделить на следующие категории:

- *облигатные элементы* — структурные локусы, количество и расположение которых в геноме постоянны. Среди них по характеру экспрессии могут быть выделены гены «домашнего хозяйства» (*housekeeping genes*) — их продукты необходимы для обеспечения жизнедеятельности клеток всех типов. Тканеспецифические гены как облигатные элементы обеспечивают специализированные функции в определенных типах клеток и тканей на определенных стадиях онтогенеза;

- *факультативные элементы* — к ним относятся вирусы, В-хромосомы, амплифицированные копии ДНК;
- *мобильные (мигрирующие) генетические элементы* включают транспозоны и ретротранспозоны (рис. 47). Транспозоны перемещаются в геноме с участием комплекса белков, обеспечивающих активность фермента *транспозазы*, которая узнает подвижный элемент и обеспечивает его перенос на новое место. Они составляют 3 % генома человека и представлены в нем 300 000 копиями 7 разных классов. Другой класс подвижных элементов — это так называемые *ретротранспозоны*. В отличие от транспозонов, они не вырезаются из хромосомы. Их перемещение основано на транскрипции ретротранспозона (синтезе РНК на матрице ДНК), за которой следует «обратная транскрипция» — синтез нити ДНК (к ДНК) на РНК. Обозначения разных мобильных элементов:

1. LINE (long interspersed elements) — семейство длинных диспергированных повторов генома человека длиной каждого 6 т. п. н.; число копий  $8,5 \cdot 10^5$  на геном.
2. SINE (short interspersed elements, в том числе Alu) — семейство коротких диспергированных повторов, которые занимают в тексте ДНК генома человека в 10 раз больше места, чем белковые кодирующие последовательности. Длина каждого 100–400 п. н.; число копий  $1,5 \cdot 10^6$  на геном.
3. LTR ретротранспозоны: размер 1,5–11 т. п. н.; число копий  $4,5 \cdot 10^5$  на геном.

Клеточный геном представляет собой сбалансированную систему генов, своего рода архив генетической информации, которая необходима и достаточна для контроля всего клеточного метаболизма, самовоспроизведения, развития и морфогенеза. В него входят гены всех основных генетических процессов — конвариантной редупликации, транскрипции, трансляции, репарации, рекомбинации и сегрегации.

На пути становления геномики стояли следующие знаковые открытия:

**1966** — М. Геллер, Б. Вейс и С. Рихардсон открыли фермент ДНК-лигазу.

**1970** — Г. Темин и Д. Балтимор независимо открыли обратную транскриптазу — фермент, синтезирующий ДНК с использованием в качестве матрицы комплементарной РНК. Лауреаты Нобелевской премии 1975 г. совместно с Р. Дальбекко.

В. Арбер, Д. Натанс и Х. Смит выделили рестриктазу — фермент, разрезающий ДНК в строго определенных местах (сайтах). Лауреаты Нобелевской премии 1978 г.

1972–1973 — Пол Берг и Г. Бойр получили рекомбинантные ДНК бактериального вируса лямбда и обезьяньего вируса SV-40. Проведен первый генноинженерный эксперимент. Нобелевская премия по химии за 1980 г.

1976 — В. А. Гвоздев, Г. П. Георгиев (Россия) и Д. Хогнесс (США) открыли мобильные генетические элементы у дрозофилы.

1975–1977 — Ф. Сэнджер и независимо А. Максам и У. Гилберт разработали методы быстрого определения нуклеотидных последовательностей ДНК (*прямое секвенирование*)

Ф. Сэнджер с сотр. (Кембридж) секвенировали геном бактериофага «фи-десять-174» с кольцевой одноцепочечной ДНК размером 5386 п. н. В геноме фага установлены перекрывающиеся гены и «гены внутри генов». Вместе с У. Гилбертом и П. Бергом удостоен Нобелевской премии 1980 г. Для Ф. Сэнджера это была вторая Нобелевская премия (первая — за 1958 г. по химии).

1977 год знаменовал начало новой эры в молекулярной генетике — эры геномики.

Последующий период характеризовался увеличением числа объектов секвенирования геномов и повышение эффективности методов. Среди тех и других отметим наиболее важные:

- расшифровка последовательности ДНК митохондриального генома человека в лаборатории молекулярной генетики Кембриджского университета в 1981 (Сэнджер Ф., Андерсон С. и др.);
- К. Муллис и Р. Сайки разработали *метод полимеразной цепной реакции (ПЦР)*, с помощью которого можно быстро получать миллионы копий участков ДНК. Этот метод ускорил реализацию программы «Геном человека». Удостоены Нобелевской премии 1993 г.;
- основан Международный консорциум по секвенированию генома человека (IHGSC) (1990);
- фирмой Applied Biosystems (США) разработан первый автоматический ДНК-секвенатор, модель которого подвергается ежегодному усовершенствованию. В результате секвенирование ДНК становится простой процедурой, доступной в любой молекулярно-биологической лаборатории;
- секвенированы геномы первого высшего организма — нематоды *Caenorhabditis elegans* размером  $8 \cdot 10^7$  п. н. (1998), высшего растения арабидопсиса размером  $8 \cdot 10^7$  п. н. (1999), дрозофилы размером  $1,75 \cdot 10^8$  (2000);
- полное секвенирование генома человека Международным консорциумом (IHGSC) и частной фирмой Celera Genomics (2001). Размер генома человека составил  $3,2 \cdot 10^9$  п. н.;
- полностью секвенирован геном мыши размером  $2,2 \cdot 10^9$  п. н. (2002) и геном неандертальца (2006).

В настоящее время секвенированы геномы у более чем 1 тыс. видов организмов. Особое внимание уделяется расшифровке структуры геномов патогенных бактерий и вирусов. Нарастающее число секвенированных нуклеотидных последовательностей у разных организмов создало проблему упорядочения их хранения и создания баз данных таких последовательностей. Достижения молекулярной генетики и геномики были бы невозможны без создания технических условий для хранения и обработки огромных массивов информации. Первой электронной базой данных о секвенированных последовательностях ДНК была Los Alamos DNA Database (1980). Затем были организованы другие банки данных генетических последовательностей с адресами в Интернете:

- 1) Gen Bank, США (1982);
- 2) EMBL Data Library, Европа (1980);
- 3) EMBL Nucleotide Sequence Database, Европа (1994);
- 4) DNA Data Bank of Japan, Япония (1986);
- 5) International Nucleotide Sequence Database;
- 6) SWISS-PROT — крупнейшая база данных аминокислотных последовательностей (1986);
- 7) Система WIT (What is There) разработана компанией Интергрейт геномикс инкорпорейшн (Чикаго, США) для сравнительного анализа нуклеотидных последовательностей геномов и проведения «реконструкций» метаболизма по хромосомным последовательностям (2000);
- 8) HUMBIO — база знаний по биологии человека (1995), созданная в Институте молекулярной генетики РАН;
- 9) Компьютерная система GeneExpress включает большой перечень программ и баз данных, которые могут быть использованы при создании трансгенных растений. Разработана в Институте цитологии и генетики СО РАН.

В научной литературе по геномике встречаются различные термины (*биоинформатика*, *биология in Silico*, *вычислительная биология*), поэтому возникла необходимость разграничения между этими дисциплинами (Александров А. А., Дроздов-Тихомиров Л. Н., Шепелев В. А., 2004; Дромашко С. Е., 2006).

*Биоинформатика* — быстро развивающийся раздел теории информации, который представляет собой науку о создании банков данных, разработке удобного компьютерного интерфейса, графиков, а также программно-математических методов для анализа последовательностей и пространственных структур. Целью биоинформатики является накопление биологических знаний в форме, обеспечивающей их наиболее эффективное использование, а также построение и анализ математических моделей биологических систем и их элементов. В биоинформатике существует специ-



альный раздел — *компьютерная геномика*, решающая проблемы расшифровки генетических «текстов», хранящихся в последовательностях нуклеотидов ДНК (РНК). *Метаболомика (протеомика)* — другой раздел биоинформатики, исследующий принципиальную сторону организации метаболизма клетки и его управления со стороны генома. Ее задачами является выявление и моделирование определенной динамической структуры метаболизма, обеспечивающей поддержание гомеостаза в клетке за счет регуляторных свойств уже существующих в клетке ферментов и функционирования генома, поддерживающего существование этой структуры.

Задачами *биологии in Silico (или вычислительной биологии)* является анализ динамики систем макромолекул, молекулярно-генетических систем и процессов, создание математических моделей функционирования клеток и целых организмов, т. е. объединение наработок геномики и протеомики. С 1998 г. в Германии выходит международный научный журнал по вычислительной молекулярной биологии «*In Silico Biology*»: как электронное Интернет-издание его публикует Bioinformation Systems E. V. (Брауншвейг), а печатную версию готовит издательство IOS Press.

В структурной геномике широко используется теоретико-информационный подход для анализа генетических текстов. Сами генетические тексты — а это уже никем не подвергается сомнению — показывают сходство с компьютерными программами, например, наличие знаков начала (промотор) и конца (терминатор транскрипции) программы, выраженных «программ», меток циклов. Эти проблемы интенсивно разрабатываются на кафедре информационной биологии Новосибирского государственного университета, где читаются спецкурсы «Компьютерная геномика». В качестве примера приведем спецкурс «*Обзор задач биоинформатики, связанных с анализом и обработкой текстовых последовательностей*» (Ю. Л. Орлов). В нем сформулированы основные задачи компьютерного анализа генетических текстов:

- 1) поиск гомологии и выравнивание генетических текстов, множественное выравнивание;
- 2) статистический анализ генетических текстов, исследование структуры повторов и модели порождения символьных последовательностей, сегментация геномов;
- 3) предсказание кодирующих участков генов и *открытых рамок считывания* (ОРС — англ. *open reading frame = ORF*);
- 4) предсказание функциональных сигналов (функциональных сайтов и регуляторных районов);
- 5) анализ вторичной структуры РНК и сигналов трансляции;

---

\* Открытая рамка считывания — участок ДНК между иницирующим и стоп-кодонами.

б) анализ аминокислотных последовательностей белков, предсказание вторичной структуры, функциональных сайтов и доменов глобулярных белков по их аминокислотным последовательностям.

После рассмотрения общих вопросов геномики, ее истории, проблематики и терминологии в последующих разделах мы перейдем к конкретному описанию геномов доклеточных (вирусов и фагов), прокариот (бактерий) и эукариот.

### Ключевые слова и термины

Элементы генома (облигатные факультативные, мобильные), транспозоны и ретротранспозоны; диспергированные повторы LINE, Alu, SINE; метод полимеразной цепной реакции; автоматический ДНК-секвенатор; база данных, банки данных; биоинформатика, метаболономика, биология *in Silico*.

## Геномика вирусов и фагов. Вирусы как объект молекулярной генетики

По определению Андре Львова, вирус представляет собой инфекционное, потенциально патогенное начало, содержащее один тип нуклеиновой кислоты, репродуцирующееся за счет собственного генетического материала, неспособное к росту и делению и лишенное ферментов энергетического обмена (Lwoff A., 1957). В истории развития зоологических и ботанических наук сформировались определенные критерии, которым должны удовлетворять живые организмы. С этих позиций *вирусы нельзя считать живыми организмами*. Тем не менее вирусы обладают их главными свойствами — *способностью к репликации, мутационным изменениям и передаче этих изменений потомкам*.

По современным представлениям, вирусы являются неклеточными формами жизни, которые способны проникать в определенные живые клетки и в них размножаться. Генетический аппарат вирусов несет информацию о синтезе вирусных частиц, при этом используются биохимические предшественники, а также биосинтетические и энергетические системы клетки-хозяина. Вне клетки вирус являет собой совокупность метаболически инертных химических соединений, тогда как внутри клетки-хозяина это своеобразный живой организм, *внутриклеточный паразит на генетическом уровне*.

Открыты вирусы в 1892 г. Д. И. Ивановским при фильтрации раствора с возбудителем болезни табака через бактериальный фильтр. Термин «вирус» введен в 1898 г. М. Бейеринком. Первым наблюдал и подсчитал частицы бактериофагов в 1933 г. Макс Шлезингер в темнопольном микроскопе. Первые снимки бактериофагов сделал Гельмут Руска в 1941 г. с помощью электронного микроскопа. Этот прибор увеличивает объект в 300 000 раз, поэтому с его помощью излучение вирусов продвинулось впе-

ред. Их размер колеблется в пределах от 25 до 500 нм, т. е. вирусы являются ультрамикроскопическими организмами, а по своим размерам они стоят ближе к молекулам и атомам.

Таблица 8

Классификация вирусов — возбудителей инфекций человека  
(Зинченко, Паруль, 2005)

Семейство	Тип нуклеиновой кислоты	Важнейшие представители
<i>Poxviridae</i>	ДНК	Вирус натуральной оспы, вирус осповакцины
<i>Herpesviridae</i>	ДНК	Вирусы простого герпеса типа 1 и 2, вирус ветряной оспы и опоясывающего лишая, цитомегало-вирус, вирус Эпштейна—Барр
<i>Adenoviridae</i>	ДНК	Аденовирус человека, аденовирусы млекопитающих
<i>Parvoviridae</i>	ДНК	Латентный вирус крыс Килхема, аденовирусные сателлиты
<i>Papovaviridae</i>	ДНК	Вирус папилломы Шоупа, вирус полиомы, вакуолизирующий вирус SV-40
<i>Hepadnaviridae</i>	ДНК	Вирус гепатита В
<i>Orthomyxoviridae</i>	РНК	Вирусы гриппа А, В и С
<i>Paramyxoviridae</i>	РНК	Вирусы парагриппа 1, 2, 3 и 4, вирус кори, респираторно-синцитиальный вирус
<i>Retroviridae</i>	РНК	Вирус иммунодефицита человека, вирус саркомы Рауса
<i>Bunyaviridae</i>	РНК	Вирус Буньямвера, вирус Укуниеми
<i>Togaviridae</i>	РНК	Вирус Синдбис, вирус желтой лихорадки, вирусы клещевого энцефалита, вирус краснухи
<i>Coronaviridae</i>	РНК	Коронавирус человека, вирус бронхита птиц
<i>Reoviridae</i>	РНК	Реовирус человека, реовирус позвоночных
<i>Picornaviridae</i>	РНК	Вирус полиомиелита человека, вирус гепатита А, вирус ящура
<i>Arenaviridae</i>	РНК	Вирус лимфоцитарного хориоменингита
<i>Rhabdoviridae</i>	РНК	Вирус бешенства, вирус везикулярного стоматита

Вирусы поражают все группы живых организмов. Описано более 500 вирусов теплокровных животных, более 300 вирусов высших растений, идентифицировано свыше 30 тыс. разновидностей (штаммов, типов и т. п.). В настоящее время нет достаточного количества данных для построения эволюционной классификации вирусов. Базовыми таксонами у них являются семейства, выделение которых основано на наиболее определенных свойствах вирусов — размерах, морфологии, наличии оболочек, важных для практической диагностики вирусных инфекций. Охарактеризовано более 70 семейств, среди которых 17 включают вирусы животных и человека (табл. 8).

Вирусы как объекты исследования широко используются в молекулярной биологии и генетике, в геномной инженерии, при изучении канцерогенеза. Эксперименты на бактериофагах — вирусах, поражающих бактерии (обычно используются более краткий термин *фаги*) — и их бактериях-хозяевах привели к открытию *репарации ДНК и индукции репаративных ферментов*. Многие авторы допускают, что именно исследование фагов послужило главным стержнем, от которого возникли молекулярная биология и генетика.

## Основные свойства вирусов

### 1. Вирус и клетка-хозяин

Выбор клетки-хозяина у вирусов ограничен. Так, вирусы растений и рыб не способны поражать клетки млекопитающих, и наоборот. Вирус полиомиелита в естественных условиях поражает лишь высших приматов. Вирус бешенства способен поражать любого представителя млекопитающих, но не инфицирует растения. Возможности размножения фагов строго ограничена одним подтипом данного вида бактерий (фаготип).

### 2. Размеры, молекулярная масса ДНК/РНК

Размеры вирусов 25–500 нм, большинство видны только в электронном микроскопе. Диаметры крупных вирусов оспы и осповакцины приближаются к 250 нм. Диаметр головки фагов составляет 50 нм, а отростка 100–200 нм. Масса вирусной ДНК  $1 \cdot 10^6$ – $200 \cdot 10^6$ , вирусной РНК от  $10^6$  до  $15 \cdot 10^6$  Да.

### 3. Формы существования. Строение

Установлены две формы существования вирусов: *вирионы* — покоящиеся внеклеточные формы, и внутриклеточные репродуцирующие формы как комплекс «вирус-клетка». Простые вирусы состоят из нуклеиновой кислоты и белковой оболочки — *капсида*. Используется также термин *нуклеокапсид* — структурная единица простого вируса, состоящая из нуклеи-

новой кислоты и капсида. То же, что и вирион. Более сложные вирусы помимо белков капсида и нуклеиновой кислоты могут содержать *липопротеиновую мембрану*, углеводы и неструктурные белки — *ферменты*.

#### 4. Взаимодействие вируса и клетки

Различают три основные формы таких взаимодействий:

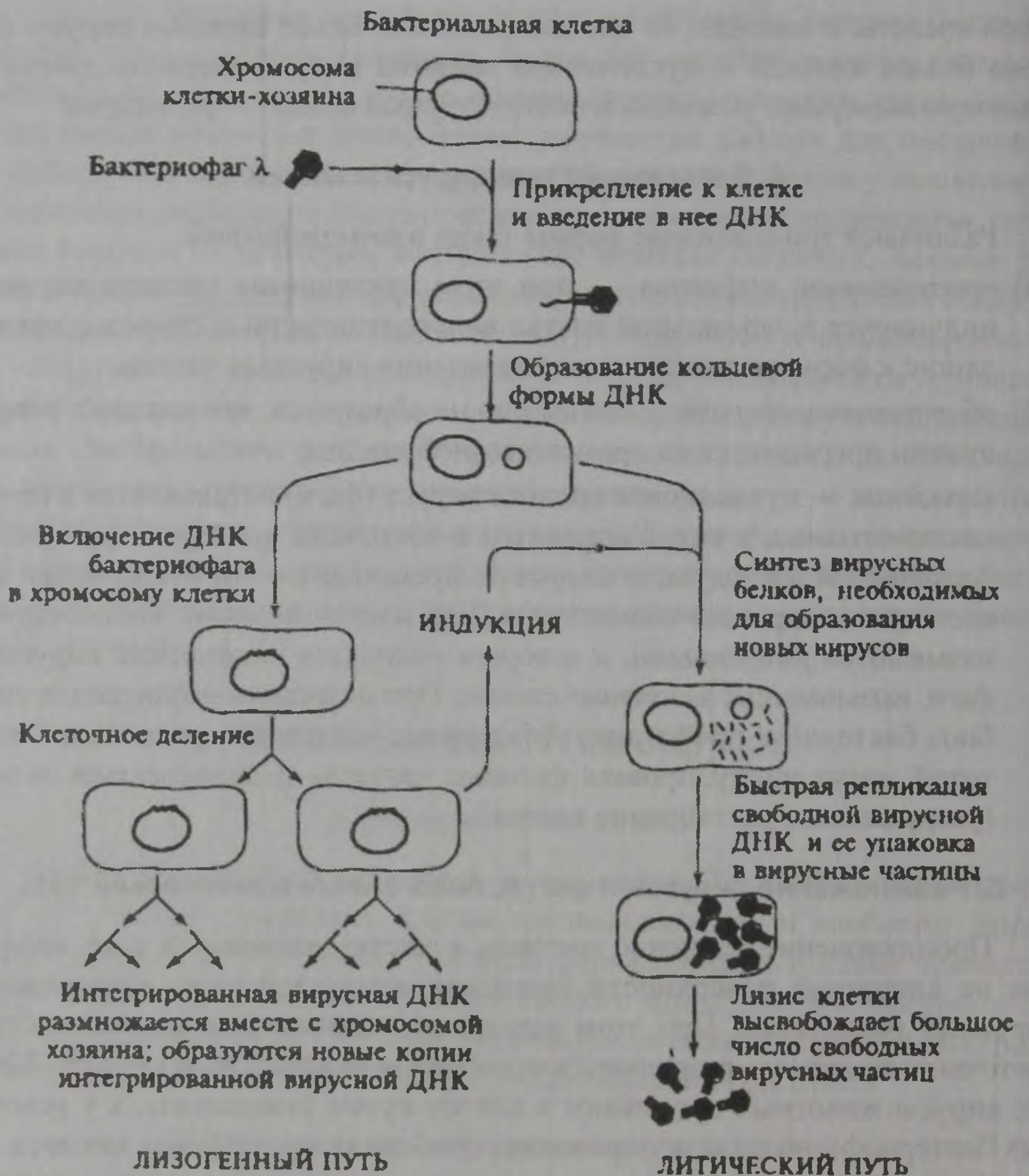
- а) *продуктивная инфекция* — при этом нуклеиновая кислота вириона индуцирует в зараженной клетке вирусспецифичные синтезы, приводящие к формированию нового поколения вирусных частиц;
- б) *абортивная инфекция* — потомство не образуется, так как цикл репродукции прерывается на промежуточной стадии;
- в) *виrogenия* — нуклеиновая кислота вируса (фага) встраивается в геном клетки-хозяина, а вирус находится в состоянии *провируса (профага)*. Автономной репродукции вируса не происходит, а его нуклеиновая кислота реплицируется совместно с ДНК клетки-хозяина. Такие вирусы называются *умеренными*, к которым относятся онкогенные вирусы и фаги, вызывающие *лизогению клеток*. При лизогении наблюдается симбиоз бактерий с профагами. Лизогенные клетки с определенной частотой могут продуцировать фаговые частицы и подвергаться *лизису* (разрушение и растворение клеток).

#### 5. Размножение вирусов и фагов. Лизогенный и литический путь

Проникновение вирусной частицы в клетку начинается с ее *адсорбции* на клеточной поверхности благодаря взаимодействию клеточных и вирусных рецепторов. При этом капсид приобретает чувствительность к клеточным протеазам, разрушается освобождая нуклеиновую кислоту. Многие вирусы животных проникают в клетку путем пиноцитоза, а у некоторых бактериофагов в клетку проникает свободная нуклеиновая кислота.

Размножение вирусов можно проследить на примере широко используемых в экспериментах специфических для *E. Coli* Т-фагах и фаге  $\lambda$  (лямбда). Отросток фага прикрепляется к специфическому рецептору клетки, его специфический фермент *лизоцим* образует отверстие в клеточной оболочке, а ДНК фага из его головки инъецируется в клетку. Внутри фаговой частицы ДНК находится в линейной форме с «липкими концами», которые взаимно комплементарны и могут спариваться друг с другом. При попадании в клетки *E. Coli* геном фага  $\lambda$  замыкается в кольцо и сшивается с помощью бактериального фермента ДНК-лигазы.

После проникновения ДНК фага в клетку описано два альтернативных пути инфекции — *литический путь* или *лизогенный путь* (рис. 49). В течение латентного периода при литическом пути инфекции геном фага заставляет синтезировать фаговую ДНК и белковую оболочку фага, они

Рис. 50. Жизненный цикл бактериофага  $\lambda$ 

соединяются и образуют фаговые частицы. Через определенное время после прикрепления фага к поверхности клетки оболочка бактерии разрывается (*клетка лизируется*), вновь образованные частицы освобождаются и начинается новый цикл их развития.

При лизогенном пути замкнутые в кольцо молекулы ДНК фага интегрируются в кольцевую хромосому клетки-хозяина. Такие клетки называются *лизогенными*, а встроенная ДНК фага — *профагом*. ДНК профага ведет себя как нормальная составная часть клетки, в ходе клеточных делений образуются новые копии фаговой ДНК.

ДНК-вирусы в животных клетках могут размножаться литическим путем (*пермиссивные клетки*). В непермиссивных клетках размножение вирусов блокируется, а вирусная ДНК размножается: один способ — интеграция в геном клетки-хозяина и репликация вместе с ним, а при другом способе ДНК вируса образует *плазмиду* — реплицирующуюся кольцевую молекулу ДНК.

### 6. Устойчивость вирусов к факторам окружающей среды

Вирусы животных вне клетки-хозяина инактивируются, однако наблюдаются большие различия у разных форм. Вирус полиомиелита сохраняет жизнеспособность в сточных водах в течение недель. Большинство вирусов погибает при пастеризации за несколько минут. В то же время вирус гомологичной сывороточной желтухи сохраняется после кипячения. Вирус табачной мозаики годами сохраняются в высушенных тканях растений.

Многие вирусы устойчивы к низким температурам и остаются живыми при  $-76^{\circ}\text{C}$  в течение года. Хранят вирусы в твердой углекислоте. Облучение ультрафиолетом приводит к быстрому разрушению всех вирусов, так как нуклеиновые кислоты очень чувствительны к действию лучей. Фаги не обезвреживаются в малых концентрациях дезинфицирующих веществ, они резистентны к антибиотикам, хлороформу и ферментным ядам. Исследование вирусов имело огромное значение для возникновения и развития молекулярной генетики. За заслуги в области изучения лизогении французский микробиолог и вирусолог Андре Мишель Львов удостоен Нобелевской премии в 1965 г. За исключительные заслуги в деле изучения фагов три патриарха молекулярной генетики — Макс Дельбрюк, Альфред Д. Херши и Сальвадор Э. Лурия — удостоены Нобелевской премии в 1969 г.

## Репликация генома и экспрессия генов вирусов

Геном вирусов внутри вирионов может быть представлен ДНК или РНК, которые могут быть одно- и двуцепочечными, кольцевой или линейной формы (рис. 51).

Размер и сложность строения генома вирусов разных групп варьируют в широком диапазоне — от наличия 4 генов у РНК-содержащего фага Q  $\beta$  до 250 генов у ДНК-содержащего вируса оспы. Геном вирусов может состоять из одной или нескольких молекул нуклеиновой кислоты. Двуцепочечная РНК реовирусов состоит из 10 молекул (или сегментов), а геном вирусов с одноцепочечной РНК может быть цельным (ретровирус) или сегментированным (вирус гриппа) (рис. 52). Геномы ДНК-содержащих вирусов позвоночных состоят из одной линейной или кольцевой молекулы ДНК, с наличием одной или двух цепочек. Геном вируса гепатита В представля-



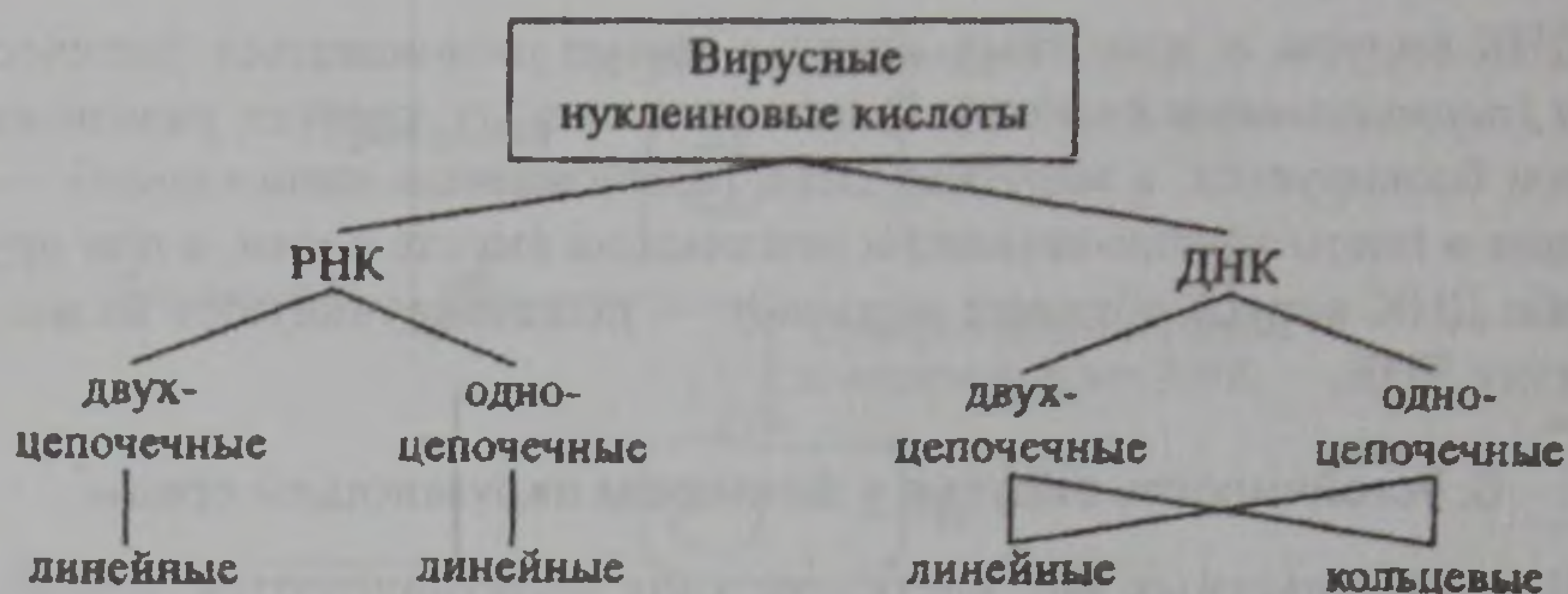


Рис. 51. Геномные нуклеиновые кислоты вирусов

ет собой кольцевую молекулу двуцепочечной ДНК, тогда как геномы вируса гепатитов А и С состоят из однонитевых РНК с положительной полярностью (+) РНК.

В настоящее время разработана «Классификационная система Балтимора», в которой учитывается два параметра — тип находящейся в вирионе нуклеиновой кислоты, а также стратегия ее репликации в клетке хозяине (табл. 9).

Все вирусы РНК разделяются на две группы по полярности ее молекулы — позитивные и негативные. Позитивные РНК, обозначаемые как (+РНК), способны в клетке-хозяине транслироваться рибосомами, т. е. выполнять функцию мРНК. Геном с негативной РНК, обозначаемый как (-РНК), не способен выполнять функцию мРНК, а РНК этого типа служит матрицей для синтеза мРНК.

Таблица 9

Система Балтимора по классификации вирусных геномов  
(Зинченко, Паруль, 2005)

Геномная нуклеиновая кислота	Сокращенное обозначение	Группа Балтимора
Двухцепочечная ДНК	dsДНК	I
Одноцепочечная ДНК	ssДНК	II
Двухцепочечная РНК	dsРНК	III
Одноцепочечная РНК позитивной полярности	ss(+)-РНК	IV
Одноцепочечная РНК негативной полярности	ss(-)-РНК	V
Одноцепочечная позитивная РНК, в цикле репликации которой имеется стадия обратной транскрипции	ss(+)-РНК	VI
Двухцепочечная ДНК, в цикле репликации которой имеется стадия обратной транскрипции	dsflНК	VII


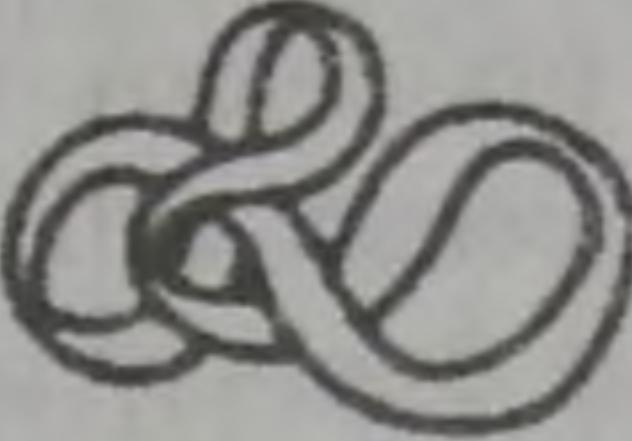
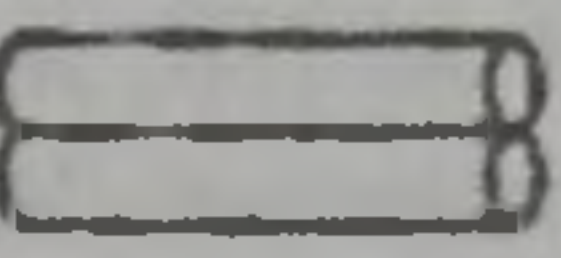
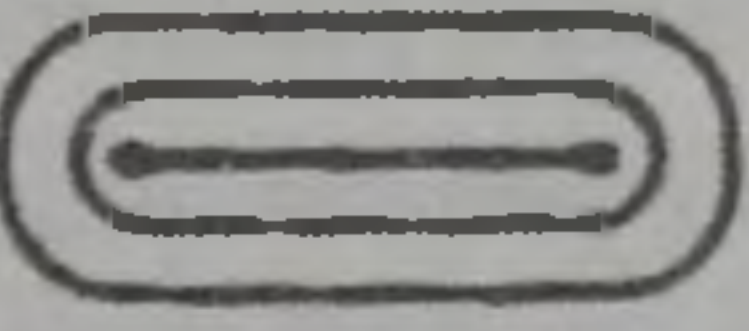
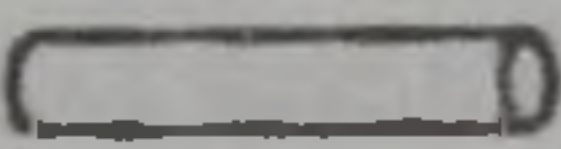

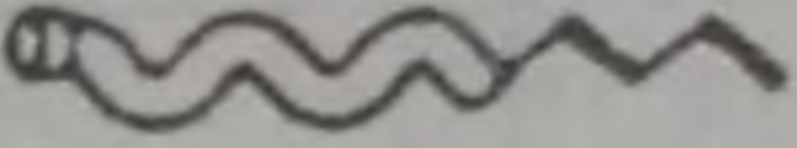
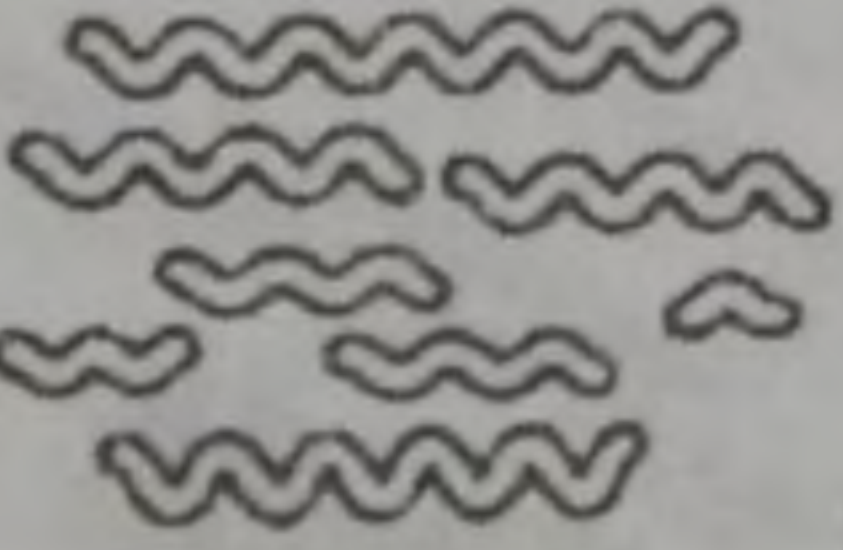
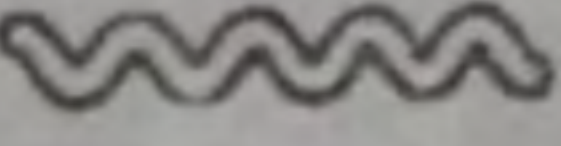
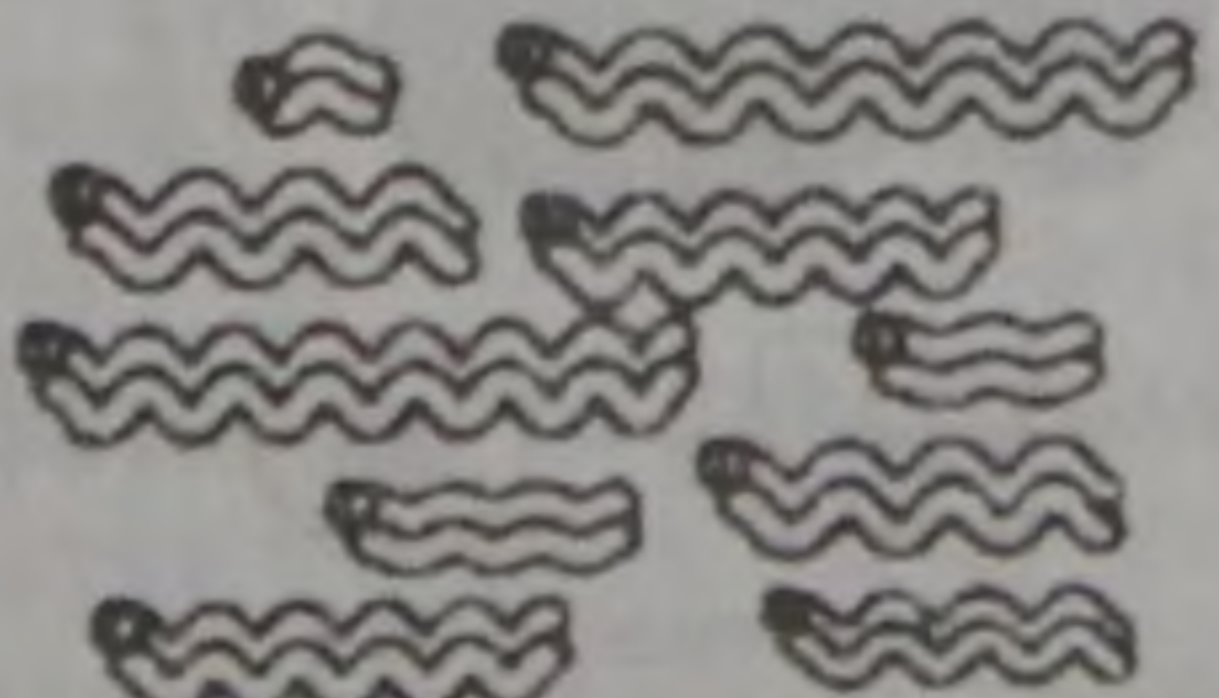
Вirus (примерный размер генома в тыс. н.п. или тыс. н.)	ДНК-содержащие вирусы	Структура генома
Вирус гепатита В (3)		Частично одноцепочечная кольцевая ДНК
Паповавирусы (5)		Сверхспиральная замкнутая кольцевая ДНК
Аденовирус (36) Вирус герпеса (200)		Линейный дуплекс
Вирус осповакцины (200)		Линейный дуплекс с ковалентно сшитыми концами
Парвовирус (2)		Одноцепочечная линейная ДНК
Фаг φX174 (5,3) Фаг M13 (6,0)		Одноцепочечное кольцо
<b>РНК-содержащие вирусы</b>		
Ретровирус (10)		Одноцепочечный диплоидный геном
Вирус гриппа (16)		8 разных одноцепочечных молекул
Поллиовирус (7) ВТМ (7000)		Одноцепочечная молекула
Реовирус (30)		10 разных двуцепочечных молекул

Рис. 52. Структура генома вирусов (Коницев, Севастьянова, 2003)

Репликация дочерних вирусных геномов после проникновения вируса в клетку, а также их транскрипция и трансляция (синтез собственных белков) имеет ряд особенностей:

- во-первых, в клеточном ядре и цитоплазме отсутствуют ферменты, необходимые для транскрипции вирусных ДНК-геномов, т. е. для синтеза м-РНК на вирусных РНК. В цитоплазме также нет ферментов, которые способны транскрибировать вирусную ДНК. Поэтому ДНК-содержащие вирусы могут проникать в ядро клетки и использовать хозяина для считывания вирусных мРНК. Все остальные вирусы способны создавать собственные ферменты для синтеза мРНК;
- во-вторых, белоксинтезирующий аппарат эукариот может транслировать только моноцистронные мРНК, а внутри полицистронных молекул мРНК не распознает инициации. Вирусы же способны синтезировать моноцистронные мРНК, а также мРНК для кодирования большого полипротеина, который затем расщепляется на отдельные белки.

По способам осуществления репликации, транскрипции и трансляции все вирусы разделяются на семь групп Балтимора.

### Вирусы I группы Балтимора (двухцепочечная ДНК)

Вирусы эукариот размножаются в ядре, или в цитоплазме используя клеточные ферменты. В некоторых случаях процесс репликации и транс-

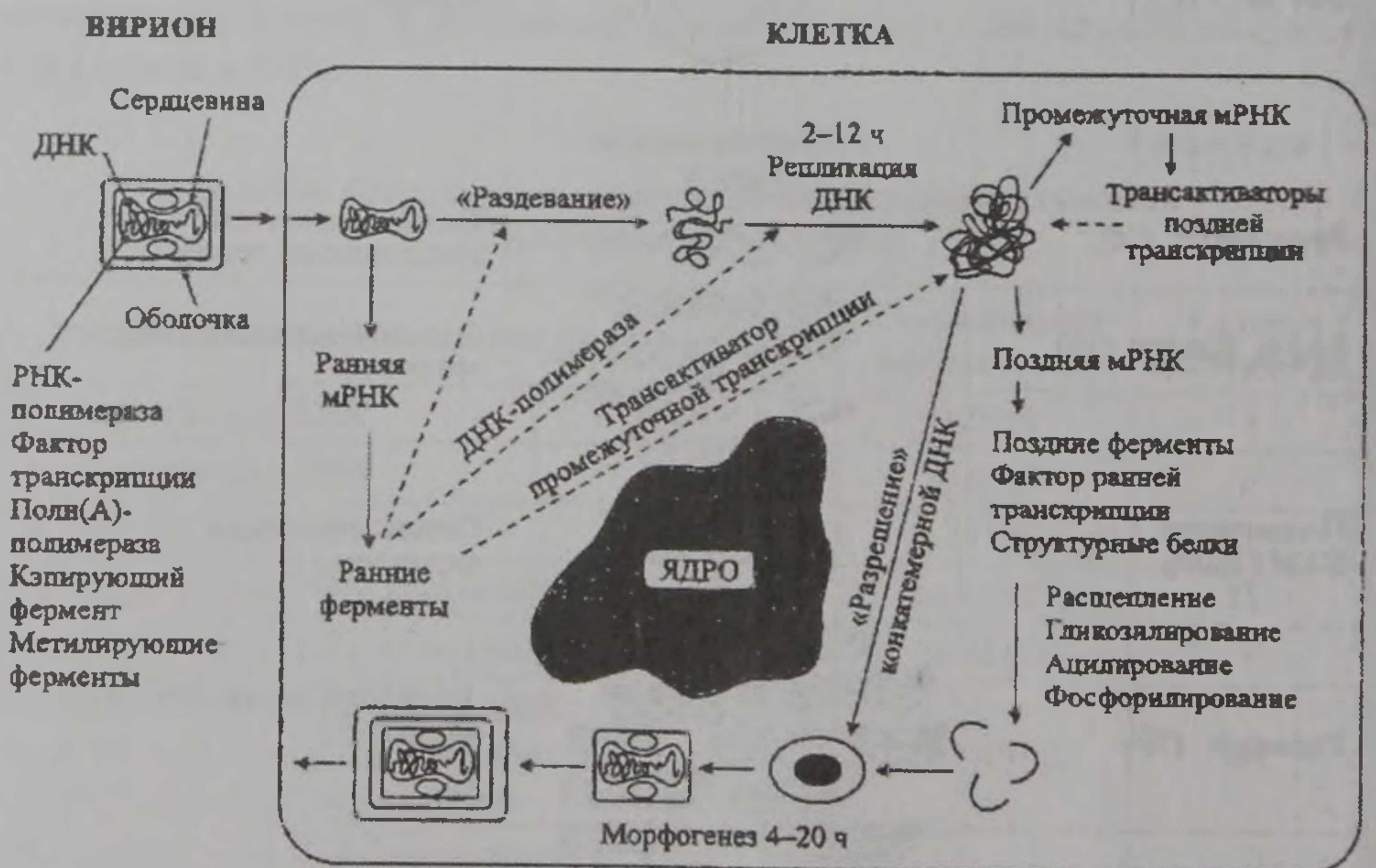


Рис. 53. Цикл репродукции вируса осповакцины (Щелкунов, 2004)

крипции обслуживают клеточные и собственные ферменты. Репликация осуществляется по схеме ДНК ДНК. К этой группе относятся вирусы герпеса, оспы и др. (рис. 53).

### **Вирусы II группы Балтимора (одноцепочечная ДНК)**

Репликация всех вирусов этой группы протекает в ядре. После начала инфекции при участии клеточных ДНК (+ ДНК) транскрибируется только одна из нитей вирусной ДНК.

### **Вирусы III группы Балтимора (Двуцепочечная РНК)**

Большинство вирусов этой группы имеет сегментированный геном, т. е. состоят из нескольких разных молекул. Все фрагменты генома находятся в составе одной вирусной частицы. Геномы вирусов этой группы реплицируются в цитоплазме клетки-хозяина с помощью вирусной РНК-зависимой РНК-полимеразы. После транскрипции моноцистронных мРНК происходит трансляция вирусных белков. Цепочки (+) и (-) РНК вступают в комплекс и образуют двунитевый (+/-) РНК-геном, который упаковывается в белковую оболочку. К этой группе относятся реовирусы.

### **Вирусы IV группы Балтимора (одноцепочечная (+) ДНК)**

Вирусы этой группы размножаются в цитоплазме. К ним относятся фаг Q<sub>β</sub> (кью-бета), а также вирусы табачной мозаики, полиомиелита, клещевого энцефалита. Их геном кодирует несколько белков, в том числе репликазу (РНК-зависимую РНК-полимеразу), которая способна синтезировать молекулы РНК без участия ДНК. На матрице геномной (+) РНК синтезируется комплементарная (-) РНК. Изолированная из вирионов «голая» (+) РНК также инфекционна. После проникновения в клетку геномная РНК функционирует как мРНК, связывается с рибосомами и транслируется, образуя *полипротеины*. На заключительной стадии из накопившихся вирусных белков и молекул дочерних (+) РНК формируются вирионы.

### **Вирусы V группы Балтимора (одноцепочечная (-) РНК)**

К этой группе принадлежат вирусы гриппа, кори, бешенства, желтой карликовости картофеля и др. Эти вирусы содержат собственную РНК-зависимую РНК-полимеразу — фермент, упакованный в вирионе в удобной для доставки в клетку форме. Она катализирует образование полноразмерной (+ РНК), которая служит матрицей для синтеза на ней геномной (-) РНК потомства.

### Вирусы VI группы Балтимора (одноцепочечная (+) РНК)

Вирусы этой группы в своем репликативном цикле имеют стадию обратной транскрипции и ДНК в качестве промежуточного продукта. Они представлены сем. *Retroviridae*, представители которого инфицируют только позвоночных. Важнейшие представители семейства: вирус иммунодефицита человека (ВИЧ), некоторые возбудители злокачественных новообразований (онкогенные вирусы) — вирус саркомы Рауса (птиц), мышинный вирус рака молочной железы, вирус лейкоза мышей (вирус Молони), вирус лейкоза обезьян, Т-лимфотропный вирус человека и др. В состав вирионов ретровирусов входят 5–8 различных полипептидов. Внутри капсида имеются несколько десятков молекул *обратной транскриптазы*, молекулы тРНК и две идентичные геномные одноцепочечные молекулы РНК — *уникальный случай диплоидного вирусного генома*.

В процессе инфекции на геномной РНК ретровируса с помощью обратной транскриптазы синтезируется двухцепочечная кольцевая ДНК-копия (рис. 54), которая затем интегрируется в геном клетки-хозяина, формируя так называемый *провирус*. ДНК провируса под контролем клетки реплицируется вместе с ДНК хозяина при делении. Провирусные гены также транскрибируются в ядре клетки транскриптазой хозяина в (+) РНК-транскрипты. Это используется в качестве геномов дочерних вирусных частиц, а также транслируется для синтеза их структурных белков и ферментов. Самосборка вирусных нуклеопротеидов происходит в цитоплазме, на клеточной оболочке которой вирусные частицы созревают и отделяются почкованием. Провирус в течение долгого времени может не проявлять себя и передаваться потомкам клетки-хозяина «вертикально».



Рис. 54. Цикл репродукции вирусов VI группы Балтимора

### Вирусы VII группы Балтимора (двухцепочечная ДНК)

Вирусы этой группы в своем репликативном цикле имеют стадию обратной транскрипции и РНК в качестве промежуточного продукта (рис. 55).

Они получили название ретроидные вирусы, а их представителями являются вирус гепатита В человека, вирусы гепатита сурков, земляных белок, сусликов, пекинских уток, а также вирус мозаики цветной капусты.

ДНК и полимераза вириона мигрируют в ядро клетки-хозяина, где ДНК превращается в *сверхспирализованную молекулу*. Она транскрибируется с образованием молекул (+) РНК. Под влиянием обратной транскриптазы сначала синтезируется (-) нить ДНК, а на ней тот же фермент строит (+) нить. Вирусная ДНК может интегрироваться в геном клетки, но этот процесс не обязателен. Если он имеет место, то ДНК вируса расщепляется на фрагменты и встраивается в разные участки клеточной ДНК.

Реализация генетической информации, закодированной в последовательности нуклеотидов молекулы ДНК (или РНК) у вирусов, имеет следующие особенности:

1. Несегментированные (однокомпонентные) и сегментированные (фрагментированные) геномы различаются по механизму транскрипции. Для однокомпонентных геномов установлена транскрипция *полицистронных мРНК*, трансляция которых ведет к образованию нефункциональных полипротеинов. С помощью специфических клеточных и вирусных протеаз полипротеины расщепляются на функционально активные белки. При транскрипции сегментированных геномов образуются *моноцистронные мРНК* с последующим синтезом разных белков в разных количествах.

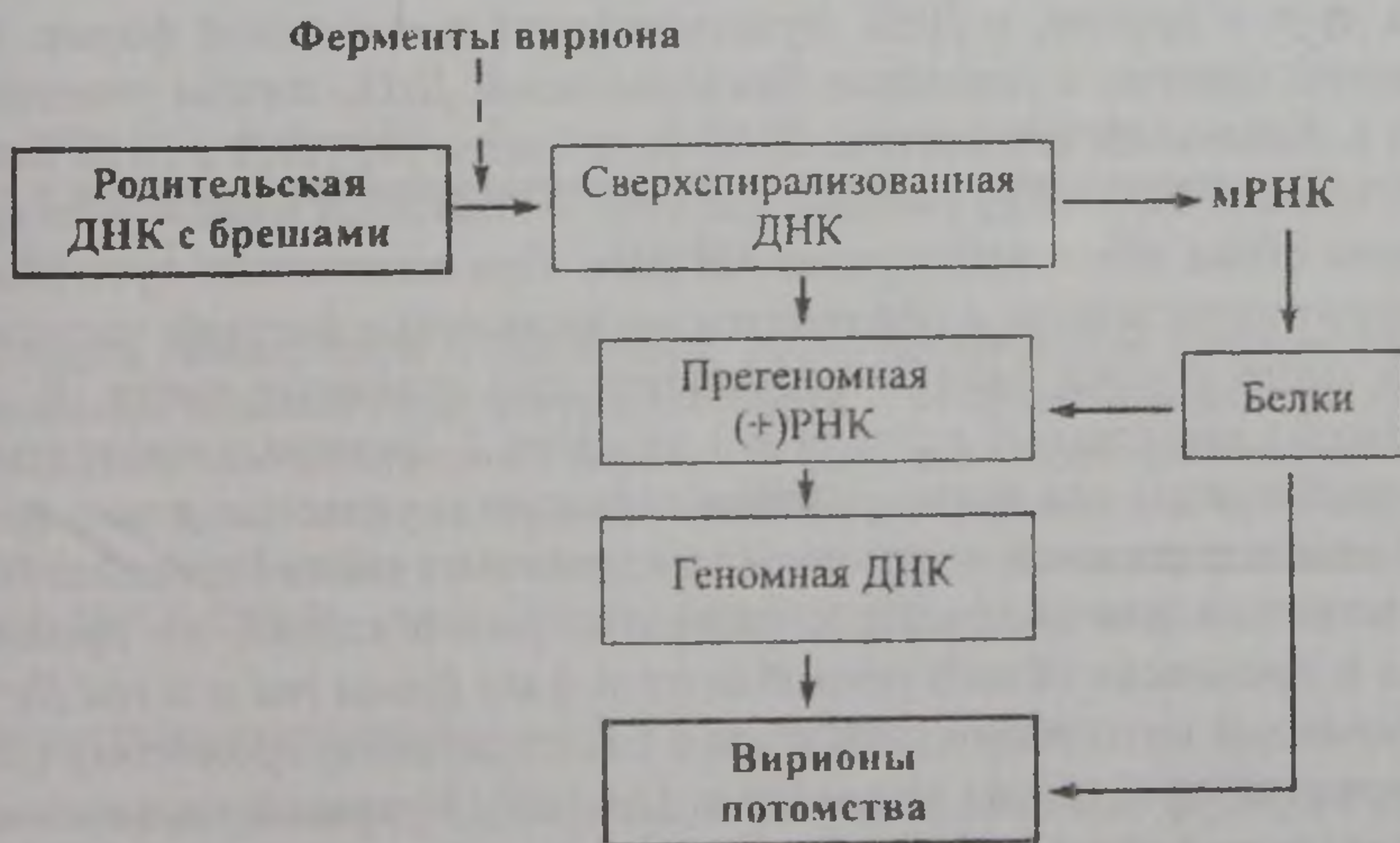


Рис. 55. Цикл репродукции вирусов VII группы Балтимора

2. Для обеспечения гармоничного взаимодействия между вирусным геномом и клеткой-хозяином вирусные мРНК несут распознаваемые клеткой информационные сигналы. К ним относятся сайты связывания с рибосомой; последовательности полиаденилирования; способность белков ДНК-содержащих вирусов связываться с участком «*ориджин репликации*» и тем самым стимулировать ДНК-полимеразу клетки к репликации вирусного генома.
3. Для простых вирусов отмечена зависимость от клеточных факторов, необходимых для репликации геномов. Сложные вирусы, кодирующие многие белки для участия в синтезе ДНК, от клеточных факторов зависят мало.

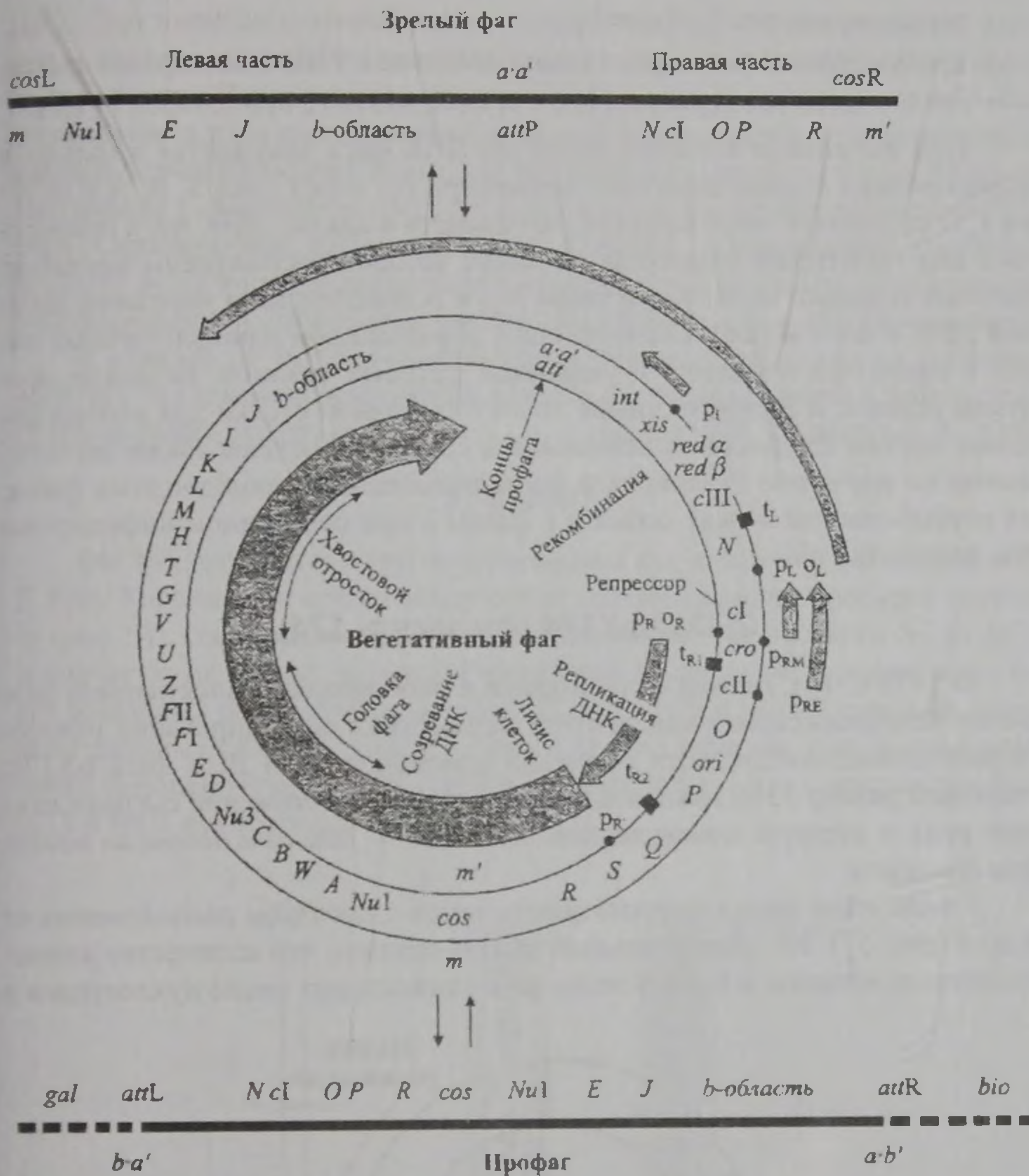
## Характеристика вирусных геномов

### 1. Фаг (колифаг) $\lambda$ лямбда

Белки головки и хвостового отростка кодируются геномом фага и синтезируются в инфицированных клетках кишечной палочки *E. coli*. Фаг  $\lambda$  является умеренным фагом и, в зависимости от условий, развивается в клетке либо по литическому, либо по лизогенному пути. В зрелых вирионах фага ДНК находится в виде двухцепочечной линейной молекулы размером 48.502 п. н. Определена полная нуклеотидная последовательность генома и проведено картирование всех более 60 генов (рис. 56).

Внутри фаговой частицы ДНК имеет линейную форму: 5'-концы каждой цепи, обозначенные как *cos R* и *cos L*, содержат взаимокomплементарные GC-обогащенные участки длиной 12 нуклеотидов. Эти комплементарные «липкие концы» сразу после попадания в клетку бактерии спариваются друг с другом, и ДНК функционирует в кольцевой форме. Район ковалентно сшитых с помощью бактериальной ДНК-лигазы участков *cos R* и *cos L* называется *cos*-сайтом. ДНК фага  $\lambda$  интегрируется в ДНК бактериальной хромосомы между сайтами *gal* и *bio* и находится в состоянии *профага*. Такова схема лизогенного развития фага. При литическом пути развития происходит лизис клеток и образуются инфекционные фаговые частицы.

На карте генома фага  $\lambda$  выделяются три основные части. В левую часть входят все гены от сайта N и I до сайта J. Белковые продукты этих генов необходимы для формирования капсидов и упаковки в них фаговой ДНК. Гены центральной части включают район от сайта J до сайта N. Они несущественны для литического развития фага в клетке, но принимают участие в процессах общей рекомбинации фага (гены *red  $\alpha$*  и *red  $\beta$* ), сайт-специфической интеграции ДНК фага в бактериальную хромосому (*ген int*) и исключения профага из хромосомы (*ген xis*). К правой части отнесены все остальные, контролирующие элементы, необходимые для репликации фаговой ДНК (гены O и P) и для лизиса клеток (гены S и R).



**Рис. 56.** Карта генома фага  $\lambda$ . (Щелкунов, 2004) Прописными буквами обозначены жизненно важные гены фага, строчными — несущественные для литического развития; заштрихованные стрелки — единицы транскрипции фагового генома (скриптоны), ширина стрелок отражает интенсивность транскрипции соответствующих районов генома; толстая штриховая линия (*gal* и *bio*) — хромосомная ДНК *E. Coli*; P — промотор; O — оператор; T — терминатор транскрипции

Следует отметить также следующие сайты генома:

- промоторы транскрипции генов литического развития  $P_L$ ,  $P_R$  и  $P_{RI}$ , при этом транскрипция имеет направление с промотора  $P_L$  влево через ген *N*, а с промотора  $P_R$  вправо через ген *cro*;



- терминирующие  $T_L$ ,  $T_{R1}$  и  $T_{R2}$ ;
- продукт гена  $N$  регулирует взаимодействие с РНК-полимеразой клетки;
- ген *cro* является дерепрессором транскрипции с промоторов  $P_L$  и  $P_R$ .

При попадании в клетки молекула ДНК фага замыкается в кольцо, а транскрипция осуществляется с промотора  $P_{R1}$  через гены  $S$ ,  $R$ ,  $A$  и до гена  $J$ . В результате многократной репликации в клетке ДНК фага приобретает вид гигантской молекулы, по длине которой расположены несколько фаговых геномов. Затем белки генов  $N_{U1}$  и  $A$  обеспечивают нарезание фаговой ДНК в *cos*-сайтах с последующим образованием одноцепочечных нитей с «липкими концами» и упаковкой фаговых геномов. Белковые продукты генов  $R$  и  $S$  способствуют лизису клеточной стенки для выхода фаговых частиц следующего поколения в среду. Проведены также эксперименты по изучению *in vitro* ряда близкородственных лямбдоидных фагов, их рекомбинантов между собой и с фагом  $\lambda$  при совместном инфицировании клеток *E. coli*.

## 2. Фаг $\phi X174$ (фи-десять 174)

В 1978 г. Ф. Сэнджер с соавторами в лаборатории молекулярной биологии Кембриджского университета осуществил секвенирование нуклеотидной последовательности геномной одноцепочечной ДНК фага  $\phi X174$ , имеющей размер 5386 нуклеотидов. Это научное достижение сыграло важную роль в истории молекулярной генетики — оно ознаменовало начало эры геномики.

Геном этого фага содержит девять генов с порядком расположения от  $A$  до  $J$  (рис. 57). Но сравнительный анализ показал, что количество аминокислотных остатков в белках этого фага превосходит число нуклеотидов в

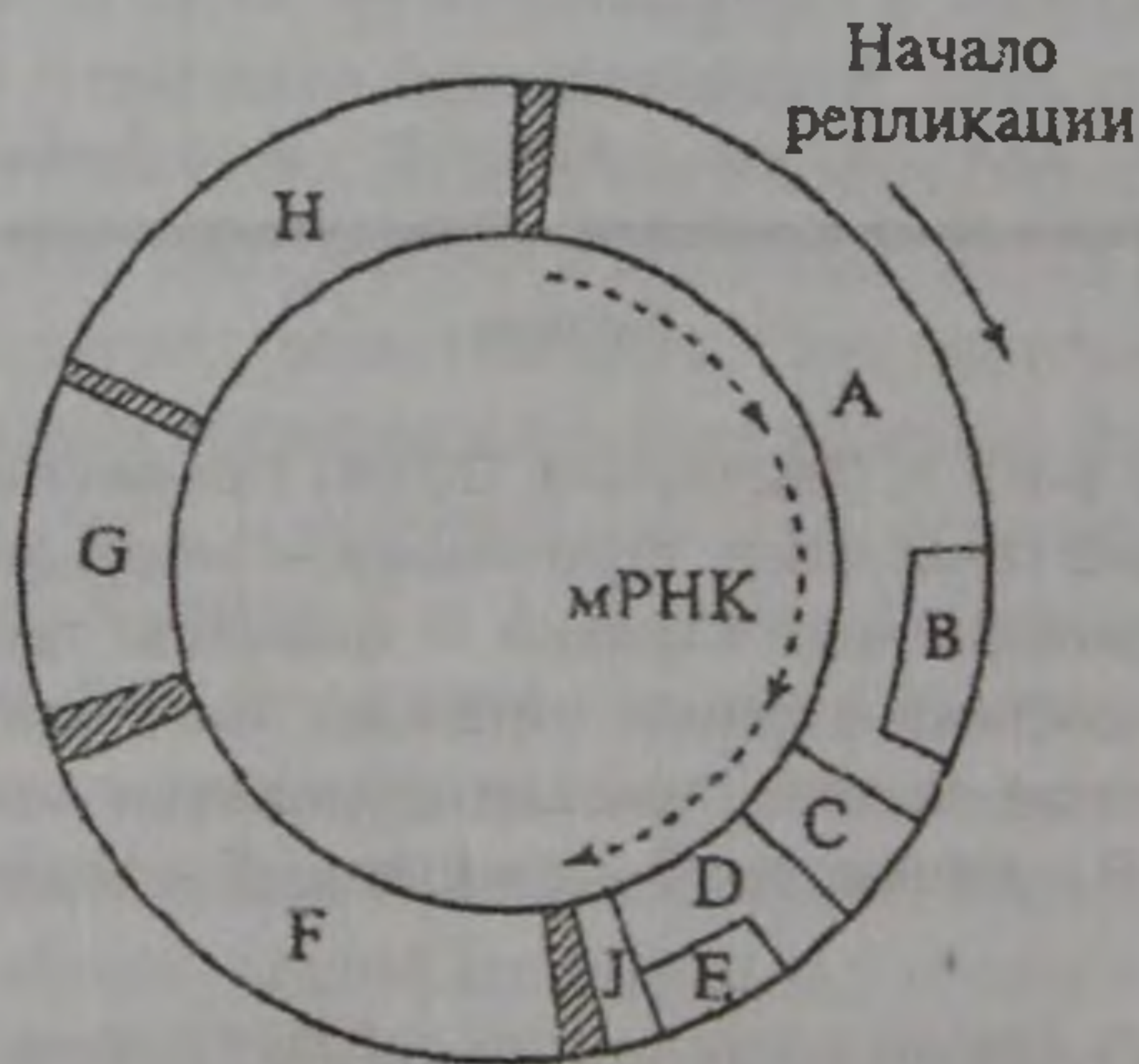


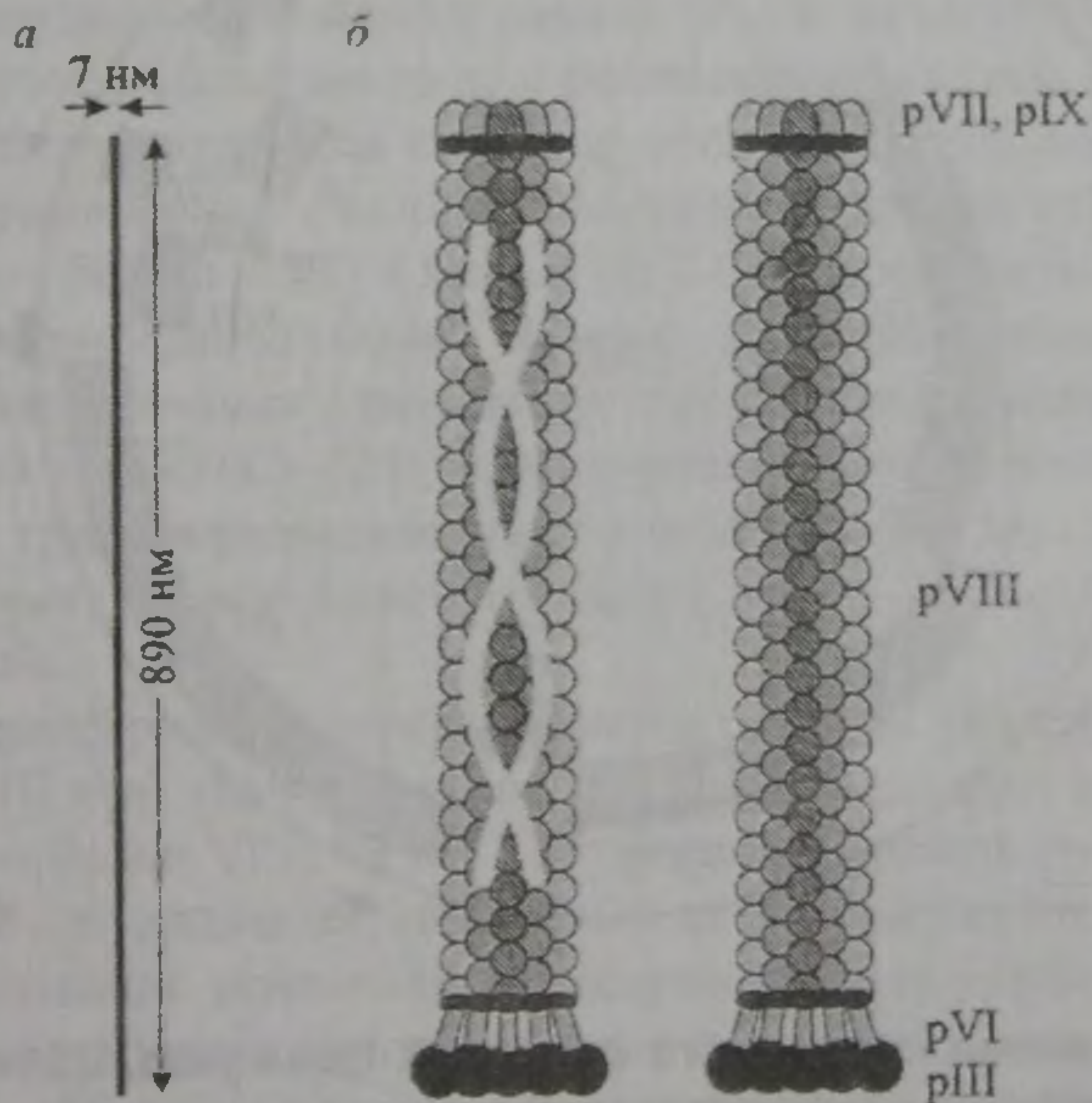
Рис. 57. Карта генома фага  $\phi X174$ . (Сэнжер, 1977) Буквами обозначены гены. Ген  $B$  локализован внутри гена  $A$ , а ген  $E$  — внутри гена  $D$ . Заштрихованы не-транскрибирующиеся участки

ДНК. Это привело к открытию важного молекулярно-генетического явления — «перекрывания генов», или «гены внутри генов». На карте генома видно, что ген В расположен внутри гена А, а ген Е — внутри гена D. Было показано в ряде случаев, что стартовый кодон одного гена может перекрывать терминирующий кодон другого гена. Это увеличивает число рамок считывания и, в конечном счете, компенсирует нехватку нуклеотидов в ДНК фага. Ограниченное количество ДНК в геноме обуславливает необходимость ее экономного использования для кодирования белков, обеспечивающих инфекционность фаговых частиц и их способность к воспроизведению.

В отличие от фагов, большой геном человека теоретически позволяет обойтись без такого перекрывания. Однако и в геноме человека, хоть и относительно редко, перекрывание генов также имеет место (Тарантул В. З., 2003).

### 3. Фаг M13

Фаг M13 относится к группе *нитевидных фагов* (наряду с фагами fd и f1) *E. Coli*. Капсид фага представляет собой напоминающую пробирку трубочку (рис. 58), состоящую из 2700 копий основного оболочечного белка pVIII и закрытую на концах четырьмя или пятью молекулами каждого из четырех минорных оболочечных белков pIII, pVI, pVII и pIX. В капсиды упакованы длинные (900 нм), тонкие (7 нм) и гибкие вирионы этих фагов с кольцевой одноцепочечной ДНК (оц ДНК) длиной 6407 нуклеотидов для фага M13 и 6408 нуклеотидов для фагов fd и f1.

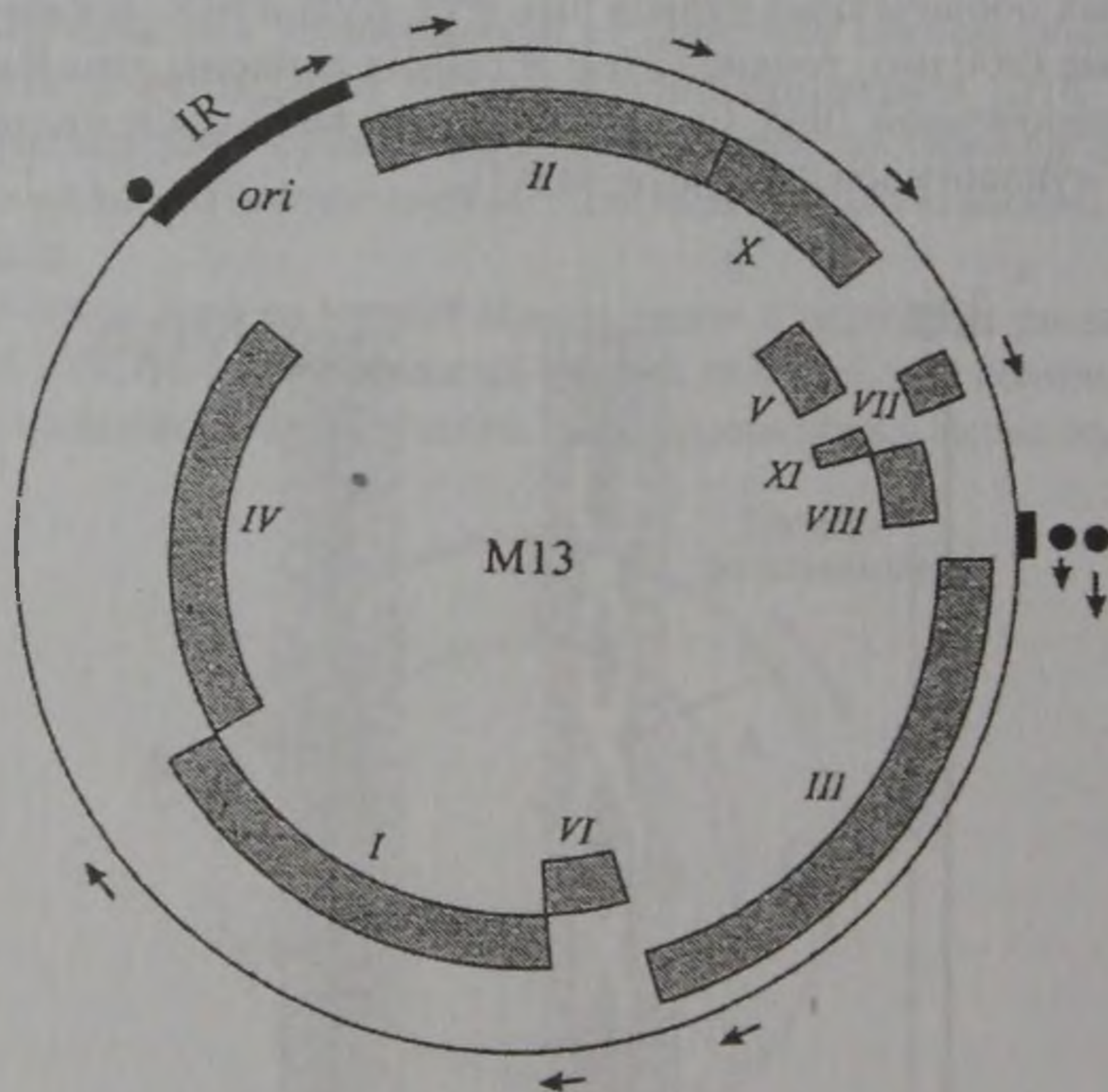


**Рис. 58.** Схема частицы нитевидного фага: а — соотношение размеров частицы; б — схема упаковки белков оболочки (внутри изображена цепь ДНК)

Нитевидные фаги адсорбируются на тонких, напоминающих волоски структурах — *F-пилях* *E. coli* и инфицируют только мужские ( $F^+$ ) бактериальные клетки. Одноцепочечная ДНК, являющаяся (+)-цепью, проникает в цитоплазму клетки, где с помощью бактериальной полимеразы происходит синтез второй, (-)-цепи. Образуется двухцепочечная *репликативная форма (РФ)* фагового генома.

В межгенной области IR генома находятся участки  $ori(-)$  и  $ori(+)$  (рис. 59).

На молекулах РФ осуществляется транскрипция и последующая трансляция фаговых генов. Репликация фаговой ДНК происходит по модели катящегося кольца с высвобождением (+)-цепи, концы которой лигирует белок pII. Всего образуется 200–300 копий РФ на клетку. Новосинтезированные оболочечные белки фага встраиваются в плазматическую мембрану клетки. Инициация сборки и экспорта вириона происходит после взаимодействия специфической нуклеотидной последовательности (так называемый сигнал упаковки) в IR-области (+)-ДНК с комплексом pI-pIV и минорными оболочечными белками pVII и pIX. Присоединение белков pVI и pIII к концам формируемых вирионов завершает процесс их сборки.



**Рис. 59.** Карта генома нитевидного фага M13. (Щелкунов, 2004) Гены обозначены серыми буквами и римскими цифрами, межгенные области — черными блоками. Стрелками показаны промоторы и направление транскрипции, черными кружками — терминаторы транскрипции

Образующиеся фаговые частицы экскретируются клетками без лизиса последних. За одну генерацию образуется несколько сотен фагов на клетку. В геноме фагов существует межгенная область IR размером 500 нуклеотидов. В этот сегмент может быть помещена вставка чужеродной ДНК без нарушения способности самого фага к репликации.

Двухцепочечная репликативная форма ДНК легко выделяется из инфицированных клеток, а затем используется в качестве векторов в геной инженерии. Векторные производные нитевидных фагов используют для экспрессии чужеродных генов в клетках *E. Coli*.

#### 4. Вирус SV-40

Этот вирус является эукариотическим ДНК-содержащим онкогенным вирусом. Он относится к роду *Polyomavirus* семейства *Polyomaviridae*, т. е. мелким по размерам, двухцепочечная ДНК которого реплицируется в ядре клетки-хозяина. Размножение вируса SV-40 может происходить либо вегетативно по литическому пути, либо в интегрированном в хромосомы состоянии (*непермиссивные клетки*). Природным хозяином SV-40 является макак резус. В непермиссивных клетках ДНК вируса интегрируется в геном и обуславливает онкогенную трансформацию клеток.

Геном вируса имеет размер 5243 п. н. Хорошо изучена репликация его генома, регуляция экспрессии в различных клетках млекопитающих. Установлено, что гены вируса SV-40 в значительной степени перекрываются. Промоторные последовательности ранней ( $P_E$ ) и поздней ( $P_L$ ) транскрипции перекрывается с областью начала репликации (рис. 60).

При ранней транскрипции синтезируется пре-мРНК, из которой в результате *альтернативного сплайсинга* образуются зрелые мРНК, кодирующие два белка — большой (Т) и малый (t) Т-антигены. Большой Т-антиген является многофункциональным белком, поддерживающим состояние *онкогенной трансформации* у разных культур клеток млекопитающих. Внедрение вирусного генома в ДНК клетки-реципиента не является сайтспецифическим, а трансформированное состояние клеток может сохраняться длительное время. Малый t-антиген также может служить трансформирующим агентом.

Поздняя транскрипция осуществляется с другой цепи вирусной ДНК. Молекулы мРНК при этом образуют два класса молекул, с них считываются вирионные белки VP1, VP2 и VP3, формирующие вирусные частицы. К участку ДНК *ori* прилегает тандемный повтор размером 72 п. н., который получил название *усилителя транскрипции, или энхансера*. Энхансеры обнаружены у многих других вирусов и в геномах клеток эукариот в форме коротких (30–100 п. н.), тандемно повторенных последовательностей. Они могут активировать транскрипцию с промоторов, расположенных на расстоянии от сотен до 10 т. п. н.

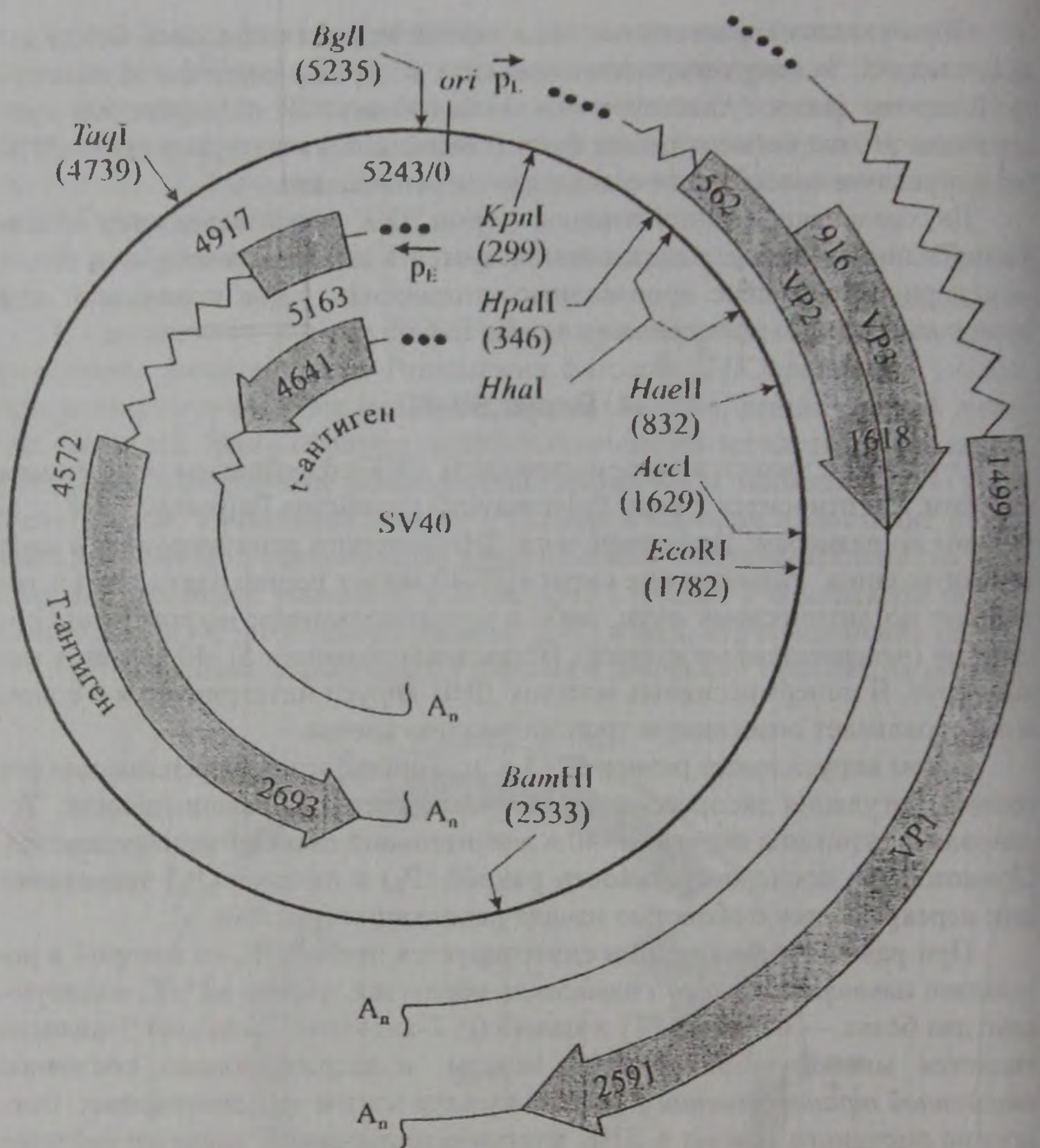


Рис. 60. Карта генома вируса SV-40. Толстые стрелки — последовательности вирусных мРНК, кодирующие белки; зигзагообразные линии — интроны молекул мРНК;  $A_n$  — участки полиаденилирования. Цифрами отмечено положение соответствующих точек на физической карте ДНК (в парах нуклеотидов)

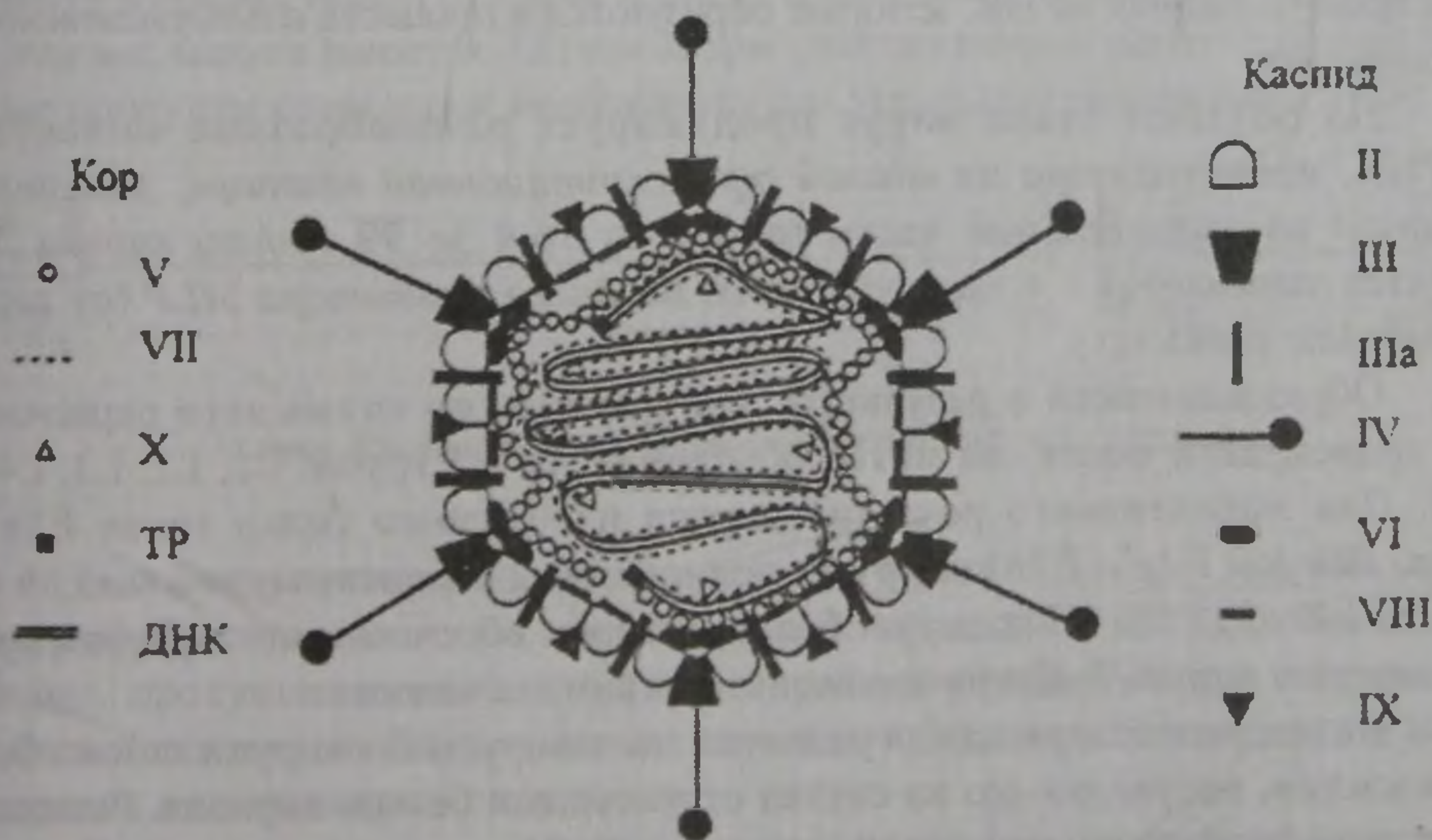
Репликация ДНК SV-40 инициируется в единственном участке *ori* (размер 65 п. н.) и происходит в двух направлениях (по часовой и против часовой стрелки). Этот и другие признаки вируса SV-40 определило его значение как замечательную модель для изучения регуляции экспрессии генов, сплайсинга пре-мРНК и других процессов и эукариот.

## 5. Аденовирусы

Эта разновидность вируса относится к семейству Adenoviridae, состоящее из двух родов — Mastadenovirus (вирусы разных видов млекопитающих) и Aviadenovirus (вирусы разных видов птиц). Они впервые выделены из аденоидов и миндалин детей. Аденовирусы могут вызывать заболевания у человека респираторной, центральной нервной систем, желудочно-кишечного тракта, мочеполовых путей, глазные болезни. Среди аденовирусов человека идентифицировано 50 серотипов, а также 7 подгрупп по содержанию GC-пар ДНК, способных вызывать агглютинацию эритроцитов. Хорошо изучены аденовирусы типов 2 и 5 (Ad2 и Ad5) человека с гомологичной последовательностью ДНК. Установлена их способность интенсивно размножаться только в организме или культуре клеток естественного хозяина.

Белковый капсид вириона состоит из 7 структурных белков (рис. 61), в него упакован геносодержащий кор (сердцевина). В состав кора входят 4 структурных белка.

Линейная двухцепочечная молекула ДНК аденовирусов имеет размер около 36 т. п. н. На концах молекулы находятся инвертированные повторяющиеся последовательности ITR (от англ. inverted terminal repeats) размером 102–103 п. н. В 1987 г. расшифрована полная нуклеотидная последовательность молекулы ДНК Ad2 (35 937 п. н.). При биохимическом и генетическом изучении аденовирусов Ad2 и Ad5 была выяснена картина функционирования их геномов на уровне транскрипции (рис. 62).



**Рис. 61.** Вирион аденовируса (схематичное изображение). Римскими цифрами обозначены полипептиды; TR — терминальный белок

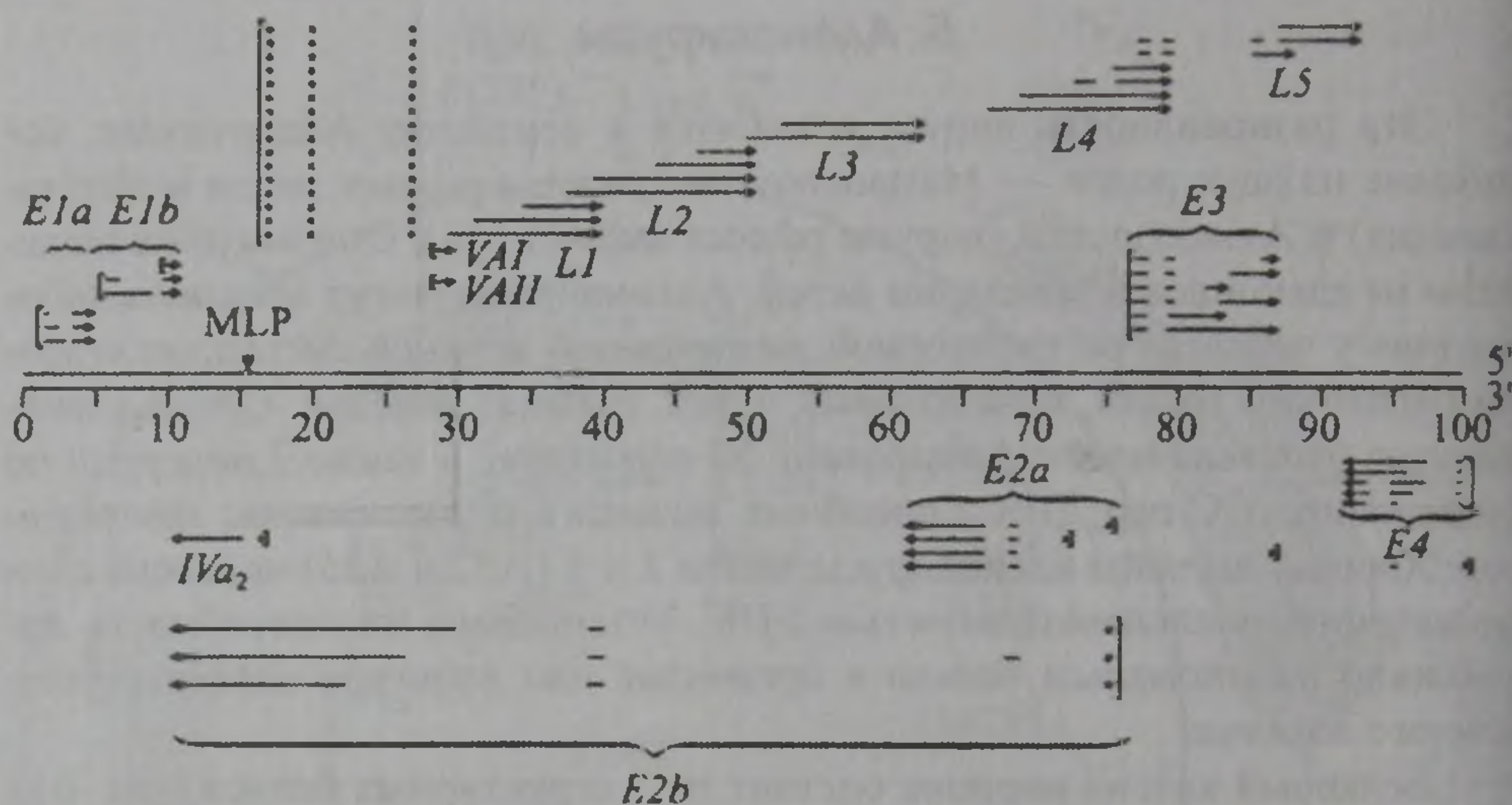


Рис. 62. Транскрипционная карта генома аденовирусов Ad1 и Ad5. (Щелкунов, 2004) Точка начала транскрипции каждой мРНК обозначена вертикальной квадратной скобкой. Горизонтальными стрелками с разрывами обозначены экзоны, из которых составлены соответствующие зрелые молекулы РНК

Можно выделить ранний и поздний этапы цикла развития вируса после инфицирования клетки. Вначале с вирусной ДНК считываются шесть транскрипционных единиц E1a, E1b, E2a, E2b, E3, E4. В них закодированы наборы разных мРНК, которые образуются в процессе *альтернативного сплайсинга*.

На позднем этапе вирус продуцирует разнообразные молекулы мРНК, происходящие из *единой транскрипционной единицы*, локализованной на значительной части генома от 16,4 до 99 единиц карты. Ее синтез начинается с т. наз. *основного позднего промотора MLP* (от англ. major late promoter).

Образовавшиеся в результате альтернативного сплайсинга первичного транскрипта молекулы мРНК объединяются в 5 групп: L1, L2, L3, L4 и L5. Для эффективного развития вируса необходимы белки генов E1a и E1b. Локусы E2a и E2b кодируют полипептиды с молекулярной массой от 72 до 140 кДа. Ген E3 кодирует белки, которые обеспечивают *молекулярную мимикрию* вируса Ad2 при заражении организма человека.

На поздней стадии цикла развития аденовирусы блокируют синтез белков клетки, направляя его на синтез структурных белков вириона. Репликация линейной двухцепочечной молекулы ДНК аденовирусов и сборка вирионов происходит в ядре клетки. Продуктивная инфекция вирусами заканчивается *лизисом клеток*.

## 6. Герпесвирусы

Это крайне разнообразные по биологическим свойствам вирусы семейства *Herpesviridae*. К важнейшим представителям семейства относятся у человека 8 видов: вирус простого герпеса типа 1 (HSV-1, или HHV-1) и типа 2 (HSV-2, или HHV-2), вирус ветряной оспы, или опоясывающего лишая (VZV-2, или HHV-3), вирус Эпштейна—Барр (EBV, или HHV-4), цитомегаловирус человека, герпесвирусы человека типов 6, 7, 8. У различных видов животных выделено более 70 видов герпесвирусов. Характерным свойством этих вирусов является способность оставаться в организме зараженного хозяина в латентном состоянии. Но некоторые из них, например HSV-1 и HSV-2, эффективно размножаются и быстро вызывают лизис клетки.

Все представители этого семейства имеют в составе вириона крупную двухцепочечную линейную молекулу ДНК размером 120–220 т. п. н., которая в ядре клетки упаковывается в белковый капсид и окружается оболочкой, которая образуется из ядерной мембраны.

В 1988 г. установлена полная нуклеотидная последовательность одного из штаммов вируса простого герпеса 1 — его длина составляет 152.260 п. н. геном HSV-1 представлен двумя сегментами — длинным (L) и коротким (S) (рис. 63). Они связаны ковалентно, при этом каждый из них содержит области уникальных последовательностей  $U_L$  и  $U_S$  соответственно, которые, в свою очередь, фланкированы инвертированными повторами  $TR_L$   $IR_L$   $IR_S$   $TR_S$ . L-сегмент по размеру включает 82 % вирусной ДНК. Сегмент «а» повторяется в геноме HSV-1 трижды и имеет длину 400 п. н. Предполагается, что в геноме вируса имеется 72 гена и три участка начала репликации. Белковые продукты семи генов необходимы для успешной репликации ДНК.

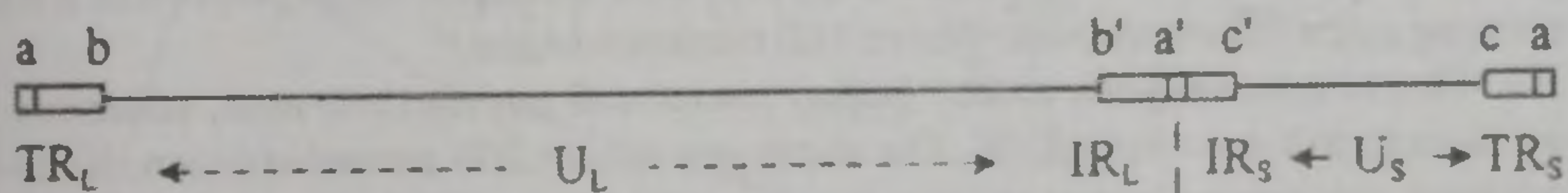


Рис. 63. Структура генома вируса простого герпеса

Гены вирусов герпеса объединяют в 5 групп ( $\alpha, \beta_1, \beta_2, \gamma_1, \gamma_2$ ). Их характерной чертой является то, что их экспрессия инициируется по каскадному типу: продукты  $\alpha$ -генов запускают транскрипцию  $\beta$ -генов (экспрессия генов  $\beta_1$  генов начинается раньше, чем генов  $\beta_2$ ), а продукты  $\beta$ -генов необходимы для транскрипции  $\gamma$ -генов.

Вирусы простого герпеса и ветряной оспы *нейротропны*, т. е. они находятся в нервной ткани зараженного организма на протяжении всей жизни в латентном состоянии. ДНК HSV-1 в латентном состоянии имеет коль-



цевую форму. Для других герпесвирусов характерна *лимфотропность*, т. е. они переходят в латентное состояние после инфицирования лимфоцитов.

## 7. Поксвирусы

Поксвирусы относятся к самым крупным и сложноорганизованным представителям империи доклеточных. Их геном имеет большую длину (175–230 т. п. н.) и кодирует около 200 полипептидов. Семейство Poxviridae состоит из двух подсемейств, патогенных для позвоночных (Chordopoxvirinae) и для беспозвоночных (Entomopoxvirinae). Семейство Chordopoxvirinae состоит из 8 родов. Наибольший интерес исследователей прикован к роду ортопоксвирусов, включающему патогенные для человека виды — вирус натуральной оспы, вирус оспы обезьян, вирус оспы коров и вирус осповакцины. Изучение этой группы направлено на картирование генов и определение их функций, выявление генов патогенности и их изменчивости, разработку живых поливалентных вакцин. Приоритет в изучении молекулярной генетики, геномики и биотехнологии этих вирусов завоевала группа исследователей под руководством профессора С. Н. Щелкунова в Государственном научном центре вирусологии и биотехнологии «Вектор». В 1992 г. эта группа вручную методом Максама—Гилберта осуществила секвенирование генома вируса натуральной оспы (размер 186 т. п. н.). К 1995 г. стали известны последовательности нуклеотидов ДНК наиболее крупных вирусов осповакцины (192 т. п. н.) и цитомегаловируса (229 т. п. н.).

Жизненный цикл поксвирусов протекает в цитоплазме клетки (рис. 52). При этом вирусный геном кодирует полную ферментную систему транскрипции, обеспечивающую синтез *mРНК* в *полиаденилированной, кэпированной и метилированной форме*. Вирионы зрелого внутриклеточного вируса содержат две мембранные структуры, а методом электрофореза в их составе идентифицировано более 100 полипептидов.

Геном поксвирусов имеет форму линейной двухцепочечной, ковалентно замкнутой концами ДНК. Он кодирует около 200 полипептидов. Кодирующие последовательности генов непрерывны, поэтому сплайсинг отсутствует. Экспрессия генов вируса начинается с *ранней экспрессии* сразу после вхождения вириона в клетку (*обозначается E*). для транскрипции *поздних генов (L)* необходима предшествующая репликация вирусной ДНК, а также экспрессия *ранних и промежуточных генов (I)*.

## 8. Ретровирусы

Ретровирусы — это обширная группа самых разнообразных вирусов семейства Retroviridae, включающего 7 родов. В состав рода Lentivirus входят патогенны иммунодефицита человека 1 и 2, иммунодефицита обезьян, крупного рогатого скота и кошек (отдельные виды). Отнесение всех этих

вирусов в единую классификационную группу основано на одинаковом способе репродукции. Геномы представителей этой группы имеют ряд общих черт:

- вирусная частица содержит две идентичные геномные одноцепочечные молекулы РНК длиной от 3,5 до 9 т. н. (это редкий, если не уникальный, случай диплоидного вирусного генома);
- РНК представляет собой (+) нить, при этом молекулы кэпированы на 5'-конце и полиаденилированы на 3'-конце;
- в молекулах РНК имеются уникальные 5'-(RU<sub>5</sub>)-3'-(RU<sub>3</sub>)-геномные концевые избыточности длиной несколько десятков нуклеотидов (рис. 64).

В состав вирионов ретровирусов входят 5–8 полипептидов, в том числе 1–3 гликопротеина оболочки. Внутри капсида кроме геномных молекул РНК имеются несколько десятков молекул обратной транскриптазы и молекулы тРНК.

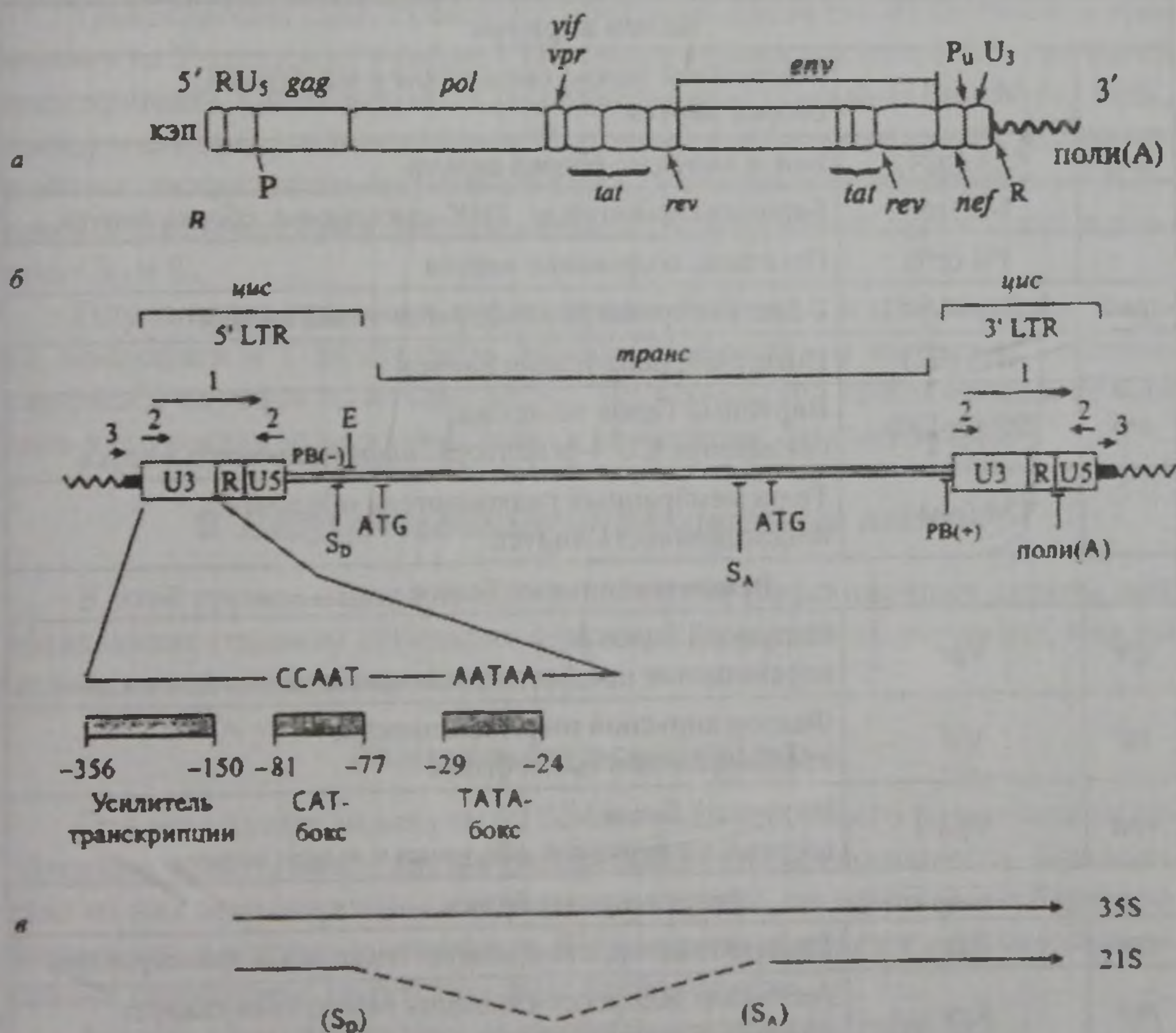


Рис. 64. Структура геномной РНК ретровируса (а), ДНК провируса (б) и единиц ее транскрипции (в)

В цитоплазме клетки на геномной РНК ретровируса с участием ревертазы синтезируются двухцепочечная копия ДНК. Она интегрируется в геном клетки-хозяина и образует так называемый *провирус*. Синтез молекулы ДНК на вирусной РНК инициируется *мРНК-праймером* на специфичном участке РВ(-) в 5'-концевой области вирусной РНК. На образовавшейся *цепи (-) ДНК* с участка РВ(+) синтезируется вторая комплементарная цепь (+) ДНК. На 5'- и 3'-концах образуются длинные концевые повторы LTR. Образовавшиеся молекулы ДНК находятся в ядре клетки в линейной или кольцевой форме. После интеграции вирусной ДНК в хромосому ее структура полностью соответствует схеме строения *транспозонов*.

Таблица 10

Белки, закодированные в геноме ВИЧ-1  
(Коничев, Севастьянова, 2003)

Ген	Белок	Функция
<b>Белки вириона</b>		
	МА (p17)	Матричный белок; связывание с мембраной, сборка вируса
<i>gag</i>	СА (p24)	Белок капсида; сборка вируса
	NC (p9)	Белок нуклеокапсида; РНК-связывание; сборка вируса
	PR (p11)	Протеаза; созревание вируса
<i>pol</i>	RT (p66/p51)	Обратная транскриптаза; репликация вируса
	IN (p32)	Интеграза; репликация вируса
<i>env</i>	SU (gp120)	Вирусный белок оболочки; связывание CD 4-рецептора, инфекционность вируса
	TM (gp41)	Трансмембранный гликопротеин оболочки; инфекционность вируса
<b>Вспомогательные белки</b>		
<i>vpr</i>	Vpr	Вирусный белок R; перемещение предынтеграционного комплекса в ядро.
<i>vif</i>	Vif	Фактор вирусной инфекционности; предынфекционные события
<i>vpr</i>	Vpr	Вирусный белок U; созревание вирусной оболочки и выход вируса
<b>Регуляторные белки</b>		
<i>tat</i>	Tat	Трансактиватор; стимулирует элонгацию транскрипции
<i>rev</i>	Rev	Регулятор экспрессии поздних генов; стимулирует появление несплайсированных мРНК в цитоплазме
<i>nef</i>	Nef	Отрицательная регуляция CD4; регулирует сигнальные пути Т-клеток

В настоящее время расшифрована структура генома вируса иммунодефицита человека 1 (ВИЧ-1). В нем закодированы белки, представленные в табл. 10. Геном содержит 9 генов, но в ходе *процессинга* вирусными и клеточными ферментами синтезируется 14 белковых продуктов.

Ген *gag* кодирует большой *белок-предшественник* всех основных белков вириона, из него протеаза вырезает конкретные белки. Кодирующий участок *pol* детерминирует синтез обратной транскриптазы — полифункционального фермента, способного синтезировать ДНК на матрице РНК или ДНК. Этот же ген кодирует еще *интегразу*, которая осуществляет включение ДНК вируса в геном клетки-хозяина.

Ген *env* кодирует гликопротеины вирусной оболочки, которая способна захватывать часть цитоплазматической мембраны клетки.

Представлены также гены вспомогательных и регуляторных белков: белки Tat и Rev обеспечивают значительное увеличение синтеза вирусных белков при переходе провируса из латентной в активную форму.

Транскрипция ДНК провируса протекает после его интеграции в хромосому на 5'-концевом участке LTR, где находятся промотор и терминатор транскрипции. Синтезируются полноразмерные геномные молекулы РНК, претерпевающие затем сплайсинг. При этом из полноразмерной РНК выщепляются последовательности *генов gag* и *pol*, а ген *env* сближается с 5'-концевой последовательностью вирусной РНК. Сплайсинг происходит в участках  $S_D$  и  $S_A$ .

Ретровирусы способны инфицировать эмбрионы и кроветворные клетки, макрофаги и Т-лимфоциты. Многие ретровирусы вызывают лейкозы, саркомы и опухоли молочных желез. Их небольшой геном позволяет вводить в его состав чужеродные гены и экзогенные фрагменты ДНК.

## 9. Вирусоподобные инфекционные агенты

К этой группе относятся неканонические инфекционные агенты, или обладающие геномом небольшого размера (*сателлиты*, *вириоды*), или им полностью не обладающие (*прионы*).

### Сателлиты (вирусы-сателлиты)

Это небольшие молекулы РНК, которым для своего размножения необходимо присутствие в клетке специального вируса-помощника. Большинство из них ассоциированы с вирусами растений, но некоторые с бактериофагами или с вирусами животных. К последним относятся *аденоассоциированные сателлиты*, сопутствующие аденовирусам.

Можно отметить типичные свойства сателлитов (табл. 11):

- геном состоит из одноцепочечной РНК размером 500–2000 нуклеотидов;
- первичная структура генома не сходна с геномами вирусов-помощников;

- они вызывают специфические заболевания у растений, но при инфицировании растений только вирусом-помощником болезнь не обнаруживается.

### Вироиды

Другими неканоническими инфекционными патогенами являются ви­роиды, возбудители 35 инфекционных болезней растений (вероятно, и жи­вотных). Прототипной виroidной инфекцией является *веретеновидность клубней картофеля*. Геном представляет собой молекулу односпиральной, ковалентно замкнутой РНК, с участками из двухспиральных сегментов. Размер молекулы 200–400 нуклеотидов, М. м.\* 120–170 кДа. По ряду харак­теристик виroidы отличаются от сателлитов (табл. 11).

Вироиды не кодируют никаких белков и реплицируются при помощи РНК-полимеразы клетки-хозяина. Предположительно, процесс протекает по механизму «*катящегося кольца*» с последующим *аутокаталитическим расщеплением и лигированием*.

Разные формы виroidов различаются по первичной структуре, что может лежать в основе их разделения на отдельные виды. В то же время в центральной части виroidной РНК имеется общая последовательность нуклеотидов — *центральный консервативный домен*, важный для процес­са репликации (рис. 65).

Показано сходство вирионов по первичной структуре с некоторыми нук­леотидными последовательностями генома эукариотической клетки-хозяи­на — *с интроном, в локусе между 5,8S и 25S рРНК*. Поэтому не исключает­ся, что виroidы могут нарушать *посттранскрипционный процессинг* РНК-инфицированной клетки.

Таблица 11

Характеристика сателлитов и виroidов

Характеристика	Сателлиты	Вироиды
Для репликации необходим вирус-помощник	Да	Нет
Геном кодирует белки	Да	Нет
Геном реплицируется при помощи:	ферментов вируса-помощника	РНК-полимера-зы клетки-хозяина
Место репликации	То же, что и вируса-помощника	Ядро

\* М. м. — молекулярная масса в кило-Дальтонах (кДа).

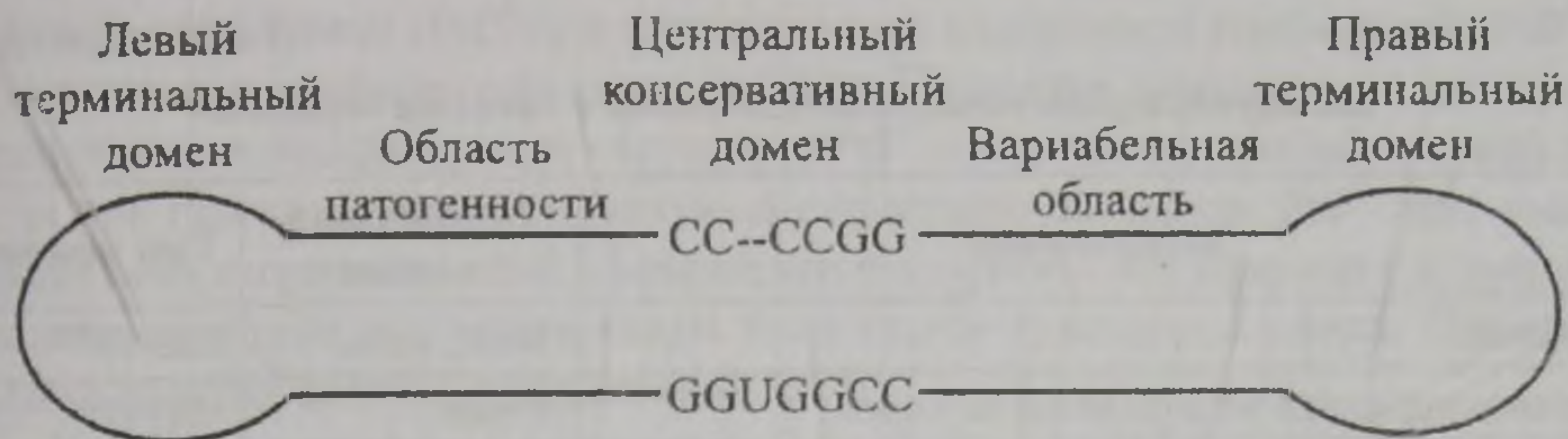


Рис. 65. Схематическая структура молекулы вирионной РНК (Зинченко, Паруль, 2005).

Описан уникальный *вирус гепатита дельта (ВГД)*, у которого РНК имеет свойства и вириона, и сателлита. Вирусом-помощником, обеспечивающим его репликацию, является *вирус гепатита В (ВГВ)*. ВГД находят в клетках, инфицированных ВГВ. Он покрыт защитной оболочкой из липида и белка ВГВ. Из клеток, инфицированных ВГВ/ВГД, выделяют частицы — капсид ВГВ, содержащий ковалентно замкнутую кольцевую РНК ВГД. В отличие от других вирионов, РНК ВГД кодирует *белок дельта-антиген*, который является фосфопротеином. Как полагают, геном ВГД реплицируется путем РНК-зависимого синтеза РНК с использованием *ДНК-зависимой РНК-полимеразы*. ВГД усиливает патогенные эффекты ВГВ.

### Прионы

Прионы представляют собой инфекционные агенты, вызывающие трансмиссионные спонгиозные (от «спонги» — губка) энцефалопатии человека и животных. Среди них: скрепи овец (от англ. scarp — скоблить, тереть), губчатая энцефалопатия крупного рогатого скота (коровье бешенство), трансмиссивная энцефалопатия норки, куру человека и др. (табл. 12). Сам термин «прион» — это *слово-акроним*, образованное по аналогии со словом вирион (от англ. proteinaceous infection particle). При всех прионных заболеваниях в мозгу у больных накапливаются отложения белка *амилоида (парамилоида)*, при этом ткань мозга принимает пористую форму наподобие губки.

Инфекционным агентом прионов является *белок-конформер (изоформа)*, получивший название PrP<sup>Sc</sup> (от англ. Scarpie). Прузинер (1984) установил, что в геноме млекопитающих имеется *ген prp*, кодирующий белок, по первичной структуре идентичный PrP<sup>S</sup>. В нормальных клетках его обозначают PrP<sup>C</sup> (рис. 66). Прионные гены кртированы у человека (20-я хромосома) и у мыши (2-я хромосома).

Таблица 12

## Международная номенклатура типов в таксоне «прионы»

Заболевание	Природный хозяин	Тип приона
Скрепи	Овцы, козы	ShePrP <sup>Sc</sup>
Трансмиссивная энцефалопатия норок	Норки	MkPrP <sup>Sc</sup>
Хроническая изнуряющая болезнь	Мулы, лоси	MDePrP <sup>Sc</sup>
Губчатый энцефалит коров (ГЭК)	Кошки	FePrP <sup>Sc</sup>
Энцефалопатия экзотических (редких) животных	Ниала, куду	NyaPrP <sup>Sc</sup>
Куру	Человек	HuPrP <sup>Sc</sup>
БКЯ	Человек	HuPrP <sup>Sc</sup>
нвБКЯ	Человек	HuPrP <sup>Sc</sup>
Синдром Гертсмана — Штреусслера — Шейнкера (ГШШ)	Человек	HuPrP <sup>Sc</sup>
Семейная смертельная бессонница (ССБ)	Человек	HuPrP <sup>Sc</sup>

Прионовый белок — это гидрофобный и термостабильный белок (инактивируется только при автоклавировании в течение 30 мин при 135° С) с длиной полипептидной цепи 253–254 а. о. и с М. м. 27–30 кДа. Мутанты, у которых ген *prp* отсутствует или поврежден, не восприимчивы к прионовой инфекции. Физиологические функции прионового белка сводятся к защите нейронов от оксидативного стресса.

Между PrP<sup>C</sup> и прионом не обнаружено никаких химических различий. Возникло предположение, что *фундаментальное отличие* между инфек-

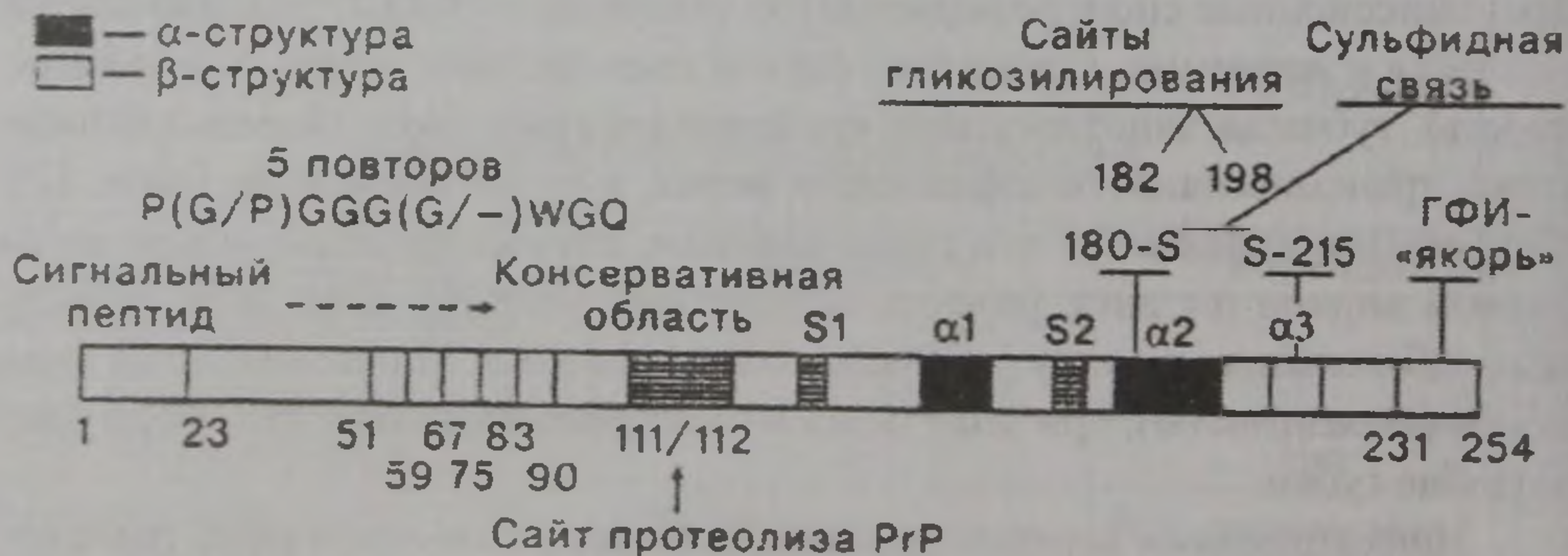


Рис. 66. Карта прионового белка (Супотницкий, 2007). Координаты: 1–23 — сигнальный пептид; 51–90 — глициновый октаповтор; 111–112 — сайт протеолиза в консервативной последовательности; 231–254 — С-концевая сигнальная последовательность. S1 и S2 — β-домены, ГФИ — глюкофосфатидинозитол, мембранный «якорь»

ционной изоформой ( $\text{PrP}^{\text{Sc}}$ ) и непатогенной эндогенной изоформой ( $\text{PrP}^{\text{C}}$ ) заключается в конформации этих молекул. Считается доказанным, что пространственная конфигурация структуры  $\text{PrP}^{\text{C}}$  представляет собой  $\alpha$ -спираль, тогда как прион имеет складчатую  $\beta$ -структуру. Молекула  $\text{PrP}^{\text{C}}$  при взаимодействии с прионом сама превращается в прион. Это приводит к амплификации прионов и исчезновению пула белка-предшественника. Процесс носит характер цепной реакции (рис. 67).

Конверсия изоформ прионного белка — это необратимое превращение  $\text{PrP}^{\text{C}}$  в  $\text{PrP}^{\text{Sc}}$  в ходе инфекции вследствие нарушения кинетически контролируемого равновесия между изоформами прионного белка, которое в норме регулируется как  $\text{PrP}^{\text{C}} \rightleftharpoons \text{PrP}^{\text{Sc}}$ . Медленный процесс конверсии значительно ускоряется в присутствии большого количества конвертирующей изоформы или экзогенного приона.

При межвидовом заражении наблюдается пониженная инфекционность чужого приона. Это обусловлено снижением способности белка  $\text{PrP}^{\text{Sc}}$  передавать свою инфекционность белку  $\text{PrP}^{\text{C}}$ . Первичная структура при-

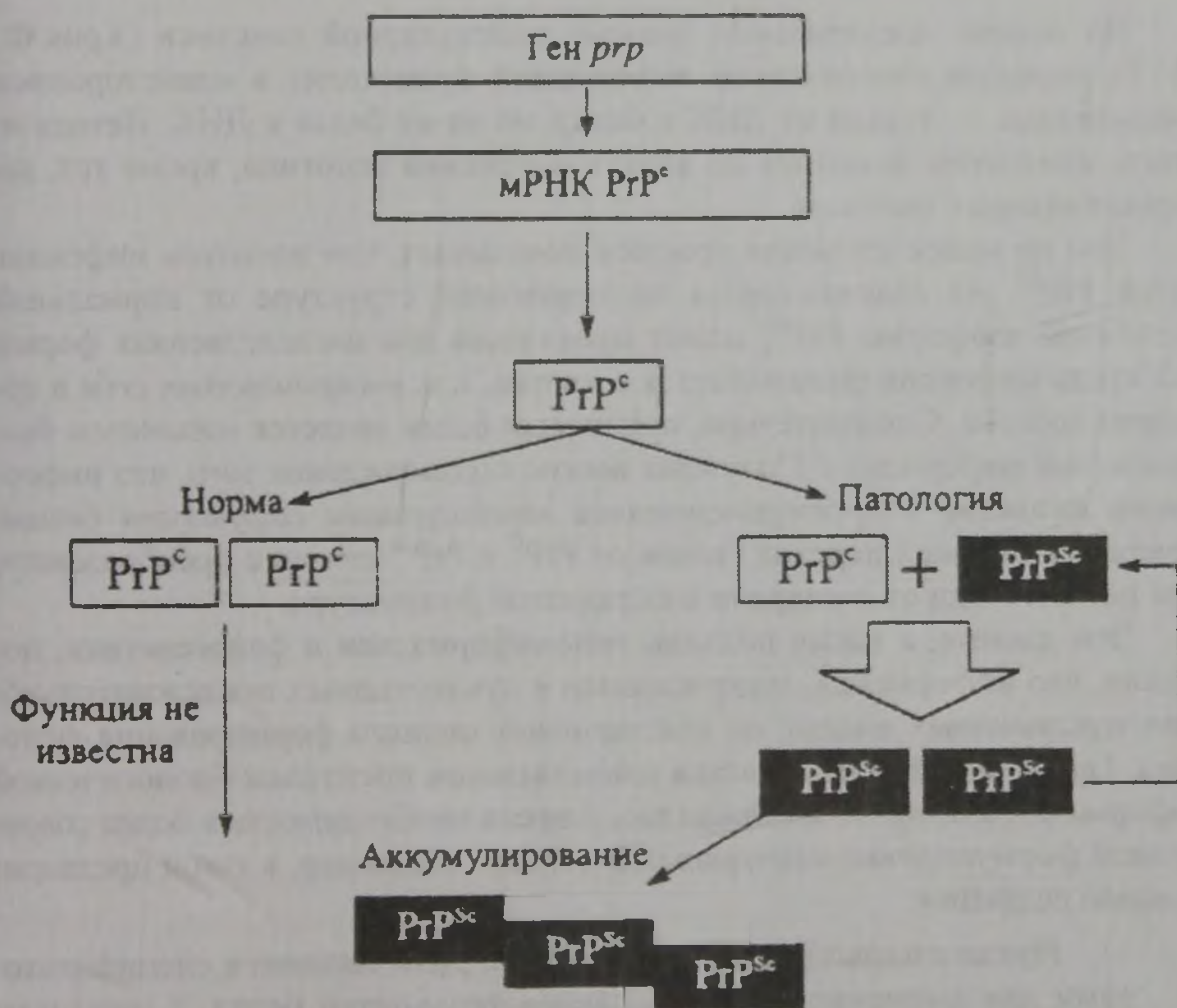


Рис. 67. Амплификация приона в ходе конверсии белка  $\text{PrP}^{\text{C}}$



онов высоко консервативна, с незначительными различиями у разных видов млекопитающих. Но межвидовые барьеры на пути передачи все же существуют. Получены данные о существовании прионов у пекарских дрожжей *Saccharomyces cerevisiae*, связанное с нехромосомно наследуемыми детерминантами у этого вида. В итоге можно сформулировать следующее определение:

*Прионы* — это белки, способные посредством взаимодействия друг с другом передавать свое *конформационное состояние* от одной молекулы к другой. В таком значении понятие «прионы» может быть распространено на белки дрожжей и аналогичные белки других эукариот.

В настоящее время изучение прионов приобретает не только теоретический, но и все более прикладной характер в связи со вспышками эпизоотий среди сельскохозяйственных животных в европейских странах. Изучается также возможность передачи этих болезней от животных к человеку.

## Заключение

На основе «центральной догмы» молекулярной генетики (Крик Ф., 1957), передача генетической информации происходит в одностороннем направлении — только от ДНК к белку, но не от белка к ДНК. Исходя из этого, изменения фенотипа не влекут изменения генотипа, кроме тех, которые связаны с геномом.

Тем не менее изучение прионов показывает, что носитель инфекции белок PrP<sup>Sc</sup>, не отличающийся по первичной структуре от нормальной клеточной изоформы PrP<sup>C</sup>, может проявиться как наследственная форма. Носитель инфекции размножается в клетке, т. е. *воспроизводит себя* в организм-хозяине. Следовательно, прионовый белок является *носителем биологической информации*. Получены веские подтверждения того, что информация заложена в *пространственной конфигурации структуры белков*. Конформационный переход белков от PrP<sup>C</sup> к PrP<sup>Sc</sup> связан с преобразованием полипептида от  $\alpha$ -спирали к складчатой  $\beta$ -структуре.

Эти данные, а также подходы геноинформатики и фенотипетики, показали, что информация, содержащаяся в нуклеотидных последовательностях нуклеиновых кислот, не обеспечивает полного формирования фенотипа. Геномная ДНК не является единственным носителем биологической информации в онтогенезе. Очевидно, созрела необходимость в более современной формулировке «центральной догмы», например, в такой предварительной редакции:

Нуклеотидные последовательности ДНК являются спецификаторами для аминокислотных последовательностей белка, а структура первых не определяется структурой вторых (Sarkar S., 2000; Каменская Н. А., 2006).

В классический (домолекулярный) период развития зоологии и ботаники сформулированы критерии понятия «живые организмы», которым вирусы как представители доклеточных не удовлетворяют, так как они не способны воспроизводиться самостоятельно. С позиции современных молекулярно-генетических представлений можно предполагать о существовании *континуума организмов возрастающей сложности*, в котором прионы и вирусы занимают место между живым и неживым.

### Ключевые слова

Вирион, капсид, бактериофаги, лизогения, лизис клетки, сателлиты, вириоды, прионы, секвенирование, секвенатор, биология *in silico*, биоинформатика, геноинформатика, протеомика, электронные базы данных.

## Геномика прокариот

Большинство фундаментальных фактов молекулярной биологии получено при использовании микроорганизмов в качестве биологического объекта. Установлена роль ДНК, как материального носителя наследственности, расшифрован генетический код, изучены принципы репликации, транскрипции и трансляции. Все современные технологии манипуляций с нуклеиновыми кислотами целиком базируются на использовании ферментов, полученных из микроорганизмов (например, рестриктазы, полимеразы). Возможности изучения этого мира далеко не исчерпаны... Современные молекулярно-биологические методы в исследовании микроорганизмов могут привести к неожиданным открытиям и принести весомые плоды в виде неисчерпаемого резервуара новых генов и новых принципов организации живых систем.

Академик В. Г. Дебабов

Прокариоты таксономически составляют подимперию Prokarya, которые наряду с эукариотными организмами (подимперия Eucarya) входят в состав империи клеточных Cellulata. Взаимоотношения между этими основными группами организмов отражены на рис. 38, 39 и 40, а в табл. 2 представлено их сравнение по молекулярно-биологическим признакам. Исследования выдающегося американского микробиолога Карла Вёзе и его школы (Woese C. et al., 1977, 1990) на генах малой субъединицы рибосомной РНК (16S рРНК) из различных организмов привели к разделению микроорганизмов-прокариот на отдельные *домены (надцарства)*: Eubacteria — истинные бактерии (зубактерии) и Archaeobacteria — архебактерии (архен). Отдельно стоят ядерные организмы — домен Eucarya.

Каждый домен (надцарство) содержит по несколько царств: археи — царства Euryarchaeota (метаногенные археобактерии), Crenarchaeota (экстремальные термофилы), и Kogarchaeota (обитатели горячих источников). У эубактерий под влиянием молекулярных исследований филогенеза разные авторы выделяют от 9 до 16 филогенетических групп. За последнее столетие описано около 5000 видов бактерий, а это сравнительно немного, например, по сравнению с насекомыми, у которых описано около 500 000 видов. Одной из причин может служить методология работы с микробами: прежде чем их охарактеризовать, они должны быть переведены в чистую лабораторную культуру, что часто оказывается трудным или даже невозможным. Методы прочтения последовательностей генов 16S рРНК может привести к увеличению в несколько раз крупных таксонов.

Это и не удивительно, так как именно микроорганизмы составляют большую часть биомассы планеты, а их вклад в биологическое и генетическое разнообразие жизни на Земле превышает 90 %. Они являются основными производителями кислорода, обеспечивают плодородие почвы, находятся в основе многих пищевых цепей. Карл Вёзе выразил это такими словами: «...В фундаментальном смысле биосфера — это бактосфера» (1996).

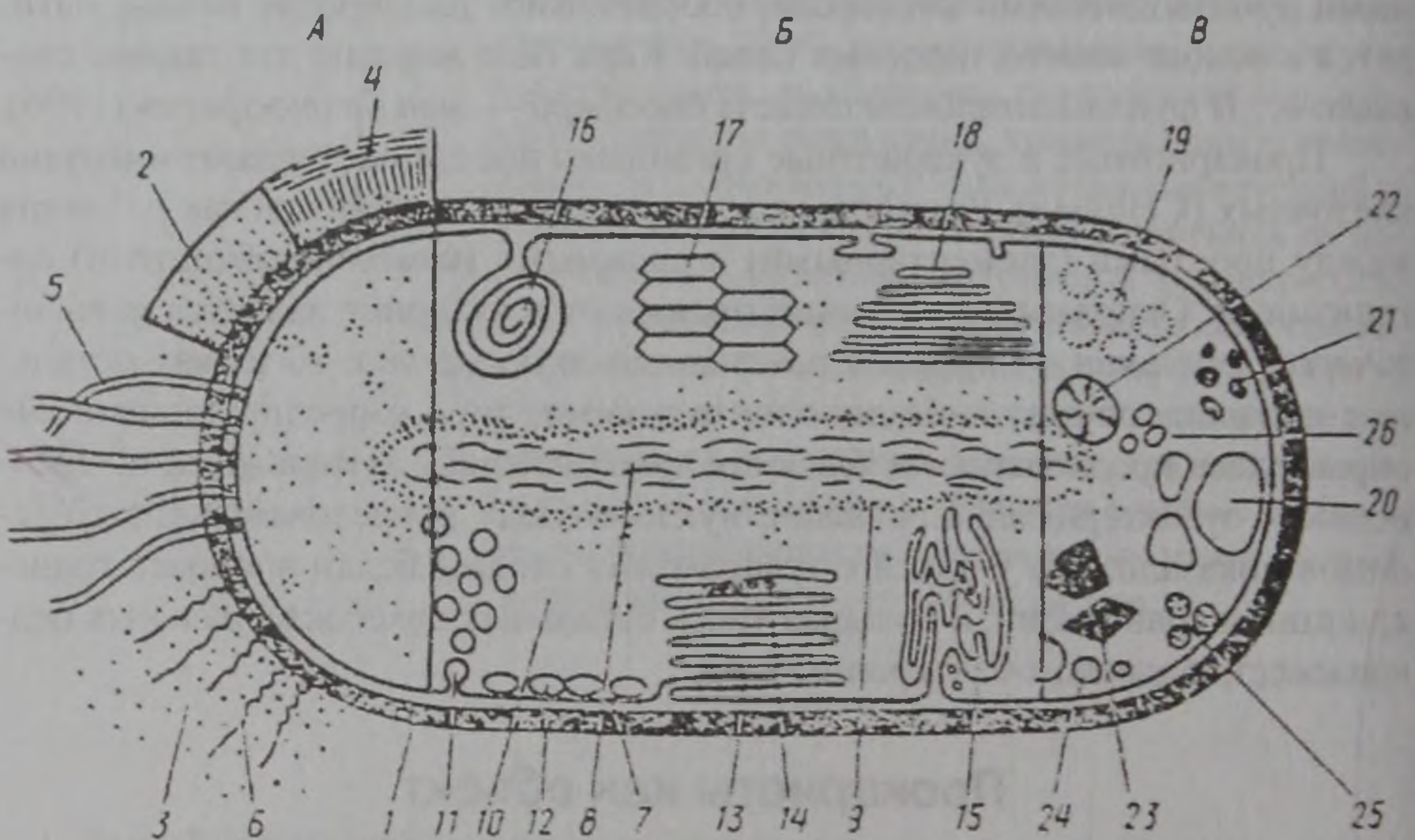
Прокариотные и эукариотные организмы вместе составляют империю клеточных (Cellulata). Различия между ними можно провести как различия между простыми (элементарными) и сложными (комбинированными) организмами. Основным в различии прокариот и эукариот является *цитологическое различие в строении генетического аппарата*, тогда как остальные признаки служат либо дополнительными, либо коррелирующими. Не оправдалась предполагаемая близость археобактерий с эукариотами по сравнению с эубактериями. Сравнение нуклеотидных последовательностей геномов показало, что у архей с эукариотами сходны белки аппарата транскрипции и трансляции, а большая часть остальных археобактериальных белков имеет сходство с бактериальными.

## Прокариоты как объект молекулярно-генетических исследований

Большая часть микроорганизмов представляют собой одноклеточные формы с большим диапазоном размеров. В среднем линейные размеры бактерий лежат в пределах 0,5–3,0 мкм. Самые мелкие бактерии группы микоплазм имеют диаметр клеток 0,1–0,15 мкм, т. е. размер, близкий к теоретическому пределу клеточного уровня организации жизни. Диаметр клеток нитчатой бактерии *Beggiatoa alba* равен 50 мкм, а длина клетки спирохеты достигает 250 мкм.

Химическая и ультраструктурная организация клетки прокариот имеет ряд особенностей (рис. 68). Внешний липопротеидный слой протопла-

ста бактерий называется *цитоплазматической мембраной (ЦПМ)*. Структуры, расположенные снаружи от ЦПМ (клеточная стенка, капсула, жгутики, ворсинки), называют *поверхностными структурами*. В клеточную оболочку входят все слои, располагающиеся с внешней стороны от ЦПМ. *Протопласт* состоит из ЦПМ и цитоплазмы. Генетический аппарат прокариот представлен двойной замкнутой нитью ДНК (бактериальная хромосома = нуклеоид), не отделенной от цитоплазмы мембраной. Гистоны и митотический аппарат отсутствуют, нет интронов. Отсутствуют органеллы, свойственные эукариотам — хлоропласты, митохондрии. Кольцо ДНК закреплено в одной точке на внутренней стороне клеточной мембраны. Деление нуклеоида происходит после репликации нити ДНК, при этом расхождение дочерних нуклеоидов обеспечивается ростом клеточной мембраны. Нуклеоид представляет собой тонкую, неправильную, фибриллярную сеть из ДНК, часть располагающуюся параллельно осевой линии клетки. ДНК-связывающей структурой выступает мезосома.



**Рис. 68.** Комбинированное изображение прокариотной клетки. А — поверхностные клеточные структуры и внеклеточные образования: 1 — клеточная стенка; 2 — капсула; 3 — слизистые выделения; 4 — чехол; 5 — жгутики; 6 — ворсинки; Б — цитоплазматические клеточные структуры: 7 — ЦПМ; 8 — нуклеотид; 9 — рибосомы; 10 — цитоплазма; 11 — хроматофоры; 12 — хлоросомы; 13 — пластинчатые тилакоиды; 14 — фикобилисомы; 15 — трубчатые тилакоиды; 16 — мезосома; 17 — аэросомы (газовые вакуоли); 18 — ламеллярные структуры; В — запасные вещества: 19 — полисахаридные гранулы; 20 — гранулы поли- $\beta$ -оксимасляной кислоты; 21 — гранулы полифосфата; 22 — цианофициновые гранулы; 23 — карбоксисомы (полиэдральные тела); 24 — включения серы; 25 — жировые капли; 26 — углеводородные гранулы (по Schlegel, 1972)

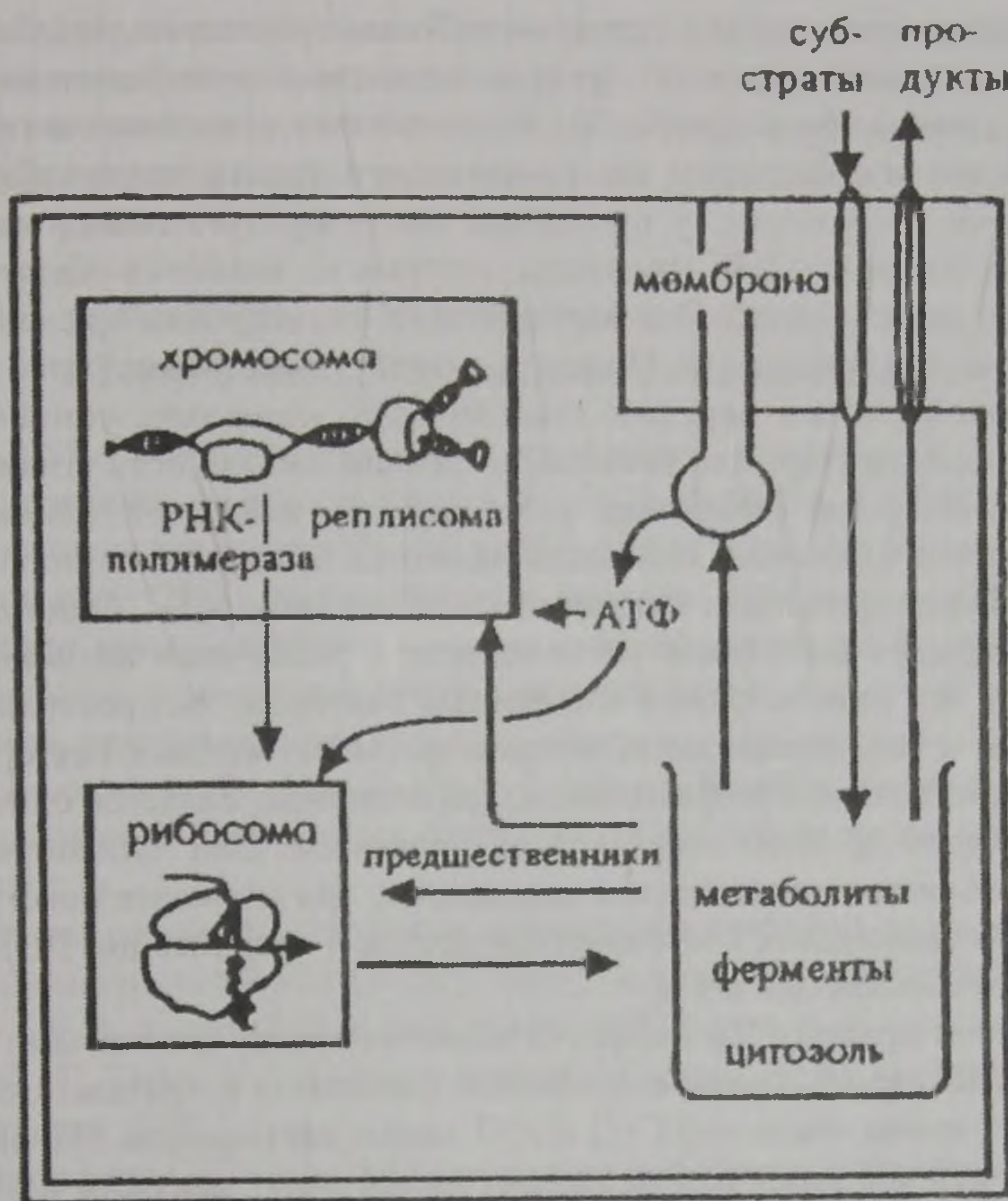


Рис. 69. Основные подсистемы прокариотной клетки

Г. А. Заварзин (2001) рассматривает прокариотную клетку как систему, состоящую из взаимодействующих подсистем. Схематически они выглядят следующим образом (рис. 69):

1. Геном — хромосома, содержащая генетическую информацию для собственной репликации за счет азотистых оснований и, во-вторых, информацию для синтеза белка.
2. Аппарат синтеза — это рибосома, представляющая РНК-содержащий аппарат синтеза белка из аминокислот-предшественников.
3. Цитозоль — растворимая часть цитоплазмы, включающая сеть метаболических путей с обслуживающими их ферментами.
4. Мембрана включает энергодающий аппарат синтеза АТФ и транспортной системы, осуществляющие взаимодействие клетки с внеклеточной средой.

Благодаря различиям в строении рибосом эубактерии, архебактерии и эукариоты проявляют разную чувствительность к антибиотикам — ингибиторам синтеза белка (табл. 3). Это позволяет применять антибиотики, действующие на эубактерии, для селективного накопления архебактерий.

Помимо хромосомы, у прокариот могут присутствовать небольшие кольцевые молекулы ДНК-плазмиды, которые не являются обязательными для генома данного вида. Они варьируют по размеру и по численности от одной до многих десятков. Плазмиды определяют важнейшие функции клетки: способность к передаче генетического материала, устойчивость к лекарственным препаратам, тяжелым металлам, способность утилизировать токсичные вещества (например, углеводороды, нафталин, толуол, и др.). Благодаря этим свойствам плазмиды являются мощными факторами адаптации микроорганизмов к меняющимся условиям среды. Одна плазида может содержать до 5 генов устойчивости к различным антибиотикам и переносить эти гены в клетки патогенных бактерий. Распространение устойчивости к различным антибиотикам форм патогенных бактерий, препятствующее лечению инфекционных заболеваний, является одной из самых серьезных проблем современной медицины. Они используются как удобные объекты в генетической инженерии, для изучения молекулярных механизмов важнейших клеточных процессов — репликации ДНК, транскрипции, рекомбинации и т. д.

Согласно правилу Чаргаффа, отношение между азотистыми основаниями в ДНК строго стехиометрические (находятся в кратных отношениях). В то же время число пар Г+Ц и А+Т может варьировать. По числу преобладания тех или иных пар различают организмы с высоким или низким молекулярным содержанием ГЦ (ГЦ, моль %), величины которого обычно приводятся для конкретного вида.

## Структурная геномика прокариот

Термин «геном» появился примерно 80 лет назад. Он произведен от слов Gene и chromosOME и обозначает полный набор генов и хромосом. Термин «геномика» появился в 1986 г. и относится к науке, занимающаяся картированием и секвенированием генов. В настоящее время геномику подразделяют на две составные части — *структурную и функциональную*. Задача последней — развитие новых технологий для использования структурной информации.

Результаты проведенного в последние годы секвенирования нуклеотидных последовательностей, а также генетического и физического картирования позволяют рассматривать структуру *полных геномов прокариот*, выявить отдельные гены, определить структуру оперонов, провести поиск семейств родственных генов. Полные нуклеотидные последовательности

генов дают теоретическую возможность реконструировать метаболизм биохимически неохарактеризованного вида на основе нуклеотидной последовательности его генома, по наличию тех или иных ферментативных функций и структурного сходства между генными продуктами изучаемого вида и белками, функция которых известна.

### 1. Размеры, нуклеотидный состав геномов и оперонная организация генов прокариот

До 1955 г. в микробиологии было принято считать, что хромосомы прокариот имеют линейную форму. Но Э. Л. Вольман и Ф. Жакоб в опытах по картированию генов у *E. coli* (1955) установили, что хромосома представлена нитью ДНК, погруженную в цитоплазму и замкнутую в кольцо. В настоящее время методами прямого анализа физической структуры показано наличие у некоторых видов бактерий кроме кольцевой хромосомы одного или нескольких репликонов, которые названы *хромосомами или мегаплазмидами* (табл. 13). Геном у бактерии *Borrelia tumefaciens*, возбудителя клещевого спирохетоза, представлен линейной хромосомой размером 1 млн п. н.

Размеры генома у прокариот показывают большой диапазон колебаний. Установлен самый маленький геном у микоплазмы *M. genitalium* размером 580 т. п. н. и самый большой геном 9500 т. п. н. у *Mucosoccus xanthus* (табл. 14, 15). Геном *E. coli* имеет размер 4,6 млн п. н.,  $M. m.$   $3 \cdot 10^9$  Да и длиной молекулы ДНК 1,5 мм.

Содержание GC-пар в ДНК прокариот колеблется в пределах 23–72%. У некоторых бактерий (*E. coli*, *M. genitalium*, *Bacillus subtilis*, *Helicobacter pylori* и др.) установлено преимущественное присутствие G по сравнению с C в матричной (ведущей) нити ДНК от области начала репликации *ori* до

Таблица 13

Формы хромосом у прокариот

Бактерия	Геном
<i>Agrobacterium tumefaciens</i> CSS	1 линейная хромосома; 1 кольцевая хромосома; 2 плазмиды
<i>Bacillus cereus</i> F0836 76	1 кольцевая хромосома; 1 мегаплазида
<i>Borrelia melioides</i>	2 кольцевые хромосомы
<i>Leptospira interrogans</i>	1 кольцевая хромосома; 1 мегаплазида
<i>Rhizobium meliloti</i>	1 кольцевая хромосома; 2 мегаплазмиды
<i>Rhodobacter sphaeroides</i>	2 кольцевые хромосомы
<i>Rhodococcus facians</i>	1 линейная хромосома; линейная гтлазида
<i>Streptomyces ambofaciens</i>	1 линейная хромосома



участка терминации *ter*. Напротив, у других видов — цианобактерии и архебактерии (*Methanococcus*, *Methanobacterium*, *Archaeoglobus*) — асимметрия нуклеотидного состава не наблюдается. Наличие асимметрии может указывать на различия в репликативном синтезе ведущей (матричной) и ведомой (смысловой нитей).

Порядок расположения генов в бактериальной хромосоме не случаен. Гены одной определенной функции располагаются один за другим и организованы в оперон с совместной транскрипцией на одну РНК (рис. 22Б). У разных видов гены и опероны расположены различным образом. При сравнении геномов *E. coli* и *B. subtilis*, имеющих 1000 общих генов, установлено, что 100 генов находятся в составе одних и тех же оперонов, хотя порядок генов в составе оперонов не всегда одинаков. Анализ полностью секвенированных геномов бактерий показал, что 16 кластеров генов у эубактерий остаются консервативными, из них у архебактерий 8 также консервативны.

Таблица 14

Характеристики геномов, определенные по их полной нуклеотидной последовательности (Боринская, Янковский, 2004)

Видовое название	Домен	Размер генома (п, н.)	% кодирующих последовательностей	% GC	Количество ORF	Характеристика (для термофилов указана оптимальная температура роста $T_{opt}$ )
<i>Mycoplasma genitalium</i>	В	580 070	89,7	32	479	Вызывает воспаления мочеполовых путей, грам+*
<i>Mycoplasma pneumoniae</i>	В	816 394	88,7	40	677	Возбудитель пневмонии, грам+*
<i>Borrelia burgdorferii</i>	В	910 725 <sup>#</sup> 533 000 <sup>##</sup>	93	28,6 <sup>#</sup> от 23,6 до	853 <sup>#</sup> 430 <sup>##</sup>	Возбудитель клещевого спирохетоза (болезнь Лайма), грам-. Хозяева: грызуны, иксодовые клещи, человек
<i>Chlamidia trachomatis</i>	В	1 042 519 + плазмида 7493		41,3	894	Вызывает трахому и воспаление мочеполовых путей, грам-**
<i>Rickettsia prowazekii</i>	В	1 111 523	76	29,1	834	Возбудитель сыпного тифа, грам-
<i>Treponema pallidum</i>	В	1 138 006	92,9	52,8	1041	Возбудитель сифилиса, грам-
<i>Chlamidia pneumoniae</i>	В	1 230 230		40,6	1073	Возбудитель пневмонии, грам-**
<i>Aquifex aeolicus</i>	В	1 551 335 плазмида 39 456 (2 копии)	93	43,4 36,4	1512 32	Хемолитотроф, микроаэрофил, термофил, $T_{opt}$ 85°C, грам-

Окончание Таблицы 14

Видовое название	Домен	Размер генома (п, н.)	% кодирующих последовательностей	% GC	Количество ORF	Характеристика (для термофилов указана оптимальная температура роста $T_{opt}$ )
<i>Methanococcus jannaschii</i>	A	1 660 000 плазмиды: 58 000 16 000		31,4 28,2 28,8	1738 44 12	Анаэроб, метаноген, обитает в глубоководных термальных источниках, $T_{opt}$ 85°C
<i>Pyrococcus horikoshii</i>	A	1 738 505	90,7	41,9	2061	Анаэроб, гипертермофил, $T_{opt}$ 98°C
<i>Methanobacterium thermoautotrophicum</i>	A	1 751 377			1855	Термофил, $T_{opt}$ 65°C
<i>Haemophilus influenzae</i>	B	1 830 137		38	1073	Вызывает отиты, ОРЗ, возбудитель менингитов, грам–
<i>Thermotoga maritima</i>	B	1 860 725	95	46	1877	Один из древнейших видов зубактерий, $T_{opt}$ 65°C
<i>Archeoglobus fulgidus</i>	A	2 178 400	92,2	48,5	2436	Метаболизирует серу, $T_{opt}$ 83°C
<i>Synechosystis</i> sp. PCC6803	B	3 573 470	87		3168	Фотосинтезирующая цианобактерия, грам–
<i>Bacillus subtilis</i>	B	4 214 810	87	43,5	4100	Автотроф, грам–
<i>Mycobacterium tuberculosis</i> H37Rv	B	4 411 529	91	65,6	4000	Возбудитель туберкулеза, грам–
<i>Escherichia coli</i> K-12	B	4 639 221	88,6	50,8	4288	Прототрофная энтеробактерия, грам–
<i>Saccharomyces cerevisiae</i>	E	12 068 000 (16 хромосом)			5885	

## Примечания.

A — археи, B — бактерии, E — эукариоты. Обозначения грам+ и грам– указывают тип клеточной стенки.

# Данные относятся к линейному геному *B. burgdorferii*.

## Суммарные данные по 11 плазмидам *B. burgdorferii*, нуклеотидная последовательность которых определена. Всего в клетке боррелии может присутствовать до 20 различных плазмид.

\* Относится к ветви грам+ бактерий, не имеет клеточной стенки.

\*\* Не имеет муреинового слоя, клеточная стенка грам– типа.

Сохранение генов в составе одного оперона в ходе эволюции может давать некоторое преимущество. Оперонная организация генов необходима для обеспечения координированной экспрессии генов. С другой стороны,

кластеризация генов, выполняющих одну функцию, повышает вероятность их совместного распространения путем горизонтального переноса видам, утратившим существенную функцию. Оперонная организация генов у термофильных бактерий (растут при температуре выше 50° С) отличается от других бактерий. Тогда как у зубактерий объединены в опероны гены одного и того же метаболического пути, у термофильной бактерии *A. aeolicus* они разбросаны по геному. В опероны объединены гены электротранспорт-

Таблица 15

## Секвенирование геномов патогенных организмов

Организм	Вызываемое заболевание	Размер генома, млн пн	Секвенирующее учреждение	Год публикации
<b>Секвенированные геномы</b>				
<i>Haemophilus influenzae</i>	Менингит, пневмония		TIGR <sup>1</sup>	1995
<i>Mycobacterium tuberculosis</i>	Туберкулез	4,4	Сэнгер-центр, Институт Пастера	1998
<i>Campylobacter jejuni</i>	Диарея	1,6	Сэнгер-центр	2000
<i>Escherichia coli</i> 0157 (два штамма)	Пищевое отравление	5,5	Университет Висконсина, Японский консорциум	2000
<i>Vibrio cholerae</i>	Холера	4,0	TIGR	2000
<i>Mycobacterium leprae</i>	Проказа	33	Сэнгер-центр	2000
<i>Neisseria meningitidis</i>	Менингит	2,3	TIGR, Сэнгер-центр	2000
<i>Streptococcus pneumoniae</i>	Пневмония	2,2	TIGR	2001
<i>Yersinia pestis</i>	Чума	4,7	Сэнгер-центр	2001
<i>Salmonella typhi</i> (CT18)	Брюшной тиф	4,5	Сэнгер-центр	2001
<i>Shigella flexneri</i>	Дизентерия	4,6	Университет Висконсина	2003
<b>Геномы в процессе секвенирования</b>				
<i>Plasmodium falciparum</i>	Малярия	30	Консорциум по геному малярии	
<i>Leishmania major</i>	Лейшмания	33,6	Сэнгер-центр, Европейский консорциум	
<i>Trypanosoma brucei</i>	Африканская сонная болезнь	25	TIGR, Сэнгер-центр	
<i>Aspergillus fumigatus</i>	Грибковая инфекция	30–35	Сэнгер-центр, TIGR, Институт Пастера, Университет Саламанки, Университет Нагасаки	
<i>Bacillus anthracis</i>	Сибирская язва	4,5	TIGR	

Примечание. Источники информации: [www.tigr.org](http://www.tigr.org); [www.sanger.ac.uk](http://www.sanger.ac.uk).

<sup>1</sup> Институт по исследованию генома (The Institute for Genome Research).

ной цепи, субъединиц гидрогеназы, транспортной системы, рибосомных белков и др. Сходная ситуация и у автотрофной археобактерии *M. jannaschii*. Эти микроорганизмы имеют мультифункциональные слитые белки, что возможно связано с их экстремальной термофильностью.

Ранее предполагалось, что оперонная организация генов имеет место только у прокариот. Тем не менее было установлено, что в геноме нематоды *Caenorhabditis elegans* (размер 100 млн п. н.), содержащем 14 000 генов, около 25 % входят в состав полицистронных оперонов с небольшим расстоянием между генами (около 100 п. н.).

## 2. Структуры репликации, выявление ORF, интроны и интеины

Полные нуклеотидные последовательности разных видов дают возможность определять структуры генома, связанные с процессами репликации, транскрипции, трансляции, регуляции и т. п. Репликация бактериальной хромосомы начинается в точке *ori C* и продолжается в обоих направлениях до участка терминации репликации *ter C*. У большинства видов участки *ori C* и *ter C* делят кольцевую хромосому на две почти равные реплихоры.

Область начала репликации во многих бактериальных геномах имеет консервативную структуру (рис. 21, 22). Не идентифицирован участок *ori C* в геноме *H. pylori*, в котором нумерация нуклеотидной последовательности начата с повтора (AGTGATT)<sub>26</sub>, содержащего кодоны терминации трансляции во всех рамках в любую сторону. В геноме *V. burgdorferii* положение области начала репликации определено ровно посередине линейной хромосомы в локусе гена *dna A*.

Направление транскрипции большинства генов бактерий с высокой скоростью роста совпадает с направлением репликации. В геноме *V. subtilis* такое совпадение показано у 75 % предсказанных генов, у *V. burgdorferii* — у 66 % генов транскрибируется от центра к концам молекулы. Асимметрия распределения кодирующих участков просматривается в геноме археобактерий *Methanococcus jannaschii* (рис. 70). Структура области инициации репликации ДНК археобактерий неизвестна. Отсутствие асимметрии пар GC в их геномах может указывать на наличие множественных участков инициации репликации.

Часть хромосомы возле *ori C* более консервативна по сравнению с участком терминации. Степень дивергенции генов *E. coli* и *Salmonella thyphimurium* возрастает по мере удаления от локуса *ori C*. Частота спонтанных рекомбинаций (выщепление из хромосомы профага лямбда) в локусе *ter C* возрастает в десятки тысяч раз. Также сравнение геномов видов хламидий *Chl. trachomatis* и *Chl. pneumonia* показало, что район *ter C* содержит намного больше перестроек, чем остальной геном. Ведущая и ведомая нити ДНК отличается не только преимущественным присутствием в

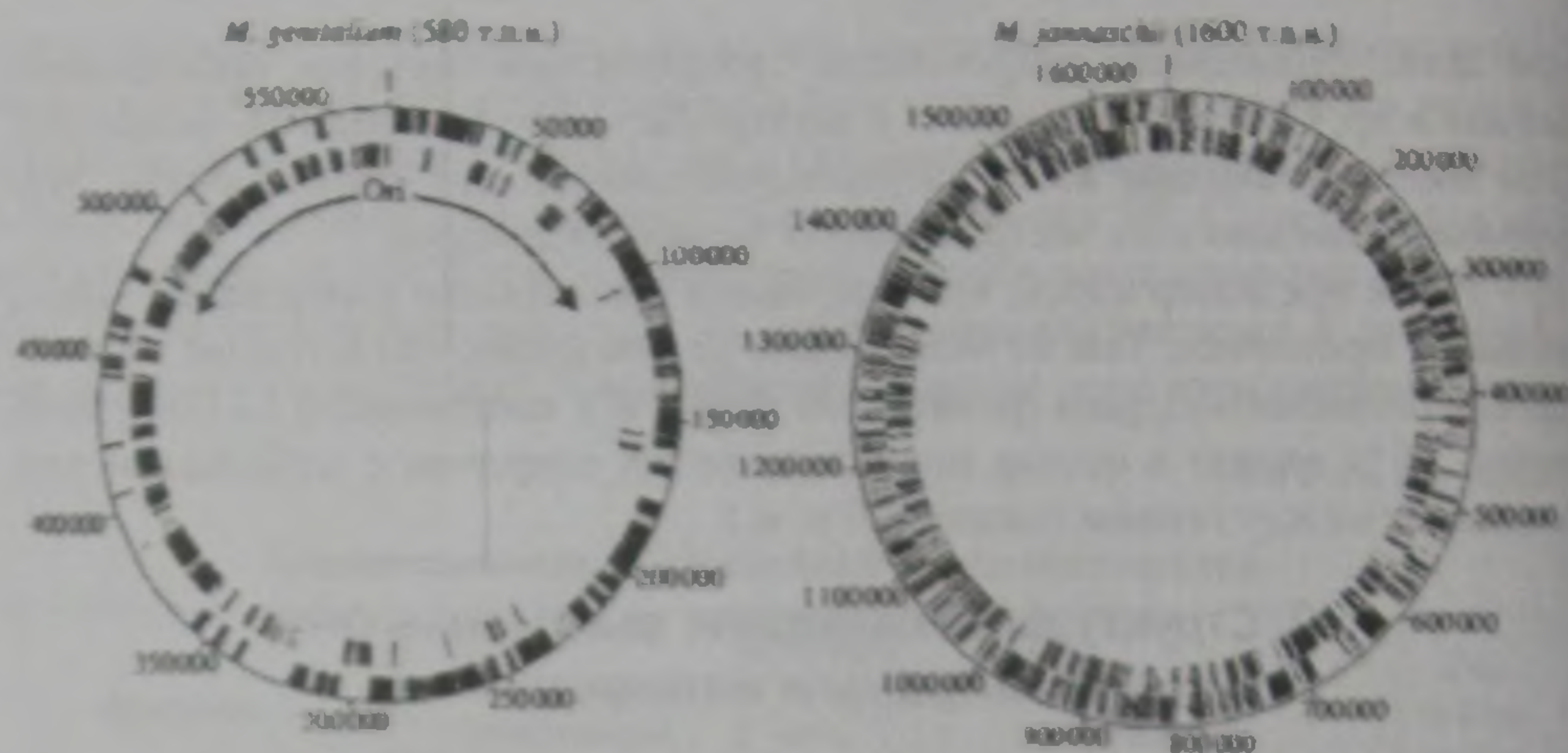


Рис. 70. Распределение кодирующих последовательностей в геномах *M. genitalium* и *M. jannaschii*. Кодирующие районы указаны на наружном кольце. Для *M. genitalium* стрелки, начинающиеся после *ori*, указывают направление репликации

ведущей нити нуклеотида  $G_4$  но и по частоте встречаемости некоторых олигомеров (коротких, по 6–10 нуклеотидов, сегментов одноцепочечной ДНК). Например, в ведущей нити ДНК *E. coli* более часто встречаются так называемый Chi-сайт (октамер GCTGGTGG).

В настоящее время возникла довольно интересная ситуация, когда для большого числа генов известна последовательность нуклеотидов, но о функции этих генов или ничего неизвестно, или известно очень мало. Крупнейшей базой данных аминокислотных последовательностей является SWISS-PROT, существующая с 1986 г. Банки нуклеотидных последовательностей ДНК и аминокислотных последовательных белков, снабженные соответствующими компьютерными программами, обеспечивают возможность сравнивать расшифрованные последовательности генов и их белковых продуктов с уже известными, что позволяет предсказывать функции анализируемых белков. Установленную нуклеотидную последовательность генома при помощи компьютерных программ транслируют в шести рамках считывания и таким способом выявляют *открытые рамки считывания ORF* (англ. *Open reading frame, ORF*). Средний размер ORF в геномах прокариот соответствует примерно 300 аминокислотных последовательностей (а. о.).

Следующим шагом в анализе геномов является сопоставление их состава по функциям. При этом идентифицируют компоненты, общие для всех организмов, общие для данной группы видов и уникальные, специфичные только для данного организма. Так, среди 4288 ORF *E. coli*, аннотированных в геноме, ранее были описаны 1853, а функции части остальных были установлены на основе сходства первичных структур с известными генами других видов. Однако функции 40 % ORF остаются неидентифи-

фицированными даже у такого хорошо изученного представителя прокариот, как *E. coli*. Такая же доля неизвестных функций и у других видов: у *H. influenzae* 42 %, у *T. pallidum* 45 % (табл. 16). У архебактерий *P. horikoshii* неизвестны функции 1655 ORF, из них 453 имеют сходство с последовательностями в базах данных, а 1202 (более 50 %) — уникальны, что указывает на эволюционную дистанцию между этим видом и другими, гены которых секвенированы. Необходимо отметить, что несмотря на более чем столетний период изучения клетки и ее метаболизма, биохимикам и молекулярным биологам до сих пор не известны функции почти половины клеточных белков.

Таблица 16

Количество ORF у эубактерий и архебактерий  
(Боринская, Янковский, 2004).

Видовое название	Всего ORF	Функция известна	Функция неизвестна	
			есть гомологи в базах данных	нет гомологов в базах данных
<i>Af. genitalium</i>	479	468 (97 %)		
<i>M. pneumoniae</i>	677	603 (89 %)		
<i>B. burgdorferii</i>	853*	59 %	12 %	29 %
	430**	70 (16 %)	100 (23 %)	250 (58 %)
<i>Ch. trachomatis</i>	894	604 (68 %)	35 (4 %)	255 (28 %)
<i>R. prowazekii</i>	834	63,7 %	12 %	24,8 %
<i>T. pallidum</i>	1041	577 (55 %)	177 (17 %)	287 (28 %)
<i>Ch. pneumoniae</i>	1073	636 (60 %)	251 (23 %)	186 (17 %)
<i>A. aeolicus</i>	1512	849 (56 %)	256	407
<i>bL pylori</i>	1590	1091 (69 %)		
<i>M. jannaschii</i>	1738	38 %		
<i>P. horikoshii</i>	2061	406 (19,7 %)	453 (22 %)	1202 (58,3 %)
<i>H. influenzae</i>	1073	58 %	20 %	335 (22 %)
<i>T. mantima</i>	1877	1014 (54 %)	407 (22 %)	373 (20 %)
<i>A. fulgidus</i>	2436	1797		639
<i>B. subtilis</i>	4100		42 %	
<i>M. tuberculosis H37Rv</i>	3924	40 %	44 %	16 %
<i>E. coli K-12</i>	4288	62 %		

\* — число ORF в хромосоме.

\*\* — суммарное число ORF в 11 плазмидах.

Приняв в качестве модельного объекта *E. coli*, ряд авторов предложили классифицировать белки кодирующие их гены по *трем функциональным группам: энергообмен, информация и коммуникация (внутри- и внеклеточная)*. В группе коммуникации количество белков увеличивается по мере повышения уровня организации. Такая функциональная классификация генов и белков дает возможность сравнивать между собой геномы различных видов и их метаболических путей.

В генах архебактерий и зубактерий, в частности, в генах тРНК, обнаружены *интроны — последовательности, автокаталитически вырезающиеся при созревании мРНК*, которые считали типичными для эукариот. У прокариот обнаружены также *интеины — участки, кодирующие самосплайсирующиеся полипептидные цепи*. Впервые самосплайсируемый белок обнаружен у пекарских дрожжей *S. cerevisiae*. Так был открыт новый механизм процессинга, в котором участвует *посттрансляционное вырезание внутреннего пептида*, с последующим лигированием концевых пептидов. Позднее было показано, что такой сплайсинг белка проходит *автокаталитически*. По аналогии с интронами, такие автокаталитически вырезаемые участки белка и кодирующие их участки ДНК названы *интеинами*, а остающиеся сплайсируемые фрагменты по аналогии с экзонами — *экстеинами*.

Подобно некоторым интронам (интроны группы I), которые кодируют эндонуклеазы, последовательности, кодирующие интеины, являются мобильными генетическими элементами. Предполагают, что они и интроны группы I возникли сходным образом: путем инвазии гена фермента эндонуклеазы в последовательности ДНК, кодирующие маленькие элементы белкового или нуклеинового сплайсинга. В настоящее время описано несколько десятков интеинов, кодируемых генами архебактерий, зубактерий и эукариот. Открытие интеинов добавляет еще один уровень сложности в проблему перевода полной последовательности нуклеотидов генома в полный набор белков, кодируемых этим геномом. Для понимания всех совокупностей связей последовательностей нуклеотидов с аминокислотными последовательностями белков необходимо проводить сопоставление информации о полной структуре ДНК вида со структурой всех его РНК (*международная исследовательская программа Scription*) и всех его белков (*программа Proteom*).

### **3. Паралогичные и ортологичные гены. Сравнение геномов. Минимальный размер генома прокариот**

В процессе выявления сходства нуклеотидных последовательностей возникают ситуации, когда одному гену у одного вида соответствует несколько генов у другого вида, или группе генов одного вида соответствует

группа генов другого вида. Возникает вопрос о гомологических отношениях между генами — ортологическими и паралогическими.

Ортологические гены происходят от предкового гена, содержащегося в геноме вида, от которого образовались сравниваемые виды. Следовательно, их эволюция отражает эволюцию видов, в генофонде которых они присутствуют. Паралогичные гены — это потомки дублированного предкового гена. Они эволюционируют в пределах одного и того же вида, а их эволюция отражает изменения, накопившиеся с момента дубликации предкового гена.

Паралоги могут эволюционировать и приобретать новые функции, родственные исходной, а ортологи сохраняют в ходе эволюции одну и ту же функцию. Паралоги составляют существенную часть в геномах прокариот: около 50 % генов у *E. coli* и *B. subtilis*, 25 % — у *M. genitalium*, 30 % — у *H. influenzae*, 16 % — у *M. jannaschii* (рис. 71).

Наиболее многочисленны паралоги в семействе генов АТФ-связывающей субединицы АВС-трансфераз: у *E. coli* их более 70, а у *T. maritima* — 67. Это отражает большое количество разнообразных субстратов, которые способны транспортировать прототрофные бактерии. У разных видов прокариот разные гены амплифицированы в разной степени, а увеличение их копийности связано с экологией данного вида и ее специфичностью. Изучение генов-паралогов является важным аспектом изучения прокариотных геномов, эволюции генов путем дубликации, обеспечивающей адаптацию видов к меняющимся условиям окружающей среды.

Группа ученых из NCBI (National Center for Biotechnology Information, США) разработали наиболее информативный способ объединения ортологичных генов в кластеры. Основой метода является выявление в геномах

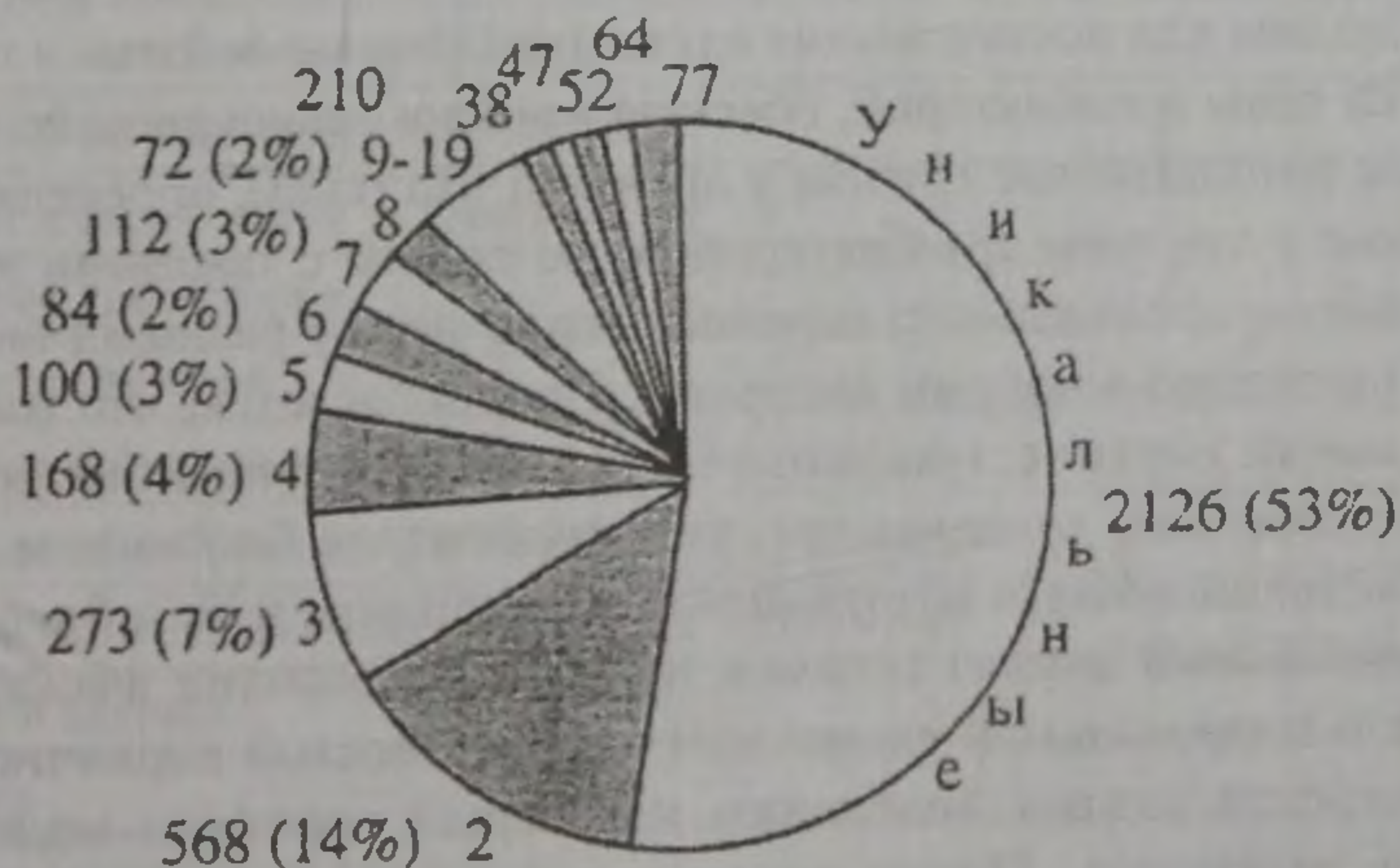


Рис. 71. Распределение числа гомологичных генов в геноме *B. subtilis*. вдоль окружности указана копийность генов, снаружи — число генов



*групп ортологичных генов* — метод COG (англ. cluster of ortologous genes). Был проведен компьютерный анализ аминокислотных последовательностей 17 967 белков 7 видов микроорганизмов, сгруппированных в пяти неродственных группах:

- эукариоты (*S. cerevisiae*),
- грамположительные бактерии (*M. genitalium*, *M. pneumoniae*),
- грамотрицательные бактерии (*E. coli*, *H. influenzae*),
- археобактерии (*M. jannaschii*),
- цианобактерии (*Scenedocystis* sp.).

Кластер выделяли в тех случаях, когда ортологи имелись не менее чем в трех группах из пяти.

38 % проанализированных генов, которые кодировали 6.814 из 17.967 белков, составило 720 *кластеров ортологичных групп*. В группы объединялись потомки одного предкового гена. Кластеры и входящие в них гены контролируют протекание *консервативных функций клетки*. Общее число входящих в кластеры ортологичных групп насчитывалось около 1000. к ним относятся в том числе группа «генов домашнего хозяйства» и гены шапероны, участвующие в сворачивании белков. Гены этих групп отвечают за стабильность основных физиологических процессов в клетке и локализованы в больших и малых геномах без существенных изменений. У патогенных бактерий даже в этих консервативных группах часть генов была утрачена, а особенно часто не представлены кластеры генов, контролирующих процессы метаболизма. В то же время повсеместно представлены гены компонентов трансляции. 114 таких кластеров, относящихся к транскрипции и трансляции, составляют *универсальное ядро жизни*. Это в два раза меньше того минимального набора генов, который, как предполагают, необходим для *поддержания клеточной формы жизни*.

Многие гены археобактерий, обеспечивающие транскрипцию и трансляцию, имеют ортологичные группы у дрожжей. По генам, определяющим метаболические пути, гены археобактерий более сходны с таковыми эубактерий.

Явление горизонтального переноса генов между разными видами прокариот установлено многими авторами. Предполагается, что имеет место горизонтальный перенос эукариотических генов в геном бактерий — например, *перенос гена холинкиназы*, участвующего в биосинтезе полисахаридов и обеспечивающего патогенность у микоплазм и *H. influenzae*.

Сравнительный анализ геномов прокариот позволил высказать предположение о минимальной сложности гипотетической первичной клетки. Для организации клетки необходим некоторый минимум молекулярных структур и процессов. Путем сравнения геномов выявлено, что минимальная клетка, способная к автономной жизнедеятельности и самовоспроизведению, должна была бы содержать не менее 250–300 наиболее

существенных генов. Группа японских исследователей (Tomita et al., 1998) показала, что можно обеспечить все основные метаболические потребности клетки, трансляцию и репликацию РНК-генома в системе со 127 генами. При этом клетка должна быть лишена архива ДНК, репарационных и других важных систем защиты и помехоустойчивости, что делает ее эволюционно беззащитной.

В качестве модели минимальной живой клетки была использована микоплазма *M. genitalium* с геномом размером 0,58 Мб, 470 генов. Минимальный набор генов определялся путем сравнения всех белков грамположительной *M. genitalium* и грамотрицательной *H. influenzae*. Бактерии этих видов — паразиты с геномами, кодирующими 468 и 1073 белков, соответственно, эволюция которых протекала путем утраты генов от общего предка. Было высказано предположение, что общие для этих микроорганизмов гены могут составить минимальный набор. Сравнение этих геномов выявило 240 ортологов, а всего в минимальный набор было включено 256 генов, кодирующих белки с известными функциями (табл. 17). Таким организмом, утратившим почти все анаболические пути, может быть *только паразит*, зависящий в обеспечении метаболитами от организма-хозяина. У микроорганизмов с маленькими геномами, как у микоплазмы, все гены минимального набора должны быть необходимыми и существенными.

Таблица 17

Функции белков, соответствующих минимальному набору из 256 генов  
(Коничев, Севастьянова, 2003)

Функция	Число белков
Преобразование энергии	28
Транспорт и метаболизм аминокислот	11
Транспорт и метаболизм нуклеотидов	20
Транспорт и метаболизм углеводов	5
Метаболизм кофакторов	8
Метаболизм липидов	6
Трансляция и биогенез рибосом	94
Репликация, транскрипция, рекомбинация, репарация	35
Структурная функция (белки наружной мембраны)	7
Секреция и адгезия	5
Шапероны	13
Транспорт неорганических ионов	4
Предсказана гипотетическая функция	15
Функция неизвестна	4

Теоретические расчеты о величине минимального набора генов экспериментально подтверждены на *B. subtilis*. В геноме этой бактерии в произвольном порядке инактивировали разные гены, а в результате была определена та фракция генов, инактивация которых делает бактерию неспособной к образованию колоний. Размер существенной части генома этого вида составил 318 т. п. н., что соответствует 250–300 бактериальных генов.

Минимальный размер генома способен обеспечить автономность жизнеобеспечения клетки. Размер геномов и их функциональный состав различны у разных прокариот с разными функционально-энергетическими потребностями:

- фототрофы для роста нуждаются в свете,  $\text{CO}_2$  и неорганических ионах;
- прототрофы нуждаются в органическом углероде, источнике энергии (глюкоза и т. д.) и неорганических ионах;
- гетеротрофы растут на сложных средах;
- облигатные внутриклеточные паразиты и органеллы (хлоропласты и митохондрии) стоят отдельно.

Организмы с небольшими геномами предположительно могли возникнуть двумя различными эволюционными путями — «снизу вверх» и «сверху вниз». Возникновение и развитие генома и генетического аппарата по первому пути могли претерпеть представители РНК-мира до появления *гипотетических прогенов* как исходного предка всех целлюлат. Согласно представлениям К. Вёзе, прогеномы были предками архебактерий и зубактерий, они уже имели ДНК и РНК, а их гены имели как экзоны, так и интроны.

Другой путь эволюции «сверху вниз» прослеживается у видов с возрастающими метаболическими потребностями — у паразитов. Вероятно, таким путем возникли микоплазмы из грамположительной бактерии с размером генома 1700–2000 т. п. н. значительное сокращение генетической информации произошло в процессе паразитического образа жизни. В ее геноме отсутствуют многие гены биосинтеза аминокислот, имеется очень мало генов синтеза витаминов, предшественников нуклеиновых и жирных кислот. Все эти вещества микоплазма получает от хозяина. Число генов рРНК и тРНК минимальное. Такова геномная цена, уплаченная за поддержание паразитизма. Минимальные размеры генома 600–800 т. п. н. в разных независимо возникших линиях микоплазм представляют, по-видимому, низший предел для размера генома у клеточных форм организмов.

#### 4. Геномы прокариот в процессе функционирования и эволюции

Геномы современных прокариот и их структура сформировались в ходе эволюционного процесса под влиянием различных эволюционных механизмов:

- генных мутаций, дупликаций и амплификаций небольших участков генома с последующей специализацией генов-паралогов);
- делеций, инверсий и транслокаций сегментов генома;
- горизонтального переноса генов (ГПГ).

Адаптации к условиям сред обитания также обеспечиваются посредством изменения общих характеристик геномов — их GC-состава, суперспециализации ДНК, а также наличием особых генов, способствующих приспособлению микроорганизмов к тем или иным условиям. Установлены примеры таких приспособлений.

- А. Адаптация термофилов к обитанию при высокой температуре связана с повышенным содержанием полярных аминокислот и с присутствием фермента обратной гиразы, вводящего в ДНК позитивные супервитки.
- Б. Белки *H. pylori* содержат удвоенное, по сравнению с *E. coli* и *H. influenzae*, количество основных аминокислотных остатков аргинина и лизина. Как и другие паразитические бактерии, *H. pylori* имеет ограниченный метаболический репертуар.
- В. К генам экологической специфичности относятся также гены *вирулентности патогенных бактерий*, кодирующие токсины, адгезины, инвазины и другие факторы вирулентности, находящиеся в дискретных сегментах генома — так называемых «островках патогенности». Так, в геноме *H. pylori* локализовано 32 гена поринов и адгезинов, а у возбудителя урогенитальных инфекций *M. genitalium* 10 % генома составляют гены адгезинов и поверхностных антигенов.

### (1) Амплификация участков генома

Характерной особенностью прокариотических геномов является амплификация генов. В настоящее время она рассматривается не как мутация, а как *динамическое состояние геномов*, связанное с адаптацией популяций бактерий к меняющимся условиям среды. Как установлено по нуклеотидным последовательностям, характер амплификации различных частей генома противоречит идее эволюции путем дупликации целых предковых геномов с их последующей дивергенцией. Экспансия гомологичных генов внутри генома связана с физиологической и экологической адаптацией разных видов прокариот.

### (2) Перестройки генома

Перестройки генома могут происходить как в процессе эволюции, так и в ответ на изменение физиологических условий в процессе жизнедеятельности клетки.

В зависимости от физиологических условий описаны перестройки у цианобактерий *Anabaena* и *Nostoc*. В отсутствие источников азота (нитраты, аммоний) у них образуются *гетероцисты* — специализированные неделиющиеся клетки, способные к фиксации атмосферного азота. С участием сайт-специфической рекомбиназы у них происходит вырезание фрагментов ДНК из кодирующих участков генов азотофиксации *fdxN*, *nifD* и *hurL*.

У *B. subtilis* перестройка генома происходит во время спорообразования. При этом из фактора транскрипции сигма К, контролирующего экспрессию специфических генов, вырезается фрагмент ДНК — так называемый *skin* — элемент размером 48 032 п. н., который содержит 57 ORF.

Филогенетические транслокации сегментов генома можно проследить у микоплазм. Все 470 генов *M. genitalium* имеют гомологов в геноме *M. pneumoniae* (679 генов). Белки этих видов гомологичны на 67 %, что отражает их близкое родство. Геномы микоплазм подразделяют на 6 сегментов с консервативным порядком генов в них, но расположение сегментов в геномах имеет различную последовательность. Возможно, что эти различия являются результатом транслокации. У *E. coli* штамм W3110 отличается от штамма K12 наличием инверсии сегмента размером 17 % от всего генома. Частота других инверсий возрастает по мере приближения к локусу *ter C*.

Транслокации и инверсии сегментов генома изменяют порядок и ориентацию генов относительно направления репликации, что оказывает влияние на их функционирование. Локализация генов в разных участках хромосомы имеет важное значение для адаптации к определенным условиям роста.

### (3) Консервативная и оперативная части генома

В. А. Геодакян сформулировал принцип, в соответствии с которым эволюционирующие системы подразделяются на две части: *консервативную*, сохраняющую эволюционные «достижения», и *оперативную* (*поисковую*), быстро меняющуюся в ходе эволюции в ответ на изменения среды. В бактериальном геноме такие системы соотносятся следующим образом:

- консервативная часть хромосомы локализована возле участка *ori C*,
- оперативная часть хромосомы расположена в районе *ter C*.

Район *ter C* может играть роль «кузницы» новых генов — находящиеся здесь копии генов быстро дивергируют, приобретая новые функции. При перемещении гена в более консервативные части генома скорость его эволюции снижается.

### (4) Горизонтальный перенос генов (ГПГ)

В эволюции прокариотических геномов определенную роль играет ГПГ. Этот процесс происходит между родственными и неродственными

бактериями, а вероятно также между эукариотами и прокариотами. Гены, появившиеся в геноме в результате ГПГ, принято называть *ксенологами*. Конкретные примеры ГПГ:

- фермент *дегидролипоамиддегидрогеназа* архебактерии *Halobacterium halobium* имеет сходство с соответствующим ферментом грамположительных бактерий (50 % гомологии аминокислотных последовательностей) в большей степени, чем с подобными ферментами трех других видов архебактерий (25 % гомологии);
- три гена сульфатредуктаз архебактерий *A. fulgidus* составляют оперон, соответствующий найденному у эубактерий;
- у паразитических бактерий обнаружены ферменты эукариотического типа;
- в геноме хламидии обнаружено около 30 генов эукариотического типа.

Эти и другие наблюдения могут послужить аргументами в пользу точки зрения, на основе которой горизонтальный перенос генов происходит между представителями всех трех доменов жизни.

#### (5) Попытки установления филогенетического древа

Скорость изменения разных частей генома в процессе эволюции различна. Установлено, что наиболее консервативны последовательности, кодирующие белки. Порядок расположения генов является вторым по уровню консервативности. Значительная изменчивость и наибольшая скорость эволюции характерна для регуляторных участков.

Для определения эволюционных отношений между видами используют сравнение нуклеотидных и аминокислотных последовательностей. При построении филогенетического древа живых организмов исходят из предположения, что общий предок имел определенную последовательность, которая постепенно менялась в дивергирующих ветвях в результате единичных замен или делеций. Чем больше различаются последовательности двух видов, тем больше эволюционная дистанция между ними. Однако существующие методы не позволяют однозначно определить эволюционные дистанции. При филогенетических построениях важное значение приобретает учет наличия паралогов, возможность ГПГ и различия в скорости изменения разных генов. Наибольшее значение для построения филогенетического древа приобрело сравнение генов 16S рРНК как наиболее консервативных последовательностей в геноме.

Делаются попытки реконструировать последовательность возникновения основных клеточных процессов в эволюции эубактерий, архебактерий и эукариот. Так как участвующие в трансляции белки и ферменты метаболизма (ферменты цикла трикарбоновых кислот, гликолиза, метаболизма нуклеотидов) являются общими у всех доменов, предполагается, что

трансляция и метаболические реакции возникли до разделения трех основных доменов. Вероятно, трансляция возникла еще до перехода от мира РНК к миру ДНК. Транскрипция появилась позже трансляции, а еще позднее возникла репрессия генов и компактизация ДНК, которые способствовали увеличению размеров генома.

## Характеристика геномов прокариот

### 1. *Haemophilus influenzae* (возбудитель менингита, пневмонии)

Секвенирование нуклеотидной последовательности полного генома самостоятельно существующего организма этой грамотрицательной бактерии проведено в 1995 г. Институтом геномных исследований (TIGR — The Institute for Genome Research, США). Размер генома 1.830.137 п. н. Для него характерно относительно низкое содержание GC-пар (38 %), но при этом найдено 7 протяженных участков с более высоким содержанием GC-пар (около 50 %). При анализе нуклеотидной последовательности обнаружен *предполагаемый ориджин репликации* (область начала репликации), состоящий из 280 п. н.; 6 оперонов рРНК, 54 гена тРНК для всех 20 аминокислот. Эти данные позволили составить кольцевую карту хромосомы *H. influenzae*. вычислено 1743 открытых рамки считывания (ORF), но для 736 не удалось выявить функции кодируемых ими белков. Для 78 % ORF *H. influenzae* обнаружена гомология с имеющимися в базах данных последовательностями других организмов.

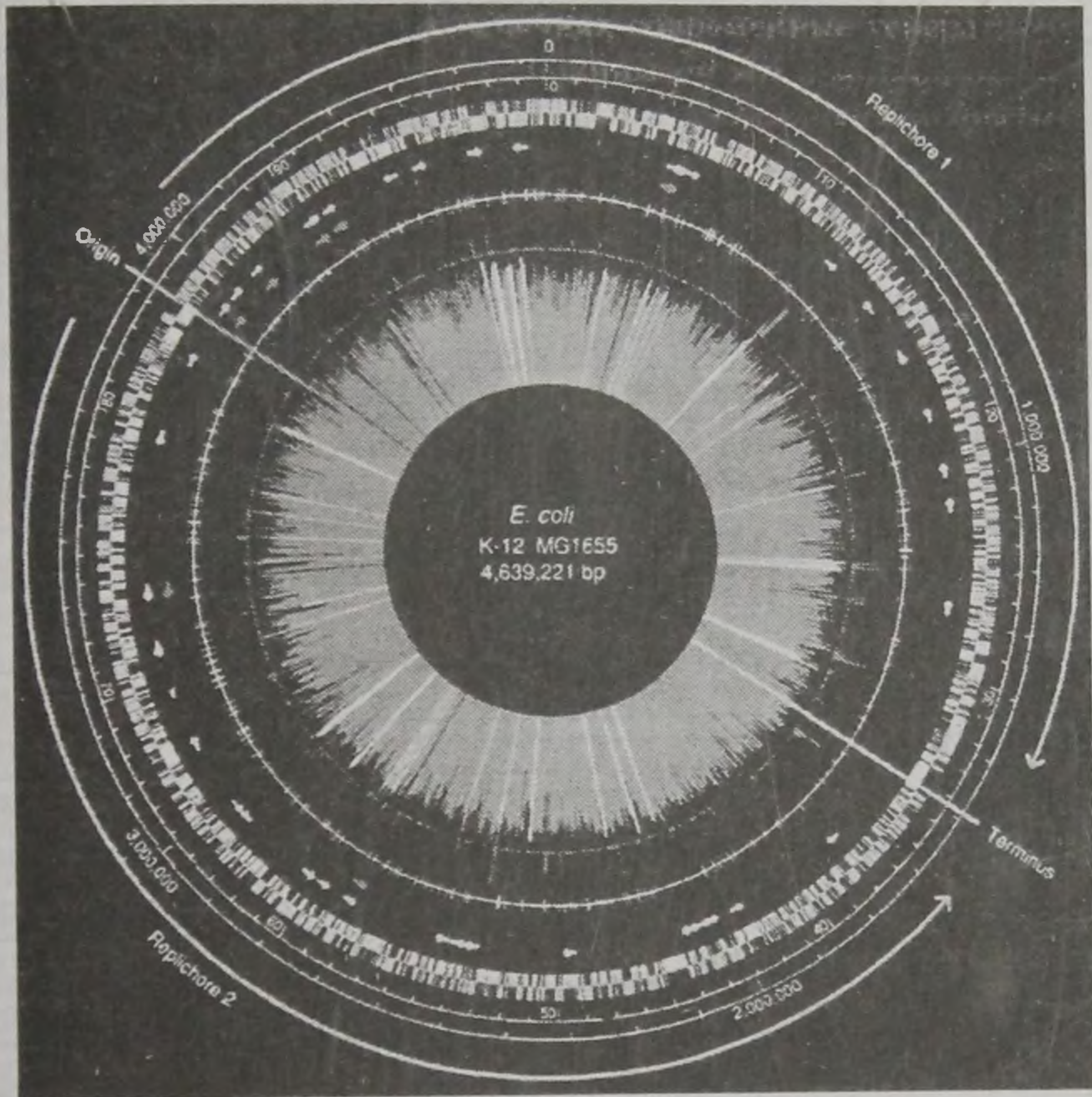
### 2. Кишечная палочка *Escherichia coli*

Эта грамотрицательная бактерия является самой изученной клеткой из всех существующих. В ней впервые обнаружена F-плазмида полового фактора и другие плазмиды. Большинство изученных фагов также имеют в качестве хозяина *E. coli*. именно эта бактерия, ее плазмиды и фаги явились первыми объектами генетической инженерии, на которых достигнуты сенсационные результаты по клонированию и экспрессии чужеродных генов.

*E. coli* является прототрофной энтеробактерией — условным патогеном, так как некоторые ее штаммы вызывают заболевания у людей (пищевые отравления, энтериты). Геномные исследования проводились на разных штаммах (K-12 MG 1655, W3110, 0157), при этом установлены различия по размеру и структуре геномов.

Результаты полного секвенирования генома *E. coli* K12 (штамм MG 1655) опубликованы в 1997 г. (Blattner F. R. et al., 1997). Размер генома составляет 4.639.221 п. н. эта последовательность отвечает кольцевой генетической карте *E. coli* K12, калиброванной на 100 мин по времени конъюгационного переноса (рис. 72). Начало отсчета карты выбрано между генами *las T* и *thr L*.

Содержание GC-пар в геноме составило 50,8%. Компьютерный анализ выявил 4288 потенциальных ORF, принадлежащих как действительно существующим, так и гипотетическим генам. Для 38% ORF (цистронов) не удалось определить функции кодируемых ими белков.



**Рис. 72.** Общая структура генома бактерии *E. coli* K12 (Blattner et al., 1997). Начало (origin) и конец (terminus) репликации изображены зеленой линией, а реплихоры 1 и 2 — голубыми стрелками. Шкала указывает координаты карты в парах нуклеотидов и в минутах конъюгационного переноса (фактически — в центисомах, т. е. в процентах от полной ДНК). Распределение генов показано на двух внешних окружностях: оранжевые полосы — гены, локализованные на одной цепи ДНК, желтые — на противоположной цепи. Красные стрелки указывают локализацию и направление транскрипции генов рРНК, зеленые — генов тРНК. Следующая окружность иллюстрирует положение повторов REP в виде отдельных тонких штрихов. Следующая окружность показывает сходство между белками профагов и *E. coli*, которое мы здесь не обсуждаем. Наконец, в центре представлена радиальная картина (похожая на оранжевый солнечный взрыв) распределения так называемого индекса адаптации кодонов (CAI), который отражает степень экспрессии отдельных генов



Геном *E. coli* плотно нагружен генами (88,5 %), а межгенные участки занимают относительно малую долю (11 %). Самый большой цистрон содержит 7149 п. н. (2383 кодона), функция его неизвестна. Средний размер цистрона 951 п. н. (317 кодонов). Средний интервал между цистронами — 118 п. н. При этом межгенные интервалы в большинстве своем содержат различные функциональные сайты, т. е. выполняют регуляторные функции. Кроме того, цистроны не содержат интронов — внутренних некодирующих участков. Известно, что цистроны выделяются в ДНК и мРНК начальными и конечными знаками пунктуации, внесенными в генетический код. У *E. coli* для реальных и потенциальных белков стартовыми триплетами являются AGT (3542 случая), GTG (612), TTG (130), ATT (1), CTG (1). У 405 генов показано перекрывание стартового и терминирующего триплетов: ATGA (224 случая), TAATG (98), TGATG (48), GTGAA (28), TAGTG (4), TTGA (3). Сравнение белков (ORF) *E. coli* и *H. influenzae* (в расчет принималась 30 %-ная идентичность аминокислот на 60 % всей последовательности белка) показала, что 1130 из 1743 потенциальных белков *H. influenzae* совпадают с таковыми *E. coli*.

Сложность молекулярно-генетической системы управления и метаболической сети *E. coli* можно охарактеризовать следующим образом:

Длина ДНК генома, Мб	4,6
Полное число генов	4909
Число цистронов	4288
Число кодируемых ими ферментов	804
Число метаболических реакций	988
Число метаболических путей	123
Число химических веществ, участвующих в метаболизме	1303
Число фракций т-РНК (генов т-РНК)	79 (86)
Число регуляторных белков	60

Более подробная классификация цистронов по 22 функциональным классам представлена в табл. 18. Здесь около 25 % клеточных ресурсов связана с метаболизмом малых молекул; 13 % — с метаболизмом макромолекул и 20 % — с клеточными структурами и процессами. В метаболизме малых молекул ключевую роль играют синтез, распад и преобразование нуклеотидов (58 цистронов), аминокислот (131), энергетические процессы (243), транспорт (146), центральный промежуточный метаболизм (188) и другие процессы. Например, системы выполняющие основные генетические процессы, содержат следующее число генов (% генома):

- Репликация, рекомбинация и репарация ДНК — 115 (2,68 %);
- Транскрипция, синтез, метаболизм и модификация РНК — 55 (1,28 %);

Таблица 18

Распределение цистронов и белков *E. coli* по 22 функциональным классам (Blattner et al., 1997)

Функциональный класс	Число белков	%
Регуляторная функция	45	1,05
Предполагаемая регуляторная функция	133	3,10
Структура клетки	182	4,24
Предполагаемые мембранные белки	13	0,30
Предполагаемые структурные белки	42	0,98
Фаги, транспозоны, плазмиды	87	2,03
Транспортные и связывающие белки	281	6,55
Предполагаемые транспортные белки	146	3,40
Энергетический метаболизм	243	5,67
Репликация, рекомбинация, модификация и репарация ДНК.	115	2,68
Транскрипция, синтез, метаболизм и модификация РНК	55	1,28
Трансляция, посттрансляционная модификация белков	182	4,24
Клеточные процессы, включая адаптацию и защиту	188	4,38
Биосинтез кофакторов, простетических групп и носителей	103	2,40
Предполагаемые шапероны	9	0,21
Биосинтез и метаболизм нуклеотидов	58	1,35
Биосинтез и метаболизм аминокислот	131	3,06
Метаболизм жирной кислоты и фосфолипидов	48	1,12
Катаболизм соединений углерода	130	3,03
Центральный промежуточный метаболизм	188	4,38
Предполагаемые ферменты	251	5,85
Другие известные гены (генные продукты и фенотипы неизвестны)	26	0,61
Гипотетические, неклассифицированные, неизвестные	1632	38,06
Всего	4288	100

- Трансляция и посттрансляционная модификация белки — 182 (4,24 %) + 21 ген рРНК + 86 генов тРНК.

Для генома *E. coli* и других энтеробактерий характерно присутствие управляемых единиц транскрипции — *оперонов*. Первые опероны были открыты именно у *E. coli*: *lac*-оперон, контролирующий сбраживание сахара лактозы; *trp*-оперон, контролирующий синтез аминокислоты триптофана. Важной особенностью оперонов является наличие обратной связи между

концентрацией контролируемого метаболита и наработкой ферментов его синтеза или распада. Всего в геноме *E. coli* выявлено и предсказано 2584 оперона. Среди них:

- 73 % содержат 1 цистрон,
- 16 % — 2 цистрона,
- 4,6 % — 3 цистрона,
- 6 % — 4 цистрона и более.

Все они имеют не менее 1 промотора — начального знака транскрипции.

Геном *E. coli* содержит 2 функциональные единицы репликации, названные Ф. Блаттнером *реплихорами* (рис. 71). Общее двустороннее начало репликации (origin) локализовано на участке 84,5 мин конъюгационного переноса и занимает 250 п. н. В этой зоне инициируется двусторонняя репликация. Реплихор 1 ориентирован по часовой стрелке, реплихор 2 — против. Оба процесса заканчивается на противоположном участке генетической карты, ~ 34–35 мин, где каждый из них имеет свой отдельный ориентированный терминальный знак (ter) — T1 и T2. Геном *E. coli* K12 содержит также большое число необязательных (факультативных) включений — профагов, плазмид и транспозонов. Они подвижны, способны к внедрению в геном и выщеплению из него. Наиболее изучено это для умеренного фага  $\lambda$  (лямбда) и полового фактора (плазмиды) F, которые в линии K12 отсутствуют.

### 3. Сенная палочка *Bacillus subtilis*

*B. subtilis* — автотрофная грамположительная непатогенная бактерия. Она, как и многие виды бацилл, активно используется микробиологической промышленностью для производства ферментов, антибиотиков, аминокислот.

Секвенирование полного генома этого вида (штамм 168) проводилось усилиями международного консорциума, в который входили европейские, японские и южноевропейские лаборатории. Размер генома *B. subtilis* составил 4.214.810 п. н., содержание GC-пар — 43,5 %. Вычислено 4100 ORF, для 42 % из них имеются гомологи в базах данных. Определено 87 % кодирующих последовательностей.

Бурному развитию молекулярной генетики *B. subtilis* способствовала генно-инженерная методология, которая позволяет изучать процессы регуляции экспрессии целевых генов. В ходе генно-инженерных экспериментов на *B. subtilis* было обнаружено, что гены различных грамположительных бактерий, как правило, экспрессируются в клетках данной бактерии. В то же время оказалось, что подавляющее большинство генов гра-

мотрицательных бактерий, в том числе *E. coli*, не функционируют в *B. subtilis*, хотя ее гены в *E. coli* экспрессируются. Поскольку генетический код универсален, полученные результаты могут указывать на различие регуляторных последовательностей нуклеиновых кислот, а, следовательно, и на различие ферментных систем транскрипции и трансляции у грамположительных и грамотрицательных бактерий. В экспериментах было установлено, что для возможности правильной и эффективной экспрессии чужеродного гена в клетках *B. subtilis* он должен иметь *бациллоподобную последовательность промотора*, а на синтезируемой мРНК в участке связывания рибосом должна находиться достаточно протяженная последовательность, комплементарная 3'-концу 16S рибосомой РНК бацилл. Схожим образом организованы гены других грамотрицательных бактерий, что и обуславливает их экспрессию в клетках бацилл.

#### 4. Актиномицеты рода *Streptomyces*

Стрептомицеты являются почвенными грамположительными бактериями, которые в своем жизненном цикле проходят несколько стадий дифференцировки от прорастания споры до формирования воздушного мицелия. Продуктами вторичного метаболизма стрептомицетов являются антибиотики, и 70 % всех производимых промышленностью антибиотиков синтезируются именно бактериями этого рода. Важной особенностью стрептомицетов является то, что они не патогенны, ни для человека, ни для животных. Известен лишь один вид (*S. scabies*), патогенный для растений.

Геном бактерий рода *Streptomyces* представлен одной молекулой кольцевой двуцепочечной ДНК размером около  $10^4$  т. п. н. (в 2,5 раза протяженней, чем у *E. coli*). Для стрептомицетов характерно очень высокое (70–73 %) содержание ДНК GC-пар. (У бактерий рода *Escherichia* этот показатель составляет 50 %, а у *Bacillus* — 32–62 %). Многие штаммы стрептомицетов имеют естественную *систему генетического обмена через конъюгацию*, направляемую половыми факторами. Хотя механизм этого процесса полностью невыяснен, проведенные исследования позволяют сделать заключение, что конъюгационная система *Streptomyces* генетически устроена намного проще, чем F-фактор *E. coli*. К наиболее изученным генетически относятся штаммы *S. coelicolor* A3 (2) и *S. lividans* 66. Для них получено и охарактеризовано большое число различных мутантов.

### Заключение

Изучение бактериальных геномов продолжается около 50 лет, но существенный вклад внесла информация, полученная за последние 10 лет при анализе полных нуклеотидных последовательностей. К настоящему

времени полному секвенированию подвергнуты геномы более 400 видов прокариот, некоторые примеры приведены в табл. 13, 14, 15. Данные о размерах геномов, количестве генов и их плотности показывают, что качественные различия между геномами прокариот и простейших эукариот не столь велики. Методы структурного, функционального и компьютерного анализа, разработанные при изучении геномов прокариот, заложили основу для анализа сложных эукариотических геномов, в том числе генома человека.

Анализ структуры геномов прокариот показал неприменимость общей концепции эволюции геномов, разработанной для эукариот. *Дупликация полных геномов* или их протяженных участков в качестве основы эволюции геномов у прокариот не просматривается. В то же время у них имеет место дупликация и амплификация небольших фрагментов генома, сравнимых по размеру с генами. Сравнение геномов бактерий предоставляет доводы в пользу того, что в эволюции существует горизонтальный перенос генов. Полученные результаты имеют важное прикладное значение для биотехнологии и медицины, а также вносят существенный вклад в расширение наших фундаментальных знаний о природе живого.

### Ключевые слова

Прокариоты, зубактерии, архебактерии, молярное содержание GC-пар (GC, моль %), молярное содержание AT-пар (TA, моль %), грамположительные бактерии, грамотрицательные бактерии, реплихоры, участки *ori C* и *ter C*, открытая рамка считывания — ОРС (анг. open reading frame — ORF), интроны и экзоны, интеины и экстеины, сплайсинг мРНК, самосплайсинг полипептидной цепи, паралогичные и ортологичные гены, минимальный набор генов, опероны, дупликация и амплификация генов.

## Геномика эукариот

К подимперии эукариот (Eucaryota) относятся представители империи клеточных организмов (Cellulata), у которых ядро полностью обособлено от цитоплазмы, имеется ядерная оболочка, а ядерная ДНК расположена в форме линейных хромосом (рис. 73). Описание эукариот представлено в главе 7 (рис. 38), а молекулярно-биологические признаки приведены в табл. 2.

У эукариот (грибов, водорослей, простейших, высших растений, животных) различают две группы: одноклеточные и многоклеточные организмы. Чередование поколений у них происходит при размножении, которое в разных таксонах может осуществляться как половым, так и бесполом путем.

В аспекте геномики в данной главе будут рассмотрены несколько видов из разных таксономических групп эукариот, геномы которых изучены более досконально. К ним относятся пекарские дрожжи *Saccharomyces cerevisiae*, травянистое растение резуховидка *Arabidopsis thaliana*, нематода *Caenorhabditis elegans*, плодовая мушка *Drosophila melanogaster*, рис *Oryza sativa*, домовая мышь *Mus musculus* и человек *Homo sapiens*. (табл. 19).

Таблица 19

Примеры секвенирования полных геномов разных типов организмов

Организм	Размер генома, млн пн	Предполагаемое число генов	Год завершения секвенирования
Бактерия ( <i>Haemophilus influenzae</i> )	1,8	1743	1995
Дрожжи ( <i>Saccharomyces cerevisiae</i> )	12,1	6102	1996
Круглый червь ( <i>Caenorhabditis elegans</i> )	97	19 099	1998
Растение ( <i>Arabidopsis thaliana</i> )	100	25 000	2000
Насекомое ( <i>Drosophila melanogaster</i> )	180	13 061	2000
Человек ( <i>Homo sapiens</i> )	3000	35 000–45 000	После 2003

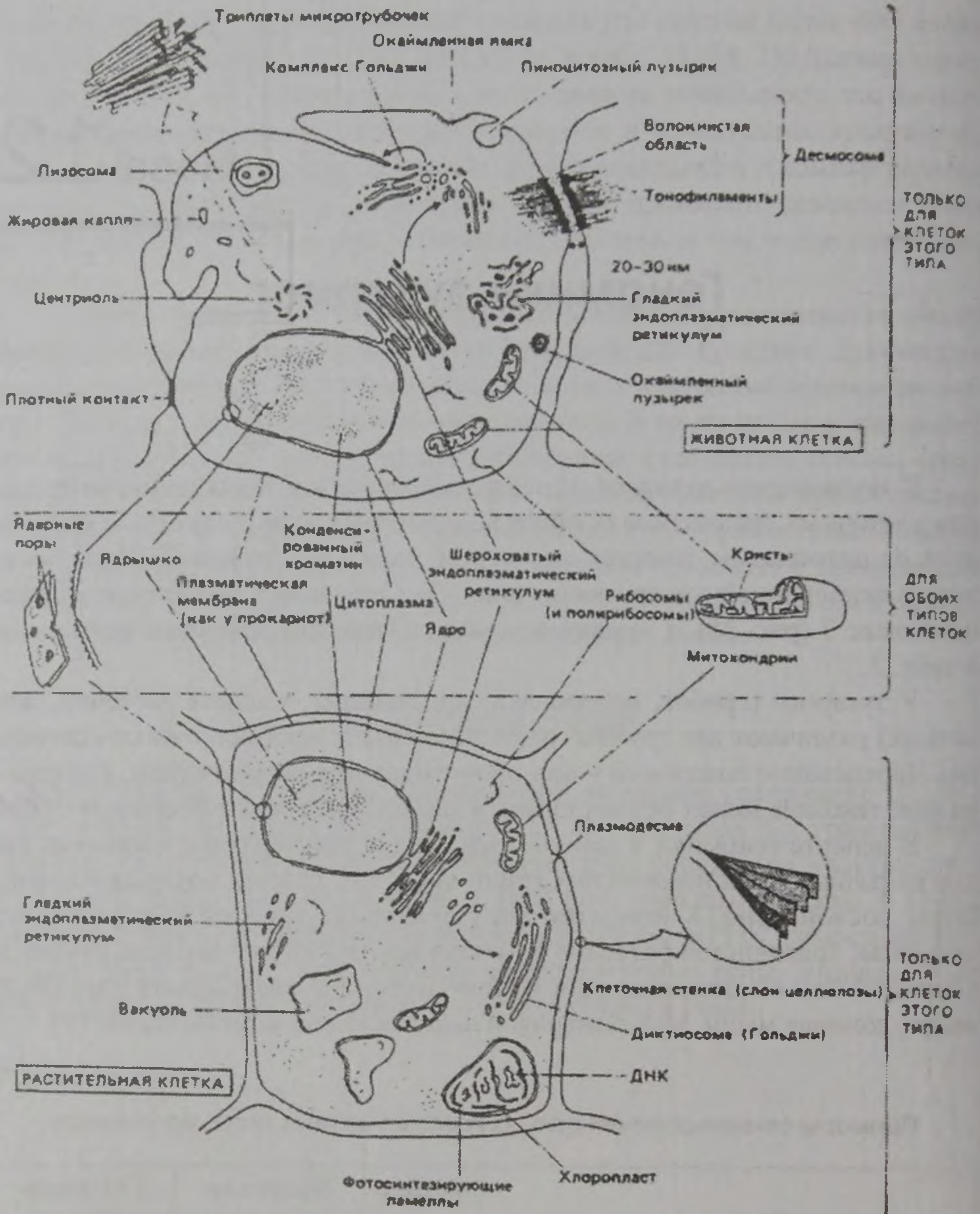


Рис. 73. Схематическое представление эукариотических клеток (животной и растительной) (Петухов и др., 2007)

### Геном пекарских дрожжей *Saccharomyces cerevisiae*

Этот вид одноклеточных грибов относится к классу Ascomycetes (сумчатые грибы). Дрожжевая клетка содержит ядро, окруженное ядерной мембраной, и другие органеллы — митохондрии, плазмиды двух типов.

Данный низший эукариотический организм является наиболее изученным в биохимическом и генетическом аспекте. Для него применимо большинство методов работы с бактериальными культурами. У пекарских дрожжей подробно изучен клеточный цикл, включающий почкование диплоидных и гаплоидных клеток с наличием митоза (рис. 74). Клетки гаплоидных штаммов экспрессируют *гены половых феромонов  $a$  и  $\alpha$  (альфа)*, определяющих клетки противоположных типов спаривания. При их копуляции образуются диплоидные клетки ( $2n = 32$ ). В определенных условиях диплоидные клетки претерпевают мейоз, в результате которого образуются 4 гаплоидные аскоспоры. После прорастания каждая спора дает начало отдельному гаплоидному клону со *специфическим генотипом и фенотипом*. К недостаткам дрожжей как генетического объекта относятся малые размеры хромосом — они практически невидимы в световом микроскопе, что исключает возможность проведения цитогенетических исследований. Генетический анализ мутантных штаммов позволил составить подробную генетическую карту по 16 группам сцепления с локализацией более 400 генов.

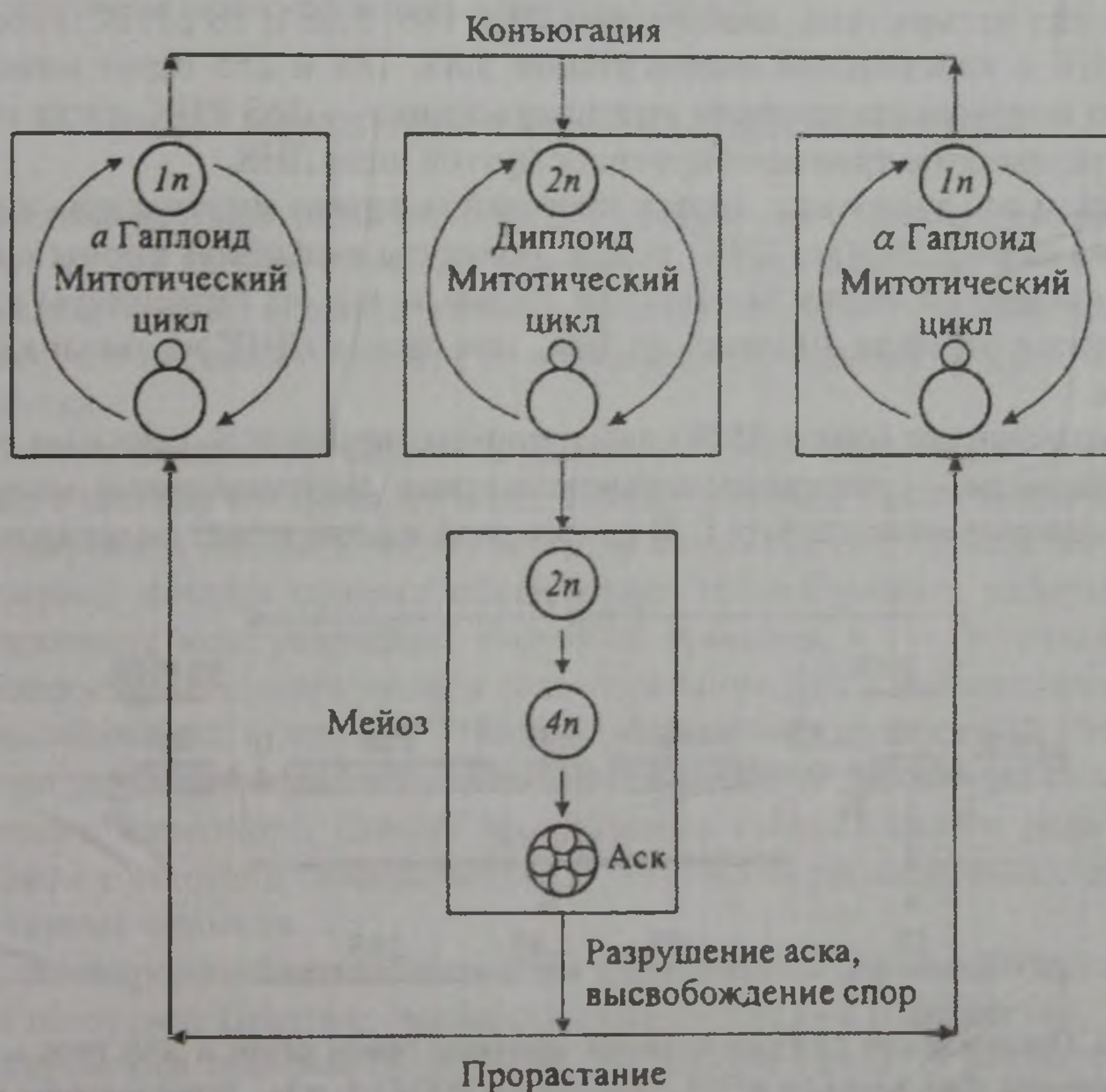


Рис. 74. Жизненный цикл *Saccharomyces cerevisiae*

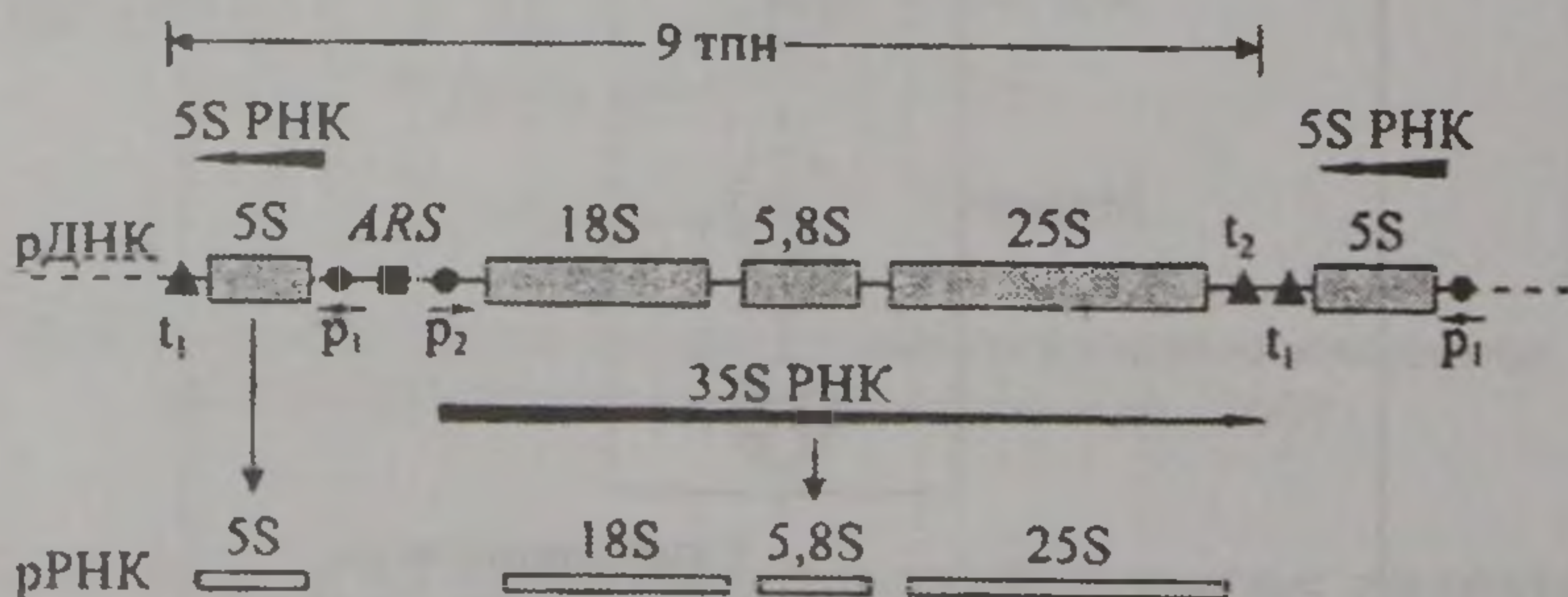


Дрожжи *S. cerevisiae* стали первым эукариотическим организмом, у которого в 1996 г. была определена полная нуклеотидная последовательность. Суммарное содержание ДНК в гаплоидном ядре составило 12 млн 68 тыс. п. н., М. м.  $1,2 \cdot 10^{10}$  Да. Это количество всего в 3 раза превышает содержание ДНК у бактерии *E. coli* (4.639 млн п. н.). У многих представителей эукариот размер генома превышает размер *S. cerevisiae* на 2–3 порядка величин. Хромосомы дрожжей по содержанию нуклеотидов находятся в пределах от 250 до 2200 тыс. п. н. Предполагаемое число генов составило 6102, а при анализе нуклеотидной последовательности выявлено 6034 потенциальных ORF.

Относительно небольшой размер генома дрожжей объясняется небольшим числом генов (4 %) с интронами, а также низким содержанием повторяющихся последовательностей в составе ДНК. В то же время на хромосоме XII идентифицирован кластер генов рРНК с многократно повторяющимися последовательностями ДНК (рис. 75). Таких тандемно повторяющихся единиц генов рРНК приходится по 100–140 копий на гаплоидную клетку. Длина каждой единицы составляет 9 тыс. п. н., при этом в нее входят четыре гена, кодирующие 25S, 18S, 5,8S и 5S рРНК. Рибосомные РНК с константами седиментации 5,8S, 18S и 25S берут начало из общего высокомолекулярного предшественника — 35S РНК, тогда как 5S рРНК независимо транскрибируется с другой цепи ДНК.

Как и все эукариоты, клетки дрожжей содержат митохондрии с митохондриальной ДНК (мт ДНК) в виде молекулы кольцевой формы с молекулярной массой около 50 мДа (или 75 тыс. п. н.). На гаплоидную клетку приходится от 10 до 40 копий мтДНК, при этом мтДНК не связана с гистонами.

Большинство (около 75 %) лабораторных штаммов *S. cerevisiae* содержат плазмиды — внехромосомные кольцевые двухцепочечные молекулы ДНК, которые названы Scp 1. В гаплоидной клетке может содержаться от



**Рис. 75.** Организация повторяющейся единицы генов рРНК и 35S РНК — предшественника трех классов рРНК (Щелкунов, 2004);  $t_1$  и  $t_2$  — терминаторы транскрипции рибосомных оперонов; ARS — область начала репликации ДНК. Направление транскрипции показано стрелками

30 до 200 копий Scp 1. Функции плазмид в клетках дрожжей не выяснены, но предполагают, что жизненно важных функций клетки Scp 1 не кодирует. Длина ДНК Scp 1 равна 6318 п. н. На ней картированы четыре гена — FLP, REP1, REP2 и RAF, белковые продукты которых участвуют в поддержании плазмиды.

После полной расшифровки генома дрожжей *S. cerevisiae* С. Филдс с сотрудниками (1997) предприняли широкие исследования по выявлению генов дрожжей, белковые продукты которых взаимодействуют между собой *in vitro* в масштабе полного генома. Для охвата всего генома необходимо было клонировать 6102 ORF дрожжей с помощью нового методического подхода с использованием полимеразной цепной реакции (ПЦР). В реальном эксперименте с помощью ПЦР удалось успешно амплифицировать 6040 ORF *S. cerevisiae* (99,1 %). Проведенный глобальный анализ С. Филдсом с большой группой сотрудников (2000) *белок-белковых взаимодействий* для подавляющего большинства потенциальных ORF дрожжей показал возможности экспериментальной молекулярной генетики в изучении функционирования сложных геномов как на уровне транскрипции генов, так и на уровне белок-белковых взаимодействий.

## Геном нематоды *Caenorhabditis elegans*

Мелкий круглый червь (нематода) *C. elegans* имеет размер около 1 мм, состоит из 959 клеток, имеет 302 нервные клетки (нейрона), живет около двух недель. Объект обладает замечательным свойством — *полной прозрачностью*, что позволяет следить за поведением и судьбой каждой отдельной клетки.

Секвенирование *C. elegans* началось в 1990 г. и протекало с темпом анализа 1 млн пар нуклеотидов в год. При сохранении такого темпа на расшифровку всего генома понадобилось бы около 100 лет. Однако именно в этот период начался процесс объединения усилий разных лабораторий, формирование международных геномных проектов, а также создавались высокопроизводительные методы секвенирования ДНК. В ходе восьмилетних исследований в декабре 1998 г. опубликована серия статей (*Science*, 1998) по результатам расшифровки строения генома этого первого многоклеточного животного. Данное исследование явилось своего рода предвестником того, что в ближайшие годы достижима расшифровка строения всего генома человека.

Секвенирование генома нематоды проводилось двумя исследовательскими центрами: Центром геномного секвенирования (Вашингтон, США) и Сенгеровским центром (Кембридж, Великобритания). В работе принимало участие большое число авторов, поэтому журнал не стал публиковать их список, и отослал читателей в Internet, а коллектив ученых назвал

«Консоциум секвенаторов *C. elegans*». Этот коллектив внес значительный вклад в расшифровку генома человека.

Размер генома *C. elegans* составляет 97 млн п. н., предполагаемое число генов 19.099, из них идентифицировано 36 % (7000 генов). Проводился также компьютерный анализ соотношения в геноме кодирующих и некодирующих областей. Экзоны и интроны занимают примерно равные доли — 27–26 % соответственно. Оставшиеся 47 % генома приходится на повторы, межгенные участки и т. д., т. е. на ДНК с неизвестными функциями. При сравнении данных по нематоды с дрожжевым геномом и геномами растений, насекомых и человека выясняется, что в ходе эволюции доля кодирующих участков в расчете на весь геном резко уменьшается — у дрожжей она очень высока, у человека очень мала. Л. Л. Киселев (2000) предложил, что эволюция эукариот от низших форм к высшим сопряжена с «разбавлением генома», так как на единицу длины ДНК приходится все меньше информации о структуре белков и РНК и все больше информации «ни о чем», то есть для нас непонятной и непрочитанной. Это одна из больших загадок эволюции.

В геноме нематоды обнаружено большое количество *высококонсервативных доменов*, участков полипептидных цепей белка или определенных участков хромосом (ДНК). Их существенная часть находится в белках, вовлеченных в межклеточные взаимодействия и в регуляцию транскрипции. Среди таких доменов у нематоды обнаружено 650, характерных для трансмембранных хеморецепторов и 140 трансмембранных рецепторных доменов, родственных родопсину; 410 доменов, встречающихся у эукариотических протеинкиназ; множество различных Zn-фингерных доменов, компонентов факторов транскрипции; 90 гомеобоксных доменов; 70 доменов, характерных для фактора роста эпидермиса и столько же доменов иммуноглобулинового типа. В эволюции наблюдается использование перетасовки готовых доменов и создание новых функций организма за счет новых комбинаций стандартных модулей. По мере усложнения организма увеличивается количество многодоменных белков.

Эукариоты обладают множеством генов, кодирующих присущие только им продукты, обуславливающие специфику эукариотической клетки: ядро, цитоскелет, органеллы, хромосомную организацию генетического материала и т. д. Большинство соответствующих генов в бактериальной клетке отсутствует, в частности, гены, кодирующие элементы цитоскелета. С появлением в эволюции такой организации эукариотической клетки возникает изменение аппарата транскрипции, в частности, появляются регуляторные элементы — Zn-фингерные домены (компоненты факторов транскрипции). При переходе от бактерий к дрожжам и далее, к нематоды, меняется содержание генов, кодирующих Zn-фингерные белки. У бактерий и археобактерий отсутствуют большие семейства Zn-связывающих доменов,

имеющихся у эукариот. Дрожжи и другие эукариоты, с увеличением размера генома и усложнением структуры клетки, содержат специальные белки для импорта и экспорта цинка из клетки.

Расшифровка геномов бактерий, дрожжей и нематоды дает биологам реальную возможность сравнения не отдельных генов или генных ансамблей, а целых геномов. Например, многие уникальные гены нематоды в отличие от генов дрожжей могут быть связаны с межклеточными взаимодействиями, возникающими у многоклеточного организма. У человека число генов в два раза больше, чем у нематоды. Предполагается, что часть генома человека должна иметь гомологов среди генов дрожжей и червя. При этом функции генов нематоды изучать несравненно легче, чем аналогичные гены у человека. Гены нематоды можно легко изменить путем мутаций или убить, прослеживая при этом изменения свойств организма. Таким путем можно выявлять биологическую роль генных продуктов у представителя низших многоклеточных с последующей экстраполяцией данных на другие организмы.

## Геном плодовой мухи *Drosophila melanogaster*

Плодовая, или уксусная, муха *D. melanogaster* принадлежит к семейству *Drosophilidae* из отряда двукрылых насекомых *Diptera*. Это очень маленькая мушка, величиной около 3 мм, с ярко-красными глазами и серым телом. К настоящему времени описано более тысячи видов, относящихся к этому семейству. Генетически более изучены *D. melanogaster*, *virilis*, *funnebris*. Дрозофила — один из прекраснейших объектов, обладающий всеми качествами, необходимыми для успешного проведения генетического анализа и исследований. Она широко используется в генетических экспериментах, начиная с 1909 г. в американской школе генетики Т. Х. Моргана. Ряд теоретических вопросов генетики — искусственное получение мутаций и природа гена, определение пола и локализация генов в хромосомах, проблема осуществления гена, генетика популяций, механизм расо- и видообразования и многие другие проблемы — интенсивно изучаются на дрозофиле и дали важные результаты для решения не только вопросов частной генетики, но и общей биологии и эволюции видов.

В конце 70-х гг. одновременно в СССР — группой советских генетиков во главе с академиком Г. П. Георгиевым, и в США — группой американских генетиков во главе с Д. Хогнессом, были открыты мобильные генетические элементы (МГЭ) у дрозофилы. С этого момента исследования МГЭ существенно ускорились. Мир МГЭ оказался велик и многолик — различные МГЭ были обнаружены также у дрожжей, млекопитающих, включая человека, и других объектов.

Таблица 20

Геном *Drosophila melanogaster* [Adams et al., 2000]

Показатель	Размер	%
ДНК эухроматиновой части генома	$1,2 \times 10^8$ п. н.	100
Кодирующая часть генома	$6,0 \times 10^6$ п. н.	5
Число локализованных генов по результатам секвенирования	13 600	
Некодирующая часть генома	$1,14 \times 10^8$ п. н.	95
Геномная система МГЭ	$1,1 \times 10^7$ п. н.	10
Число дисков на цитологической карте поли- тенных хромосом	5100	
Число сегментов на карте Бриджеса:		
1-й уровень	102	
2-й уровень	612	
Размер сегментов 1-го уровня	190–1820 тыс. п. н.	

Международная программа полного секвенирования генома дрозофилы представила к 2000 г. полную последовательность ДНК с указанием положений всех известных генов, других генов с неизвестными пока функциями (ORF), регуляторных сайтов и других функциональных субструктур (Adams et al., 2000). Без учета прицентромерного гетерохроматина геном дрозофилы содержит 120 млн п. н. ДНК с числом предполагаемых генов ~13.600, кодирующих различные типы РНК и белков (табл. 20). По более поздним уточненным данным, размер генома с прицентромерным гетерохроматином, составил 180 млн п. н. с числом генов 13.061 (Щелкунов С. Н., 2004). Методы оценки предполагаемого числа генов и их белков допускают некоторые неоднозначности в интерпретации. Поэтому оцененные числа генов не являются окончательными. Оценки некоторых ORF в качестве генов могут уточняться.

Непосредственное участие в кодировании принимает около 5 % суммарной ДНК генома, а остальные 95 % считаются некодирующими: это *интроны, межгенные промежутки, содержащие знаки управления — энхансеры, инсуляторы* и другие участки, возможно, выполняющие иные функции. Приведенные в списке гены достаточно хорошо изучены, они секвенированы и клонированы. Для половины секвенированных генов установлены их функции или сходство с известными функционирующими генами, выявлены их белковые продукты, а для многих — механизмы управления. Другая половина секвенированных генов пока функционально не описана, а на основе их ORF проводится изучение по базам данных.

Особенностью дрозофилы являются огромные политенные хромосомы клеток слюнных желез с характерной поперечной исчерченностью в виде 5100 окрашенных дисков, или *участков плотной упаковки*. Порядок расположения дисков рассматривается как своеобразная хромосомная карта — *карта Бриджеса*, которая устойчива при наследовании и характерна для вида. Эта карта разделена на 102 сегмента 1-го уровня, 612 сегментов 2-го уровня и более мелкие сегменты. Цитологическая карта Бриджеса охватывает плечи больших хромосом, содержащие 75 % гаплоидного набора ДНК и организованные в диски. Остальные 25 % ДНК образуют районы прицентромерного гетерохроматина, где дисковая организация отсутствует.

Геномная система МГЭ дрозофилы обобщена в табл. 21. В геноме содержится 50 различных семейств МГЭ, они занимают 10 % ДНК генома. Средний размер копии МГЭ составляет 3–5 тыс. п. н., а каждая копия МГЭ в среднем окружена фрагментом геномной ДНК длиной до 50 тыс. п. н. число копий МГЭ в геноме дрозофилы оценивается в 3–5 тыс.

На основе этих данных Л. А. Васильева и В. А. Ратнер (2000) предложили концепцию, в основе которой лежит представление о МГЭ генома как о *геномной системе*, на долю которой приходится существенная эволюционная роль. Специфические сайты локализации МГЭ каждого семейства заполнены лишь на 25–30 %, поэтому МГЭ имеют достаточный простор для перемещений. Сайты локализации МГЭ в плечах больших хромосом не случайны, а сосредоточены в окрестностях функционирующих генов и их регуляторных зон.

Копии МГЭ способны модифицировать функции менделеевских генов и полигенов посредством энхансеров. *Энхансеры* — это функциональные сайты, знаки управления, узнаваемые регуляторными белками и активирующие транскрипцию смежных генов и полигенов. Зона влияния энхансера может простираться до 50–100 тыс. п. н. вдоль цепи ДНК, т. е. в зоне действия системы МГЭ находятся практически все гены генома.

Таблица 21

Геномная система МГЭ дрозофилы

Показатель	Характеристика
Число семейств МГЭ	50
Число копий МГЭ на семейство	1–100
Средний размер копии МГЭ	3000–5000 н. п. (10 %)
Средний размер фрагмента ДНК на 1 копию МГЭ	30 000–50 000 н. п. (100 %)
Средний размер кодирующего фрагмента на 1 копию МГЭ	1500–2500 н. п. (5 %)
Средний размер кодирующей части гена	1000 н. п.
Среднее число генов на 1 копию МГЭ	2–3
Среднее число копий МГЭ на геном	3000–5000

При сопоставлении МГЭ геномов родственных видов дрозофилы выявлены случаи «горизонтального» переноса МГЭ между видами. Вид *D. melanogaster* приобрел семейство Р-элементов совсем недавно, в XX веке, от других видов дрозофил. Это своеобразное «заражение» произошло у популяций Нового Света, тогда как линии, взятые из природных популяций в начале века и позднее в Старом Свете, не содержали Р-элементов.

Структура относительно протяженного «среднего» генома классического генетического объекта — плодовой мухи *D. melanogaster* — расшифрована при участии частной американской фирмы «Селера Дженомикс». При сопоставлении этих данных с результатами изучения генома дрозофилы, полученными мировым научным сообществом за девять десятилетий XX века, оказалось, что большинство генов, входящих в описанный континг\*, выявлены ранее. Однако он охватывает лишь 2/3 от общей протяженности генома, его эухроматическую часть. Остальная гетерохроматическая треть состоит из различных повторяющихся последовательностей, не поддающихся расшифровке.

А. В. Зеленин, Е. Д. Бадаева, О. В. Муравенко (2001) уточнили вопрос о значении понятия «полное секвенирование генома». В ходе многих исследований, в том числе при создании программы «Геном человека», оно подразумевало установление линейной последовательности всех нуклеотидов генома. Такое представление адекватно в отношении «сверхмелких» геномов вирусов и прокариот, а также и применительно к геному нематоды *C. elegans* (97 млн п. н.), в котором определено линейное расположение 97 % всех нуклеотидов. Приблизительно такая же картина наблюдается и в отношении генома представителя крестоцветных — арабидопсиса (120 млн п. н.), для которого известна первичная структура 94 % генома. Такой уровень секвенирования геномов принято называть «по существу завершенным» (*essentially completed*).

Секвенирование генома дрозофилы открывает другой уровень оценки (размер генома 180 млн п. н.). Составленный континг, содержащий 13.061 генов, охватывает лишь эухроматическую часть (120 млн п. н.). остальные 33 % генома относятся к протяженным гетерохроматическим районам, секвенирование которых на современном уровне знаний практически невозможно и вряд ли целесообразно. Такой уровень секвенирования генома принято называть «в основном завершенным» (*substantially completed*). Таким образом, выясняется, что для всех геномов, в том числе и для малых, в обозримом будущем будет ставиться вопрос не о полной нуклеотидной последовательности, а лишь об эухроматической части, а также о составлении генного континга, т. е. определении линейного расположения генов в индивидуальных хромосомах и целом геноме.

---

\* Континг — совокупность перекрывающихся смежных сегментов ДНК.

## Особенности исследований геномов высших растений

Исследования по структурной геномике растений столкнулись с большими методическими трудностями, связанными с особенностями высших растений. Они включают специфику структуры хромосом, а также расположения гетерохроматина и его количества в геноме.

1. Огромные размеры геномов. Приведем примеры в сравнении с «мелкими» и «средними» геномами:
  - кишечная палочка *E. coli* — 4,6–5,5 млн п. н.;
  - дрожжи *S. cerevisiae* — 12,1 млн п. н.;
  - круглый червь *C. elegans* — 97 млн п. н.;
  - мелкое цветковое растение резуховидка класса двудольных *A. thaliana* — 120 млн п. н. (без гетерохроматиновых повторов);
  - плодовая муха *D. melanogaster* — 180 млн п. н.;
  - рис *Oryza sativa* — 388,8 млн п. н.;
  - человек *Homo sapiens* — 3 млрд п. н.;
  - рожь посевная *Secale cereale*, культурный ячмень *Hordeum vulgare*, диплоидная пшеница *Triticum monosocum* — 6–7 млрд п. н.;
  - твердая (гексаплоидная) пшеница *T. aestivum* — 16 млрд п. н.;
  - лилейные (*Lilium*) — 50–60 млрд п. н.
2. Большое число полиплоидных форм.
3. Обогащенность геномов растений (до 99 %) некодирующей, т. е. не содержащей структурных генов ДНК, что серьезно затрудняет секвенирование, картирование генов и создание карт контигов и генных контигов.
4. Слабое, по сравнению с геномами дрозофилы и человека, морфологическое, генетическое и физическое картирование хромосом.
5. Большое число «мелкохромосомных» видов, а также видов растений с неизвестным числом хромосом и количеством ДНК.

В геномике растений используются два направления хромосомных технологий:

1. Клонирование и секвенирование геномов.
2. Изучение геномов на базе сравнительной геномики.

Стратегия секвенирования геномов основана на двух методических подходах. *Первый подход, классический*, сводится к предварительному разбиению генома на части и получению генетических и молекулярных маркеров. После постепенного изучения различных участков хромосом (*clone to clone technology*) их клонируют, секвенируют отдельные клоны и составляют локальные контиги. Конечный этап состоит в установлении полной нуклеотидной последовательности участков хромосом и целых



хромосом. Исследования по программе «Геном человека» проводили по классическому пути. Геном «похромосомно» разбивали на 24 части. Чистые фракции индивидуальных хромосом получали с помощью *метода проточной флуорометрии*. Из этого материала создавали *хромосомоспецифические клонотекы*, а также на основе участков хромосом (плечей, локусов, бэндов) — *плечо-, локус- и бэнд-специфические клонотекы*.

С помощью этого же классического подхода международные консорциумы осуществили полное секвенирование геномов первого эукариотического организма — дрожжей *S. cerevisiae*, первых многоклеточных организмов — круглого червя *C. elegans* и двудольного растения *A. thaliana*.

*Второй подход* основан на одномоментном разбиении генома на части (*shotgun technology*) и последующими этапами:

- клонирование полученных фрагментов в одном из вирусных векторов,
- получение значительного числа отдельных клонов,
- их частичное секвенирование и составление контига,
- заключительный этап — установление полной последовательности нуклеотидов в геномной ДНК.

Для получения хромосом- и локус-специфичных клонотек растений пока не удается использовать подходы, применяемые в геномике человека и животных. до сих пор не разработана методика получения *чистых фракций индивидуальных хромосом* растений с помощью сортировки в потоке. У некоторых видов хозяйственно-важных растений — риса, ржи, ячменя и сахарной свеклы — клонотекы ДНК получают методом лазерной микродиссекции. Для видов с мелкими хромосомами необходимо создать метод *дифференциального окрашивания*, который позволит в дальнейшем клонировать полученную ДНК. Вызывает интерес *G-подобное окрашивание*, которое не разрушает ДНК, позволяет *идентифицировать индивидуальные хромосомы* или локусы, подобно тому, как это делается в отношении хромосом человека.

Растение арабидопсис, геном которого подвергнут секвенированию, не имеет никакого хозяйственного значения. Однако его геном уже используется в качестве «базового» или «справочного» генома. С этой же целью успешно используются в сравнительной геномике растений уже имеющиеся сведения о геноме риса. Примером может служить установленное с помощью сравнительного картирования локализация гена чувствительности к яровизации VRn-A1 на хромосоме 5A гексаплоидной пшеницы, а на хромосоме 3 риса — соответствующего ему гена Hd-6. Развитие геномики растений будет связано с установлением структурных аналогий и генетических взаимосвязей отдельных локусов и хромосом различных геномов, и будет основываться на знаниях о тотальной структуре геномов арабидопсиса и риса. Возможно, к числу «справочных» организмов может добавиться и какое-нибудь растение с крупным геномом, например, ячмень,

геном которого больше генома риса в 10 раз при примерно одинаковом числе генов.

## Геном резуховидки (арабидопсиса) *Arabidopsis thaliana*

Это небольшое растение, однолетнее, реже двулетнее травянистое, стебель высотой 4,5–7 см. имеет большое число подвидов, разновидностей и рас. Род *Arabidopsis* из семейства крестоцветных Brassicaceae (Cruciferae) состоит из 13 видов с ареалом в умеренных странах Северного полушария. Число хромосом  $2n = 10$ . В статье «Поиски ботанической дрозофилы» (Титова Н. Н., 1935) это растение рекомендовано наиболее подходящим претендентом на звание «ботанической дрозофилы» из-за короткого жизненного цикла и удобства при использовании в качестве объекта генетики.

Таблица 22

### Геном *Arabidopsis thaliana*

Показатели	Размер
Размер генома, п. н.	115.409.949
Размер хромосомы 1, п. н./число генов	29.105.111/6.543
Размер хромосомы 2, п. н./число генов	19.646.945/4.036
Размер хромосомы 3, п. н./число генов	23.172.617/5.220
Размер хромосомы 4, п. н./число генов	17.549.867/3.825
Размер хромосомы 5, п. н./число генов	25.953.409/5.874
Состав оснований (% GC)	35,0
Число генов на геном	25.498
Средний размер гена, п. н.	2.013
Средняя длина пептида, п. н.	434
Экзоны	
Количество на геном	132.982
Средний размер экзона, п. н.	251
Среднее число экзонов на 1 ген	5.14
Интроны	
Количество на геном	107.484
Средний размер интрона, п. н.	169
Протеом	
Всего белков	25.498
Среднее число белков на 1 хромосому	5.100

Для проведения секвенирования генома и идентификации генов в 1990 г. был основан Международный проект по изучению генома *Arabidopsis*, результаты секвенирования опубликованы в 2000 г. (*Nature*. V. 408. 14 Dec. 2000). Состав генома представлен в табл. 22.

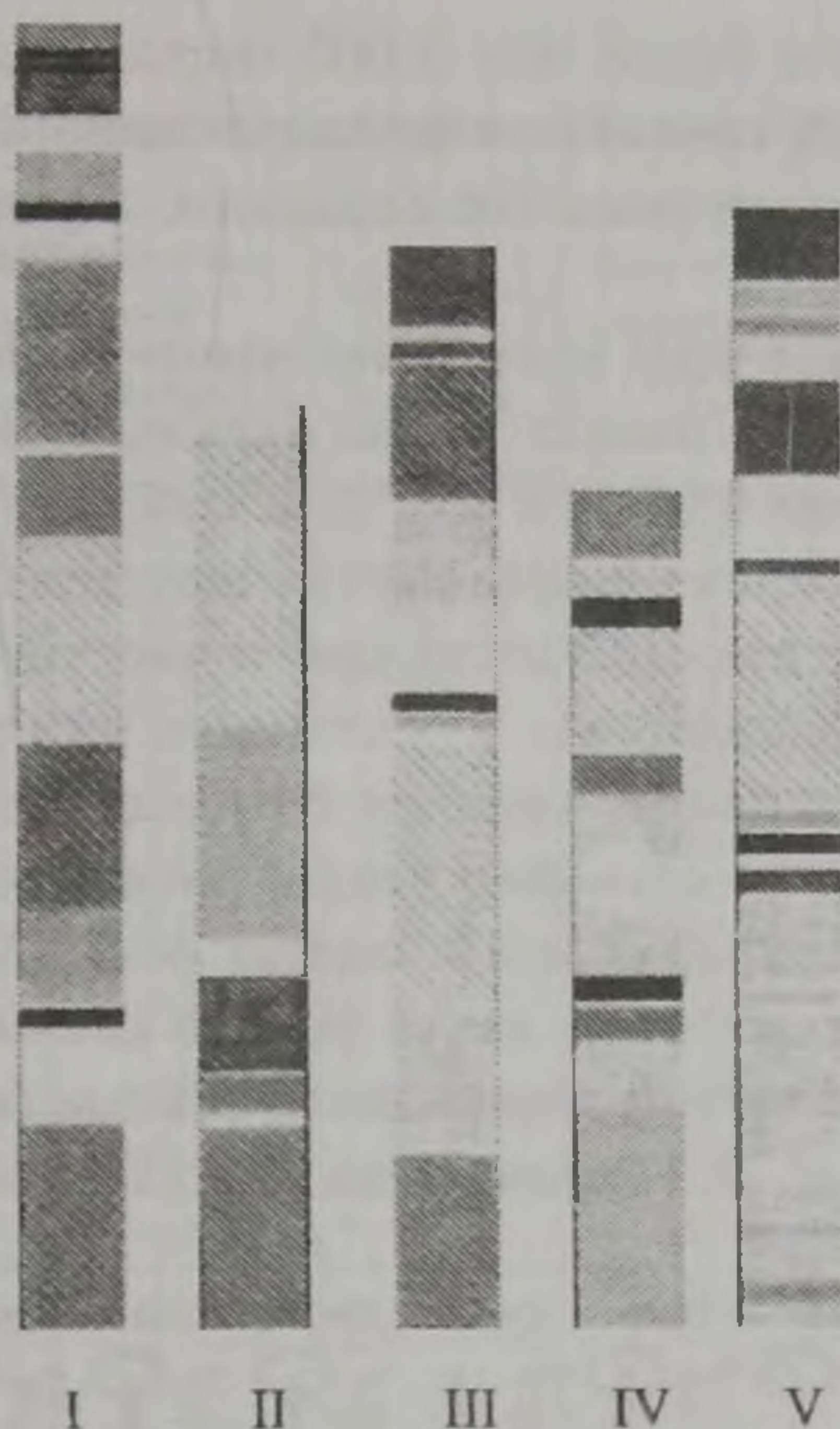
Количество ДНК в геноме *Arabidopsis* составило 115.409.949 п. н., с числом генов 25.498. средний размер гена равен 2.013 п. н., а средний размер пептида — 434 п. н.

В геноме содержится 35 % GC-пар. Количество экзонов на 1 ген было 5,14 при среднем размере экзона 251 тыс. п. н. Средний размер интрона составил 169 п. н. Протеом *Arabidopsis* включает 25.498 белков, а одна хромосома в среднем кодирует 5100 белков. В геноме у *Arabidopsis* имеется также около 80 хлоропластных и около 60 митохондриальных генов, кодирующих белки. Большой интерес представило открытие у этого вида несколько десятков генов, имеющих сходные последовательности нуклеотидов с генами наследственных болезней человека, например, синдром Менке (Вильсона). Геном *Arabidopsis* имеет небольшие размеры по сравнению с видами злаков — у риса геном в 3 раза больше, у кукурузы в 20 раз, а у пшеницы в 120 раз, по сравнению с геномом *Arabidopsis*.

Среди отсеквенированных в настоящее время геномов, геном *Arabidopsis* имеет наибольшее количество предсказанных 25.498 генов: число генов у *C. elegans* 19.099, у *Drosophila* — 13.061. По плотности генов близки геномы *Arabidopsis* и нематоды, у дрозофилы она более низкая. Характерные для арабидопсиса значительная протяженность tandemных дупликаций генов и дупликаций сегментов, возможно, объясняют большой размер генома и число генов у этого вида. В центромерных районах хромосом, содержащих большие сегменты монотонных повторов 5S рРНК, полное секвенирование не проводилось (рис. 76).

Для оценки сходства и отличий генного набора арабидопсиса с другими эукариотами, проводили его сравнение с известными базами данных и их функциональной классификацией белковых последовательностей. На основе сходства последовательностей белков с таковыми других видов классифицировали около 70 % генов арабидопсиса.

Анализ микроструктуры геномов растений позволяет сравнивать детали строения, а также степень родства арабидопсиса и культурных растений, давно претерпевших эволюционную дивергенцию. Установлено большое сходство с другим крестоцветным *Capsella rubella* (пастушья сумка) по составу и организации генома, по количеству многочисленных дупликаций, по длине экзонов и положению интронов. Показан высокий уровень (85 %) сходства положения нуклеотидов в кодирующих участках генома капусты (*Brassica*) с ортологичными генами арабидопсиса. Отмечены также отдельные сегменты генома арабидопсиса с участками генома томата (пасленовые) и риса (злаковые).



**Рис. 76.** Хромосомный набор *Arabidopsis thaliana*. Одинаково окрашенные сегменты указывают на наличие дупликаций, белые участки несут уникальные последовательности, заштрихованы центромерные повторяющиеся элементы. Теломерные повторы и ядрышковые организаторы не указаны.

Широкое распространение в геномах растений тандемных генных дупликаций, делеций отдельных генов и небольших групп смежных генов позволили сделать предположение о том, что ключевым механизмом эволюции микроструктуры геномов растений является *неравный кроссинговер*.

### Геном риса посевного *Oryza sativa* L.

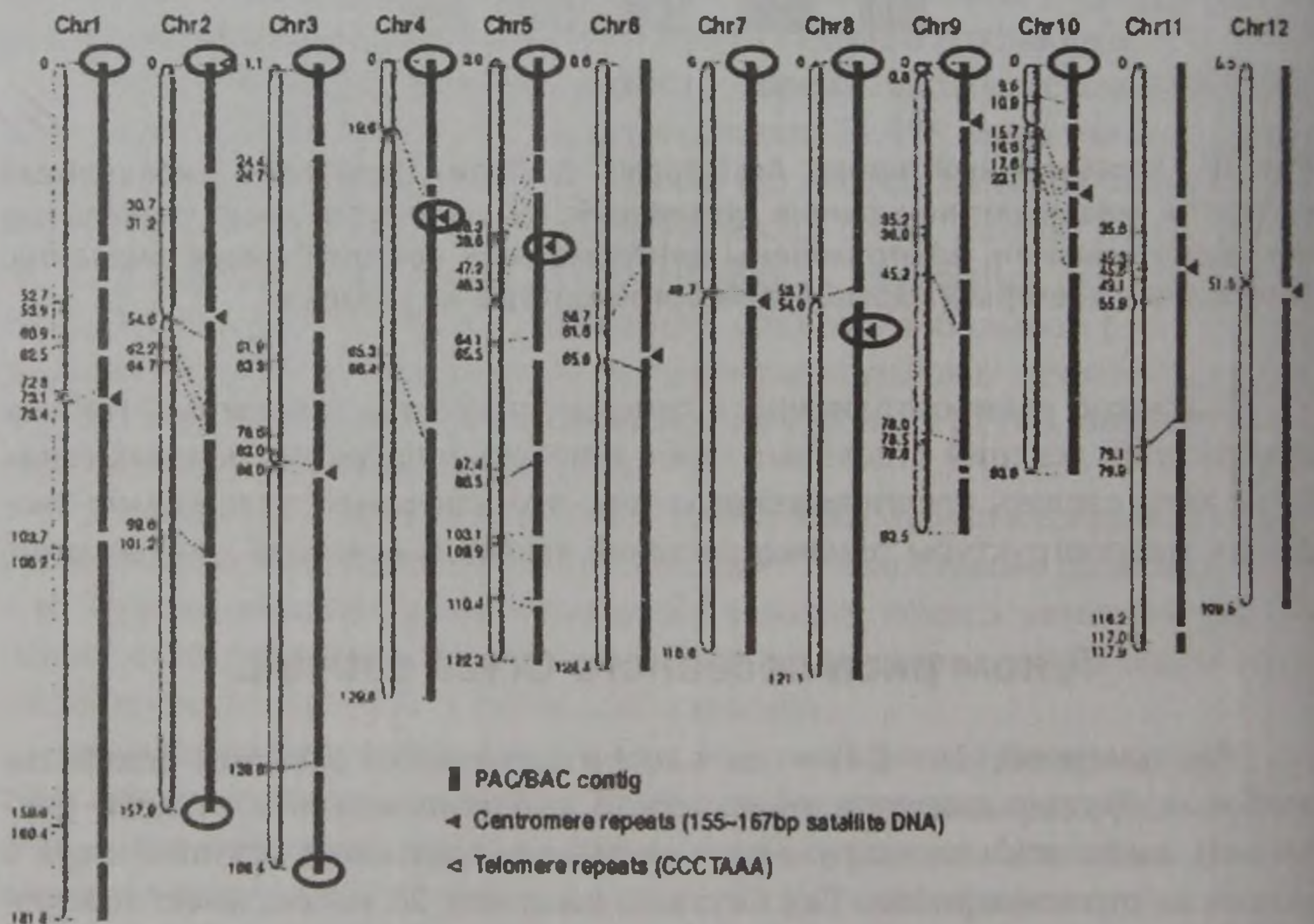
Рис посевной ( $2n = 24$ ) — однолетнее автогамное растение семейства злаковых. Возделывается в тропической, субтропической и отчасти умеренной зонах земного шара. Более половины населения земного шара в основном питается рисом. Род *Oryza* L. содержит 28 видов, имеет пантропическое распространение. В культуру издревле были введены только 2 вида — *O. sativa* в Азии (Китай и Индия) и *O. glaberrima* в Африке (Жуковский П. М., 1971).

Для декодирования генома риса в 1991 г. в Японии была составлена Программа по секвенированию генома риса (RGP), рассчитанная на срок 21 год. Исследования протекали успешно, и к 1997 г. создана молекуляр-

но-генетическая карта с более чем 3.000 маркерами ДНК. Тогда же был основан Международный проект по секвенированию: 12 хромосом риса были поделены между 10 участвующими странами, а ученым Японии отошли хромосомы 1, 2, 6, 7, 8 и 9.

К 2005 г. было закончено конструирование точной физической карты риса с общей длиной 370,7 млн п. н. она включала 95,3 % от размера полного генома, оцениваемого в 388,8 млн п. н. (рис. 77).

Размер генома оказался меньше прогнозируемого ранее. Все пары нуклеотидов определены с высокой степенью точности. Участки генома, обогащенные повторами одинаковых нуклеотидов или транспозонами, будут подвергнуты дополнительному анализу с применением других методических приемов. Результаты секвенирования риса представлены в табл. 23. Размеры 12 хромосом варьируют в пределах от 22,7 до 43,3 млн п. н., т. е. почти в 2 раза. Плечи хромосом несут разное количество гэпов (gaps) — от 1 до 6. Степень от секвенирования всего генома составила 95,3 %, а отдельных хромосом — от 74,3 (хромосома 9) до 99,5 % (хромосома 8).



**Рис. 77.** Физическая карта риса *O. sativa* ssp. *japonica*, cultivar *Nipponbare* (Sasaki, 2005). Левая и правая вертикальные линии представляют генетическую и физическую карты для каждой хромосомы. Треугольник на каждой физической карте указывает положение центromеры. Окруженный участок центromеры соответствует секвенированным контигам. Теломеры, обозначенные кружками, подтверждают наличие специфического теломерного повтора CCCTAAA

Таблица 23

Результаты секвенирования хромосом риса (Sasaki T., 2005)

Хромо- сома	Отсеквени- ровано п. н.	Гэпы * в плечах хромосом		Теломер- ные гены, млн п. н.	Центро- мерные гены, млн п. н.	рДНК, млн п. н.	Общая длина, млн п. н.	Степень обработ- ки, %
		Число	Общая дл. млн п. н.					
1	43.260.640	5	0,33	0,06	1,40		45,05	96,0
2	35.954.074	3	0,10	0,01	0,72		36,78	97,7
3	36.189.985	4	0,96	0,04	0,18		37,37	96,8
4	35.489.479	3	0,46	0,20	0,00		36,15	98,2
5	29.733.216	6	0,22	0,05	0,00		30,00	99,1
6	30.731.386	1	0,02	0,03	0,82		31,60	97,2
7	29.643.843	1	0,31	0,01	0,32		30,28	97,9
8	28.434.680	1	0,09	0,05	0,00		28,57	99,5
9	22.692.709	4	0,13	0,14	0,62	6,92	30,53	74,3
10	22.683.701	4	0,68	0,13	0,47		23,96	94,7
11	28.357.783	4	0,21	0,04	1,90	0,25	30,76	92,2
12	27.561.960	0	0,00	0,05	0,16		27,77	99,2
Всего	370.733.456	36	3,51	0,81	6,59	7,20	388,82	95,3

\* Гэпы (*gaps*) — пробелы, разрывы на карте секвенирования участка генома. Для каждой хромосомы риса гэпы отмечены на физической карте (рис. 76)

При проведении секвенирования и расшифровки генома риса проводилось генное моделирование, изучались дубликации генов в хромосомах, хромосомные дубликации, распределение простых повторяющихся последовательностей (SSR — simple sequence repeat), распространение мигрирующих элементов, точек вставки эндогенных ретротранспозонов риса Tos 17, а также инсерций геномных последовательностей хлоропластов и митохондрий. На основе компьютерной генной программы *F gene SH*, общее число генов, включая мигрирующие элементы, составило 37.664 гена. Средняя плотность генов определена 9,9 тыс. п. н. на 1 ген, а у арабидопсиса она была 4,5 тыс. п. н. на 1 ген, т. е. в 2 раза выше, чем у риса. Причиной более низкой плотности генов у риса может быть большая длина интронов.

Повторное аннотирование нуклеотидной последовательности генома риса осуществлялось также методом использования полной длины кДНК. Такой прием убедительно показал, что в геноме риса имеется 30 тыс. генов, т. е. меньше, чем прогнозировалось с применением ранее программы *F gene SH*. Последние достижения молекулярной биологии внесли некоторые уточнения в само понятие «ген». Считается общепринятым, что ген представляет последовательность ДНК, которая транскрибируется в белок.

Оказалось, в белки транслируется не все РНК, так как некоторые из них являются регуляторами транскрипции других генов. Эти РНК, не кодирующие белки, предложено *классифицировать как гены*. В геноме риса, как и у других эукариот, локализовано большое количество мигрирующих элементов — транспозоны ДНК и РНК занимают до 35 % генома риса.

В геномном проекте по рису конструировались плазмидные РАС-библиотеки и бактериальные ВАС-библиотеки кДНК. Так, на основе сорта Nirponbaga получено около 30 тыс. кластеров последовательностей кДНК. Кроме того, использовались частичные последовательности кДНК, так называемые EST (expressed sequences tag — короткие секвенированные последовательности ДНК, изолированные из библиотеки кДНК) риса и родственных видов злаков — сорго, пшеницы, ячменя и кукурузы. Наряду с другими методами, применялся прием аннотирования для определения биологической и функциональной роли нуклеотидных последовательностей генома. В результате, из 30 тыс. генов определены функции 19.969 аннотированных генов, а также 131 гена небелковой РНК.

Осуществление Международного проекта по секвенированию генома риса (IRGSP) является огромным научным и практическим достижением молекулярной генетики и геномики. Ведущий специалист этого проекта Такуи Сасаки (Sasaki T., 2005) оценил это достижение как *«приобретение сокровища для человечества («...is very important treasure for human beings»)». Эти результаты будут использоваться в практике селекции, сравнительной геномике растений для сравнения функциональных участков генома. Наряду с арабидопсисом, рис теперь будет включен в число «справочных» организмов для реализации последующих геномных проектов на злаковых растениях с крупными геномами.*

## Геном домашней мыши *Mus musculus* L.

Мелкие мыши без заметно выраженной специализации. Род насчитывает более 30 видов. Представители большинства европейских и средиземноморских таксонов имеют одинаковое диплоидное число хромосом  $2n = 40$  и не изменяющийся по внешней морфологии хромосом кариотип (Соколов В. Е., 1969).

В качестве лабораторных животных мыши издавна служили предметом особо пристального внимания исследователей, поскольку они в больших количествах размножаются для биологических и медицинских опытов и находятся под наблюдением всей или большей части своей жизни. Мыши отличаются высокой плодовитостью, длительность их жизни меньше, чем у других лабораторных животных, благодаря чему имеется возможность получить три поколения в год. В настоящее время известно более 200 важнейших и наиболее распространенных линий и сублиний мы-

шей. Все они выведены генетиками почти исключительно во всемирно известной Джексонской лаборатории в США (The Jackson Laboratory, Bar Harbor, Maine) (Медведев Н. Н., 1990). Эта лаборатория возглавляет проект «Информатика генома мыши», который включает шесть баз данных: «Экспрессия генома мыши», «Биология новообразований у мыши», «Фенотип мыши» и др. В последней из них представлены фенотипические характеристики основных линий мышей. В базе данных геномов мыши описано более 1000 спонтанных мутаций, из них 128 охарактеризованы по нуклеотидным последовательностям ДНК. 56 мутантных генов (~45 %), вызывающих наследственные аномалии, имеют гомологичные гены у человека с соответствующими аномалиями. Из этих 56 мутаций 35 первоначально были обнаружены у мышей, а 21 — у человека, затем у мышей.

Переломным моментом развития геномики явилось опубликование в 2001 г. основных характеристик первого («чернового») варианта генома человека. На очередь встала проблема правильного и точного извлечения информации, закодированной у человека в 3 миллиардах нуклеотидов. Эта информация представляет собой каталог генов всех белков и РНК, а также регуляторных элементов, обеспечивающих экспрессию генов, структурных элементов, которые управляют функционированием хромосом и эволюционной историей человека. На этом уровне интенсивно используются общедоступные базы данных GenBank, EMBL, DDBJ и др., содержащие полные геномы сотен видов, генные и полипептидные последовательности, транскрипты, EST, SNP и т. д.

Другим реальным приближением к секретам генома человека является сравнительная геномика, так как в организации геномов человека и грызунов (мышь, крыса) обнаружен определенный параллелизм: многим кодирующим участкам генома человека соответствует гомолог в геноме мыши. Мышь является непревзойденной модельной системой для изучения генетики млекопитающих и болезней человека. За столетний период изучения генетики этого вида получены множество линий и сублиний, сотни спонтанных мутаций, разработаны методы изучения мутагенеза.

Секвенирование генома мыши проводилось Mouse Genome Sequence Consortium, в составе исследовательских центров США и Великобритании, а результаты опубликованы в декабре 2002 г. (Nature, 2002, vol. 420). В исследованиях использовались самки линии C57BL/6J. Размер генома мыши был на 14 % меньше генома человека — 2,5 млрд н. по сравнению с 2,9 млрд н. Секвенированию подвергались эухроматические районы генома, включая X- и Y-хромосомы. Размеры 20 хромосом варьировали от 55 мегабаз (млн н.) (хромосома 19) до 183 мегабазы (хромосома 1). Суммарный размер эухроматиновой части генома составляет 2.373 мегабазы (2.372 млн п. н.), а с учетом гетерохроматиновых сегментов — 2.493 мегабазы.



Около 90 % размера геномов мыши и человека представляли районы с сохранившейся синтенией, которая у обоих видов отражала порядок генов, полученный от общего предка. На нуклеотидном уровне у человека 40 % генома выстроены в том же порядке, как и в геноме мыши. Такой порядок расположения нуклеотидов в большинстве случаев отражает ортологические генные последовательности, также сохранившиеся в линиях от общего предка. Проект «Энциклопедия генома мыши» анализируется в Центре геномных наук RIKEN (Иокогама, Япония). По первоначальной оценке этого Центра, в геноме мыши имеется около 37 тыс. транскрипционных единиц, 20 тыс. которых представляют белок-кодирующие гены.

Сравнительные данные по геномам мыши и человека представлены в табл. 24. Предсказанное число транскриптов РНК у мыши составляет 28–29 тыс., у человека при первой оценке оно было 44,8 тыс., а при последующей (сентябрь 2002 г.) — 27 тыс. Число генов предсказано у мыши около 22 тыс., у человека первично — около 23 тыс. Среднее число экзонов на транскрипт и общее содержание экзонов на геном у обоих видов оценивалось около 200 тыс. При сравнении геномов этих видов выявлены значительные различия по локализации и количеству делеций и инсерций (вставок).

В процессе аннотирования генома мыши получено большое количество коротких секвенированных последовательностей к ДНК (EST). У разных линий мышей проводилось сравнение «снипсов» — SNP (single nucleotide polymorphisms — единичные нуклеотидные замены, полиморфизмы). Обна-

Таблица 24

Сравнение геномов мыши и человека

Признаки генома	Мышь		Человек	
	Первая оценка	Расширенная (+кДНК)	Первая оценка февр. 2001	Последующая сент. 2002
Предсказанные транскрипты	28.097	29.201	44.860	27.048
Предсказанное число генов	22.444	22.011	31.778	22.808
Известные кДНК	13.591	12.226	14.882	17.152
Вновь предсказанные	8.853	9.785	16.896	5.656
Среднее число экзонов/транскрипт	8,2 (ф. 6)	8,4 (ср. 6)	4,2 (3)	8,7 (6)
Всего предсказано экзонов	191.290	213.562	170.211	198.889

ружены длинные сегменты с разным уровнем полиморфизма, в результате создана база данных из 2 млн SNP в геномах 48 линий мышей. Проводится сравнительная линейная идентификация генных локусов, ответственных за болезни, а также за детерминацию количественных признаков.

Генетический анализ домашней мыши явился ценным вкладом в сравнительную геномику. Изучение других видов грызунов, обезьян, хищных — показывает, что геномы млекопитающих отличаются высокой степенью консервативности. Все это является основой для реализации проектов по картированию и аннотированию геномов сельскохозяйственных животных — свиней, крупного рогатого скота, овец, лошадей.

### • Ключевые слова

Эукариоты одноклеточные и многоклеточные, размножение пекарских дрожжей, нематода и ее особенности, геном плодовой мухи, мигрирующие генетические элементы, (МГЭ), ботаническая дрозофила резуховидка (резушка), особенности генома домашней мыши.

## Геном человека

В начале 70-х гг. прошлого века началось интенсивное развитие методов молекулярных исследований, разработаны генно-инженерные способы клонирования фрагментов ДНК и быстрые методы секвенирования. Стремительный характер развития молекулярной генетики привел к возникновению в конце 80-х гг. программы «Геном человека», в которой планировалось к 2005 г. завершить определение полной последовательности всех 3 млрд нуклеотидных звеньев, составляющих весь геном. К программе присоединились ведущие лаборатории Великобритании, Франции, Германии, Японии и России (всего более 100 центров и 1100 ученых разных стран). В США были созданы Национальный институт исследования генома человека во главе с Френсисом Коллинзом, а также на базе частной фирмы Celera Genomics Институт геномных исследований (TIGR — The Institute for Genome Research) во главе с известным ученым Крэйгом Вентером.

Первоначально программа планировалась на 15 лет, а ее стоимость оценивалась в 3 млрд долларов США: цена 1 шага, т. е. установление положения одной пары нуклеотидов в цепи ДНК, составляла тогда около 1 доллара. Однако крупные технические и методические усовершенствования позволили автоматизировать процесс секвенирования, сделать его более эффективным, быстрым и экономичным. 15 февраля 2001 г. в журнале *Nature* (International Human Genome Sequencing Consortium, 2001) и 16 февраля в журнале *Science* (Venter et al., 2001) о результатах расшифровки и анализа генома человека объявили две указанные группы исследователей. Та и другая представляли собой огромные исследовательские коллективы: так, число авторов в журнале *Nature* превышало 4 тысячи, а в журнале *Science* — более 450. Результаты двух исследовательских центров во многом совпадают, хотя имеются некоторые отличия. Полностью секвенирование генома человека завершилось к апрелю 2003 г. — к 50-летию юбилею открытия двойной спирали ДНК. Информация по всем результатам этих исследо-

ваний и дочерним направлениям науки широко представлены в Интернете. На сегодняшний день секвенировано около 90 % генома в черновом виде, а 30 % — в окончательном (т. е. нуклеотиды выверены с точностью до 99,99 %). Все это способствовало окончанию гонки, продолжавшейся несколько лет, и наступила **постгеномная эпоха**.

В 1977 г. были разработаны методы прямого определения последовательности нуклеотидов в ДНК, что знаменовало **начало эры секвенирования**. Вскоре произошла революция и в вычислительной технике. Исследователям из разных областей науки стали доступны ЭВМ различных классов: **началась эра компьютерной генетики**. Несколько лет спустя появились персональные компьютеры, и компьютерная генетика вошла в молекулярно-биологические лаборатории.

Компания Celera Genomics в рамках проекта по секвенированию генома человека использовала пять образцов человеческой ДНК от двух женщин и троих мужчин: один афроамериканец, один китаец, один испано-мексиканец и двое европейцев. Международный консорциум HGSFC использовал в исследованиях ДНК семи разных людей, т. е. каждый из двух соперничавших коллективов брал ДНК из своего источника. На основании данных секвенирования этих образцов ДНК была рассчитана консенсусная последовательность генома человека длиной 2,91 млрд п. н.

## Программа «Геном человека»: цели и методы

Ставились три основные цели программы:

- создание точной генетической карты,
- создание физической карты,
- секвенирование всего генома человека.

### 1. Создание генетической карты генома

Генетические карты показывают линейное расположение маркерных сайтов, расстояние между которыми выражают в *сантиморганах* (условная единица частоты рекомбинации гомологичных хромосом в профазе мейоза). Расстояние между локусами равно 1 сМ, если частота рекомбинации между ними составляет 1 %, что соответствует физическому расстоянию около 1 млн п. н., или 1 мегабазе. При этом частота рекомбинации, а, следовательно, и реальная длина 1 сМ, варьируют в разных частях генома, поэтому на генетических картах реальное (физическое) расстояние является ориентировочным показателем.

Для создания генетической карты необходимо найти в геноме большое количество *полиморфных генетических маркеров*, а для анализа сцепления между этими маркерами установить их взаимное расположение,

используя большое количество семей. Для этого имеются различные типы ДНК-полиморфизмов, а именно:

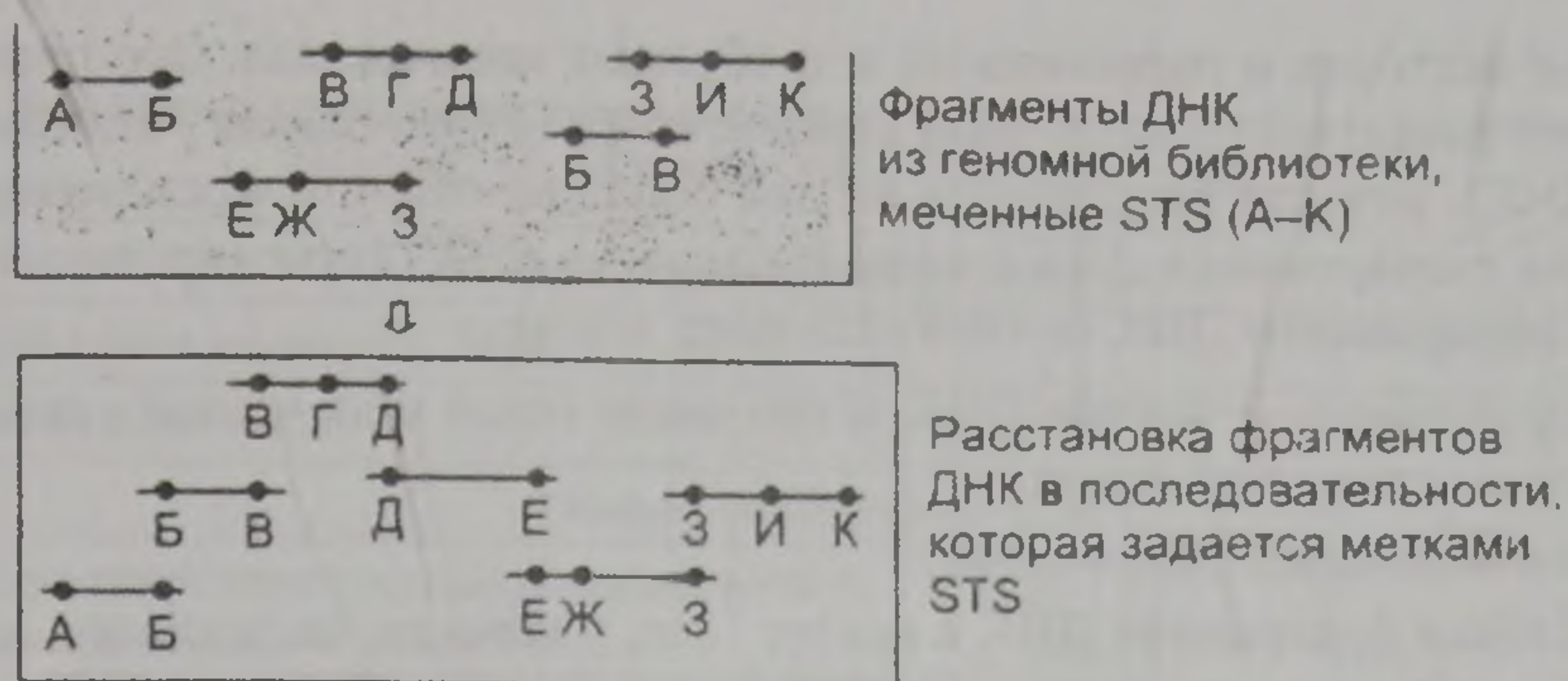
- полиморфизм длины рестрикционных фрагментов (ПДФ);
- полиморфизм, обусловленный варьирующим числом тандемных повторов: три-, тетра- и пентануклеотидных, или VNTR-полиморфизм;
- полиморфизм мини-сателлитов, состоящих из тандемно повторяющихся единиц длиной 20–70 п. н.;
- полиморфизм микросателлитов из повторяющихся единиц длиной 2–5 п. н.;
- однонуклеотидный SNP-полиморфизм, обусловленный изменением одного нуклеотида в определенных последовательностях ДНК у разных людей («снипсы»).

Все перечисленные полиморфные маркеры проявляют менделевское наследование, а все вместе дают возможность плотно маркировать геном человека. Среднее расстояние между маркерами составило менее 1 сМ (менее 1 млн п. н.). Взаимное расположение маркеров устанавливалось на основе коллекции клеточных линий от членов нескольких сотен семей на протяжении трех поколений. В каждой хромосоме устанавливалось и картировалось взаимное расположение 10–15 полиморфных маркеров. В настоящее время в геноме человека локализовано около 6 тыс. полиморфных ДНК-маркеров, среднее расстояние между которыми равно 1 сМ. На таком участке может локализоваться около 10 генов.

## 2. Создание физической карты генома

Физическое картирование — это определение места локализации и взаиморасположения генов, участков кДНК или каких-либо генетических маркеров (например, STS-маркеров, сайтов узнавания рестриктазами и т. д.), присутствующих в исследуемой молекуле ДНК. Физические карты всего генома, отдельной хромосомы или ее сегмента строятся на основе прямого исследования генетического материала. Они дают представление о реальном расположении генов по длине ДНК. Расстояние между ними и фланкирующими их маркерами выражается в числе п. н., что облегчает их идентификацию, изучение и секвенирование.

Необходимое маркирование генома проводят с помощью секвенированных фрагментов размером несколько сот п. н. — STS-маркеров, которые легко идентифицируются методом ПЦР (*полимеразной цепной реакции*). Получено более 50 тыс. разбросанных по геному STS — при физическом картировании их используют как маркеры перекрывающихся сегментов ДНК (рис. 78). Перекрывающиеся фрагменты ДНК, ранее клонированные в искусственных бактериальных (BAC-) или дрожжевых (YAC-) хромосомах, используют для построения так называемых *контиг* (*contigs*) —



**Рис. 78.** Использование STS как маркеров перекрытия фрагментов ДНК при построении физической карты (Гинтер, 2003). Вверху показаны фрагменты ДНК из геномной библиотеки, в которых найдены специфические STS, внизу — взаимное расположение фрагментов ДНК, определенное по расположению STS

протяженных отрезков ДНК. Контиг — это набор упорядоченных перекрывающихся клонов ДНК, охватывающих всю хромосому или ее участок, измеряемый в числе п. н.; служит основой для построения окончательного варианта физической карты.

При окончательной оценке размеров генома человека и идентификации генов учитывается существенная разница масштабов генетических, физических (молекулярных) и цитологических карт (табл. 25). Установлено, что 1 сМ на генетической карте соответствует приблизительно одному «бэнду». Бэнд — это полоса, видимая в световом или люминисцентном микроскопе после окраски хромосом. В свою очередь один «бэнд» прометафазной хромосомы соответствует 1 млн п. н. Средние размеры одного гена составляют 100–200 тыс. п. н., следовательно, при физическом картировании реально определяется участок из 5–10 генов-кандидатов.

Таблица 25

Оценка размеров генома человека

Метод оценки	Размер
Физический	$3 \times 10^9$ пар нуклеотидов (п. н.)
Генетический	3000–3500 сантиморганид (сМ)
Цитогенетический	2500 – 3000 Q-«бэндов» в прометафазной хромосоме

### 3. Секвенирование полного генома человека

Созданию высокоэффективных методов определения и расшифровки нуклеотидной последовательности ДНК-секвенирования — способствовало появление возможности клонирования целевых фрагментов ДНК в

составе векторов и наработка их в требуемых количествах. Это привело к бурному развитию генетической инженерии (или *технологии рекомбинантных ДНК*), которая разделяется на две большие области исследования — *методы клонирования ДНК и методы анализа ДНК* (секвенирования).

Клонирование ДНК *in vitro* включает этапы:

- 1) получение фрагментов ДНК, в том числе генов и их частей с помощью ферментов рестрикции;
- 2) рекомбинация фрагментов;
- 3) вставка фрагментов ДНК в вектор (фаг, плаزمид, космида и т. д.);
- 4) трансформация организма хозяина с помощью вектора;
- 5) отбор клонов, интересующих исследователя.

Секвенирование ДНК заключается в определении последовательности нуклеотидов во фрагменте. Эта процедура позволяет расшифровать структуру гена, установить природу мутаций в нем, определить аминокислотную последовательность в молекуле белка, кодируемую этим геном на основе генетического кода, и т. д. Наиболее распространенным методом секвенирования является *дидезокси-метод Ф. Сенгера*: для терминации синтеза цепи ДНК при этом используют *дидезоксинуклеотиды* dATP, dCTP, dGTP и dTTP.

Ручные методы секвенирования не позволяли справляться с большим объемом работ, поэтому возрастающие масштабы получения информации о последовательности нуклеотидов ДНК у разных видов организмов привели к созданию систем автоматизации. В 1987 г. фирмой Applied Biosystems (США) был создан первый автоматический *ДНК-секвенатор*. Такие роботы-секвенаторы управляются с помощью специально созданных компьютерных программ сбора и анализа данных. Каждый год такие автоматы совершенствуются, они превосходят человека по производительности. Процесс секвенирования становится простой процедурой и проводится в любой молекулярно-биологической лаборатории.

### **«Черновой» (первый) вариант генома человека**

Основные показатели «чернового» варианта генома суммированы в табл. 26.

Из всего генома секвенировано около 90 % последовательности нуклеотидов в *эухроматических районах хромосом*, для которых характерно слабая степень окрашивания красителем Гимза и флуорохромами. Размер этой части составляет 2,9 млрд п. н. Наибольшие методические трудности при секвенировании показали следующие участки, остающиеся нерасшифрованными:

- сателлитная ДНК теломерных и околоцентромерных районов хромосом;

Таблица 26

Основные характеристики первого («чернового») варианта человека  
[Nature, 15 February 2001, Vol. 409, № 6822]

<b>Предварительные оценки</b>	
Общая длина молекулы ДНК	1,5–1,7 м
Число нуклеотидов	$3,3 \times 10^9$
<b>Общая характеристика</b>	
Просеквенировано (установлена первичная последовательность нуклеотидов)	90 %
Допустимая частота ошибок	$1 \times 10^4$
Частота ошибок секвенирования хромосом 21 и 22	$1 \times 10^6$
Несеквенировано	10 %
<b>Общая структура ДНК генома</b>	
Повторяющиеся последовательности	45–50 %
Транскрибируемая часть составляет:	
Всего	28–30 %
транскрибируется в РНК	23–25 %
транслируется до белков (экзонная часть гагома)	5 %
Кодирует синтез всех белков организма	1,2 %
«Паразитическая» ДНК (LTR, SINE, LINE, Transposones)	45 %
Короткие повторы (микросателлитная ДНК)	3 %
Длинные геномные повторы	5 %
<b>Генетический полиморфизм</b>	
Идентичность геномов разных индивидуумов	99,9 %
Межиндивидуальная вариабельность	0,1 %
Общее число однонуклеотидных замен (SNP)	$3,2 \times 10^6$
Число «значимых» (внутригенных) SNP	$1,5 \times 10^6$
<b>Число генов</b>	
Хромосома 21	225
Хромосома 22	525
Всего определено	
Human Genome Project	31 780
Celera Genomics	38 114
Всего идентифицировано	
Human Genome Project	22 000
Celera Genomics	26 000
Картировано на хромосомах [OMIM, 2002 December]	14 065



- сильно спирализированные области интеркалярного (внутрихромосомного) *гетерохроматина*, способные ярко окрашиваться красителем Гимза и флуорохромами;
- небольшие интерстициальные фрагменты *гэпы* (*gaps*).

Установлено, что только около 1,2 % всей ДНК генома кодирует структуру белков и РНК, а более 80 % ее длины представляют собой повторяющиеся нуклеотидные последовательности разной длины, которые не транскрибируются. На основе сравнительного компьютерного анализа геномов домашней мыши и человека было высказано предположение о том, что в повторяющихся участках геномной ДНК возможно закодирована информация, связанная с программой онтогенеза. Общее количество однонуклеотидных замен SNP достигает 3,2 млн на геном, причем почти половину из них (~ 1,5 млн.) составляет *смысловую (экспрессирующуюся) часть генома*. Они играют важную роль в генетическом полиморфизме человека, в различиях рас и этносов, отдельных индивидов, в молекулярной диагностике наследственных болезней. Общее число генов человека оценивается 31–38 тыс., из них идентифицировано 22–26 тыс., а на индивидуальных хромосомах картировано 14.065 генов. К 2003 г. идентифицированы, клонированы *in vitro* и изучены на присутствие мутаций гены всех наиболее распространенных наследственных заболеваний.

Размер генома оценен в 3.164,7 млн п. н. Гены сильно отличаются друг от друга по размерам. Усредненная длина гена человека составляет около 30 тыс. п. н. (содержит 9 экзонов и 8 интронов). Самый большой ген белка *миодистрофина* имеет размер 2,7 млн п. н. (этот белок обеспечивает устойчивость и эластичность мышечных волокон при сокращениях). Самые короткие гены содержат два десятка нуклеотидов, например, гены *эндорфинов* — полипептидов, вызывающих ощущение удовольствия. О функциях почти половины всех генов ничего не известно; многочисленные повторяющиеся последовательности нуклеотидов не кодируют белки, но играют важную роль в поддержании структуры и функции хромосом и генов. Светлые участки хромосом, обогащенные GC-парами, содержат много генов, тогда как участки темно окрашенные, богаты AT-блоками и содержат мало генов. Области генома с кодирующими участками (генами) отделены друг от друга протяженными участками не кодирующей ДНК. Структурные (смысловые) гены отделены от бессмысленной «эгоистической» ДНК протяженными последовательностями из GC-пар длиной около 30 тыс. п. н., которые регулируют функциональную активность генов.

Показано, что у человека синтезируется во много раз большее количество белков, чем у нематоды и дрозофилы. Это объясняется особенностями транскрипции генов, химической модификации белков и сплайсинга, когда с одного гена может считываться информация о синтезе несколь-

ких белков с различными функциями. Несмотря на большое сходство белков человека с белками нематоды, дрозофилы и даже растений, для человека более характерны многочисленные и разнообразные семейства белков, ответственных за процессы раннего онтогенеза и иммунитет. Из 1278 семейств белков, закодированных в геноме человека, позвоночным животным свойственны 94 семейства. Проблемы идентификации генов в секвенированных последовательностях генома показали, что наши представления о структуре генов и их регулирующих элементах явно *недостаточны и неполны*.

Представляет интерес плотность генов в просеквенированных геномах разных видов, т. е. число генов на участке  $10^6$  п. н. У пекарских дрожжей эта величина равна 483 гена, у нематоды — 197, у дрозофилы — 117, у резуховидки (*Agabidopsis*) — 221 ген. У человека в хромосоме 21 это число 7 генов, в хромосоме 16 — 16 генов, а в среднем в секвенированной части генома — 12 генов. Некоторое количество генов по-видимому получено человеком от бактерий в результате *горизонтального переноса*. Очевидно бактериальный геном мог служить прямым донором генов для позвоночных. У человека, по сравнению с другими просеквенированными видами, чаще имеет место *альтернативный сплайсинг* в генах, т. е. для образования молекул функциональной мРНК реализуется больше путей объединения экзонов.

Секвенирование генома человека позволило провести сравнение физических и генетических карт по 5282 полиморфным локусам, положение которых по длине хромосом определено в сантиморганидах и в миллионах пар нуклеотидов. Средняя частота рекомбинаций (кроссинговера) возрастает по мере уменьшения длины хромосомных плеч. Средняя частота рекомбинаций в длинных плечах составляет 1 сМ, а в наиболее коротких — 2 сМ на 1 млн п. н. Частота рекомбинации уменьшается вблизи центромер, а в терминальных участках размером 20–35 млн п. н. возрастает. Это явление наиболее заметно на мейотической карте хромосом у мужчин, в частности, в *псевдоаутосомной области хромосом X и Y*. Физический размер этой области составляет 2,6 млн п. н., а ее генетическая длина достигает 50 сМ. В результате кроссинговер в этой области происходит практически в каждом акте мейоза, т. е. частота рекомбинации превышает среднее значение более чем в 20 раз.

В последнее время большой интерес проявляется к *альфа-сателлитным ДНК*, установленных в геномах всех приматов и человека. Их анализ является необходимым этапом на пути изучения эволюции хромосом. Они представляют собой тандемно повторяющиеся последовательности ДНК с элементарной единицей длиной 171 п. н. (Казаков, 2000). Все альфа-сателлиты гибридизируются между собой и имеют степень гомологии около 60 %. Альфа-сателлиты расположены в участках центромерного хроматина всех хромосом человека, где они занимают крупные блоки длиной от

200 до 5000 тыс. п. н. Такие блоки обнаружены и в других участках хромосом. Альфа-сателлитные последовательности человека имеют сложную трехуровневую организацию и играют важную роль в организации центромеры и определении ее функции в процессе митоза.

К настоящему моменту точность определения генов на неизвестной нуклеотидной последовательности компьютерными методами не превышает 70 %. Знание полных геномов дает возможность осуществить их исчерпывающий анализ, что позволит описать виды организмов на разных уровнях. Развитие в этой области проходит очень успешно, поэтому высказываются шутливые опасения насчет ближайшего будущего: через несколько лет будет выяснено всё, что только можно выяснить, и делать уже будет нечего (Александров, Дроздов-Тихомиров, Шепелев, 2004).

### «Геномизация» человечества

Расшифровка генома человека и других видов спровоцировала создание и интенсивное развитие совершенно новых направлений биологии, теории эволюции, генетики, антропологии, медицины, фармакологии и лингвистики. В отличие от классической молекулярной биологии и генетики, геномика имеет колоссальный диапазон воздействия на сопредельные науки, на общество в целом, включая биоэтические проблемы, а также социальную концепцию Русской Православной Церкви. Все это позволило внести новый аспект рассмотрения методологических и практических достижений геномики, влияющих на социальную и интеллектуальную жизнь современного человека, придав ему интригующее название «Геномизация человечества» (Баранов, Киселев, 2005). Связанный с геномикой неиссякаемый объем информации позволяет надеяться, что естествознание стоит в преддверии получения ответа на извечный вопрос: «Что такое жизнь?» (Свердлов, 2003). Необходимо дать краткое изложение основных направлений исследований, возникших недавно под влиянием достижений геномики.

*Сравнительная геномика* использует методы анализа молекулярных механизмов путем сравнения геномов разных видов, их генов и генных продуктов в разных органах и тканях. Изучение действия генов в геномах разных видов основано на признании того, что жизненно важные регуляторные функции в ходе эволюции сохранились у организмов далеко отстоящих таксономически групп. Самые элементарные клеточные функции, такие как базовый метаболизм, репликация ДНК, транскрипция, трансляция остались консервативными со времени расхождения одноклеточных дрожжей и бактерий. Но консервативность распространяется на значительно большее число генов. В сумме 60 % из предсказанных белков человека имеют сходство с белками других видов, геномы которых просекве-

нированы. Информация о регуляции клеточного цикла у млекопитающих получена путем сравнения с аналогичным процессом у дрожжей. Сравнительный компьютерный анализ геномов инициирует создание универсальной системы классификации живых организмов, своего рода «Периодическую систему геномов», аналогичную Периодической системе химических элементов Д. И. Менделеева. Французский математик Андри Пуанкаре (Пуанкаре, 1983) в качестве естественной предложил понимать *параметрическую систему*, в которой положение объекта задается одним (как число протонов в системе Менделеева), двумя (как координаты на географической карте) или несколькими параметрами.

*Транскриптомика* нацелена на масштабный анализ всех мРНК, кодирующих белки, и выяснение того, где, когда и в каких условиях транслируются разные гены. Изучается координированная работа генов, образование первичных транскриптов, процессы сплайсинга и формирования зрелых мРНК. Исследование «*транскриптома*» методом *микрочипов* дает возможность установить различия между экспрессией генов в разных тканях, проанализировать характер экспрессии в разные периоды болезни, классифицировать белки секретлируемые и связанные с мембранами на основе положения их мРНК.

Задачей *протеомики* является исследование синтеза и функций набора белков в разных клетках на разных стадиях развития организмов. При этом охватываются все синтезируемые в клетке белки, выясняется их строение, количество, локализация. Кроме первичной структуры, другой задачей протеомики является предсказание пространственного сворачивания (фолдинга) белка по его аминокислотной последовательности, так как функционирование белков зависит от их пространственной структуры. В настоящее время из более 30 тыс. генов на физической карте генома человека функционально изучены 11 тыс. генов. Предполагается, что выяснение функций всего генома человека позволит идентифицировать в *протеоме* 200–300 тыс. белков. В 2002 г. создан проект «Протеом плазмы крови» с участием 35 лабораторий из 13 стран. Методом масс-спектропии было идентифицировано 9504 белка. Быстро развивается также протеомика тромбоцитов, отвечающих за свертывание крови.

С протеомикой непосредственно связаны исследования обмена веществ в клетке — метаболизма. Это направление названо *метабономика*, в задачи которой входит выявление и моделирование структуры метаболизма, обеспечивающего поддержание гомеостаза в клетке за счет регуляторных свойств ферментов и функционирования генома. В свою очередь, устойчивый метаболизм, как замкнутая система согласованно протекающих биохимических реакций, обеспечивает считывание и реализацию генетической информации, а через них — собственное воспроизведение. Эти исследования о жизни клетки, отраженной в структуре и динамике ее

метаболизма, являются менее разработанными, но не менее важными, чем исследования генетической информации.

Практический потенциал «геномизации» жизни человечества наиболее ярко проявляется в биомедицинских исследованиях, где интенсивно развивается новый раздел медицинской науки — *геномная (молекулярная) медицина* (рис. 79). Большие успехи достигнуты в идентификации генов моногенных заболеваний, наследование которых подчиняется законам Менделя. Это относится к генам, ответственным за муковисцидоз (кистозный фиброз поджелудочной железы), рак молочной железы, другие виды наследственной предрасположенности к опухолям, мышечной дистрофии Дюшенна; ген, ответственный за преждевременное старение (синдром Вернера); гены, ответственные за нейрофиброматоз, и др.

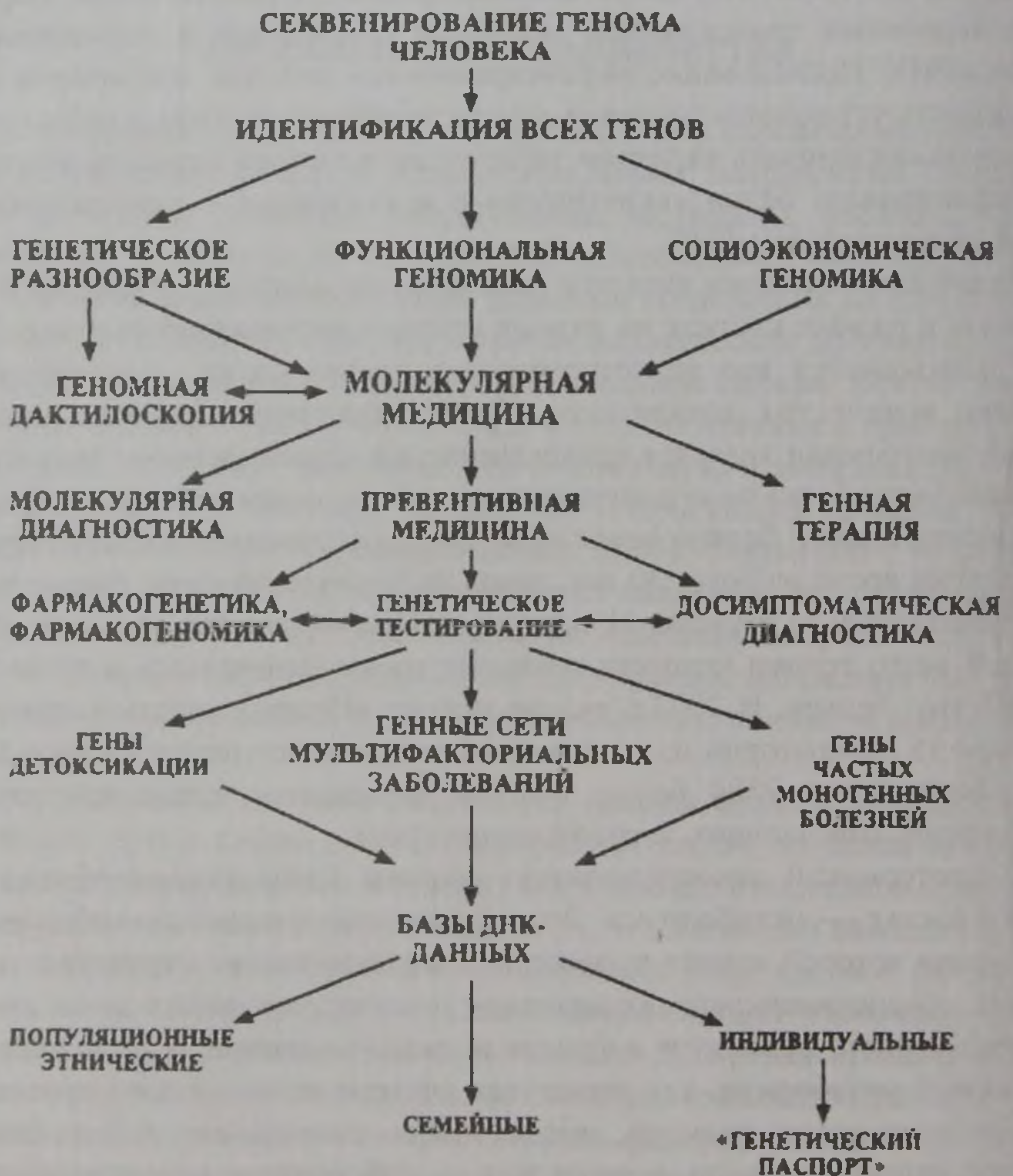


Рис. 79. Новые направления в исследовании генома человека

Более распространены и наносят больший экономический ущерб обществу сложнонаследуемые (мультифакториальные) заболевания, не подчиняющиеся менделеевским законам наследования: атеросклероз, гипертония, психиатрические заболевания, диабет, ревматоидный артрит и др. В результате достижений геномики расширились возможности изучения этой многочисленной группы болезней.

Изучение генетического полиморфизма в отношении восприимчивости к лекарствам привело возникновению еще одного направления — *фармакогеномики (фармакоантропологии)*, объединяющей генетический полиморфизм фармакокинетических и фармакодинамических процессов. Современная медицина для лечения пациентов опирается на популяционно-статистические данные, без учета индивидуальных особенностей. Поэтому в задачи фармакогеномики входит систематизация вариабильности генов, контролирующих реакции организма на условия среды, а также на действие лекарственных препаратов. К ним относится большая группа генов:

- контролирующих метаболизм ксенобиотиков;
- репарацию нуклеиновых кислот;
- контроль клеточного цикла;
- контроль программируемой клеточной смерти, и др.

Сюда же примыкает проблема генетических межэтнических различий в восприимчивости к лекарствам.

Накапливается все больше данных в пользу того, что к генетически детерминированным личностным особенностям относятся уровень интеллекта (показатель IQ), самостоятельность и зависимость, активность и пассивность, мнительность и тревожность, экстравертность и интровертность, чувствительность или толерантность к стрессам, альтруизм и эгоизм, агрессивность или сексуальность (Вельков, 2007).

Геномная медицина ускорила открытие механизмов возникновения и развития многих патологий. В ближайшем будущем можно ожидать, что единая медицинская наука раскроет молекулярные механизмы патологий и объединит их в систему взаимодействий:

Гены (ДНК) → РНК → Белки → Метаболиты → Физиологические процессы → Психические и психиатрические особенности → Ментальные и интеллектуальные характеристики

Геномика по-новому поставила проблему происхождения человека и его эволюции. В этногеномных исследованиях установлены генетические маркеры, различающиеся у разных рас, этнических групп и народностей, что внесло новые веяния в этнографию, археологию и лингвистику. Этногеномика — это своеобразный мост между естественными и гуманитар-

ными науками. Человек разумный как биологический вид сформировался около 200 тыс. лет назад. Секвенирование генома человека позволило образно представить длинную историю человечества как бы записанной в геноме *Homo sapiens*, а по геномным маркерам достаточно реально проследить миграции народов.

К этногеномике вплотную примыкает *палеогеномика*, которая позволяет реконструировать события давнего прошлого, например, охарактеризовать геном вымершего ближайшего родственника современных людей — человека неандертальского *Homo neanderthalensis*.

На основе медицинской геномики создана молекулярная диагностика наследственных болезней, их прогнозирование и генотерапия. Возможности генной диагностики расширились в результате картирования однонуклеотидных замен (полиморфизмов SNP). Идентифицированы гены, с которыми

## ПЕРСОНАЛЬНАЯ ГЕНЕТИЧЕСКАЯ ФОРМА

(индивидуальная база данных)

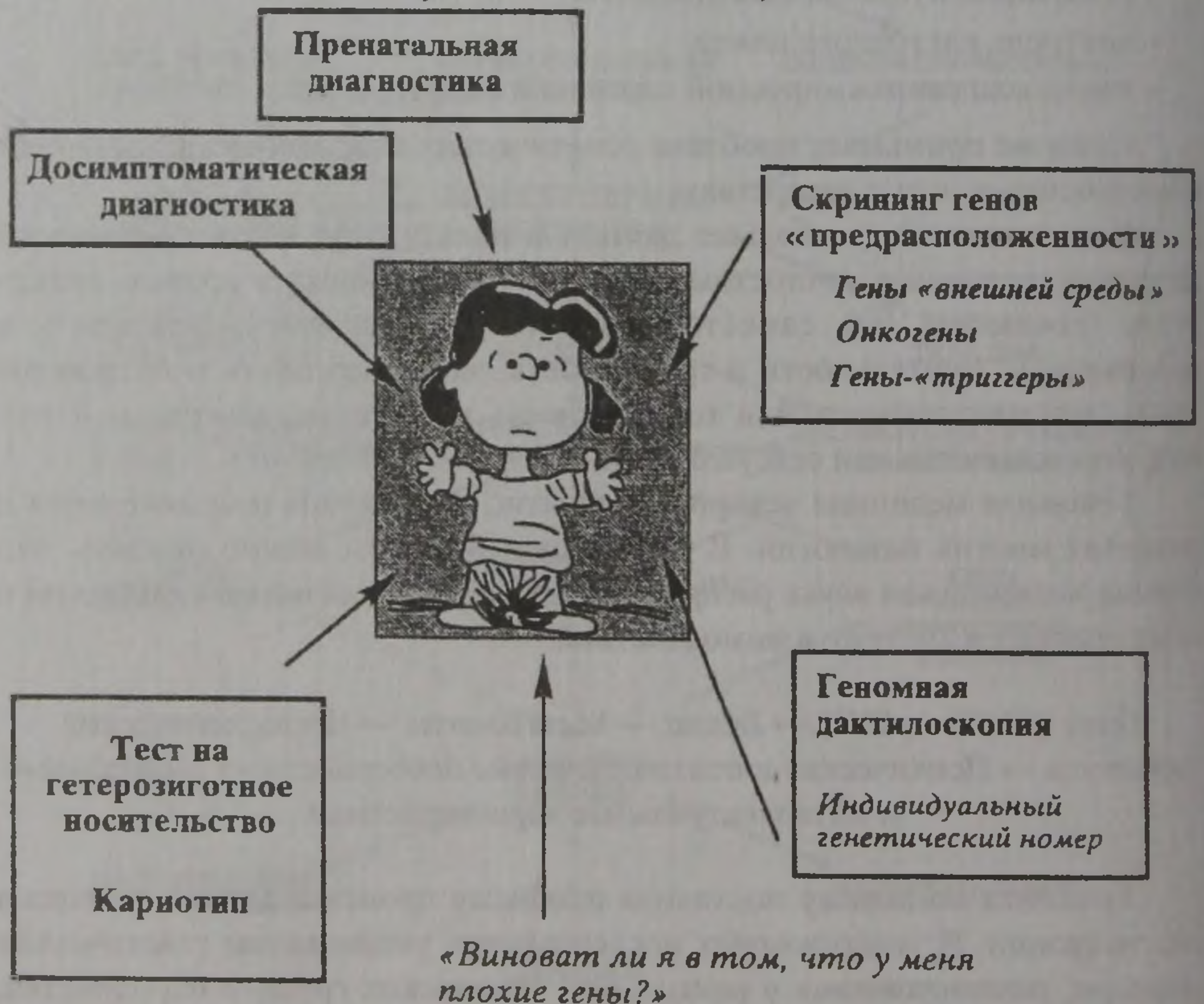


Рис. 80. Персональная генетическая форма (Баранов, Киселев, 2005)

связана этиология более чем 6 тыс. наследственных болезней. Диагностические центры широко применяют в своей практике молекулярные методы:

- для диагностики генных болезней;
- для выявления гетерозиготного носительства патологических мутаций в семьях высокого риска;
- досимптоматической диагностики болезней с поздней манифестацией;
- для идентификации личности (геномная дактилоскопия).

Используются также результаты исследования генов предрасположенности и бессимптомного носительства мутаций генов наиболее частых наследственных болезней (гемофилия, муковисцидоз, фенилкетонурия и др.), информация о кариотипе пациента. На основе тестирования в Лаборатории пренатальной диагностики наследственных болезней НИИ акушерства и гинекологии им. Отта РАМН (Санкт-Петербург) разработан «генетический паспорт», упрощенным вариантом которого является персональная генетическая форма (рис. 80).

Для интерпретации результатов в медицинской практике используется компьютерная программа SESAM (Франция). Она дает возможность врачу-генетику получить объективную информацию о предрасположенности субъекта к тому или иному заболеванию, судить о его медицинском прогнозе, аргументировано проводить медико-генетическое консультирование.

Проникновение биотехнологии и геномики в медицинскую практику ускорило создание *биоэтики и этических принципов* деятельности генетической службы. Они сформулированы в рекомендациях Всемирной Организации Здравоохранения (ВОЗ, Женева, 16 сентября 1997 г.). в нашей стране разработан законопроект «О генно-инженерной деятельности» применительно к человеку. Большое значение имеет оценка биоэтических проблем в послании «Социальная концепция Русской Православной Церкви» (2000), где дается оценка основным направлениям прикладной геномики и репродуктивной медицины и их совместимости с Православием.

### Ключевые слова

Программа «Геном человека» — цели и методы; однонуклеотидный SNP-полиморфизм; генетическая карта генома; физическая карта генома; контигов карта; бэнд; клонирование ДНК; ДНК-секвенатор; эухроматические районы хромосом; гетерохроматические районы хромосом; альфа-саттелитная ДНК; геномизация; сравнительная геномика; транскриптомика; протеомика; метаболомика; геномная медицина; фармакогеномика; биоэтика; генетический паспорт.



### Эра геномики: This is the End of the Beginning (англ. выражение: Это конец начала)

В последнее десятилетие определена полная нуклеотидная последовательность геномов у большого числа видов вирусов, прокариот (зубактерий и архебактерий), ряда низших эукариот и модельных представителей растений и животных. У многих объектов секвенированы *субгеномы* хлоропластов и митохондрий. Сравнительный анализ отдельных групп генов и целых геномов оказал революционизирующее влияние на развитие всех основных направлений современной биологии, филогенетики, таксономии и теории эволюции. Осуществлено создание концептуальной схемы универсального филогенетического древа жизни (Woese et al., 1990; Woese, 2000), состоящего из трех доменов — зубактерий, архебактерий и эукариот, которые ведут свое начало от *единого гипотетического корневого предка* — *прогенома*. Итоговым достижением эры геномики явилось секвенирование полного генома человека, реакцией на которое явилось начало процесса так называемой «*геномизации человечества*».

Дальнейшее развитие геномных исследований и их методологии происходит по следующим направлениям:

- структурная (описательная) геномика;
- функциональная геномика и биоинформатика;
- сравнительная (эволюционная) геномика;
- экологическая геномика и нутригеномика;
- метагеномика.

Далее мы представим краткую характеристику каждого из этих направлений.

## 1. Структурная (описательная) геномика

В задачи этого направления входит:

- секвенирование полных геномов;
- анализ структуры генома;
- оценка структуры изохоров — участков ДНК длиной около 300 т. п. н., различных по таксономическому положению организмов, но сходных по нуклеотидному составу;
- выделение и описание генов;
- определение нуклеотидной асимметрии комплементарных нитей ДНК.

Метод секвенирования представляет собой грандиозную техническую задачу по структурному анализу ДНК. Его осуществили Международный консорциум (International Human Genome Sequencing = IHGSC) во главе с Френсисом Коллинзом, а также коммерческая организация Celera Genomics во главе с Крэйгом Вентером. Они использовали разные методические подходы. Селера применила *метод дробления* (англ. *shotgun*). При этом ДНК дробили на небольшие фрагменты, у каждого из которых по несколько раз определяли последовательность нуклеотидов. Затем с помощью компьютерной программы высраивали целую хромосому из фрагментов. Тактика Международного консорциума сводилась к тому, что фрагменты молекулы ДНК выстраивали и делали карту, по которой проводилось секвенирование.

Основополагающая единица наследственности — *ген* — кодирует структуру белковой молекулы, отсюда оформилось название «*структурный ген*». Выяснено, что не все гены кодируют белки (полипептидные цепи). К ним относятся гены, кодирующие так называемые стабильные РНК — это гены рибосомной РНК (рРНК) и транспортной РНК (тРНК). Еще одна разновидность представлена *регуляторными элементами* — участками ДНК, необходимыми для транскрипции структурных генов. К этим участкам присоединяются либо РНК-полимеразы, либо специальные белки, регулирующие транскрипцию. Имеются следующие представители регуляторных элементов:

- *промоторы* (к ним присоединяются РНК-полимераза, чтобы начать транскрипцию);
- *терминаторы* (на таких участках РНК-полимераза кончает транскрипцию);
- *операторы* (к ним присоединяются белки-репрессоры, выключающие работу РНК-полимеразы);
- *энхансеры и сайленсеры* (участки ДНК, к которым присоединяются особые белки, изменяющие скорость транскрипции и, тем самым, скорость синтеза белков).

## 2. Функциональная геномика и биоинформатика

По состоянию на декабрь 2007 г. определены уточненные размеры разных геномов: геном человека содержит 31 897 генов, у дрожжей 7547 генов, у резушки Таля (*Arabidopsis thaliana*) имеется 29 388 ORF, у круглого червя (*Caenorhabditis elegans*) 23 399 генов, а геном мыши содержит генов на 6 тыс. больше по сравнению с геномом человека (Goodacre, 2007; <http://eucenes.org/>). Знание размера усредненного варианта человеческого генома, по образному выражению Л. Л. Киселева (Киселев, 2004), можно сравнить лишь с самым общим описанием конструкции автомобиля, включающей мотор, ходовую часть, колеса, руль, сиденья, обивку, бензин с маслом. Это описание не дает понимания взаимодействия частей и возможности управления автомобилем. Так же и геномика в настоящее время меняет свою ориентацию и осуществляет переход от *структурной геномики к функциональной*, основной задачей которой является «*прочитать геном*», установить как управляются и работают гены. Для этого потребуются еще десятилетия. Наступившая эпоха *постгеномики* стала эрой *функциональной геномики и биоинформатики*, а ее ключевыми словами — «*белок*» и «*протеом*» (вместо «ДНК» и «геном»).

Функциональная геномика использует математические методы для предсказания локализации гена или регуляторного участка в нуклеотидной последовательности. Биоинформатика дает об этом исходную информацию с вероятностью правильного предсказания около 85 %. Наличие предсказанного гена затем проверяется экспериментально: предполагаемый участок гена «вырезается» из ДНК и проверяется, кодирует ли он синтез белковой молекулы. В современных базах данных сосредоточена информация о нескольких миллиардах пар нуклеотидов (геномов) и аминокислотной последовательности белков у человека и других живых организмов. Компьютерная программа профессора Н. А. Колчанова (Институт цитологии и генетики СО РАН) позволяет разобратся в это океане информации и находить начала генов, их окончания и регуляторные участки. Другие программы — М. А. Ройтберга (Roytberg, 1992), а также М. С. Гельфанда и А. А. Миронова (Гельфанд, 1998) — осуществляют поиск *собственно генов*, то есть расшифровывают нуклеотидные последовательности.

Изменчивость набора белков организма — его *протеома* зависит от двух процессов: во-первых, на каждой стадии онтогенеза функционируют разные части генома; во-вторых, большая часть транскриптов РНК у человека и высших организмов претерпевают *альтернативный сплайсинг*. При этом несколько экзонов, содержащихся в пре-мРНК, могут объединяться в разных комбинациях с образованием различных мРНК. Это создает возможность синтеза разных белков на базе одной кодирующей последовательности (гена).

Если в гене имеется  $N$  интронов, то в ходе сплайсинга происходит  $N$  стадий вырезания интронов и сшивания экзонов. При этом экзоны одного гена могут стыковаться в разной последовательности, а некоторые экзоны могут удаляться вместе с соседними интронами. В результате этих событий один и тот же ген может кодировать целое семейство генов, сходных, но функционально разных. Имеются данные, что один ген может кодировать до 40 тыс. генов (Modrek, Lee, 2002). Это удивительное явление иллюстрируют два примера.

Ген дрозофилы содержит 95 альтернативных экзона. Он кодирует один из белков рецептора аксона. За счет альтернативного сплайсинга этого гена может образовываться 38 016 различных мРНК.

Ген *slo* у человека регулирует работу внутреннего уха. Белок этого гена присутствует в ворсинках, которые отвечают за распознавание высоты звука. Он содержит 35 экзонов, 8 из которых могут присутствовать или отсутствовать в зрелой мРНК. Возможное число вариантов сплайсинга равно  $8! = 40\,320$  ( $8!$  читается математически *8-факториал*). Из этих вариантов экспериментально идентифицировано 500. Существование других пока не установлено: ведь природа может не реализовывать все возможные варианты. При данном множественном сплайсинге разные типы волосяных клеток внутреннего уха реагируют на звуки разных частот от 20 до 20 тыс. герц. Различия клеток в восприятии частоты определяется свойствами альтернативных сплайс-форм гена *slo*.

Экспериментально показано, что у человека 74 % генов работают с использованием альтернативного сплайсинга, а возможность продуцирования различных белков оценивается равной  $10^6 = 1$  миллион (Oh et al., 2004). Выделить и охарактеризовать все эти белки невозможно. Нуклеотидная последовательность гена человека и эукариот не дает точной информации о том, какой белок он кодирует. Современное определение понятия «ген эукариот» можно представить в следующей формулировке (Киселев, 2004):

- это длинная, случайная, не кодирующая последовательность нуклеотидов, в которой расположены экзоны, способные после вырезания из транскрипта этого гена и их объединения в строго определенной очередности кодировать определенную функцию.

### 3. Сравнительная (эволюционная) геномика.

Задачи исследований:

- сравнение геномов у разных видов;
- анализ устойчивости отношений синтении (сравнение больших сегментов геномов с одинаковым порядком расположения генов у разных видов);
- полиморфизм в структуре генома.

Геномика расширила арсенал методов молекулярно-филогенетической систематики и показала, что молекулярная эволюция геномов — это изменения макромолекул, позволяющие судить о генотипическом родстве или различиях таксонов. Бурное развитие геномных реконструкций при изучении нуклеиновых кислот внесло принципиальные изменения в идеи систематики. Анализ генов 16S рРНК у многих видов привел к *трехдоменной системе организмов*. Открытие домена архебактерий стало крупнейшим достижением молекулярно-филогенетической систематики прошедшего столетия, за которым последовало создание *современной мегасистематики* (гл. 7).

Сравнительное изучение геномов представителей трех доменов жизни стимулировало исследования по реконструкции последовательных этапов возникновения основных клеточных процессов в эволюции. У всех доменов общими являются белки, участвующие в трансляции, а также ферменты метаболизма (ферменты цикла трикарбоновых кислот, гликолиза, метаболизма нуклеотидов). Отсюда сделано предположение, что процессы трансляции и метаболические реакции возникли раньше разделения трех основных доменов. Трансляция появилась, вероятно, еще до перехода от «мира РНК» к «миру ДНК». Позднее трансляции возникла транскрипция, еще позже появилась репрессия генов и компактизация ДНК, что способствовало увеличению размеров генома (Боринская, Янковский, 1999).

Исследование геномов создало новые перспективы в реконструкции процессов появления и эволюции жизни как планетарного явления, а также *природы прогенома*, который представлял собой «вариабельную коллекцию клеточных и субклеточных организмов, достаточно свободно обменивавшихся генетической информацией» (Woese, 2000).

Геномные исследования разграничили и охарактеризовали три ключевых направления эволюционного процесса:

1. *Вертикальная эволюция (эволюция «вверх»)* характеризуется наследованием базового набора предковых ортологичных генов, гомологичных у разных видов. Процессы дупликаций, мутаций и рекомбинаций приводят к образованию паралогичных генов (гомологичных в одном геноме), что ведет к увеличению набора белков и диапазона фенотипических вариаций у близких видов.
2. *Редуционная эволюция (путь «вниз»)* приводит к сокращению числа генов, утрате функций, путей метаболизма, органелл и т. д. в результате приспособления к условиям эконизи. Яркие примеры дают большинство патогенных бактерий, когда облигатные паразиты используют метаболические системы клетки «хозяина» и не нуждаются в собственных аналогичных системах.
3. Эволюционное направление на основе *горизонтального (латерального) переноса генов* между близкими и филогенетически отдаленными видами, принадлежащими даже к разным доменам (*латеральная ге-*

номика). Горизонтальный перенос генов, особенно у зубактерий и архебактерий, является одним из главных механизмов видообразования (гл. 8).

#### 4. Экологическая геномика

Задачи исследований:

- изменения генома в процессе приспособления организма к биотическим и абиотическим факторам среды;
- решение экологических проблем при изучении циклов питания, трофических связей, фенотипической пластичности;
- идентификации генов и генных сетей, определяющих адаптацию организмов к температурным изменениям, солености, инсектицидам;
- геномика питания (нутригеномика).

Благодатными объектами экологической геномики являются архебактерии. Они впервые были обнаружены в экстремальных условиях, свободных от представителей других доменов жизни (Морозова, 2005). Экстремальные виды архебактерий разделяют на следующие категории:

- термофилы (устойчивые к температурам 45–113 °С);
- психофилы (размножающиеся при температурах от –10 до +15 °С);
- ацидофилы (резистентные к средам с рН 1–5);
- алкалофилы (репродуцирующиеся при рН 9–11);
- барофилы (выдерживающие давление до 700 атм.);
- галлофилы (способные к выживанию в 25–30 % NaCl);
- ксерофилы (обитающие в необычайно сухих условиях) (Madigan, Marrs, 1997).

Геном архей состоит из двухцепочечной кольцевой ДНК длиной  $(5 - 40) \times 10^5$  п. н. и кольцевых плазмид. До настоящего времени определены нуклеотидные последовательности 18 полноразмерных геномов архей, а также более 5000 плазмид. Наиболее изучены гены рибосомных РНК (2230 нуклеотидных последовательностей), ДНК-зависимой ДНК-полимеразы (287 видов) и РНК-полимеразы (109 видов) (GenBank (<http://www.ncbi.nlm.nih.gov>)). Описано 9 различных морфотипов фагов термофильных архей. Уникальные особенности фагов архей позволили ввести 4 новых семейства вирусов с геномными двухцепочечными ДНК линейной кольцевой формы (Prangishvili, 2003; Peng et al., 2004).

Устойчивость ДНК к денатурации в экстремальных условиях обитания архей обеспечивается ДНК-связывающими белками — аналогами гистонов эукариот, полиаминами, относительно высоким содержанием нуклеотидных остатков G и C у термофилов или высоким внутриклеточными

концентрациями солей у некоторых галлофилов. Клеточные полиамины, стабилизирующие ДНК и вторичные структуры РНК, разнообразны по структуре у экстремофильных архей (Hamana et al., 2003).

Универсальный генетический код дополнен кодоном UGA для 21-й аминокислоты *селеноцистеина* у некоторых архей, бактерий, эукариот и даже человека. Недавно у метанобактерий и некоторых зубактерий открыта 22-я аминокислота *пирролизин*, кодируемая кодоном UAG, ранее считавшимся стоп-кодоном (Vöck et al., 1991; Krzycki, Chan, 2002).

Защита археобактериальных белков от денатурации в экстремальных условиях обеспечивается посредством увеличения содержания *неполярных аминокислотных остатков* (например, для белков *Pyrococcus furiosus* соотношение неполярных и полярных остатков составляет 3:1) и обилия *шаперонов*. Термостабильные белки теплового шока (шапероны семейства Hsp 60 и Hsp 70) при связывании их с другими белками делают их также устойчивыми к нагреванию. У термофильных археобактерий рода *Pyrodicticum*, живущих вблизи вулканов, 80 % цитоплазматических белков — шаперонины, состоящие из нескольких единиц — шаперонов семейства Hsp 60 (Stechman, Cavalier-Smith, 2004).

К этому направлению геномики относится также *геномика питания* (*нутригеномика*), исследования которой направлены на выяснение влияния тех или иных веществ, поступающих с пищей, на экспрессию генов. Среди многих биологически активных веществ, влияющих на экспрессию генов, особо отмечен *селен*, который влияет на уровень экспрессии 154 генов человека и оказывает антираковое действие. В птицеводстве органический селен в виде обогащенных селеном дрожжей «Сельплекс» является кормовой добавкой и используется для поддержания оптимального селенового статуса птиц, их высокие продуктивные и воспроизводительные качества (Сурай, 2006). Другой препарат, *селенит натрия*, является лекарством для лечения селеновой недостаточности.

В течение 13 лет проводились наблюдения на 250 тыс. пациентов (США, Европа, Япония) по влиянию веществ, содержащихся в овощах и фруктах, на здоровье. Было доказано, что эти вещества значительно снижали риск заболевания инсультом (He et al., 2006).

Исследования в областинутригеномики и фармакогеномики приведут к разработке компьютерных программ, которые с учетом набора генов конкретного человека позволят выявить наиболее оптимальный рацион питания и эффективные лекарства для лечения индивидуума.

## 5. Метагеномика

В задачи исследований входит:

- изучение геномов микроорганизмов из природных выборок (вирусов, бактерий и протистов);

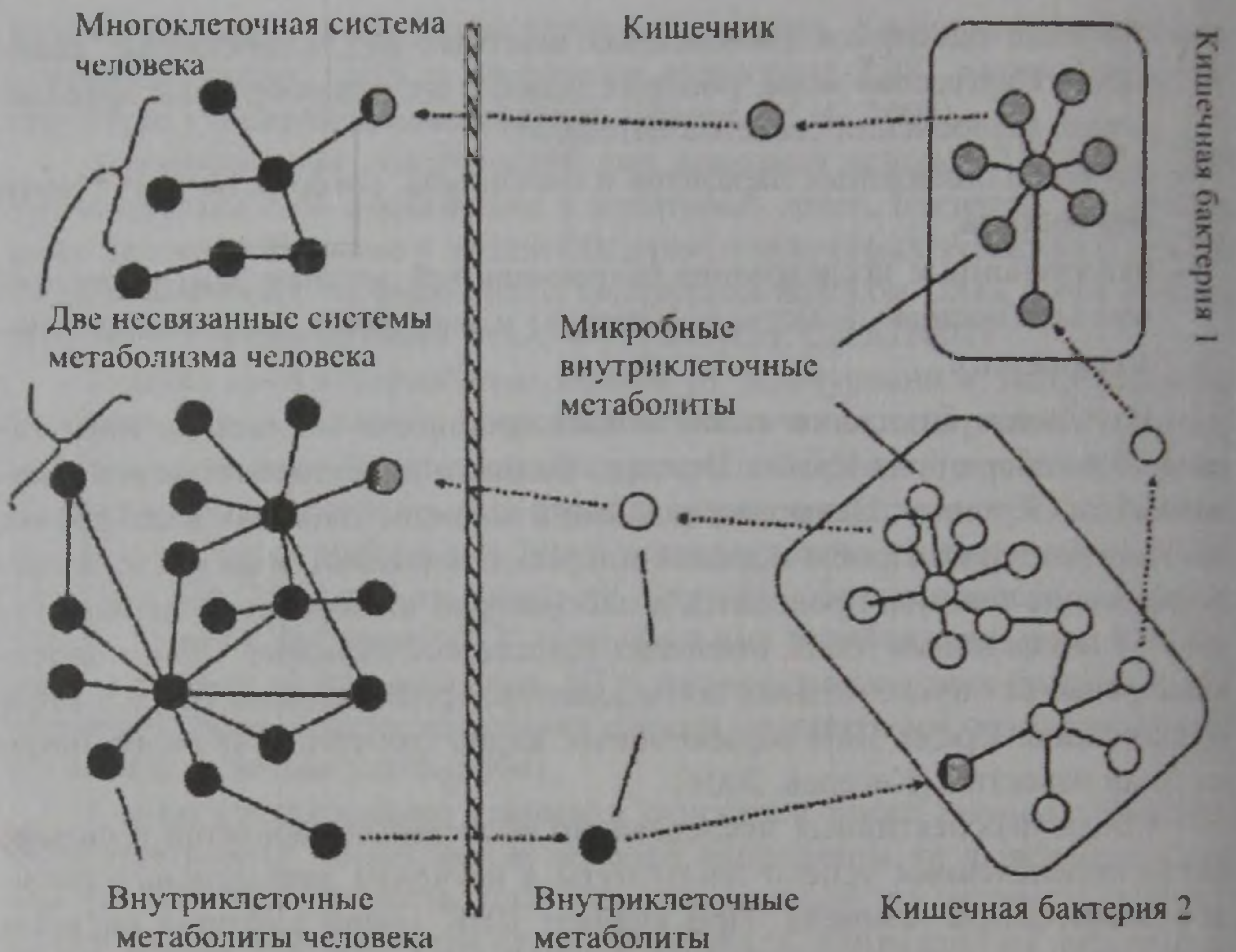
- изучение биопленок (окисленных шахтных вод, планктонных сообществ Саргассова моря, рисовых полей с их разнообразием археобактерий, захоронений скелетов китов);
- изучение облигатных паразитов и симбионов, которые не могут жить вне хозяина;
- палеогеномные исследования сохранившихся остатков вымерших животных (пещерный медведь, мамонт) и вымерших антропоидов (неандерталец).

Изучением биопленок планктонных сообществ Саргассова моря занимается лаборатория Крэйга Вентера, бывшего директора геномной фирмы Celera Genomics. Целью исследований является массовая дешифровка геномов бактерий. Пробы морских микробов добываются на месте, а секвенирование геномов проводится в лаборатории на корабле. Основная задача — поиск новых генов, имеющих прикладное значение. Объем просеквенированных нуклеотидных последовательностей составил более 1 млрд нуклеотидов. Среди 1800 обработанных видов бактерий 148 видов науке не были известны (Киселев, 2004).

Среди перспективных исследований облигатных паразитов и симбионтов определенные успехи достигнуты в изучении *метагенома* кишечной микрофлоры человека. При анализе ДНК малой единицы рибосом установлено, что кишечная микрофлора человека содержит  $10^{13}$ – $10^{14}$  микроорганизмов, более 1000 видов бактерий, то есть это количество в 10 раз превосходит число клеток в организме человека (Gill et al., 2006). По приблизительным оценкам, *метагеном* этого *микробиома* насчитывает число генов, по меньшей мере, в 100 раз больше, чем в геноме человека. Этот микробиом можно рассматривать в качестве *добавочного органа*, достигающего у взрослого человека веса 1 кг и представляющего форму *мутуализма (симбиоза)*, а обитающую у хозяина микрофлору как *комменсальную* (Backhed et al., 2005). Микробиом, в том числе кишечная флора, бактерии на коже, в дыхательных путях и урогенитальном тракте, играет важную роль в поддержании здоровья людей. На мышах и крысах показано, что виды *Lactobacillus* оказывали антидиабетическое воздействие при диабете (Yadav et al., 2006).

Взаимодействие микробиома и человека можно представить как *своеобразную гибридизацию человек X микробы*, а человек является неким *сверхорганизмом (англ. superorganism)* (Sekirov, Finlay, 2006). В связи с огромным разнообразием микробиомов наблюдается значительная вариабельность индивидов. Матка человека является стерильной, поэтому новорожденный ребенок имеет стерильный желудочно-кишечный тракт. Получение комменсальной микрофлоры сверхорганизмом достигается в первый год жизни ребенка преимущественно за счет бактерий материнского организма.





**Рис. 81.** Взаимодействие микробиома и генома человека.

Левая часть — многоклеточная система человека с внутриклеточными метаболитами (черные кружки).

Правая часть — кишечник с бактериями 1, 2. Внутриклеточные метаболиты бактерий — белые и серые кружки.

Левую и правую части разделяет стенка кишечного эпителия.

Происходит перекрестный обмен метаболитами, кодируемыми геномом человека и геномом бактерий: один из них, бактериальный, переходит через барьер кишечника и используется человеком, тогда как другой метаболит, от человека, абсорбируется кишечным микробом 2 (Goodacre R., 2007)

Поэтому человек, являясь им при рождении на 100 %, за счет бактерий становится им на 90 % — так возникает поистине сложный сверхорганизм!

На рис. 81 схематически представлена метаболическая сеть сверхорганизма. Секретируемые внутриклеточные метаболиты человека изменяются в процессе обмена и реабсорбируются через клетки стенки кишечника. Метаболиты бактерий оказывают активное воздействие на метаболиты человека.

\* \* \* \*

Подводя итог всему описанному в заключении, можно поставить вопрос: какие направления биологии XXI в. стимулировали проведенные геномные исследования?

Расшифровка последовательностей генома, как первый уровень его понимания, явилась огромным творческим моментом для всей биологии. Ведущие ученые Российской академии наук делают прогноз, что биология XXI в. будет направлена на изучение организации простейших живых систем — отдельных клеток. На этом пути будут создаваться компьютерные модели живой клетки, а первый шаг уже сделан — им явилось учреждение Международного консорциума по проекту полной компьютерной модели клетки кишечной палочки *E. coli*. Об экспериментах в этом направлении заявила теперь уже всемирно известная коммерческая фирма «The Institute for Genome Research» (TIGR). В основе этих исследований будет геном бактерии *Mycoplasma genitalium* с наименьшим по размеру геномом, состоящим из 517 генов. Это станет основой для рождения биологии XXI в. как *интегральной основы естествознания*, с ее комплексом идей и методов генетики, молекулярной и клеточной биологии, физиологии, эволюционики, информатики, математики и физики.

# Именной указатель

## А

Абрамсон Н. И. 125, 127  
Айала Ф. (Ayala) 124  
Акифьев А. П. 8  
Алексеев В. Н. 100  
Алиханян С. И. 8  
Алтухов Ю. П. 128  
Андерсон В. 12  
Андрусенко С. Ф. 21  
Арбер В. (Arber) 12, 157  
Астауров Б. Л. 8

## Б

Бадаева Е. Д. 230  
Баев А. А. 9  
Балтимор Д. (Baltimore)  
157  
Баранов В. С. 250, 254  
Барышников Н. В. 96  
Берг П. (Berg) 10, 131  
Бердников В. А. 147  
Бидл Д. (Beadle) 8  
Блаттнер Ф. (Blattner)  
215, 216  
Бойер Х. (Boyer) 10  
Бокуть С. Б. 42  
Бриджес К. (Bridges) 228,  
229  
Брок Т. (Brock) 132  
Бэтсон У. (Bateson) 7

## В

Вавилов Н. И. 7  
Ван Ниль К. Б. (Van Niel)  
102  
Вант-Хофф И. Г. (van't Hoff)  
19  
Ванюшин Б. 53  
Васильева Л. А. 229  
Вейсс С. (Weiss) 9  
Вейнер А. 11  
Вейсман А. (Weissmann)  
17  
Вилмут Я. (Willmuth) 12  
Вентер Дж. (Venter) 242,  
257  
Вёзе (Воез) К. (Woese) 36  
Вернадский В. И. 16, 17,  
22  
Вольман Э. Л. (Wollman)  
199  
Волькенштейн М. В. 21,  
128  
Воронцов Н. Н. 103, 104,  
128

## Г

Гамов Г. А. (Gamov) 73  
Гвоздев Г. А. 10, 158  
Геккель Э. 102  
Геллер М. 10, 158  
Гельфанд М. С. 258

Геодакян В. А. 212  
Георгиев Г. П. 158, 227  
Герасимович Н. В. 42  
Гилберт У. (Gilbert) 10,  
158  
Гольданский В. И. 18  
Голубовский М. Д. 147  
Гриффит Ф. (Griffith) 135  
Гудэйкр Р. (Goodacre) 264  
Гурвич Дж. 9

## Д

Дайхоф М. О. (Dayhoff) 13  
Дарвин Ч. 17  
Даунс А. Л. (Dounce) 73  
Дебабов В. Г. 194  
Дельбрюк М. (Delbrück)  
70, 167  
Денисова Е. В. 21  
Дромашко С. Е. 74, 87,  
159  
Дубинин Н. П. 7, 8  
Дулитл У. Ф. (Doolittle)  
148, 153  
Дульбекко Р. (Dulbecco)  
10, 157  
Дьяков Ю. Т. 120

## Ж

Жакоб Ф. (Jacob) 11, 63,  
199  
Жуковский П. М. 235

**З**

- Заварзин Г. А. 15, 197  
Замечник П. (Zamechnik) 75  
Зеленин А. В. 230

**И**

- Иванов В. И. 96, 99  
Ивановский Д. И. 162  
Инге-Вечтомов С. Г. 7  
Иоганнсен В. (Johannsen) 7  
Ичас М. (Ičas) 73

**К**

- Кавалье-Смит Т. (Cavalier-Smith) 103, 110, 114  
Казаков А. Е. 249  
Каменская М. А. 192  
Кемп М. (Kemp) 23  
Кидуэлл М. Г. (Kidwell) 13  
Кимура М. (Kimura) 127, 128  
Киселев Л. Л. 250, 254  
Колдуэлл П. 73  
Кольцов Н. К. 17  
Коллинз В. (Collins) 241, 257  
Колчанов Н. А. 258  
Конов А. Л. 144  
Корана Х. Г. (Khorana) 9, 10, 73  
Кордюм В. А. 146, 148  
Коржинский С. И. 88  
Корогодин В. И. 99  
Корочкин Л. И. 53, 54, 67  
Корнберг А. (Kornberg) 9

- Красилов В. А. 142  
Крик Ф. Х. К. (Crick) 4, 9, 10, 73  
Кузнецов С. Л. 52  
Кузьмин В. В. 18  
Кунижев С. М. 21  
Кунин И. (Koonin) 150, 151  
Кюри П. (Cipre) 18

**Л**

- Ламарк Ж. (Lamarck) 16  
Леван А. (Levan) 9  
Левонтин Р. (Lewontin) 259  
Ледерберг Д. (Lederberg) 133, 136  
Леонардо да Винчи 23  
Лобашев М. Е. 8, 99  
Львов А. М. (Lwoff) 9, 162, 167  
Лурия С. Э. (Luria) 167  
Ляпунов А. А. 70

**М**

- Макгиннис В. 11  
Маккьюсик В. А. (Mc Kusick) 9  
Максам Э. 10  
Матеи Ю. Г. (Matthai) 73  
Медведев Н. Н. 239  
Медников Б. М. 16, 17, 22  
Мезельсон М. (Meselson) 9  
Мёллер Г. Дж. (Muller) 8, 99  
Менделеев Д. И. 251  
Мендель Г. (Mendel) 7, 60, 99  
Мейен С. В. 145, 148

- Милкютин А. А. 142  
Миронов А. А. 258  
Мона Лиза 23  
Моно Ж. (Monod) 9, 63  
Морозов Л. Л. 18  
Морозова О. В. 78, 261  
Мотульски А. Г. 128  
Морган Т. Х. (Morgan) 59, 99  
Муллис К. (Mullis) 11, 158  
Муравенко О. В. 230  
Мушкамбаров Н. Н. 521  
Мэдиган М. Т. (Madigan) 108, 109, 114

**Н**

- Надсон Г. А. 99  
Назаров В. И. 5  
Натанс Д. 12, 157  
Ниренберг М. В. (Nirenberg) 9, 10, 73

**О**

- Орлов Ю. Л. 160  
Очоа С. (Ochoa) 9, 73

**П**

- Павлинов И. Я. 102  
Пастер Л. (Pasteur) 17, 18  
Полинг Л. (Pauling) 125  
Попов В. В. 149  
Прузинер С. 12, 71, 189  
Пуанкаре А. (Poincare) 251

**Р**

- Райт С. (Wright) 8  
Ратнер В. А. 5, 16, 70, 73, 87, 156

Реди Ф. (Redi) 18  
 Ройтберг М. А. 258  
 Романовский Ю. М. 22  
 Ромашов Д. Д. 8  
 Руска Т. (Ruska) 162

**С**

Сайки Р. К. (Saiki) 11  
 Сахаров В. В. 99  
 Сасаки Т. (Sasaki) 236,  
 237, 238  
 Сведберг Т. (Svedberg),  
 35, 107  
 Свердлов Е. Д. 6, 155,  
 250  
 Свирежев Ю. М. 70  
 Сенгер (Сенджер) Ф.  
 (Sanger) 5, 10, 72, 129,  
 158, 176  
 Серебровский А. С. 8  
 Симпсон Дж. Г.  
 (Simpson) 121  
 Сингер М. 62, 63  
 Скотт Т. М. 11  
 Смирнов Ф. А. 99  
 Смит Г. (Smith) 12, 157  
 Соколов В. Е. 238  
 Спирин А. С. 80  
 Сталь Ф. (Stahl) 9  
 Стениер Р. У. (Stanier)  
 102  
 Стивенс А. 9  
 Сурай П. 262

**Т**

Татум (Тейтем) Э. Л.  
 (Tatum) 8

Тарантул В. З. 24, 46, 57  
 Тахтаджян А. Л. 108  
 Темин Г. (Temin) 157  
 Тийо Дж. Х. (Tjio) 9  
 Тимофеев-Ресовский Н. В.  
 17, 18, 70, 87, 88  
 Титова Н. Н. 233

**У**

Уилкинс М. Х. Ф.  
 (Wilkins) 9, 24, 73  
 Уоддингтон К. Х.  
 (Waddington) 134, 145  
 Уотсон Дж. (Watson) 9,  
 15, 24, 25, 73

**Ф**

Филиппов Г. С. 99  
 Филиппович Ю. Б. 45  
 Фитч У. 13, 124  
 Фишер Р. (Fisher) 8  
 Фогель Ф. 128  
 Франклин Р. 9, 24, 73  
 Фриз Г. де (Vries) 88, 89

**Х**

Хесин Р. Б. 139, 146  
 Херши (Герши) А. Д.  
 (Hershey) 8, 167  
 Хиншелвуд Ч. 73  
 Хогленд М. Б. (Hoagland)  
 75  
 Хогнесс Д. 10, 227  
 Холдейн Дж. Б. С.  
 (Haldane) 8  
 Холи Р. (Holley) 9, 10

**Ц**

Циммер К. Г. (Zimmer) 70  
 Циндер Н. Ю. (Zinder) 133

**Ч**

Чаргафф Э. (Chargaff) 25  
 Чейз М. К. (Chase) 259  
 Чернин Л. С. 8  
 Четвериков С. С. 8

**Ш**

Шаталкин А. И. 103,  
 105, 110, 112, 114, 120  
 Шелл Дж. Г. (Shull) 11  
 Шестаков С. В. 149  
 Ширмер Р. 80  
 Шлегель Г. Г. (Schegel)  
 196  
 Шлезингер М. М.  
 (Schlesinger) 162  
 Шредингер Э.  
 (Schrödinger) 70  
 Шульц Г. 80

**Щ**

Щелкунов С. Н. 145,  
 170, 175, 178, 182

**Э**

Эвери А. (Avery) 8, 136  
 Эвклид 18

**Я**

Яблоков А. В. 88

# Терминологический словарь

## А

**Автотрофы** — организмы, способные использовать углекислоту в качестве единственного источника углерода и обладающие системой ферментов для ее ассимиляции, а также способные синтезировать все компоненты клетки. В зависимости от источника энергии, используемой для восстановления  $\text{CO}_2$ , различают фотоавтотрофы (наземные зеленые растения, водоросли, цианобактерии) и хемоавтотрофы, получающие энергию за счет окисления неорганических соединений (хемосинтез). А. — продуценты органического вещества в биосфере, образующие первый трофический уровень в сообществах.

**Аденин (А)** — одно из гетероциклических пуриновых оснований, входящих в состав ДНК и РНК, комплементарно тимину и урацилу.

**Аденозинтрифосфат (АТФ, аденозинтрифосфорная кислота)** — химическое соединение, содержащее аденин, рибозу и три остатка фосфорной кислоты; образуется в реакциях субстратного и окислительного фосфорилирования и является универсальным аккумулятором энергии в живых организмах, субстратом для биосинтеза нуклеиновых кислот.

**Азотистые основания** — общее название азотсодержащих гетероциклических органических соединений, входящих в состав нуклеозидов и нуклеотидов.

**Аденовирусы** — семейство ДНК-содержащих вирусов (Adenoviridae), впервые выделенных из аденоидов и миндалин детей. Имеют кубический тип симметрии, 252 капсомера и 10 и более структурных белков.

**Аллель** — существование в определенном локусе разных альтернативных форм гена или последовательности ДНК.

**Альтернативный сплайсинг** — тип сплайсинга РНК-предшественника, который обеспечивает кодирование одним геном нескольких отличающихся по структуре полипептидов.

**Альфоидная (альфа-сателлитная) ДНК** — форма сателлитной ДНК, установленная в геномах всех приматов и человека как тандемно повторяющиеся последовательности ДНК с элементарной единицей длиной 171 п. н.; расположены в участках центромерного хроматина всех хромосом, где занимают крупные блоки длиной от 200 до 5000 тыс. п. н.

**Аминоацил-тРНК-синтетазы** — ферменты, осуществляющие ковалентные присоединения аминокислот к 2'- или 3'-ОН концам тРНК.

**Аминокислоты** — производные карбоновых кислот, у которых один или несколько атомов водорода замещены на аминогруппу. В природе выявлено большое число различных аминокислот, но во всех многочисленных белках у всех живых организмов наблюдается явление *хиральной чистоты аминокислот*: белки состоят из  $\alpha$ -аминокислот L-конфигурации (вращают плоскость поляризации света влево). К ним относятся 20 *канонических пептидогенных аминокислот*. Недавно открыты 21-я пептидогенная аминокислота селеноцистеин (кодируется триплетом UAG) и 22-я аминокислота пирролизин (кодируется триплетом UGA). Только одна из двадцати, глицин, не обладает этим свойством. Зеркальные антиподы девятнадцати аминокислот — энантиомеры (вращают плоскость поляризации света вправо — D-конфигурация). При химическом синтезе симметричных молекул вещества образуется *рацемическая смесь*, содержащая по 50 % правого и левого антипода.

**Аmplификация** — процесс образования (умножения) дополнительных копий определенных участков ДНК.

**Антикодон** — последовательность из трех нуклеотидов (триплет), занимающая постоянное положение в структуре молекулы тРНК и комплементарно взаимодействующая с кодоном мРНК в процессе трансляции.

**Антисмысловая (матричная) цепь ДНК** — одна из цепей в двухцепочечной молекуле ДНК, нуклеотидная последовательность которой комплементарна соответствующей мРНК.

**Апоптоз** — генетически запрограммированная программа самоликвидации клеток (запрограммированная клеточная смерть).

**А-форма ДНК** — противоспиральное конформационное состояние молекулы ДНК, возникающее при 75 %-ной влажности и в присутствии ионов калия, натрия или цезия.

**Архебактерии (археи)** — группа микроорганизмов с прокариотным типом строения клетки, отличающихся от бактерий (эубактерий) многими свойствами. Выделены в отдельное царство (домен) Archaeobacteria. Отличаются от эубактерий по строению мембран клеточной стенки, наличием в геноме интронов, последовательности нуклеотидов в 16S рРНК и др.

## Б

**База (банк) данных** — специализированная совокупность данных, аккумулированная и сохраняемая как исходная информационная основа для работы в определенной области науки, практики, образованная. Накоплены многочисленные компьютерные базы данных в процессе расшифровки (секвенирования) нуклеотидных последовательностей геномов различных видов вирусов, бактерий, растений, животных и человека. Такие базы данных созданы в США, Европе, Японии и других странах со свободным доступом в них через адреса Интернета.

**Бактериофаги (фаги)** — вирусы, инфицирующие бактерии. Используются в генетической инженерии в качестве векторов для переноса фрагментов чужеродной ДНК в бактерии (*E. coli* и др.).

**Белки** — биополимеры (полипептиды), состоящие из аминокислотных остатков, соединенных пептидными связями. Различают простые белки — протеины

и сложные (липо-, глико-, нуклео- и металлопротеины). Самый распространенный в природе класс органических соединений, с огромным разнообразием структур и функций. В клетках эукариот содержится более 50 000 качественно различных белков, в клетках бактерий — более 3000.

**Библиотека генов** — полный набор клонированных фрагментов ДНК, полученных в результате рестрикции тотальной ДНК, выделенной из какого-либо специфического источника:

- геномная — библиотека генов, сконструированная из геномной ДНК;
- хромосом-специфическая — библиотека генов, сконструированная из ДНК отдельных хромосом;
- к ДНК (тканеспецифическая) — библиотека генов, сконструированная из кДНК, полученной путем обратной транскрипции из мРНК, изолированной из специфической ткани;
- скрининг библиотеки генов — выделение из библиотеки генов клонов, содержащих последовательности ДНК, комплементарные зонду.

**Биотехнология** — наука о клеточных генно-инженерных и клеточных методах и технологии создания и использования генетически трансформированных микроорганизмов, клеточных культур, растений и животных с целью получения и интенсификации производства новых видов продуктов и препаратов различного назначения. Исторически биотехнология выделилась из молекулярной генетики как одно из ее направлений.

**Бокс Прибнова (TATA-бокс)** — каноническая последовательность TATAATG, находящаяся на расстоянии 10 нуклеотидов перед стартовой точкой прокариотических генов. Представляет собой часть промотора.

**Бокс Хогнесса (Hogness box)** — нуклеотидная последовательность у эукариот, расположенная за 25 нуклеотидов до сайта инициации транскрипции. Состоит из восьми нуклеотидов.

**Бэнд (bands)** — полоса на хромосоме, видимая в световом или люминисцентном микроскопе после окраски.

## В

**Вектор** — средство для переноса вставки чужеродной ДНК (фаг, плаزمид, космида и др.)

**Вектора емкость** — максимальный размер рекомбинантной ДНК (в тыс. п. н.), которая может быть клонирована в данном векторе.

**Вирион** — элементарная вирусная частица, внеклеточная (покоящаяся) форма вируса. Вирионы простых вирусов (аденовирусы, пикорнавирусы) состоят из сердцевины (нуклеоид), содержащей ДНК или РНК, и белковой оболочки (капсид). У сложных вирионов (герпесвирусы, миксовирусы) капсид заключен в дополнительную липопротеиновую оболочку.

**Вироид** — облигатный внутриклеточный паразит, представляющий собой мелкую (200–400 нуклеотидов) кольцевую одноцепочечную РНК с характерной



палочкообразной вторичной структурой. Вироиды не имеют оболочки и не кодируют какие-либо белковые продукты.

**Вирусы** — неклеточные формы жизни, способные проникать в живые клетки и размножаться только внутри этих клеток. Вирусы обладают собственным генетическим аппаратом (ДНК или РНК), который кодирует синтез вирусных частиц из биохимических предшественников, находящихся в клетке-хозяине. Они рассматриваются как внутриклеточные паразиты на генетическом уровне.

**Вставка** — последовательность ДНК, которая вставляется в клонирующий вектор с помощью рекомбинантных технологий.

**Вставка нуклеотида (инсерция)** — мутация включения в цепь нуклеотидов молекулы ДНК дополнительного нуклеотида, что искажает генетическую информацию в пределах данного гена.

**Вырожденность генетического кода** — возможность кодирования одной и той же аминокислоты несколькими вариантами нуклеотидных триплетов.

## Г

**Гаплоиды** — организмы (клетки, ядра клеток) с одинарным (гаплоидным) количеством генетического материала и одинарным набором хромосом. Гаплоидными являются гаметы.

**Ген** — это физическая (определенный участок ДНК) и функциональная (кодирует полипептид или РНК) единица наследственности. Сочетает высокую устойчивость в ряду поколений и способность к наследуемым изменениям (генным мутациям), лежащим в основе изменчивости организмов как материала для естественного и искусственного отбора. Гены эукариот содержат: единицу транскрипции (транскриптон) мРНК, интроны (для мРНК), спейсеры (для рРНК), 5'- и 3'-нетранслируемые последовательности, промотор, терминатор, энхансеры, сайленсеры;

- **структурный** — ген, кодирующий белок;
- **гены ортологичные** — гомологичные гены, детерминирующие один и тот же белок, происходят от предкового гена в геноме вида, от которого произошли сравниваемые виды. В отличие от *паралогичных генов*, их происхождение не связано с дупликациями;
- **гены паралогичные** — образуются в результате *дупликации* предкового гена с последующей *дивергенцией*; в отличие от ортологичных генов, способны эволюционировать в пределах одного вида, например, паралогичные гены  $\alpha$ -,  $\beta$ -,  $\gamma$ -цепей гемоглобина человека.

**Генетическая инженерия** — молекулярно-генетическое направление, целью которой является конструирование *in vitro* функционально активных рекомбинантных (гибридных) ДНК. Современная стратегия генетической инженерии заключается в следующем:

- в небольшую молекулу ДНК, способную реплицироваться в клетке независимо от хромосомы, т. е. в *плазмиду* или *вирусную ДНК*, встраивают фрагменты молекул ДНК любого изучаемого организма;

- образующиеся при этом гибридные молекулы ДНК вводят в чувствительные клетки прокариот или эукариот, где они реплицируются и размножаются в своем составе встроенные фрагменты ДНК;
- выявленные гибридные ДНК подвергают структурно-функциональному изучению (секвенирование фрагментов ДНК, создание принципиально новых микробных продуцентов биологически активных соединений (ферментов, гормонов, антибиотиков и др.).

**Генетическая информация** — программа свойств организма, получаемая от предков и заложенная в наследственных структурах (ДНК, РНК) в виде генетического кода. Она определяет морфологическое строение, рост, развитие, обмен веществ, психический склад, предрасположенность к заболеваниям.

**Генетическое картирование** — составление схем линейного расположения сайтов (генов), расстояние между которыми выражают в *сантиморганах (сМ)* (условная единица частоты рекомбинации гомологичных хромосом в профазе мейоза). Расстояние между ними равно 1 сМ, если частота рекомбинации составляет 1 %, что соответствует физическому расстоянию около 1 млн п. н. (1 мегабаза).

**Генетический код** — присущая всем живым организмам единая система записи наследственной информации в молекулах ДНК и РНК в виде последовательности нуклеотидов — А (аденин), Г (гуанин), С (цитозин), Т (тимин) в ДНК (в РНК вместо тимина — У (урацил)). Реализация генетического кода в клетках осуществляется в два этапа: транскрипция (синтез мРНК) и трансляция у белоккодирующих генов (синтез белков).

**Генетическое расстояние (или дистанция)** — определяется путем оценки среднего числа замен аллелей в каждом локусе, произошедших за время отдельной эволюции двух популяций или видов. Замены аллелей имеют место тогда, когда в результате мутаций аллели в отдельных локусах замещаются другими аллелями или когда сразу замещается целый набор аллелей. Генетическое расстояние (Дистанция) варьирует от нуля (когда нет никаких аллельных замен) до бесконечности; значения могут быть больше единицы, поскольку в процессе эволюции, протекающей в течение длительного времени, аллели в каждом локусе могут неоднократно полностью замещаться.

**Генное семейство** — группа генов со сходной последовательностью ДНК, эволюционировавшая от одного гена — предшественника.

**Генная терапия (генотерапия)** — лечение путем введения в ткани или клетки пациента смысловых последовательностей ДНК, или изменение гена генноинженерными методами с целью коррекции генных дефектов, либо придание клеткам новых функций, способствующих устранению патологических процессов.

**Геном** — вся ДНК — ядра, митохондрий и хлоропластов (у растений), содержащаяся в организме (клетке).

**Геномика** — раздел генетики, который ставит целью изучение геномов отдельных видов на молекулярном уровне, их структуру (структурная геномика) и функции (функциональная геномика), а также использование в генной инженерии, биотехнологии и генной терапии (медицинская геномика или геномная медицина). В настоящее время секвенированы геномы у более чем 1 тыс. видов организмов разных групп — вирусов, патогенных бактерий и эукариот.

**Геномный эпигенетический импринтинг** — механизм, обеспечивающий дифференциальную активность генов в зависимости от их родительского происхождения, т. е. от прохождения через отцовский или материнский гаметогенез.

**Геносистематика** — направление систематики, в котором родственные отношения реконструируются по молекулярно-генетическим данным (ДНК, РНК, белок), которым присваивается больший вес, чем данным классической морфологии.

Этот подход базируется на *теории нейтральной эволюции*: разработанная ею концепция молекулярных часов позволяет выделять и ранжировать таксоны по количеству геномных перестроек, которые трактуются как мера времени их эволюционного расхождения.

**Генофонд** — совокупность генов, которыми обладают особи данной популяции, группы популяций или вида, занимающих определенную территорию (акваторию).

**Гетероплазмия** — наличие в клетках и тканях смешанной популяции нормальных и мутантных молекул митохондриальной ДНК.

**Гетеротрофы** — организмы, использующие для питания только органические вещества, произведенные другими видами и, как правило, не способные синтезировать вещества своего тела из неорганических составляющих. К гетеротрофам относятся все животные, паразитарные растения, грибы и подавляющее большинство микроорганизмов.

**Гетерохроматин** — конденсированные участки хромосом, содержащие в основном некодирующие высокоповторяющиеся последовательности ДНК.

**Гибридизация (ренатурация) нуклеиновых кислот** — способность цепей ДНК и РНК к комплементарному взаимодействию с образованием двуцепочечных молекул (дуплексов) различного строения: ДНК:ДНК, ДНК:РНК, РНК:РНК.

**Гимза** — автор метода окраски хромосом, при котором выявляются G-сегменты.

**Гистоны** — консервативные белки эукариот, связывающие ДНК. Участвуют в формировании нуклеосомы, основной структурной единицы хроматина.

**Гомеобокс** — короткая высококонсервативная область ДНК длиной 180 п. н., представляющая экзоны гомеозисных генов, ответственных за эмбриональное развитие. Гомеобоксодержащие гены (их значение в развитии первым изучал американский генетик Эдвард Льюис) были найдены практически у всех организмов. Они богато представлены у млекопитающих, собраны в кластер НОХ-генов, расположены в четырех разных хромосомах у мыши и человека.

**Гомеодомен** — это продукт гомеобокса, состоящий из 60 аминокислотных остатков. С помощью этих белков осуществляется управление развитием путем включения одних и выключения других генов. Их отличительная черта — наличие в них однотипных доменов — гомеодоменов. Фрагменты гена, кодирующие разные домены белка, обязательно разделены интронами, т. е. представляют собой разные экзоны.

**Гомеозисные гены** — специфическая система генов, контролирующих эмбриональное развитие, у насекомых — сегментацию тела. При мутациях этих ге-

нов у дрозофилы наблюдалось превращение антенны в ногу (Antennapedia) или аристы в ногу (Aristapedia).

**Гомоплазмия** — наличие в клетках и тканях одного типа митохондриальной ДНК — нормальной или мутантной.

**Горизонтальный перенос генов** — проблема возможного наличия механизма переноса генетической информации от одного вида другому у прокариот и у эукариот. Имеются примеры такого переноса как в природных условиях, так и в эксперименте.

**Горячая точка** — участок ДНК, в котором частота возникновения мутаций или рекомбинаций очень велика.

**Грибы** — по современной систематике образуют отдельное царство Fungi (Mycota) подимперии эукариот, включающее от 100 тыс. до 250 тыс. видов. В настоящее время считается, что грибы по строению, характеру обмена и способу питания занимают промежуточное положение между животными и растениями и несут отдельные черты как тех, так и других. Проведено секвенирование одноклеточного вида — пекарских дрожжей *Saccharomyces cerevisiae*.

**Группа сцепления** — совокупность генов, локализованных в одной хромосоме и наследуемых совместно. Число групп сцепления у каждого вида эукариот равно числу пар гомологичных хромосом, при этом X- и Y-хромосомы являются разными группами сцепления.

**Гуанин (Г, G)** — пуриновое азотистое основание, комплементарное цитозину (Ц, C), входящее в состав ДНК и РНК.

**ГС (GC)-пар содержание** — один из фундаментальных критериев таксономии микробов. Представляет собой отношение суммы гуанина и цитозина к общей массе оснований в молекуле ДНК, выраженное в мол.%. У бактерий этот показатель колеблется в пределах 25–75 %, у простейших 20–70 %, у позвоночных 35–45 %.

## Д

**Дальтон** — единица М. м. (Да), равная массе атома водорода  $1,67 \cdot 10^{-24}$  г или  $1/12$  массы атома углерода ( $^{12}\text{C}$ ). Используется для измерения массы атомов, молекул, вирусов, клеток и их структур (хромосом, рибосом, митохондрий, хлоропластов и др.). Название дано в честь английского физика и химика Дж. Дальтона (1766–1844).

**Дезоксирибоза** — пятиуглеродный моносахарид, входящий в состав ДНК.

**Дезоксирибонуклеиновая кислота (ДНК)** — вещество наследственности («энциклопедия жизни»); единственный тип молекул, способных к самовоспроизведению в живой клетке (репликация, конвариантная редупликация) и кодированию генетической информации. Нитевидная молекула, у которой остов из чередующихся остатков дезоксирибозы и фосфорной кислоты ковалентно соединен с четырьмя азотистыми основаниями — аденином (А), Тимином (Т), гуанином (Г или G) и цитозином (Ц или C). может существовать как в однонитевой, так и двунитевой форме за счет образования водородных связей между комплементарными парами оснований по правилу А–Т, Г–Ц (G–C).

- геномная — тотальная ДНК, выделенная из любого типа клеток, хромосом или их фрагментов;
- избыточная (эгоистическая, паразитическая) — ДНК, не несущая кодирующих функций;
- комплементарная (кДНК) — однонитевая ДНК, полученная в процессе обратной транскрипции молекул мРНК;
- митохондриальная — ДНК, локализованная в митохондриях;
- плазмидная — ДНК, входящая в состав плазмиды;
- рекомбинантная (гибридная) — химерные молекулы ДНК, составленные из фрагментов разного происхождения.

**Денатурация (плавление) ДНК** — переход молекул ДНК из двуцепочечной формы в одноцепочечную в результате разрыва водородных связей между комплементарными парами оснований под воздействием высоких температур, а также в щелочной среде.

**ДНК-диагностика (генодиагностика)** — совокупность методов по выявлению изменений в структуре генома с целью диагностики наследственных заболеваний.

**ДНК-лигаза** — фермент, катализирующий образование фосфодиэфирной связи между 3'-гидроксильной группой и 5'-фосфатом соседних нуклеотидов в месте одноцепочечного разрыва дуплекса ДНК. Принимает участие в репликации, используется в генетической инженерии.

**ДНК-праймаза** — фермент, синтезирующий в процессе инициации репликации короткую РНК-затравку (РНК-праймер). ДНК-праймаза может быть отдельным ферментом (как у бактерий) или входить в качестве субъединицы в ДНК-полимеразу.

**ДНК-полимераза** — фермент, катализирующий синтез полинуклеотидной цепи из отдельных нуклеотидов с использованием другой цепи в качестве матрицы и ДНК-затравки (или РНК-затравки) со свободной 3'-ОН-группой.

**ДНК-хеликаза** — ДНК-зависимая АТФаза, использующая энергию гидролиза АТФ для расплетания двойной спирали ДНК в процессе репликации. Хеликаза однонаправлено перемещается по одной из цепей ДНК, расплетая перед собой двойную спираль, в результате чего возникает вилка ( $\gamma$ ) из двуцепочечного участка ДНК и двух одноцепочечных ветвей.

**Домен** — участок аминокислотной последовательности белка, связанный с определенной функцией.

**Дупликация** — тип мутации, заключающийся в удвоении той или иной нуклеотидной последовательности (участка хромосомы).

## Е

**Емкость вектора** — максимальный размер рекомбинантной ДНК (в т. п. н.), которая может быть клонирована в данном векторе. Например, у фага лямбда *e. v.* составляет 24 т. п. н.

**Единица Сведберга (S)** — единица измерения коэффициента седиментации, равная  $10^{-13}$  с, зависящая от массы и формы белковой частицы. Используется для оценки молекулярной массы макромолекул, в частности, размеров рибосом.

## И

**Избыточная ДНК** — ДНК, не несущая кодирующих функций (молчащая, эгоистическая, паразитическая).

**IS** — элементы (англ. *insertion sequences*) — последовательности — вставки, сегменты ДНК прокариот, способные перемещаться из одного участка в другой, содержащие лишь гены, необходимые для их собственного перемещения.

**Импринтинг геномный** — дифференциальная экспрессия генов, зависящая от того, от кого из родителей (отца или от матери) эти гены унаследованы. Причиной являются эпигенетические модификации (метилирование и др.) ДНК геномов, зависящие от происхождения гамет.

**Инвертированные повторы (обращенные П.)** — состоят из двух тождественных копий ДНК длиной около 300 п. н., ориентированных в противоположных направлениях и лежащих друг от друга на расстоянии от нуля до десятка тыс. п. н., занимают около 5 % генома. Различают палиндромы — обращенные повторы, не разделенные промежуточными последовательностями (около 1/3 всех инвертированных повторов).

**Иницирующий кодон** — кодон AUG в составе мРНК, кодирующий метионин (N — формилметионин у прокариот), с которого начинается (иницируется) синтез полипептидных цепей.

**Интрон** — транскрибируемый участок ДНК (гена) эукариот, который вырезается из первичного транскрипта при сплайсинге при образовании функциональной мРНК; последовательности, находящиеся по обе стороны от интрона (экзоны), объединяются.

**Инсерция** — мутация, при которой происходит вставка нуклеотидов в молекулу ДНК.

**Информационная (матричная) РНК (иРНК, мРНК)** — молекула РНК, образующаяся в результате транскрипции с матричной цепи ДНК.

**Интеины** — участки генома прокариот, кодирующие самосплайсирующиеся участки полипептидных цепей. Процесс заключается в автокаталитически протекающем посттрансляционном вырезании внутреннего пептида с последующим лигированием концевых пептидов.

## К

**Капсид** — белковая оболочка вирусной частицы (вириона). Морфологические субъединицы К. — капсомеры — состоят из нескольких молекул структурного белка, образуя упорядоченные спиральные или многогранные (икосаэдрические) структуры. Стабилизирует геном и защищает его от неблагоприятных внешних воздействий.

**Карта Бриджеса** — впервые описанная К. Бриджесом в 1935 г. характерная поперечная исчерченность политенных хромосом клеток слюнных желез плодовой мухи в виде 5100 окрашенных дисков (участков плотной упаковки). Порядок расположения дисков рассматривается как своеобразная хромосомная карта, которая устойчива при наследовании и характерна для вида. Она разделена на 102 сегмента 1-го уровня и более мелкие сегменты. Она охватывает плечи больших хромосом, содержащих 75 % гаплоидного набора ДНК.

**Карта рестрикционная** — диаграмма расположенная на молекуле ДНК рестриктов — сайтов узнавания рестриктазами (эндонуклеазами рестрикции), распознающими определенные нуклеотидные последовательности ДНК. Такие карты составлены для больших молекул ДНК — вирусов, прокариот, митохондриальной ДНК.

**кДНК** — комплектарная ДНК, которая образуется из мРНК с помощью обратной транскрипции. Этот тип ДНК соответствует только экзонам соответствующего гена.

**Килобаза (Кб)** — единица длины молекулы ДНК, равная тысяче пар нуклеотидов.

**Клонирование ДНК** — один из основных методов генетической инженерии. Фрагменты ДНК любого вида, полученные в процессе рестрикции, вводятся в вектор молекулярного клонирования — плазмиду, бактериофаг и др., а это позволяет размножить их в клетках бактерий или дрожжей. Метод позволяет конструировать геномные библиотеки.

**Кодон** — дискретная единица генетического кода, состоящая из трех нуклеотидов в молекулах ДНК или РНК. Из 64 кодонов 61 кодирует определенные аминокислоты, 3 стоп-кодона (UGA, UAG, UAA) определяют окончание синтеза полипептидной цепи (сигнал терминации транскрипции). Иницирующий (стартовый) кодон AUG в составе мРНК кодирует аминокислоту метионин, с которого начинается образование полипептидной цепи и процесс трансляции.

**Кодирующая цепь (смысловая) цепь ДНК** — цепь ДНК, последовательность нуклеотидов которой идентична мРНК. Комплементарная кодирующей цепь ДНК является матричной.

**Комплементарность** — взаимное соответствие в химическом строении взаимодействующих молекул, обеспечивающее образование вторичных связей между ними. Комплементарные структуры подходят друг к другу, как ключ к замку. К. последовательности оснований в двух полинуклеотидных цепях является ключевым свойством нуклеиновых кислот, а последовательность оснований в одной цепи определяет их последовательность в другой. К. нуклеотидных пар — это образование водородных связей по правилам А–Т, G–C в двунитевой молекуле ДНК и гибриде ДНК/РНК. Принцип К. является ведущим в строении нуклеиновых кислот и их биосинтезе (репликации, транскрипции). Примеры К. в других областях биологии: а) взаимодействие фермента с соответствующим субстратом; б) в иммунологии говорят о комплементарности антигена и соответствующих ему антител.

**Коллинеарность** — линейное соответствие последовательности кодонов гена аминокислотам в первичной структуре белка.

**Космида** — один из видов векторов молекулярного клонирования в генетической инженерии. В состав К. входят ген устойчивости к антибиотикам, репликон плазмиды и фрагмент фага  $\lambda$  (лямбда). Векторная емкость К. достигает 45 Кб.

**Контиги** (англ. *contigs*) — набор клонированных перекрывающихся фрагментов ДНК, полностью перекрывающих участок, например, область поиска гена.

**Кроссинговер** — взаимный перекрест хромосом, регулярно наблюдаемый в профазе мейоза. Приводит к обмену равными участками гомологичных хромосом и перераспределению в них генов. Особей с новыми сочетаниями признаков, образовавшимися в результате К., называют кроссоверами. За единицу измерения частоты перекреста принимают величину, равную 1 % кроссоверов (морганида, сантиморган). Отношение числа кроссоверов к общему числу потомков никогда не превышает 50 %. К. представляет собой один из механизмов комбинативной наследственной изменчивости.

**Кеп** (англ. *cap* — шапочка) — специфическая структура на 5'-конце мРНК эукариот. Эта структура может быть метилирована по молекуле гуанина. К. необходим для прикрепления мРНК к рибосоме.

## Л

**Лайон гипотеза** — подтвержденное предположение М. Лайон о том, что в каждой соматической клетке нормального эмбриона женского пола человека и многих млекопитающих происходит случайная инактивация материнской или отцовской хромосомы X (лайонизация).

**Лигирование** — соединение двух линейных фрагментов ДНК ферментом ДНК-лигазой фага T 4; восстановление фосфодиэфирных связей цепи ДНК.

**Лизис** — разрушение и растворение клеток под действием или других агентов.

**Лизогенная инфекция** — включение генома фага в геном клетки-хозяина (бактерии). В каждую дочернюю клетку при делении материнской попадают и фаговые гены. Воздействие повреждающих факторов среды (ультрафиолетовые лучи, ионизирующее излучение, химические реагенты) приводит к выщеплению профага из клеточной ДНК, лизису клетки и высвобождению большого числа свободных вирусных частиц.

**Липкие концы** — взаимно комплементарные одноцепочечные участки ДНК, образующиеся в результате ступенчатых разрезов двухцепочечных ДНК рестриктазами. Они обеспечивают соединение двух фрагментов ДНК за счет комплементарных взаимодействий. Линейные молекулы ДНК с Л. к. используют в генетической инженерии для конструирования гибридных молекул ДНК.

## М

**Малые ядерные РНК (мя РНК)** — короткие (от 63 до 220 н.) стабильные молекулы РНК в ядрах клеток в составе нуклеопротеиновых частиц. Участвуют в процессе сплайсинга про-мРНК у эукариот.



**Макромолекула** — полимерное вещество с молекулярной массой от нескольких тысяч до сотен миллионов дальтон (белки, нуклеиновые кислоты и т. д.).

**Маркеры генетические** — различные типы полиморфизмов ДНК и белков, которые используют в популяционных исследованиях, при картировании генов.

**Матричная РНК (мРНК = иРНК)** — молекулы РНК, состоящие из последовательностей, комплементарных экзонам генов; образуются в результате сплайсинга и концевых модификаций из молекул первичного РНК-транскрипта (пре-мРНК). Для прокариот характерна полицистронная, а для эукариот — моноцистронная мРНК. Составляет 2 % общего количества РНК в клетке.

**Мегабаза (Мб)** — единица длины молекулы ДНК, равная миллиону п. н.

**Метилирование** — химическая реакция, результатом которой является введение метила (метильной группы) в органическое соединение. В молекулярной генетике — химическая модификация ДНК путем переноса метильной группы на остатки цитозина или аденина в молекуле ДНК при участии фермента ДНК-метилтрансферазы, что приводит к образованию метилцитозина. При метилировании зоны ДНК с определенным геном включается механизм подавления транскрипции.

**Минимальный набор генов у прокариот** — путем сравнения геномов выявлено, что минимальная клетка, способная к автономной жизнедеятельности и самовоспроизведению, должна была бы содержать не менее 250–300 наиболее существенных генов систем репликации, транскрипции, репарации, рекомбинации и др.

**Мобильные генетические элементы (МГЭ)** — дискретные фрагменты ДНК, способно перемещаться по геному клеток. У прокариот к ним относятся IS-элементы (около 1 тыс. п. н.) и транспозоны (дл. 3 тыс. — 20 тыс. п. н.); у эукариот М. г. э. представлены транспозонами, ретротранспозонами и ретропозонами. Вызывают наследственные изменения в клетках, являясь причиной инсерционного мутагенеза: инактивируют или усиливают экспрессию генов, вызывают хромосомные перестройки, вносят в геном факторы нестабильности и изменчивости, что имеет огромное значение в процессе эволюции геномов.

**Молекулярные часы (молекулярные эволюционные часы)** — концепция Лайнуса Полинга (1965) о скорости молекулярной эволюции во времени, измеряемом в годах через число аминокислотных замен. Молекулярные часы и их ход носят стохастический характер, сопоставимый с радиоактивным распадом, а в качестве постоянной принимается вероятность изменений в единицу времени, подчиняющаяся распределению Пуассона. Единицей измерения принят «полинг», как величина изменений, равная  $10^{-9}$  на аминокислотный сайт в год.

**Мутация** — спонтанное или индуцированное изменение в последовательности нуклеотидов ДНК. М. бывают генные, хромосомные, геномные, цитоплазматические. Генные мутации:

- делеция — утрата сегмента размером от одного нуклеотида до хромосомного фрагмента из нескольких генов;
- динамическая (мутация экспансии) — экспансия тринуклеотидных повторов, вызывающих более 16 наследственных заболеваний;

- дупликация — удвоение сегмента ДНК размером от одного нуклеотида до хромосомного фрагмента из нескольких генов;
- инверсия — поворот на  $180^\circ$  сегмента ДНК размером от двух нуклеотидов до хромосомного фрагмента из нескольких генов;
- инсерция — вставка сегмента ДНК размером от одного нуклеотида до хромосомного фрагмента из нескольких генов;
- миссенс — замена нуклеотида в кодирующей части гена, ведущая к замене аминокислоты в белковом продукте;
- нонсенс — замена нуклеотида в кодирующей части гена, сопровождающаяся образованием стоп-кодона;
- регуляторная — мутация в 5'- или в 3'-нетранслируемых областях гена, нарушающая регуляцию его экспрессии;
- сплайсинговая — затрагивает сайты сплайсинга или создает новые сайты сплайсинга в интронных областях гена; сопровождается либо делецией смежного с мутацией экзона, либо невырезанием соответствующего интрона при процессинге первичного РНК — транскрипта;
- точечная — затрагивает от одного до нескольких нуклеотидов;
- транзиция — точечная замена пиримидина на другой пиримидин или пурина — на другой пурин;
- трансверсия — точечная замена пиримидина на пурин, и наоборот.

## Н

**Нанометр (нм)** — единица длины, равная  $10^{-9}$  м,  $10^{-3}$  мкм, или 10 ангстремам ( $\text{A}^\circ$ ).

**N-конец** — конец полипептидной цепи со свободной аминогруппой ( $\text{NH}_2$ -). С N-конца начинается нумерация аминокислотных остатков в полипептидной цепи белка.

**Нейтральности теория** — выдвинутая М. Кимура и др. в 70-х гг. XX в. эволюционная концепция, отрицающая контролирующую роль естественного отбора ввиду селективной нейтральности многих мутаций. Теория теоретически обосновывается вырожденностью генетического кода, в результате которой замена одного из оснований в кодирующем триplete может не приводить к замене аминокислоты в полипептиде (синонимические мутации). По этой теории большинство генетических преобразований происходит на основе стохастических, ненаправленных процессов на молекулярном уровне, т. е. дрейфа генов.

**Номогенез** — эволюционная концепция, выдвинутая в 1922 г. Л. С. Бергом и противопоставлена селектогенезу (дарвинизму). Это учение о внутренней запрограммированности, предопределенности эволюционного процесса, его подчиненности строгим закономерностям запрограммированности, предопределенности. Отправным моментом является изначальная целесообразность всего живого.

**Нуклеаза** — фермент, расщепляющий ДНК или РНК.

**Нуклеиновые кислоты** — природные высокомолекулярные биополимеры, мономерами которых являются нуклеотиды, связанные 3'-, 5'-фосфодиэфирной

связью. Различают два вида н. к.: дезоксирибонуклеиновые кислоты (ДНК) и рибонуклеиновые кислоты (РНК). Углеводным компонентом в ДНК являются D-2 — дезоксирибоза, в РНК — D — рибоза. В качестве гетероциклических оснований ДНК содержит два пурина: аденин (А) и гуанин (G) и два пиримидина: Тимин (Т) и цитозин (С). В РНК вместо Тимина содержится урацил (U). Цепи ДНК (полидезоксирибонуклеотиды) и РНК (полирибонуклеотиды) обладают *полярностью*, или *направлением*. Благодаря полярности каждая полинуклеотидная цепь имеет 5'-конец и 3'-конец. Н. к. — высокомолекулярные соединения, среди которых встречаются самые большие из всех известных макромолекул. Молекулярная масса ДНК хромосомы прокариот более  $2 \times 10^9$  Da, а у плодовой мушки  $9 \times 10^{10} - 11,3 \times 10^{10}$  Da.

**Нуклеонид** — бактериальное «ядро», представляющее собой у прокариот нить ДНК, погруженную в цитоплазму, замкнутую в кольцо и не связанную с гистонами. Кольцо ДНК закреплено в одной точке на внутренней стороне клеточной мембраны. Деление Н. происходит после репликации нити ДНК. Названные признаки характерны для представителей *домена зубактерий*. Другой домен включает *архей* (*архебактерий*), имеющих сходный общий план строения клетки, но различающиеся по организации ряда структур и признаков.

**Нуклеотид** — основная структурная единица ДНК и РНК, состоящая из остатков сахара, присоединенного остатка фосфорной кислоты при 3'-м или 5'-м углеродных атомах и азотистого основания. Нуклеотиды аденин (А), Тимин (Т), гуанин (G) и цитозин (С) кодируют аминокислоты в полипептидной цепи белка.

**Нуклеосома** — основной структурный элемент хромосомы. Представляет собой гистоновую сердцевину, на которую намотана ДНК длиной 145 п. н., делающая 1,75 оборота.

## О

**Облигатный** — термин, определяющий состояние, обязательное для данного организма, например, О. аэроб или О. анаэроб.

**Обратная транскрипция** — комплементарный синтез ДНК (кДНК) на матрице мРНК при участии ревертазы (обратный транскриптазы).

**Обратная транскриптаза или Ревертаза** — наиболее полно изучен энзим вируса саркомы Рауса, кодируемый геном *pol*. В генетической инженерии Р. широко используется для синтеза комплементарных молекул ДНК (кДНК) с использованием мРНК без интронов и получения библиотек (банков) кДНК.

**Однонуклеотидный полиморфизм (SNP)** — полиморфизм, обусловленный изменением одного единственного нуклеотида в определенной последовательности ДНК.

**Олигонуклеотиды** — цепочки ДНК или РНК, состоящие из нескольких (от 2 до 20) нуклеотидных остатков.

**Онкогены** — гены, обуславливающие превращение нормальных клеток эукариот в злокачественные. Действие О. реализуется посредством кодируемых ими онкобелков. О. присутствуют в вирусах, содержащих ДНК (аденовирусы, папова-

вирусы и др.) и РНК (ретровирусы), а также в геноме опухолевых клеток. Известно около 30 О., кодирующих онкобелки.

**Оперон** — группа генов у бактерий, функционально связанных между собой. Белки, кодируемые генами одного О. — это ферменты, катализирующие отдельные этапы одного метаболического пути. К наиболее хорошо изученным относятся лактозный (lac-оперон), галактозный (gal-оперон) и триптофановый (trp-оперон). В составе каждого из этих О. присутствует несколько структурных генов (цистронов), кодирующих ферменты, а также регуляторные области: промотор, оператор и терминатор.

**Ориджин репликации (сайт инициации репликации)** — локус, в котором начинается репликация ДНК.

**Открытая рамка считывания (англ. — *open reading frame, ORF*)** — последовательность нуклеотидов, не содержащая терминирующих кодонов; кодирует полипептид или белок.

**Отжиг** — процесс образования двухцепочечных молекул (ДНК — ДНК или ДНК — РНК) из одиночных полинуклеотидных комплементарных цепей.

**Отстающая цепь** — цепь ДНК, синтезируемая в процессе репликации прерывисто в виде коротких фрагментов (фрагменты Оказаки), которые затем ковалентно соединяются между собой.

**Ортологичные гены (ортологи)** — гены, произошедшие от общего предка и неизменно передающиеся в филогенезе «вертикально».

## П

**Паралогичные гены (паралоги)** — гены, возникающие в результате дупликаций и распространяющиеся «параллельно» внутри родственной дифференцирующейся группы видов.

**Первичный транскрипт (РНК-предшественник, пре-мРНК)** — первоначально синтезированная немодифицированная молекула РНК, соответствующая транскрипционной единице.

**Плазмиды** — внехромосомные факторы наследственности у бактерий и дрожжей, замкнутые в кольцо двухцепочечные ДНК молекулярной массой  $10^6$ – $10^8$  Да, способные к автономной репликации. Содержат гены двух-трех белков, крупные кодируют до 200 белков. Наличие R-факторов обеспечивает устойчивость бактерий к антибиотикам и лекарственным препаратам. П. используют в генетической инженерии в качестве векторов клонирования фрагментов ДНК.

**Позитивный геном (плюс-геном вирусов)** — однонитчатые РНК- или ДНК-содержащие геномы вирусов, выполняющие функции матрицы для синтеза новых геномов и одновременно мРНК.

**Повторы ДНК** — последовательности ДНК, многократно встречающиеся в геноме:

- инвертированные или обращенные — состоят из двух тождественных копий ДНК длиной около 300 п. н., ориентированных в противоположных направ-

- лениях и лежащих друг от друга на расстоянии от нуля до десятка тысяч п. н.; занимают 5 % генома;
- палиндром — обращенные повторы, не разделенные промежуточными последовательностями;
  - сателлитные — относительно короткие (не более 200 п. н.) высокоповторяющиеся последовательности ДНК; расположены тандемными блоками преимущественно в центромерных, теломерных и гетерохроматиновых районах хромосом; занимают около 10 % генома;
  - альфоидная ДНК — класс сателлитных ДНК размером 170–200 п. н., в составе которых обнаруживаются последовательности, специфичные для гетерохроматиновых районов разных хромосом человека;
  - микросателлитные — одно- или двухнуклеотидные тандемные повторы, дисперсно распределенные по всему геному;
  - минисателлитные — тандемные повторы размером от 3 до 20 нуклеотидов, дисперсно распределенные по всему геному;
  - STR (short tandem repeats) — тримерные и тетрамерные короткие тандемные повторы;
  - UNTR — варьирующие по числу тандемные три-, тетра- и пентануклеотидные повторы;
  - умеренные (низкокопийные) — дисперсно распределенные по геному повторяющиеся последовательности ДНК размером от сотен до тысяч нуклеотидов; составляют около 20 % геномной ДНК;
  - ALU — последовательности ДНК размером около 300 п. н., повторенные в геноме человека несколько сотен тысяч раз;
  - LINE — длинные диспергированные повторы размером до нескольких тысяч п. н.; число копий не превышает 10 тыс. на геном;
  - SINE — относительно короткие диспергированные повторы размером до 500 п. н. с числом копий до нескольких десятков тысяч; встречаются через каждые 2,2 тыс. п. н.

**Полиаденилирование** — ферментативное добавление нескольких адениновых остатков (поли-А) к 3'-концу мРНК при ее процессинге из первичного РНК-транскрипта.

**Полимеразы** — ферменты, способные использовать полинуклеотиды как матрицы и строить комплементарные полинуклеотидные цепи.

**Полимеразная цепная реакция (ПЦР)** — метод, разработанный К. Мюллисом и Р. Сайки в 1985–1988 гг., был назван изобретением века, ускорил реализацию проекта «Геном человека», а также способствовал внедрению в практику клинической генетики многих заболеваний высокоэффективных диагностических наборов нового поколения. Метод сводится к амплификации большого числа копий последовательностей ДНК (порядка миллиона) размером от 50 до нескольких тысяч п. н. ДНК последовательно нагревается и охлаждается в присутствии ДНК-полимеразы и свободных нуклеотидов. Специфический фрагмент ДНК денатурируется, гибридизируется с праймерами копируется с помощью ДНК-полимеразы.

**Последовательность Шайна—Дальгарно** — вся или часть полипуриновой последовательности AGGAGG, находящаяся на 5'-конце мРНК непосредственно перед иницирующим AUG-кодоном, комплементарная последовательности на 3'-конце 16S рРНК; принимает участие в связывании рибосомы с мРНК.

**Проект «Геном человека»** — международный научно-исследовательский консорциум, целью которого явилось создание точной генетической карты, физической карты и секвенирование всего генома человека. Членами проекта были более 50 стран, в т. ч. Россия. Была создана точная генетическая карта генома с разрешением менее 1 сантиморгана. Общая длина молекул ДНК составила  $3,3 \times 10^9$  п. н., всего определено 31–38 тыс. генов.

**Праймер** — олигонуклеотид, выполняющий роль «затравки» и иницирующий синтез полинуклеотидной цепи на ДНК- или РНК-матрице.

**Промотор** — последовательность ДНК, которая локализуется у 5'-конца гена и с которой связывается РНК-полимераза для начала транскрипции.

**Провирус** — геном ДНК-содержащего вируса или ДНК-копия РНК-содержащего вируса, которые интегрированы в хромосомную ДНК клетки-хозяина. Возникновение провирусов характерно для умеренных фагов, онкогенных и некоторых вирулентных вирусов (ВИЧ, вирус гепатита В).

**Протисты** — термин введен для обозначения микроорганизмов. Высшие протисты — эукариоты (водоросли, грибы, простейшие), низшие П. — бактерии, археи. В настоящее время термин обычно используется в значении «одноклеточные эукариоты».

**Прототрофы** — микроорганизмы дикого типа, способные развиваться на простых средах без добавления сложных органических соединений.

**Псевдоаутосомная область** — дистальный участок короткого плеча хромосомы Y, который конъюгирует с дистальным концом короткого плеча хромосомы X в мейозе у мужчин. В этом участке отмечена высокая частота кроссинговера.

**Процессинг** — «созревание» РНК, в процессе которого первичный РНК-транскрипт претерпевает ряд модификаций (удаление интронов, полиаденилирование), превращаясь в молекулу зрелой мРНК.

## Р

**Рамка считывания** — нуклеотидная последовательность, выраженная в кодирующих триплетах, начинается со стартового кодона и заканчивается нонсенс-кодоном:

- закрытая Р. с. — Р. с., внутри которой в результате мутации возникает стоп-кодон;
- открытая Р. с. (англ. *open reading frame, ORF*) — участок ДНК между иницирующим и стоп-кодонами;
- сдвиг Р. с. — мутация, приводящая к сдвигу считывания триплетов в процессе трансляции.

**Ревертаза, РНК-зависимая ДНК-полимераза.** См. Обратная транскриптаза.

**Рекомбинантная ДНК** — см. ДНК рекомбинантная.

**Репарация** — процесс исправления повреждений в структуре ДНК, вызванный ошибками репликации или повреждающими агентами внешней среды (мутагенами). Различают несколько видов репарации:

- прямая;
- эксцизионная;
- репликативная;
- рекомбинантная;
- SOS-репарация.

**Репликация (редупликация)** — процесс, в ходе которого происходит удвоение каждой из цепей двунитевой ДНК.

**Ретровирусы** — семейство РНК-содержащих вирусов. Содержат несколько фрагментов одноцепочечной линейной РНК, обратную транскриптазу. Размножаются в клетках птиц, млекопитающих, в том числе человека. Многие Р. вызывают лейкозы, саркомы и опухоли молочных желез.

**Ретротранспозоны** — класс подвижных генетических элементов. Механизм их перемещения основан на транскрипции ретротранспозона (синтезе мРНК на матрице ДНК), вслед за которой осуществляется обратная транскрипция — синтезе нити ДНК на матрице РНК.

**Рестрикция ДНК** — метод генетической инженерии, с помощью которого получают фрагменты ДНК, в том числе гены. Для Р. используют специфические ферменты вирусов и плазмид — *рестриктазы*. Полученные фрагменты используют далее для клонирования или получения рекомбинантных молекул ДНК.

**Рибонуклеиновые кислоты (РНК)** — тип нуклеиновых кислот, содержащие в качестве углеводного компонента рибозу, а в качестве гетероциклических оснований — аденин, гуанин, урацил, цитозин. В клетках всех организмов РНК участвуют в реализации генетической информации. У РНК-содержащих вирусов РНК является носителем генетической информации. Основные типы РНК:

- мРНК — матричные, или информационные;
- тРНК — транспортные;
- рРНК — рибосомные.

**Рибосомы** — внутриклеточные рибонуклеопротеидные частицы, служащие для биосинтеза белков (трансляции). Р. бактерий и эукариот имеют сходное строение и состоят из двух субчастиц — большой и малой.

Рибосомные РНК (рРНК) бактерий: 5S рРНК; 16S рРНК; 23S рРНК; рРНК эукариот: 5S рРНК; 5,8S рРНК; 18S рРНК; 28S рРНК.

## С

**Сайт** — определенное место (позиция) в молекуле ДНК:

- рестрикции — специфическая нуклеотидная последовательность длиной 4–10 п. н., распознаваемая рестрикционной эндонуклеазой.

**Сантиморган (сМ)** — единица измерения частоты рекомбинации между двумя или более локусами хромосомы. Соответствует частоте кроссинговера 1 %.

**С-конец** — конец полипептидной цепи, содержащий свободную карбоксильную группу ( -COOH).

**Сателлитные ДНК (сат ДНК)** — последовательности ДНК эукариот, для которых характерны высокоповторяющиеся последовательности, множество копий, концентрация в прицентромерном и теломерном гетерохроматине, нахождение в хромосомах в виде тандемных (друг за другом) повторов. Сателлитная ДНК не кодирует полипептиды и РНК, а выполняет структурные функции в хромосоме в митозе и мейозе.

**Секвенирование** — определение первичной структуры нуклеиновых кислот и белков, т. е. последовательности расположения нуклеотидных остатков в полинуклеотидной цепи (ДНК или РНК) или аминокислотных остатков в полипептидной цепи белка. Прямое определение первичной структуры макромолекул проводят с использованием биохимических методов. Эти методы вошли в практику С. во второй половине 70-х гг. XX в. В настоящее время методы С. полностью автоматизированы с помощью специальных автоматических приборов — секвенаторов.

**Систематика** — раздел биологии, задачей которого является выделение, описание и обозначение всех существующих и вымерших организмов, а также их классификация по таксонам (группировкам) различных рангов.

**Снипсы (однонуклеотидные полиморфизмы, англ. *single nucleotide polymorphism, SNP*)** — полиморфизм, обусловленный изменением одного единственного нуклеотида в определенной последовательности ДНК.

**Сплайсинг** — процесс превращения первичной молекулы РНК, образующейся при транскрипции гена, в зрелую молекулу РНК, выполняющую свою функцию за счет удаления из нее последовательностей, транскрибированных с интронов:

- альтернативный — различный характер «вырезания» интронов из одного и того же первичного РНК-транскрипта, обеспечивающий возможность синтеза одним геном различных белков.

**Сцепление** — состояние, когда два локуса (гена) находятся в одной хромосоме, а частота рекомбинаций между ними оказывается меньше 50 %.

## Т

**Таксон** — группа организмов, связанных той или иной степенью родства и достаточно обособленная, чтобы ей можно было присвоить определенную таксономическую категорию ранга — вид, род, семейство и т. д. Т. подразумевает конкретные биологические объекты.

**Таксономия** — теория систематики, учение о таксонах, способах их выделения, диагностики, группировки и соподчинения.

**Теломеры** — концевые участки хромосом. В большинстве клеток они укорачиваются с возрастом, выступают как факторы, определяющие продолжительность жизни особи.

**Теломераза** — фермент, достраивающий 3'-концы линейных молекул ДНК хромосом и решающий проблему «концевой недорепликации ДНК» у эукариот.



Активность Т. обнаружена в эмбриональных и опухолевых клетках, не отмечена в соматических клетках.

**Теломерная ДНК** — специализированные повторяющиеся последовательности ДНК на концах хромосом эукариот (теломерах). Теломеры построены из коротких (6–8 п. н.) многократно повторяющихся последовательностей. Теломерная ДНК человека построена из TTAGGG-блоков, у всех растений теломерный повтор универсален — TTTAGGG. Длина теломерной ДНК у разных организмов неодинакова: у человека от 2 до 20 тыс. п. н., у мышей — сотни тысяч п. н. При каждом раунде репликации хромосомы укорачиваются на 10–20 нуклеотидов, а в первую очередь сокращается длина теломерной ДНК.

**Терминатор** — участок ДНК, входящий в состав транскриптона (оперона) и служащий для остановки (терминации) транскрипции.

**Терминирующий кодон** — один из трех триплетов — UAG, UAA и UGA, вызывающих терминацию синтеза белка; их называют также бессмысленными кодонами. Наряду с 20 каноническими аминокислотами, у некоторых архей, бактерий и эукариот, включая человека, генетический код дополнен кодоном UGA для 21-й пептидной аминокислоты — селеноцистеина. Недавно у метанобактерий и у некоторых бактерий открыта 22-я пептидогенная аминокислота — пирролизин, кодируемая UAG-кодоном, также считавшимся стоп-кодоном.

**Тельце Барра** — инактивированная хромосома X в соматических клетках женщин, видимая как интенсивно окрашенная масса хроматина. Обозначается также как половой хроматин.

**Трансгенез** — введение чужеродного гена в животную или растительную клетку и его передача в ряду поколений.

**Трансдукция** — перенос бактериального гена от одного вида бактерий к другому при помощи фага. Фаг, осуществляющий этот перенос, носит название трансдуцирующего. Его геном содержит собственные гены и гены клетки-донора.

**Транскрипция** — биосинтез РНК на матрице ДНК на основе принципа комплементарности. Обычно Т. происходит с одной из двух цепей ДНК — *матричной* (комплементарная цепь носит название *смысловой*). У некоторых бактериофагов (фаг Т 4) матрицами могут служить разные цепи ДНК. Синтез цепи РНК идет в направлении от 5'- к 3'-концу молекулы:

- обратная — комплементарный синтез молекулы к ДНК, при котором роль матрицы выполняет мРНК при участии фермента ревертазы;
- альтернативная — экспрессия гена с разных промоторов в разных специализированных клетках организма в процессе онтогенетической дифференциации тканей.

**Транскриптон** — единица транскрипции, участок ДНК, на котором идет синтез РНК. Состоит из промотора с одной стороны, терминатора с другой стороны, между которыми расположен кодирующий участок. У бактерий Т. соответствует оперону.

**Трансляция** — синтез на рибосомах полипептидной цепи, в которой порядок аминокислот соответствует порядку кодирующих триплетов мРНК-матрицы.

**Транспозоны** — мигрирующие последовательности ДНК бактерий, содержащие структурные гены, контролирующие процесс транспозиции, а также гены, кодирующие дополнительные функции (например, гены устойчивости к антибиотикам и др.). Они способны к перемещению в разные участки хромосомы или внехромосомной ДНК. Они являются причиной 90% спонтанных мутаций.

**Трансстаза** — фермент, участвующий в перемещении мобильных генетических элементов из одного сайта в другой.

**Трансфекция** — искусственное введение в эукариотические клетки изолированных молекул ДНК.

**Трансформация** — поглощение изолированной ДНК бактерии-донора клетками бактерии-реципиента. В природе Т. может происходить с участием вирусов, плазмид и других мобильных генетических элементов. В генетической инженерии Т. осуществляется с использованием специальных генетических конструкций *векторов молекулярного клонирования*. Это эффективный прием биотехнологии, позволяющий получать генно-модифицированные виды животных, растений и микроорганизмов.

**Т-фаги** — группа вирулентных фагов *E. coli*, для которых характерно наличие длинного «хвоста», через который вирусная ДНК проникает в клетку хозяина. По морфологическим особенностям и параметрам жизненного цикла Т-фаги разделены на две группы — четные (Т 2, Т 4) и нечетные (Т 1, Т 3, Т 5, Т 7).

## У

**Убиквитин** — белок, выполняющий важнейшую роль в АТФ-зависимом механизме распада белков. Это небольшой белок (76 аминокислотных остатков, М=8500 Да). Участвует в регуляции клеточного цикла.

## Ф

**Фактор F** — половой фактор бактерий: плаزمида, определяющая конъюгационные свойства мужских штаммов бактерий.

**Фактор R** — плазмида, содержащая гены, контролирующие устойчивость микроорганизма к антибиотикам или другим антибактериальным средствам; способна передаваться от особи к особи.

**Фармакогеномика (фармакоантропология)** — занимается систематизацией вариабельности генов, контролирующих реакции организма на условия среды, а также на действие лекарственных препаратов.

**Физическая карта генома** — определяет места локализации и взаиморасположения генов, участков к ДНК, а также генетических маркеров (STS-маркеров, сайтов узнавания рестриктазами и т. д.), присутствующих в исследуемой молекуле ДНК. Расстояние между ними выражается в числе п. н.

**Фолдинг** — процесс формирования пространственной (третичной) структуры белков. Ф. обслуживается особыми белковыми комплексами — шаперонами.

**Фосфодиэфирная связь** — связь между нуклеотидами в нуклеиновых кислотах. Она осуществляется остатками фосфорной кислоты, которая связывает 3'-углеродной атом дезоксирибозы (или рибозы) одного нуклеотида с 5'-углеродным атомом дезоксирибозы (или рибозы) другого.

**Фототрофы** — общее название организмов, использующих свет в качестве источника энергии. К ним относятся растения и фитотрофные микроорганизмы (пурпурные и зеленые бактерии, цианобактерии, гелиобактерии).

**Фрагменты Оказаки** — короткие фрагменты ДНК длиной 1000–2000 н. у прокариот и 100–200 н. у эукариот, образующиеся в результате прерывистой репликации на отстающей цепи; впоследствии достраиваются и ковалентно соединяются в непрерывную цепь.

## Х

**Хеликаза** — фермент (мол. м. 30 кДа), выполняющий функцию расплетения двойной спирали ДНК в репликационной вилке за счет энергии гидролиза АТФ.

**Хроматин** — конденсированная хромосомная ДНК клетки в интерфазе. Х. состоит из ДНК и небольшого количества РНК, а также белков (гистонов).

**Хромосомоспецифическая библиотека** — набор клонируемых фрагментов ДНК отдельной хромосомы.

## Ц

**Царство (regnum)** — одни из самых высоких таксономических категорий в системе организмов. К середине XX в. эта точка зрения устарела. Большинство ученых признали необходимым выделение таксона более высокого ранга, чем Ц., а именно надцарства (superregnum) и даже империи (imperium). Были выделены две империи: доклеточных и клеточных. Современная мегасистематика включает все организмы в три домена: архебактерии, эубактерии и эукариоты.

**Цистрон** — генетическая единица, кодирующая структуру одного полипептида.

## Ш

**Штамм** — чистая культура микроорганизмов, выделенная из определенного источника. Разные Ш. одного и того же микроорганизма могут отличаться рядом свойств (вирулентностью, чувствительностью к антибиотикам).

**Шапероны** — белковые комплексы, способствующие процессу формирования пространственной структуры (фолдингу) синтезированных полипептидов путем ограничения их контактов с другими белками.

## Э

**Экзон** — кодирующий участок гена, сохраняющийся в молекуле зрелой мРНК.

**Экспрессия гена** — активное состояние гена, при котором происходит реализация записанной в нем генетической информации: синтез РНК и белка.

**Элонгация** — последовательное присоединение мономеров к полимерной цепи.

**Эндонуклеазы (рестриктазы)** — ферменты бактериального происхождения, разрезающие двухнитевые молекулы ДНК в местах, соответствующих специфическим последовательностям из 4–12 нуклеотидов.

**Энхансер** — регулирующий участок ДНК эукариот, который посредством специфических белков (энхансеров) усиливает транскрипцию гена.

**Эпигенетика (эпигеномика)** — изучает формирование и наследственную передачу специфического функционального состояния генома, в том числе факторы подавления и активации транскрипции. Среди этих факторов особое значение придается метилированию (присоединение метильной группы к цитозину ДНК), ацетилированию (перенос ацетильных остатков на остатки лизина в молекулах гистонов) и др.

**Эукариоты** — по современной мегасистематике на основе молекулярных реконструкций, представляют собой, наряду с архебактериями и эубактериями, высшие таксоны организмов — *дамены*. В домен эукариот входят 5 *надцарств* (Superregnum): *экскаваты* (Excavata), *ризарии* (Rhizaria), *одножгутиковые* (Unikonta), *хромальвеоляты* (Chromalveolata), *растения в широком смысле* (Plantae).

## Основная литература

- Абрамсон Н. И. и др.* Коллекция ДНК — новый путь изучения и сохранения биологического разнообразия // Сб.: *Фундаментальные зоологические исследования. Теория и методы.* М.; СПб.: КМК, 2004.
- Айала Ф.* Введение в популяционную и эволюционную генетику. М.: Мир, 1984.
- Алексеев В. И., Каминский В. А.* Прикладная молекулярная биология. М.: КомКнига, 2004.
- Алиханян С. И., Акифьев А. П., Чернин Л. С.* Общая генетика. М.: Высшая школа, 1984.
- Алтухов Ю. П.* Генетические процессы в популяциях. М.: Академкнига, 2003. С. 431.
- Баранов В. С., Киселев Л. Л.* Геном человека и молекулярная медицина // Сб.: *Геномика-медицине* / Ред. В. И. Иванов, Л. Л. Киселев. М.: ИКЦ «Академкнига», 2005.
- Бердников В. А.* Основные факторы макроэволюции. Новосибирск: Наука, 1990. С. 286.
- Бокуть С. Б. и др.* Молекулярная биология. Минск: Высшая шк., 2005.
- Боринская С. А., Янковский Н. К.* Структура прокариотических геномов // *Мол. биол.* 1999 Т. 33 (6).
- Вернадский В. И.* Химическое строение биосферы земли и ее окружения. М.: Наука, 1965.
- Волькенштейн М. В.* Физический смысл нейтралистской теории эволюции // *Ж. общ. биол.* 1981. Т. 42. С. 680–686.
- Волькенштейн М. В.* Биофизика. М.: Наука, 1988.
- Воронцов Н. Н.* Рецензия на книгу В. А. Кордюма «Эволюция и биосфера» // *Мол. биол.* 1984. Т. 18 (3). С. 855–857.
- Воронцов Н. Н.* Развитие эволюционных идей в эволюции. М.: УНЦ ДО МГУ, 1999. С. 639.
- Воронцов Н. Н.* Эволюция. Видообразование. Система органического мира. М.: Наука, 2005. С. 365.
- Гельфанд М. С.* Компьютерный анализ последовательностей ДНК // *Мол. биол.* 1998. Т. 32. С. 103–120.
- Генетика* / Ред. В. И. Иванов. М.: ИКЦ «Академкнига», 2006. С. 638.
- Геномика — медицине* / Ред. В. И. Иванов, Л. Л. Киселев. М.: ИКЦ «Академкнига», 2005. С. 392.
- Гинтер Е. К.* Медицинская генетика. М.: Медицина, 2003.

- Глазко В. И., Калашников Л. А. и др. Введение в ДНК-технологии. М.: Росинформагротех, 2001. С. 434.
- Гнатик Е. Н. Генетика человека: былое и грядущее. М.: Издательство ЛКИ/URSS, 2007.
- Гатубовский М. Д. Книга «Непостоянство генома» в аспекте концептуальной истории генетики // Мол. биол. 2002. Т. 36 (2). С. 338–346.
- Горбунова В. Н., Баранов В. С. Введение в молекулярную диагностику и генотерапию наследственных заболеваний. СПб.: Спец. Литература, 1997.
- Гусев М. В., Минеева Л. А. Микробиология. М.: МГУ, 1992.
- Дебабов В. Г. Жизнь бактерий за стенами лабораторий // Мол. биол. 1999. Т. 33 (6). С. 1074–1084.
- Домовая мышь: происхождение, распространение, систематика, поведение / Ред. В. Е. Соколов. М.: Наука, 1994.
- Драмышко С. Е. Математические и компьютерные модели в биологии: взгляд генетика. Минск: Бел. наука, 2006.
- Дубинин Н. П. Общая генетика. М.: Наука, 1976.
- Дьяков Ю. Т. Введение в альгологию и микологию. М.: МГУ, 2000. С. 192.
- Жуковский П. М. Культурные растения и их сородичи. Систематика, география, цитогенетика. Л.: Колос, 1971.
- Заварзин Г. А. Не дарвиновская область эволюции // Вестник РАН. 2000. Т. 70. С. 403–411.
- Заварзин Г. А., Колотилова Н. Н. Введение в природоведческую микробиологию. М.: Книж. дом «Университет», 2001.
- Зеленин А. В., Бабаева Е. Д., Муравенко О. В. Введение в геномику растений // Мол. биол. 2001. Т. 35. № 3.
- Зинченко А. Н., Паруль Д. А. Основы молекулярной биологии вирусов и антивирусной терапии. Минск: Высшейшая школа, 2005.
- Инге-Вечтамов С. Г. Введение в молекулярную генетику. М.: Высш. шк., 1983.
- Каменская М. А. Информационная биология. М.: Academia, 2006.
- Карасев В. А. Генетический код: Новые горизонты. СПб.: Теса, 2003.
- Кимура М. Молекулярная эволюция: теория нейтральности. М.: Мир, 1985.
- Коничев А. С., Севастьянова Г. А. Молекулярная биология. М.: Academia, 2006.
- Киселев Л. Л. Геном человека и биология XXI в. // Вестник РАН. 2000. Т. 70 (5). С. 712–724.
- Киселев Л. Л. Юбилей самой главной молекулы // РАН. 2003. Т. 73 (6). С. 496.
- Кордюм В. А. Перенос информации в биосфере и возможное эволюционное значение этого процесса // Усп. совр. биол. 1976. Т. 81 (1).
- Кордюм В. А. Эволюция и биосфера. Киев: Наукова думка, 1982.
- Конов А. Л. Сельскохозяйственная биотехнология и «горизонтальный перенос» генов // Сб.: Современная биотехнология. Мифы и реальность. М.: Тайдекс Ко, 2004. С. 156–174.
- Красилов В. А. Эволюция и биостратиграфия. М.: Наука, 1977.
- Левонтин Р. Генетические основы эволюции. М.: Мир, 1978.

- Макаров В. В. и др. Эпизоотологический словарь. М.: Колос, 2001.
- Медведев Н. Н. Практическая генетика. М.: Наука, 1966.
- Морозова О. В. Загадки архей и их фагов // Вестник ВОГиС. 2005. Т. 9 (1). С. 55–66.
- Мушкамбаров Н. Н., Кузнецов С. Л. Молекулярная биология. М.: МИА, 2003. С. 535.
- Назаров В. И. Эволюция не по Дарвину. Смена эволюционной модели. М.: КомКнига, 2005. С. 519.
- Попов В. В. Генетика и селекция животных. М.: РГАЗУ, 2004. С. 194.
- Попов В. В. Горизонтальный перенос генов как информационный фактор эволюции // Вестник охотоведения. 2007. Т. 4 (2). С. 183–194.
- Попов В. В. Молекулярно-генетическое изучение геномов // Науч. журн. «Вестник РГАЗУ», 2007. № 3 (8).
- Пуанкаре А. О науке. М.: Наука, 1990.
- Ратнер В. А. Математическая популяционная генетика. Новосибирск: Наука, 1977.
- Ратнер В. А. Молекулярная генетика: принципы и механизмы. Новосибирск: Наука, 1983.
- Ратнер В. А. Генетика, молекулярная кибернетика. Личность и проблема. Новосибирск: Наука, 2002.
- Романовский Ю. М. и др. Математическая биофизика. М.: Наука, Гл. ред. физ.-мат. лит., 1984.
- Свердлов Е. Д. Микрокосм генома // Мол. биол. 1999. Т. 33 (6).
- Свердлов Е. Д. ДНК в клетке: от молекулярной иконы к проблеме «что есть жизнь?» // Вестник РАН. 2003. Т. 73 (6).
- Свердлов Е. Д. (отв. ред.) Проблемы и перспективы молекулярной генетики. Т. 2. М.: Наука, 2004. С. 530.
- (Тимофеев-Ресовский Н. В., Циммер К. Г., Дельбрюк М.) *Timofeeff-Resovsky N. W., Zimmer K. G., Delbrück M. Über die Natur der Genmutationen und Genstruktur. Nachr. Ges. Wiss. Göttingen, 1935. Bd. 6. S. 1.*
- Тимофеев-Ресовский Н. В., Иванов В. И., Корогодин В. И. Применение принципа попадания в радиобиологии. М.: Атомиздат, 1968. С. 228.
- Тимофеев-Ресовский Н. В., Воронцов Н. Н., Яблоков А. В. Краткий очерк теории эволюции. 2-е изд. М.: Наука, 1978. С. 407.
- Уоддингтон К. Х. Основные биологические концепции // Сб.: На пути к теоретической биологии / Ред. Б. Л. Астауров. М.: Мир, 1970.
- Фогель Ф., Мотульски А. Генетика человека. Т. 2. М.: Мир, 1990.
- Хесин Р. Б. Непостоянство генома. М.: Наука, 1984. С. 472.
- Хмель И. А. Молекулярно-генетический и функциональный анализ плазмид энтеробактерий // Сб.: Проблемы и перспективы молекулярной генетики. Т. 2. / Ред. Е. Д. Свердлов. М.: Наука, 2004. С. 164–194.
- Шаталкин А. И. Высший уровень деления в классификации организмов.1. Прокариоты и эукариоты // Ж. общ. биол. 2004 а. Т. 65 (1). С. 19–38.
- Шаталкин А. И. Высший уровень деления в классификации организмов.2. Археобактерии, эубактерии, эукариоты // Ж. общ. биол. 2004 в. Т. 65 (2). С. 99–115.

- Шаталкин А. И. Животные (Animalia) в системе организмов.1. Типологические системы // Ж. общ. биол. 2005 а. Т. 66 (4). С. 275–299.
- Шаталкин А. И. Животные (Animalia) в системе организмов.2. Филогенетическое понимание животных // Ж. общ. биол. 2005 в. Т. 66 (5). С. 389–415.
- Шестаков С. В. Геномика патогенных бактерий // Вестник РАМН. 2001. № 10. С. 18–25.
- Шестаков С. В. О ранних этапах биологической эволюции с позиции геномики // Палеонтол. журн. 2003. № 6. С. 50–57.
- Шестаков С. В. Роль горизонтального переноса генов в эволюции // Доклад на теорет. семинаре «Происхождение живых систем». 2003.
- Шлегель Г. Г. История микробиологии. М.: URSS, 2002. С. 302.
- Щелкунов С. Н. Конструирование гибридных молекул ДНК. Новосибирск: Наука, 1983. С. 168.
- Щелкунов С. Н. Генетическая инженерия. Новосибирск: Сибирское университет. издат., 2004. С. 496.
- Adams M. D. et al. The genome sequence of *Drosophila melanogaster* // Science. 2000. V. 287. P. 2185–2195.
- The Arabidopsis Genome Initiative*. Analysis of the genome sequence of the flowering plant *Arabidopsis thaliana* // Nature. 2000. V. 408. P. 796–815.
- Anderson S. F. et al. Safety issues related to retroviral-mediated gene transfer in humans // Hum. Gene Ther. 1986. Vol. 6. P. 5–14.
- Avery O. T. et al. Studies on the chemical nature of the substance inducing transformation in pneumococcal types 1 Induction of transformation by a deoxyribonucleic acid fraction isolated from pneumococcus // J. Exper. Med. 1944. V. 79. P. 137–138.
- Backhed F. et al. Hostbacterial mutualism in the human intestine // Science. 2005. V. 307. P. 1915–1920.
- Blattner F. R. et al. The complete genome sequence of *Escherichia coli* K12 // Science. 1997. V. 277. P. 1453–1462.
- Brown J. R., Doolittle W. F. Archaea and the prokaryote-to-eukaryote transition // Microbiol. Mol. Biol. Rev. 1997. V. 61. P. 456–502.
- Cavalier-Smith T. A revised six-kingdom system of life // Biol. Rev. Camb. Philos. Soc. 1998. V. 75 (3). P. 203–266.
- Cavalier-Smith T. Obcells as proto-organisms: membrane heredity, lithophosphorylation, and the origin of the genetic code, the first cells and photosynthesis // J. Mol. Evol. 2001. V. 53 (4.5). P. 555–595.
- Cavalier-Smith T. The neomuran origin of archaeobacteria, the neobacterial root of the universal tree and bacterial megaclassification // Int. J. Syst. Evol. Microbiol. 2002 a. V. 52 (1). P. 7–76.
- Cavalier-Smith T. The phagotropic origin of eukaryotes and phylogenetic classification of Protozoa // Int. J. Syst. Evol. Microbiol. 2002 b. V. 52 (2). P. 297–354.
- Cavalier-Smith T. Only six kingdoms of life // Proc. R. Soc. Lond., ser. B. Biol. Sci. 2004. V. 271. P. 1251–1262.



- Cavalier-Smith T., Chao E. E.-Y.* Protalveolate phylogeny and systematics and the origin of Sporozoa and donoflagellates (phylum Myzozoa nom. nov.) // *Eur. J. Protistol.* 2004. V. 40. P. 185–212.
- Dayhoff M. O. (ed.)*. Atlas of protein sequence and structure. Vol. 5. Washington, 1978. Nat. Biomed. Res. Found.
- Doolittle W. F.* You are what you eat: a gene transfer ratchet could account for bacterial genes in eukaryotic genomes // *Trends Genet.* 1998. V. 14. P. 307–311.
- Doolittle W. F.* Lateral genomics // *Trend Cell. Biol.* 1999 a. V. 9. P. 5–9.
- Doolittle W. F.* Phylogenetics classification and the universal tree // *Science.* 1999 b. V. 284. P. 2124–2128.
- Eisen J.* horizontal gene transfer among microbial genomes: new insights from complete genome analysis // *Curr. Opin. Genet. Developm.* 2000. V. 10. P. 606–611.
- Fox G. E. et al.* The phylogeny of prokaryotes // *Science.* 1980. V. 209.
- Gill S. R. et al.* Metagenomic analysis of the human distal gut microbiome // *Science.* 2006. V. 312. P. 1355–1359.
- Goodacre R.* Metabolomics of a superorganism // *J. Nutr.* 2007. V. 137. P. 259–266.
- Griffith F.* The significans of pneumococcal types // *J. Hygiene.* 1928. Vol. 27. P. 113–139.
- Gupta R. S.* Protein phylogenesis and signature equences: a reappraisal of evolutionary relationship among archaebacteria, eubacteria and eukaryotes // *Microbiol. Mol. Biol. Rev.* 1998. V. 62. P. 1435–1491.
- Hao B. et al.* A new UAG-encoded residue in the structure of a metanogen methyltransferase // *Sciense.* 2002. V. 296. P. 1462–1466.
- Henze K., Martin W.* How do mitochondrial genes get into the nucleus // *Trends Genet.* 2001. V. 17. P. 383–387.
- <http://eugenesis.org/>
- <http://www.ncbi.nlm.nih.gov>
- International Human Genome Sequeneing Consortium.* Initial seguencing and analysis of the human genome // *Nature.* 2001. V. 409. P. 860–921.
- Jackson D. A. et al.* Biochemical method for inserting new genetic information into DNA of simian virus 40: Circular SV-40 DNA molecules containing lambda phage genes and the galactose operon of *Escherichia coli* // *Proc. Nat. Acad. Sci. USA Biol. Sci.* 1972. V. 69. P. 2904–2909.
- Jain R. et al.* Horisontal transfer among genomics: the complexity hypothesis // *Proc. Nat. Acad. Sci. USA.* 1999. V. 96. P. 3801–3806.
- Kaleko M. et al.* Resistant gene expression after retroviral gene transfer in to liver cells in vitro // *Hum. Gene . Ther.* 1991. V. 158. P. 27–32.
- Keeling P. J. et al.* The tree of eukaryotes // *Trends Ecol. Evol.* 2005. V. 20. P. 670–676.
- Kemp M.* The Mona Lisa of modern science // *Nature.* 2003. V. 421. P. 416–420.
- Koonin E. V. et al.* Horisontal gene transfer in prokaryotes. Quantification and classification // *Ann. Rev. Microbiol.* 2001. V. 55. P. 709–742.
- Kuwardina O. N. et al.* The phylogeny of Colpodellids (Alveolata) using small subunit rRNA gene sequences suggest they are the free-living sistergroup to Apicomplexans // *J. Eukaryot. Microbiol.* 2002. V. 49. P. 498–504.

- Lang B. F. et al.* Mitochondrial genome evolution and the origin of eukaryotes // *Ann. Rev. Genet.* 1999. V. 33. P. 351–397.
- Lederberg J., Tatum E. L.* Gene recombination in *Escherichia coli* // *Nature.* 1946. V. 158. P. 558.
- Madigan M. T., Mairs B. I.* Extremophiles // *Sci. Amer.* 1997. V. 4. P. 82–87.
- Madigan M. T. et al.* Brock biology of microorganisms. New Jersey: Prentice-Hall: 2000. P. 991.
- Maio B.* Satellite DNA in mouse cells infected with virus // *J. Mol. Biol.* 1971. V. 56. P. 579–582.
- Martin W.* A briefly argues case that mitochondria and plastids are descendants of endosymbionts, but that the nuclear compartments is not // *ProcRoy. Soc. London.* 1999. V. 266. P. 1387–1395.
- Martin W. et al.* An overview of endosymbiotic models for the origins of eukaryotes, their ATP-producing organelles (mitochondrials and hydrogenosomes), and their heterotrophic lifestyle // *Biol. Chem.* 2001. V. 382. P. 1521–1539.
- Mouse genome sequencing consortium.* Initial sequencing and comparative analysis of the mouse genome // *Nature.* 2002. V. 420. P. 520–562.
- Oh J. E. et al.* Evidence for the existence of hypothetical proteins in human bronchial, epithelial, fibroblast, amnion, lymphocyte, mesothelial and kidney celllines // *Amino Acids.* 2004. V. 26. P. 9–18.
- Patron N. J.* Gene replacement of fructose-1,6-bisphosphate aldolase supports the hypothesis of a single photosynthetic ancestor of chromalveolates // *Eukaryot. Cell.* 2004. V. 3. P. 1169–1175.
- Prangishvili D.* Evolutionary insights from studies on viruses of hyperthermophilic archaea // *Res. Microbiol.* 2003. V. 154. P. 289–294.
- Prusiner S. B.* Prions // *Proc. Nat. Acad. Sci. USA.* 1998. V. 95. P. 13 363–13 383.
- Ribeiro S., Golding G. B.* The mosaic nature of the eukaryotic nucleus // *Mol. Biol. Evol.*-1998. V. 15. P. 779–788.
- Roytberg M. A.* A search for common patterns in many sequences // *Comput. Appl. Biosci.* 1992. V. 8. P. 5–64.
- Sasaki T.* Impact of the complete rice genome sequence information on future basic and applied plant science research // *Report of Symposium: Inst. Rad. Breeding. Japan,* 2005. P. 1–10.
- Sekirov I., Finlay B. B.* Human and microbe: united we stand // *Nat. Med.* 2006. V. 12. P. 736–737.
- Simpson G. G.* The major features of evolution. Columb. Univ. Press, 1953.
- Stanier R. Y., Van Niel C. B.* The concept of a bacterium // *Arch. Mikrobiol.* 1962. V. 42. P. 17–35.
- Stechmann A., Cavalier-Smith T.* The root of the eukaryote tree pinpointed // *Curr. Biol.* 2003. V. 13. P. 665–666.
- Stechmann A., Cavalier-Smith T.* Evolutionary origins of Hsp 90 chaperones and a deep paralogy in their bacterial ancestors // *J. Eukaryot. Microbiol.* 2004. V. 51. P. 364–373.

- Srinivasan G. et al.* Pyrrolysine encoded by UAG in archaea: charging of UAG-decoding specialized tRNA // *Science*. 2002. V. 296. P. 1459–1462.
- Venter J. C. et al.* The sequence of the human genome // *Science*. 2001. V. 291. P. 1304–1351.
- Waddington C. H.* The strategy of the genes. London: Allen Unwinn, 1957.
- Woese C. R.* Archaeobacteria // *Sci. Amer.* 1981. V. 244.
- Woese C. R.* The universal ancestor // *Proc. Nat. Sci. USA.*, 1998. V. 25. P. 6854–6859.
- Woese C. R.* Interpreting the universal phylogenetic tree // *Proc. Nat. Acad. Sci. USA.*, 2000. V. 97. P. 8392–8396.
- Woese C. R., Fox G. E.* Phylogenetic structure of the prokaryotic domain: the primary kingdoms // *Proc. Nat. Acad. Sci. USA.*, 1977. V. 74. P. 5088.
- Woese C. R. et al.* Towards a natural systems of organisms: proposal for the domain archaea, bacteria and eucaria // *Proc. Nat. Acad. Sci. USA.*, 1990. V. 87. P. 4576–4579.
- Wollman E. L. et al.* Conjugation and genetic recombination in *Escherichia coli* K12 // *Cold Spring Symp. Quant. Biol.* 1956. V. 21. P. 141–162.
- Zuckerandl E., Pauling L.* Evolutionary divergence and convergence in proteins. In: *Evolving Genes and Proteins*. N.-Y.: Academic Press, 1965.

## Представляем Вам наши лучшие книги:



URSS

### Химия

- Ельяшевич М. А.* Атомная и молекулярная спектроскопия. В 3 кн.
- Грибов Л. А.* Колебания молекул.
- Грибов Л. А.* От теории спектров к теории химических превращений.
- Грибов Л. А., Баранов В. И.* Теория и методы расчета молекулярных процессов.
- Грибов Л. А., Баранов В. И., Эляшберг М. Е.* Основы теории безэталонного молекулярного спектрального анализа.
- Захаров А. Г. (ред.)* Научные основы химической технологии углеводов.
- Стехин А. А., Яковлева Г. В.* Структурированная вода: Нелинейные эффекты.
- Цивадзе А. Ю. (ред.)* Структурная самоорганизация в растворах и на границе раздела фаз.
- Болотин С. Н. и др.* Координационная химия природных аминокислот.
- Золотов Ю. А., Иванов В. М., Амелин В. Г.* Химические тест-методы анализа.
- Золотов Ю. А., Варшал Г. М., Иванов В. М. (ред.)* Аналитическая химия металлов платиновой группы.
- Золотов Ю. А.* Химики в других областях или на других Олимпах.
- Золотов Ю. А.* Химики еще шутят.
- Золотов Ю. А., Широкова В. И. (ред.)* Кто есть кто в российской аналитической химии.
- Золотов Ю. А.* Кто был кто в аналитической химии в России и СССР.
- Золотов Ю. А.* Делающие науку. Кто они? Из записных книжек.
- Белова В. В., Холькин А. И. (ред.)* Кто есть кто в химической технологии.
- Холькин А. И., Патрушева Т. Н.* Экстракционно-пиролитический метод.
- Мюнстер А.* Химическая термодинамика.
- Агеев Е. П.* Неравновесная термодинамика в вопросах и ответах.
- Дуров В. А., Агеев Е. П.* Термодинамическая теория растворов.
- Казенас Е. К., Цветков Ю. В.* Термодинамика испарения оксидов.
- Салем Р. Р.* Физическая химия. Начала теоретической электрохимии.
- Плесков Ю. В.* Электрохимия алмаза.
- Михайлов О. В.* Ионнообменные процессы в тонкопленочных биополимер-иммобилизованных металлосульфидах.
- Половняк В. К., Михайлов О. В., Кузнецов А. М.* Комплексы 4d-платиновых металлов с фосфор(III)- и мышьяк(III)органическими лигандами.
- Комиссаров Г. Г.* Фотосинтез: физико-химический подход.
- Комиссарова Л. Н.* Неорганическая и аналитическая химия скандия.
- Кабачник М. И., Мاستрюкова Т. А.* Межфазный катализ в фосфорорганической химии.
- Шапиро Б. И.* Теоретические начала фотографического процесса.
- Збарский В. Л., Жилин В. Ф.* Толуол и его нитропроизводные.
- Андреев Б. М. и др.* Гетерогенные реакции изотопного обмена трития.
- Илюшин Г. Д.* Моделирование процессов самоорганизации в кристаллообразующих системах.
- Галиулин Р. В.* Кристаллографическая геометрия.
- Кедров Б. М.* День одного великого открытия. (Об открытии Д. И. Менделеевым периодической таблицы элементов.)
- Аржанцев В. И., Шишов С. Е., Жабрев В. А. и др.* Белая книга по нанотехнологиям.

## Представляем Вам наши лучшие книги:



URSS

### Серия «Из наследия И. Т. Фролова»

*Фролов И. Т.* Философия и история генетики. Поиски и дискуссии.

*Фролов И. Т.* Жизнь и познание: О диалектике в современной биологии.

*Фролов И. Т.* Очерки методологии биологического исследования.

*Фролов И. Т.* Перспективы человека.

*Фролов И. Т., Пастушный С. А.* Менделизм и философские проблемы современной генетики.

*Фролов И. Т., Юдин Б. Г.* Этика науки: Проблемы и дискуссии.

### Серия «Науки об искусственном»

*Арбиб М.* Метафорический мозг.

*Шамис А. Л.* Поведение, восприятие, мышление.

*Гаазе-Рапопорт М. Г., Поспелов Д. А.* От амебы до робота: модели поведения.

*Саймон Г.* Науки об искусственном.

*Попов Э. В.* Общение с ЭВМ на естественном языке.

*Финн В. К.* Интеллектуальные системы и общество.

*Варшавский В. И., Поспелов Д. А.* Оркестр играет без дирижера: Размышление об эволюции некоторых технических систем и управлении ими.

*Редько В. Г. (ред.)* От моделей поведения к искусственному интеллекту.

### Серия «Классический университетский учебник»

*Квасников И. А.* Термодинамика и статистическая физика. В 4 т.

*Ишханов Б. С., Капитонов И. М., Юдин Н. П.* Частицы и атомные ядра.

*Кононович Э. В., Мороз В. И.* Общий курс астрономии.

*Гнеденко Б. В.* Курс теории вероятностей.

*Колмогоров А. Н., Драгалин А. Г.* Математическая логика.

*Петровский И. Г.* Лекции по теории обыкновенных дифференциальных уравнений.

### Серия «НАУКУ — ВСЕМ! Шедевры научно-популярной литературы»

*Кац Е. А.* Фуллерены, углеродные нанотрубки и нанокластеры.

*Колмогоров А. Н.* Математика — наука и профессия.

*Стинрод Н., Чинн У.* Первые понятия топологии.

*Мизес Р.* Вероятность и статистика.

*Лебег А.* Об измерении величин.

*Харкевич А. А.* Автоколебания.

*Ашкинази Л. А.* Электронные лампы: Из прошлого в будущее.

*Перельман Я. И.* Занимательная астрономия.

*Липунов В. М.* В мире двойных звезд.

*Кононович Э. В.* Звезда Солнце.

*Каганов М. И.* Электроны, фононы, магноны.

*Каганов М. И., Цукерник В. М.* Природа магнетизма.

*Тарасов Л. В., Тарасова А. Н.* Беседы о преломлении света.

*Сазанов А. А.* Четырехмерная модель мира по Минковскому.

*Хвольсон О. Д.* Теория относительности А. Эйнштейна и новое миропонимание.

*Гарднер М.* Теория относительности для миллионов.

## Представляем Вам наши лучшие книги:



URSS

### Медицина

*Гиппократ.* О природе человека.

*Нужный В. П., Рожанец В. В., Ефремов А. П.* Лекарственные растения и фитокомпозиции в наркологии.

*Халафян А. А.* Современные статистические методы медицинских исследований.

*Аронова Е. А.* Иммунитет. Теория, философия и эксперимент.

*Николлс Дж. Г. и др.* От нейрона к мозгу. Пер. с англ.

*Загородний Е. С.* Принцип работы мозга.

*Каминский Ю. Г., Косенко Е. А.* Популярно и не очень о болезни Альцгеймера.

*Леках В. А.* Ключ к пониманию физиологии.

*Леках В. А.* Прикладная медицина – постановка и решение задач.

*Агарков С. Т.* Супружеская дезадаптация.

*Кащенко Е. А.* Преждевременное семяизвержение: 65 способов продления.

*Пичуев В. П.* Гинекология и акушерство: серьезно и доступно: Тебе, женщина.

*Алексеев А. А., Ларионова И. С. и др.* Здоровье, сексуальность и вумбилдинг.

*Алексеев А. А., Ларионова И. С.* Системное мышление в практике биолога и врача.

*Вуколова З. П. и др.* Доврачебная экстренная и неотложная помощь взрослым.

*Львовская З. Д. (ред.)* ФИЗИОТЕРАПИЯ: глоссарий текстовых конвенций (испанский, английский, французский, русский). В 2 т.

*Барышников А. Ю., Шишкин Ю. В.* Иммунологические проблемы апоптоза.

*Крутько В. Н., Донцов В. И.* Системные механизмы и модели старения.

*Грегори Р. Л.* Разумный глаз. Как мы узнаем то, что нам не дано в ощущениях.

*Павлова О. В.* Основы психодерматологии.

*Павлова О. В.* Психо-нейро-иммунные взаимодействия и кожа.

*Павлова О. В.* Психо-нейро-иммунные взаимодействия при псориазе.

*Павлова О. В.* Психо-нейро-иммунные взаимодействия при атопическом дерматите.

*Харгиттаи И.* Откровенная наука. Беседы со знаменитыми химиками. Пер. с англ.

*Харгиттаи И.* Откровенная наука. Беседы с корифеями биохимии и медицинской химии. Пер. с англ.

*Журавлев И. В.* Семиотический анализ расстройств речемыслительной деятельности.

### Серия «Раскрывая тайны мозга»

*Сергеев Б. Ф.* Ступени эволюции интеллекта.

*Сергеев Б. Ф.* Парадоксы мозга.

*Сергеев Б. Ф.* Стать гением. От инстинкта к разуму.

*Сергеев Б. Ф.* Ум хорошо...

*Сергеев Б. Ф.* Высшая форма организованной материи.

### Серия «Из наследия А. Н. Шамина»

*Шамин А. Н.* История биологической химии. Истоки науки.

*Шамин А. Н.* История биологической химии. Формирование биохимии.

*Шамин А. Н.* История биологической химии. Институционализация биохимии.

*Шамин А. Н.* История химии белка.

*Шамин А. Н.* Биокатализ и биокатализаторы. Исторический очерк.

## Представляем Вам наши лучшие книги:



URSS

### Зоология, ботаника

- Формозов А. Н.* Спутник следопыта.  
*Корытин С. А.* Повадки диких зверей.  
*Корытин С. А.* Поведение и обоняние хищных млекопитающих.  
*Корытин С. А.* Человек и медведь.  
*Корытин С. А.* Животные — наркотики — человек. Тигр под наркозом.  
*Корытин С. А.* Приманки зверолова. Управление поведением зверей.  
*Панов Е. Н.* Поведение животных и этологическая структура популяций.  
*Панов Е. Н.* Механизмы коммуникации у птиц.  
*Захваткин Ю. А.* Курс общей энтомологии.  
*Захваткин Ю. А.* и др. Словарь-справочник энтомолога.  
*Кипятков В. Е.* Мир общественных насекомых.  
*Карасева Е. В., Телицына А. Ю., Жигальский О. А.* Методы изучения грызунов в полевых условиях.  
*Ахметьева Н. П.* Рядом с ирисами.  
*Лотова Л. И.* Ботаника: морфология и анатомия высших растений.  
*Лотова Л. И., Нилова М. В., Рудько А. И.* Словарь фитоанатомических терминов.

### Экология

- Басов В. М.* Задачи по экологии и методика их решения.  
*Басов В. М., Ефремова Т. В.* Практикум по анатомии, морфологии и систематике растений.  
*Красовская Т. М.* Природопользование Севера России.  
*Сердюцкая Л. Ф.* Системный анализ и математическое моделирование экологических процессов в водных экосистемах.  
*Сердюцкая Л. Ф., Яцишин А. В.* Техногенная экология.  
*Абакумова Г. М., Горбаренко Е. В.* Прозрачность атмосферы в Москве за последние 50 лет и ее изменения на территории России.  
*Калов Х. М., Стасенко В. Н. (ред.)* Доклады Всероссийской конференции по физике облаков и активным воздействиям на гидрометеорологические процессы.  
*Ахметьева Н. П., Лапина Е. Е., Лола М. В.* Экологическое состояние природных вод водосбора Ивановского водохранилища и пути по сокращению их загрязнения.  
*Чалов Р. С.* Руслведение: теория, география, практика.  
*Дмитриев Е. А.* Математическая статистика в почвоведении.  
*Таргульян В. О., Горячкин С. В. (ред.)* Память почв: Почва как память биосферно-геосферно-антропосферных взаимодействий.  
*Самсонова В. П.* Пространственная изменчивость почвенных свойств.  
*Васенев И. И.* Почвенные сукцессии.  
*Ефремов И. В.* Моделирование почвенно-растительных систем.  
*Гольева А. А.* Микробиоморфные комплексы природных и антропогенных ландшафтов.  
*Вязилов Е. Д.* Информационные ресурсы о состоянии природной среды.  
*Касаткин Р. Г.* Система морской транспортировки сжиженного газа из Арктики.  
*Свалов А. М.* Механика процессов бурения и нефтегазодобычи.  
*Судо М. М., Судо Р. М.* Нефть и углеводородные газы в современном мире.

## Представляем Вам наши лучшие книги:



URSS

Серия «Синергетика: от прошлого к будущему»

*Пенроуз Р.* **НОВЫЙ УМ КОРОЛЯ.** О компьютерах, мышлении и законах физики.

*Хакен Г.* **Информация и самоорганизация.**

*Котов Ю. Б.* **Новые математические подходы к задачам медицинской диагностики.**

*Гельфанд И. М. и др.* **Очерки о совместной работе математиков и врачей.**

*Долгоносков Б. М.* **Нелинейная динамика экологических и гидрологических процессов.**

*Малинецкий Г. Г.* **Математические основы синергетики.**

*Малинецкий Г. Г., Потапов А. Б.* **Нелинейная динамика и хаос: основные понятия.**

*Малинецкий Г. Г., Потапов А. Б., Подлазов А. В.* **Нелинейная динамика.**

*Малинецкий Г. Г. (ред.)* **Будущее России в зеркале синергетики.**

*Малинецкий Г. Г. (ред.)* **Синергетика: Исследования и технологии.**

*Арнольд В. И.* **Теория катастроф.**

*Климонтович Ю. Л.* **Турбулентное движение и структура хаоса.**

*Безручко Б. П. и др.* **Путь в синергетику. Экскурс в десяти лекциях.**

*Данилов Ю. А.* **Лекции по нелинейной динамике. Элементарное введение.**

*Трубецков Д. И.* **Введение в синергетику. В 2 кн.: Колебания и волны; Хаос и структуры.**

*Князева Е. Н., Курдюмов С. П.* **Основания синергетики. Кн. 1, 2.**

*Князева Е. Н., Курдюмов С. П.* **Синергетика: нелинейность времени и ландшафты коэволюции.**

*Быков В. И.* **Моделирование критических явлений в химической кинетике.**

*Чумаченко Е. Н. и др.* **Сверхпластичность: материалы, теория, технологии.**

*Редько В. Г.* **Эволюция, нейронные сети, интеллект.**

*Чернавский Д. С.* **Синергетика и информация (динамическая теория информации).**

*Баранцев Р. Г.* **Синергетика в современном естествознании.**

*Баранцев Р. Г. и др.* **Асимптотическая математика и синергетика.**

*Пригожин И.* **От существующего к возникающему.**

*Пригожин И., Стенгерс И.* **Время. Хаос. Квант. К решению парадокса времени.**

*Пригожин И., Стенгерс И.* **Порядок из хаоса. Новый диалог человека с природой.**

*Пригожин И., Николис Г.* **Познание сложного. Введение.**

*Пригожин И., Гленсдорф П.* **Термодинамическая теория структуры, устойчивости и флуктуаций.**

*Суздаев И. П.* **Нанотехнология: физико-химия нанокластеров, наноструктур и наноматериалов.**

**Тел./факс:**

**(499) 135-42-46,  
(499) 135-42-16,**

**E-mail:**

**URSS@URSS.ru**

**http://URSS.ru**

**Наши книги можно приобрести в магазинах:**

«Библио-Глобус» (м. Лубянка, ул. Мясницкая, 6. Тел. (495) 625-2457)

«Московский дом книги» (м. Арбатская, ул. Новый Арбат, 8. Тел. (495) 203-8242)

«Молодая гвардия» (м. Полянка, ул. Б. Полянка, 28. Тел. (495) 238-5001, 780-3370)

«Дом научно-технической книги» (Ленинский пр-т, 40. Тел. (495) 137-6019)

«Дом книги на Ладжской» (м. Бауманская, ул. Ладжская, 8, стр. 1. Тел. 267-0302)

«Гнозис» (м. Университет, 1 гум. корпус МГУ, комн. 141. Тел. (495) 939-4713)

«У Нектавра» (РГГУ) (м. Новослободская, ул. Чаюнова, 15. Тел. (499) 973-4301)

«СПб. дом книги» (Невский пр., 28. Тел. (812) 448-2355)



## Уважаемые читатели! Уважаемые авторы!

Наше издательство специализируется на выпуске научной и учебной литературы, в том числе монографий, журналов, трудов ученых Российской академии наук, научно-исследовательских институтов и учебных заведений. Мы предлагаем авторам свои услуги на выгодных экономических условиях. При этом мы берем на себя всю работу по подготовке издания — от набора, редактирования и верстки до тиражирования и распространения.



URSS

Среди вышедших и готовящихся к изданию книг мы предлагаем Вам следующие:

- Галимов Э. М. Феномен жизни. Происхождение и принципы эволюции.  
 Галимов Э. М. (ред.) Проблемы зарождения и эволюции биосферы.  
 Гнатик Е. Н. Генетика человека: Былое и грядущее.  
 Гнатик Е. Н. Генетическая инженерия человека: Вызовы, проблемы, риски.  
 Назаров В. И. Эволюция не по Дарвину: Смена эволюционной модели.  
 Яшин А. А. Живая материя: Онтогенез жизни и эволюционная биология.  
 Яшин А. А. Живая материя: Физика живого и эволюционных процессов.  
 Яшин А. А. Живая материя: Ноосферная биология (нообиология).  
 Мюллер Ф. За Дарвина.  
 Заренков Н. А. Семиотическая теория биологической жизни.  
 Заренков Н. А. Биосимметрия.  
 Фридман М. В., Фридман В. С. Логика для биологов.  
 Шлегель Г. Г. История микробиологии.  
 Мечникова О. Н. Жизнь Ильи Ильича Мечникова.  
 Смит Дж. Математические идеи в биологии.  
 Джермен М. Количественная биология в задачах и примерах.  
 Блюменфельд Л. А. Решаемые и нерешаемые проблемы биологической физики.  
 Дробышевский С. В. Эволюция мозга человека.  
 Дробышевский С. В. Предшественники. Предки? Палеоантропы.  
 Дробышевский С. В. Предшественники. Предки? Архантропы. Гоминиды, переходные от архантропов к палеантропам.  
 Косенко Е. А., Каминский Ю. Г. Клеточные механизмы токсичности аммиака.  
 Сороко Э. М. Золотые сечения, процессы самоорганизации и эволюции систем.  
 Алексеев В. И., Каминский В. А. Прикладная молекулярная биология.  
 Гамов Г., Ичас М. Мистер Томпкинс внутри себя: Приключения в новой биологии.  
 Баксанский О. Е. (ред.) Методология биологии: новые идеи.  
 Новиков Г. Г. Рост и энергетика развития костистых рыб в раннем онтогенезе.  
 Малиновский А. А. Тектология. Теория систем. Теоретическая биология.  
 Рапопорт Г. Н., Герц А. Г. Искусственный и биологический интеллект.  
 Голицын Г. А., Петров В. М. Информация и биологические принципы оптимальности.  
 Галимова М. Х. Ферментативная кинетика: Справочник по механизмам реакций.  
 Бейтсон Г. Разум и природа: неизбежное единство. Пер. с англ.  
 Бейтсон Г. Шаги в направлении экологии разума. Кн. 1–3. Пер. с англ.

По всем вопросам Вы можете обратиться к нам:  
 тел./факс (499) 135-42-16, 135-42-46  
 или электронной почтой [URSS@URSS.ru](mailto:URSS@URSS.ru)  
 Полный каталог изданий представлен  
 в интернет-магазине: <http://URSS.ru>

Научная и учебная  
литература

## Об авторе

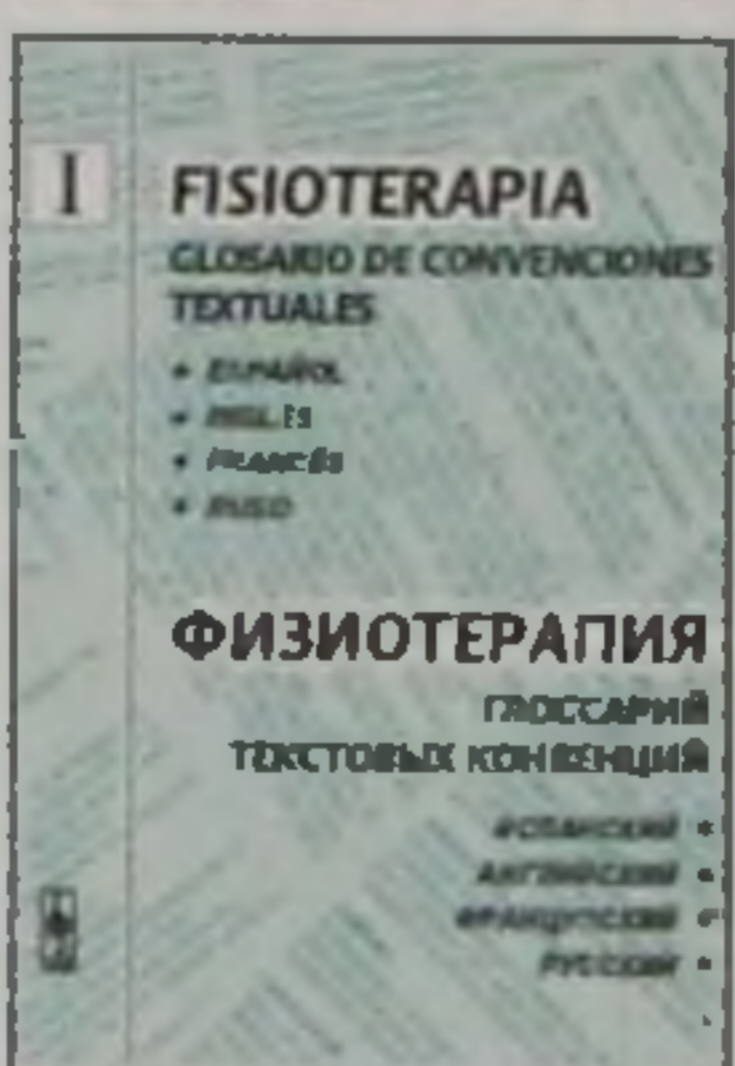
## Вадим Васильевич ПОПОВ



Закончил биологический факультет МГУ им. М. В. Ломоносова и специализировался по экспериментальной зоологии. Русский естествоиспытатель, профессор кафедры генетики и разведения животных Российского государственного аграрного заочного университета, почетный работник высшего образования Российской Федерации, действительный член Международной академии экологии и природопользования, Московского общества испытателей природы (МОИП), член-учредитель Всероссийского общества генетиков и селекционеров им. Н. И. Вавилова (ВОГИС).

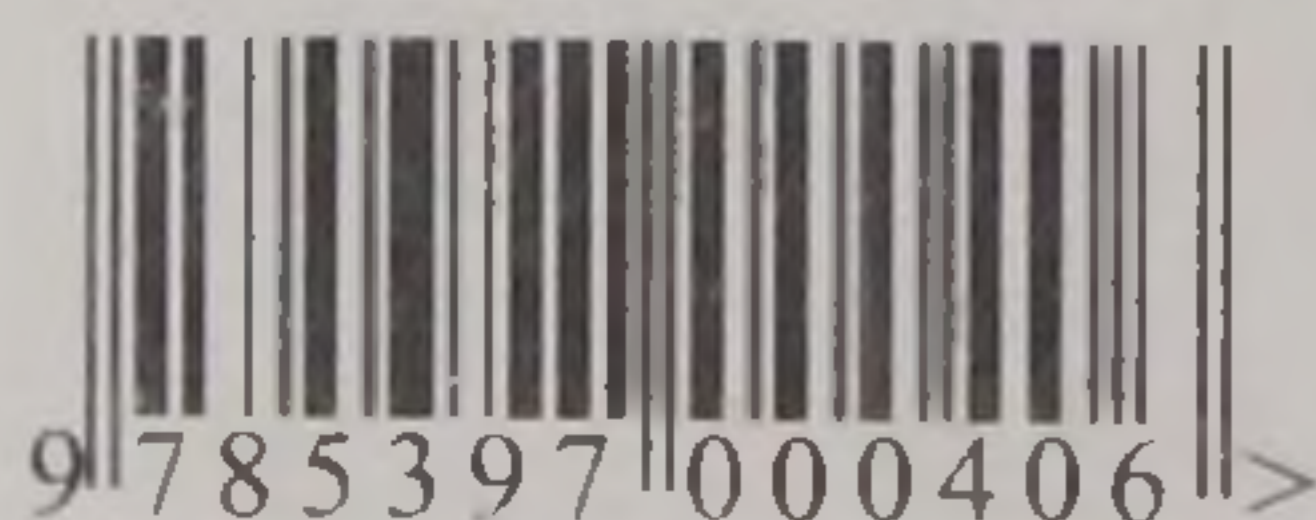
После окончания МГУ работал стажером в Институте общей генетики им. Н. И. Вавилова, научным сотрудником ВНИИ прикладной молекулярной биологии и генетики ВАСХНИЛ, НИИ лаборатории экспериментально-биологических моделей АМН СССР. Участник международных генетического, эмбриологического и ботанического конгрессов. Последние двадцать лет занимается преподавательской и исследовательской работой в вузе. Область научных интересов — биология развития, цитогенетика, популяционная генетика. Разработал генетико-экологическую типологию популяций. Автор более 100 научных трудов, в том числе учебных пособий для студентов. Им переведены с английского языка фундаментальные монографии по генетике зарубежных авторов.

Наше издательство предлагает следующие книги:



5892 ID 73179

НАУЧНАЯ И УЧЕБНАЯ ЛИТЕРАТУРА



Тел./факс: 7 (499) 135-42-16  
Тел./факс: 7 (499) 135-42-46



URSS

E-mail:  
URSS@URSS.ru  
Каталог изданий  
в Интернете:  
<http://URSS.ru>

Любые отзывы о настоящем издании, а также обнаруженные опечатки присылайте по адресу [URSS@URSS.ru](mailto:URSS@URSS.ru). Ваши замечания и предложения будут учтены и отражены на web-странице этой книги в нашем интернет-магазине <http://URSS.ru>