

Лекционный курс

Макаров Л.М.

Генные конструкции организма человека и животных

САНКТ ПЕТЕРБУРГ 2015

Генные конструкции организма человека и животных

курс		Экз.	Лекции	л/раб.	Практик	Консультации	
2	РБМ 31	3,5	20	14	16		

СОДЕРЖАНИЕ

Введение.....	4
1. Уровни организации живой природы.....	5
Понятие: живое и неживое	5
Генетика	8
Основные понятия геномики	10
Размножение ДНК (репликация)	19
Перезапись генетического текста и перевод в белковый текст	20
Транскрипция и трансляция	20
Анализ суммарной ДНК — новые сведения о структуре генома человека	29
Анализ индивидуальных элементов ДНК (генная инженерия).....	32
Определение полной последовательности нуклеотидов в ДНК человека.....	35
Секвенирование простых геномов.....	37
Правила образования информационных пакетов.....	40
Гены построены из кусков	42
Перекрытие генных текстов	45
Избирательное кодирование	47
Генная матрица	48
Генные семейства	48
Генные конструкции - текстовые конструкции.....	51
Количество генов у человека	51
Информационная оценка биосинтеза.....	53
ДНКовый текст человека перенасыщен повторами.....	53
Перевёртыши (обращенные повторы).....	54
Простые тандемные повторы (сателлиты).....	55
Диспергированные повторы.....	56
Диспергированные повторы умеют «перемещаться».....	57
Семейство повторов по имени LINE	58
Семейство повторов по имени Alu	58
Вирусы — составная часть генома человека.....	59
Нуклеотидные последовательности — «близнецы».....	63
«Эгоистичная» ДНК.....	63
Концы молекул ДНК — объект повышенного внимания	65
«Белые пятна» генома	67
Опечатки в тексте из ДНК.....	67
Молекулярные часы	70
Хромосомы - отдельные части целого	71
Хромосома 1	74
Хромосома 2	74
Хромосома 3	74
Хромосома 4	74
Хромосома 5	74
Хромосома 6	75
Хромосома 7	75
Хромосома 8	75
Хромосома 9	75
Хромосома 10	75
Хромосома 11	75

Хромосома 12	76
Хромосома 13	76
Хромосома 14	76
Хромосома 15	76
Хромосома 16	76
Хромосома 17	77
Хромосома 18	77
Хромосома 19	77
Хромосома 20	77
Хромосома 21	77
Хромосома 22	78
Хромосома X	79
Y-хромосома	79
Митохондриальный геном	79
Функциональная геномика	82
Микрочипы — новый рубеж в исследовании экспрессии генома	85
Моделирование — важный подход к пониманию функции генов	89
Протеомика	90
Медицинская геномика	94
2. Заключительные понятия по дисциплине	101
2.1 Клеточный уровень организации живой природы	101
2.2. Типы мутаций	104
Литература	105

Введение

В Природе живые организмы представлены в большом разнообразии. С течением времени изучение биологических и биофизических основ развития живых организмов становится недостаточным. Возникает естественная потребность узнать особенности возрастных изменений, особенности проявления приспособительных реакций живых организмов в среде обитания. Подобных вопросов сегодня возникает очень много. Одни из них затрагивают темы жизнедеятельности, простые ответы на которые можно получить посредством наблюдения за естественными процессами в природе. А другие вопросы, которых становится все больше на пути познания, вскрывают потребность изучения информационного базиса живых организмов.

Наличие хорошо отлаженной природной программы формирования живой материи поражает масштабностью и деталями исполнения очень сложных процессов. С этой точки зрения открытие генома, о котором так много говорили в сравнительно недалеком прошлом, становится решающим этапом в понимании творческих планов Природы.

Геномные конструкции, иначе образно говоря - программные модули работают от рождения живого организма. Учитывая наличие внешних проявлений среды обитания и, естественно, возможности существования конкретного вида живого, такие модули постоянно возобновляют тканевые структуры, организуют работу сенсорных систем, запускают сложные поведенческие процессы и целый ряд других работ, для которых важное значение имеет четкое управление.

С позиции кибернетики, как науки управления, важное значение приобретает пласт знаний заложенный Природой. Управление в живом организме строится по набору процедур. Одни процедуры имеют постоянный ритм исполнения, где результат многократно тиражируется в процессе жизнедеятельности. Другие программные модули реализуются строго в определенное время развития организма. Это так называемые биологические часы. Так происходит запуск физиологических процедур, психофизиологических реакций на факторы среды.

С точки зрения практик сегодня важно не только уметь фиксировать эти процессы, устанавливать генные альянсы - ответственные за те или иные программные модули, но и научиться понимать эти процессы.

В представленном материале для изучения в лаконичном формате изложения, адаптированном к техническому образовательному потенциалу, представлены описания сложных естественных процессов с участием генных структур. Ознакомление с такими процессами представлено в том объеме материал, который позволяет сформировать устойчивое информационное представление о геноме, как наборе программных модулей.

Следуя установившимся представлениям о программном модуле, содержащем отдельные исполнительные процедуры и циклы, формируется представление о хромосомном наборе, фрагментах ДНК. По современным представлениям - исполнительные процедуры генома, подобны тексту, обычному письменному тексту. Следуя этим представлениям многочисленные исследования генома проводятся с использованием разных по сложности информационных моделей. Реализацию этой задачи на практике проводят специалисты, знающие технологии выделения фрагментов ДНК, умеющие проводить анализ, столь важный при формировании диагноза и назначении терапевтических процедур.

1. Уровни организации живой природы

Выделяют 8 уровней.

Каждый уровень организации характеризуется определенным строением (химическим, клеточным или организменным) и соответствующими свойствами.

Каждый следующий уровень обязательно содержит в себе все предыдущие.

1. Молекулярный уровень организации живой природы

Химический состав клеток: органические и неорганические вещества, Обмен веществ (метаболизм): процессы диссимиляции и ассимиляции, поглощение и выделение энергии.

Молекулярный уровень затрагивает все биохимические процессы, которые происходят внутри любого живого организма — от одно- до многоклеточных.

Этот уровень сложно назвать «живым». Это скорее «биохимический» уровень — поэтому он является основой для всех остальных уровней организации живой природы. Поэтому именно он лег в основу классификации Живой природы на царства — какое питательное вещество является основным у организма: у животных — белок, у грибов — хитин, у растений это- углеводы.

Понятие: живое и неживое

Биология — это наука о живых организмах «науку о живых организмах», поэтому прежде всего надо определить такие понятия, как «живое» и «неживое».

Вообще, граница между живым и неживым достаточно условна. На эту тему есть много теорий и мнений...

Для начала определимся, что такое “организм”.

Это просто. Организм — это любой живой объект: животное, растение, даже бактерия!

А “неживой”? Это скалы, камни, вода...

А вирусы? Вирусы могут быть в активной — живой форме, а могут в виде кристаллов находиться под землей или в ледниках сотни лет...

В таком случае надо определить, что такое “жизнь”.

1. Почти все организмы состоят из клеток. Каждая клетка — “кирпичик” живого организма, и то как организм живет — скоординированные функции всех его клеток.

Организмы могут быть одноклеточные — некоторые водоросли или бактерии, или многоклеточные — человек, например.

А вот вирусы не считаются клеточной формой жизни. Они легко встраиваются в чужие клетки, а могут вообще кристаллизоваться и веками в таком состоянии находиться.

2. Обмен веществ

По научному — метаболизм. Во всех живых организмах происходит обмен веществ.

В животных и растениях — само собой; у вирусов, когда они в активной форме, обмен веществ тоже есть; а вот у воды и камней — нет. (Есть теория, что у камней тоже есть обмен веществ, но очень медленный)

3. Гомогенность

Живые организмы не гомогенны, они состоят из различных частей, выполняющих различные функции; т.е. сложно организованны

Неживые системы гомогенны — у воды, например, состав не меняется.

Основные признаки живых организмов

- обмен веществ (это самый главный признак)
- клеточное строение
- негомогенность строения

Организм можно считать живой системой, если у него присутствуют все три характеристики. В противном случае он считается неживым:

Живые организмы	Неживые системы
Бактерии Растения Животные Грибы	Вирусы

Справка

*Вирус (лат. *virus* — «яд») — неклеточный инфекционный агент, который может воспроизводиться только внутри живых клеток. Вирусы поражают все типы организмов, от растений и животных до бактерий и архей (вирусы бактерий обычно называют бактериофагами¹). Обнаружены также вирусы, поражающие другие вирусы (вирусы-сателлиты).*

Вирусы-сателлиты — вирусные частицы, неспособные строить капсиды самостоятельно. Они инфицируют клетки, для которых несвойственна естественная смерть от старости (например, амёбы, бактерии).

Когда клетку, зараженную вирусом-сателлитом, заражает обычный вирус, то вирус-сателлит нарушает производство вирионов этого вируса и сам начинает размножаться. Вирусы-сателлиты, по сути, являются гиперпаразитами.

Несколько слов об амёбе.

Амёбы (лат. *Amoeba*, от греч. *αμοιβή* — «изменение») — род микроскопических одноклеточных организмов. У амёб неправильная, все время меняющаяся форма. Передвигается при помощи ложноножек (псевдоподий), постоянно возникающих и исчезающих.

¹ Фаголизис – процесс разрушения фагами бактерий, обычно происходит в пораженном органе при спонтанном выздоровлении. В случае, если самовыздоровление не наступило, организму можно помочь введением соответствующего бактериофага, полученного в условиях фармацевтического предприятия. Бактериофагами называют вирусы бактерий, которые при встрече с чувствительными к ним микробными клетками проникают внутрь и вызывают их растворение (лизис). В переводе с греческого бактериофаг означает «пожиратель бактерий». В природе эти микроорганизмы весьма разнообразны и составляют самую крупную из известных на сегодняшний день групп вирусов. Фаги имеют отличное от вирусов растений, животных и человека строение, более сложное. Они встречаются в атмосферных осадках, в воздухе, в почве, в еде, на предметах, коже, шерсти животных, словом, везде, где встречаются бактерии. В последние годы возрастает интерес к бактериофагам, применяемым в качестве альтернативы антибиотикам. И это не удивительно, и те и другие препараты непосредственно влияют на возбудителей инфекций. Однако антибиотики, как известно, губительно действуют не только на патогенную флору, но и на полезную. Бактериофаги, виды которых различны, действуют избирательно, только на «свою» бактерию, оставляя от нее обломки. Собственно, поэтому ученые и не стали придумывать им названия, ведь проще называть их соответственно бактериям, на которые они воздействуют

Водятся в прудах, во влажной почве, во внутренностях животных. Размножаются бинарным делением. Клетка амёбы имеет тонкую мембрану, большое ядро, питательные и сократительные вакуоли и жировые глобулы. Выделения проходят через сократительную вакуоль. Длина до 0,5 мм.

Амеба отличается большой длиной генома. Так, геном амёбы *Amoeba dubia* состоит из 690 млрд пар нуклеотидов (для сравнения, геном человека состоит из 2,9 млрд пар).

Амебообразную форму имеют и многие другие одноклеточные, в том числе амеба-убийца *Naegleria fowleri*, которая опасна для людей тем, что поедает их мозг. Это совершенно другой организм — она принадлежит к другому классу. Эта амеба водится в теплых водоемах, прикрепившись к водорослям. Она проникает в организм через носовые каналы и питается ольфакторными клетками мозга, вызывая первичный амебный менингоэнцефалит — летальное заболевание (смертность — 97 %). В период с 1995 по 2004 года было зафиксировано 23 случая заражения амебой-убийцей *Naegleria fowleri*.

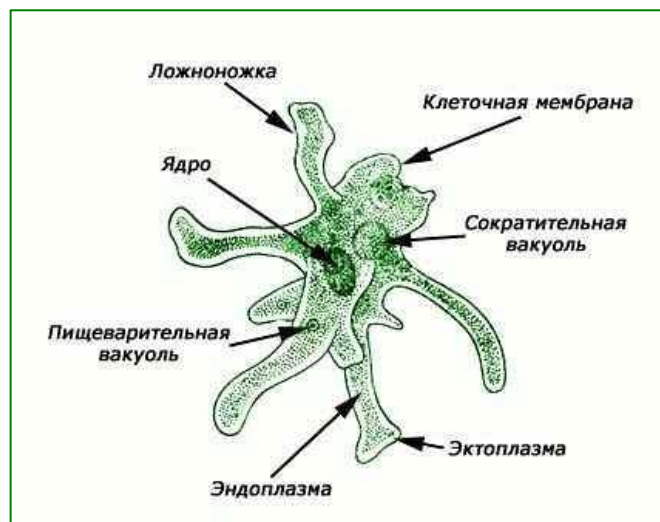


Рисунок 1. Амеба обыкновенная

Заражение амебами возможно воздушно-капельным путем в больницах и лабораториях. Также возможно инфицирование материала для прививок. Такое заражение особенно опасно для грудных детей и детей дошкольного возраста

ВАРИОНЫ - вирусные частицы состоят из двух или трех компонентов:

1. генетического материала в виде ДНК или РНК (некоторые, например мимивирусы, имеют оба типа молекул);
2. белковой оболочки (капсида), защищающей эти молекулы, и, в некоторых случаях, — дополнительных липидных оболочек. Наличие капсида отличает вирусы от вирусоподобных инфекционных нуклеиновых кислот — вироидов. В зависимости от того, каким типом нуклеиновой кислоты представлен генетический материал, выделяют ДНК-содержащие вирусы и РНК-содержащие вирусы;
3. Ранее к вирусам также ошибочно относили прионы, однако впоследствии оказалось, что эти возбудители представляют собой особые инфекционные белки и не содержат нуклеиновых кислот. Форма вирусов варьирует от простой спиральной и икосаэдрической до более сложных структур. Размеры среднего вируса составляют около одной сотой размеров

средней бактерии. Большинство вирусов слишком малы, чтобы быть отчетливо различимыми под световым микроскопом.

Вирусы являются паразитами, так как не способны размножаться вне клетки. Вне клетки вирусные частицы не проявляют признаки живого и ведут себя как частицы биополимеров. От живых паразитарных организмов вирусы отличаются полным отсутствием основного и энергетического обмена и отсутствием сложнейшего элемента живых систем — аппарата трансляции (синтеза белка), степень сложности которого превышает таковую самих вирусов.

Появление вирусов на эволюционном древе жизни неясно: некоторые из них могли образоваться из плазмид, небольших молекул ДНК, способных передаваться от одной клетки к другой, в то время как другие могли произойти от бактерий. В эволюции вирусы являются важным средством горизонтального переноса генов², обуславливающего генетическое разнообразие.

Вирусы распространяются многими способами: вирусы растений часто передаются от растения к растению насекомыми, питающимися растительными соками, к примеру, тлями; вирусы животных могут распространяться кровососущими насекомыми, такие организмы известны как переносчики. Вирус гриппа распространяется воздушно-капельным путем при кашле и чихании.

У животных вирусные инфекции вызывают иммунный ответ, который чаще всего приводит к уничтожению болезнетворного вируса. Иммунный ответ также можно вызвать вакцинами, дающими активный приобретенный иммунитет против конкретной вирусной инфекции. Однако некоторым вирусам, в том числе и возбудителям СПИДа и вирусных гепатитов, удается ускользнуть от иммунного ответа, вызывая хроническую болезнь. Антибиотики не действуют на вирусы, однако было разработано несколько противовирусных препаратов.

Генетика

Генетика человека — раздел генетики, изучающий закономерности наследования и изменчивости признаков у человека, тесно связанный с антропологией и медициной. Эту отрасль условно подразделяют на антропогенетику, изучающую наследственность и изменчивость нормальных признаков человеческого организма, и медицинскую генетику. Генетика человека связана также с эволюционной теорией, так как исследует конкретные механизмы эволюции человека и его место в природе, вместе с психологией, философией и социологией.

Геномика — раздел молекулярной генетики, посвящённый изучению генома и генов живых организмов.

Геномика сформировалась как особое направление в 1980—1990-х гг. вместе с возникновением первых проектов по секвенированию геномов некоторых видов живых организмов. Первым был полностью секвенирован геном бактериофага Ф-Х174 (5 368 нуклеотидов) в 1977 году. Следующим этапным событием было секвенирование генома бактерии *Haemophilus influenzae* (1.8 Mb) (1995). После этого были полностью секвенированы геномы ещё нескольких видов, включая геном человека (2001 год — первый черновой вариант, 2003 год — завершение проекта). Её развитие стало возможно не только благодаря совершенствованию биохимических

² Горизонтальный перенос генов (ГПГ) — процесс, в котором организм передаёт генетический материал другому организму, не являющемуся его потомком. В отличие от горизонтального, о вертикальном переносе генов говорят, если организм получает генетический материал от своего предка. В области интересов генетики основное место занимает вертикальный перенос генов. Однако в настоящее время горизонтальному переносу уделяется всё больше внимания. Искусственный горизонтальный перенос генов используется в генной инженерии.

методик, но и благодаря появлению более мощной вычислительной техники, которая позволила работать с огромными массивами данных. Протяженность геномов у живых организмов подчас измеряется миллиардами пар оснований. Например, объём генома человека составляет порядка 3 млрд пар оснований. Самый крупный из известных (на начало 2010 года) геномов принадлежит одному из видов двоякодышащих рыб (примерно 110 млрд. пар).

Разделы геномики

Структурная геномика

Структурная геномика — содержание и организация геномной информации. Имеет целью изучение генов с известной структурой для понимания их функции, а также определение пространственного строения максимального числа «ключевых» белковых молекул и его влияния на взаимодействия

Функциональная геномика

Функциональная геномика — реализация информации, записанной в геноме, от гена — к признаку.

Сравнительная геномика

Сравнительная геномика (эволюционная) — сравнительные исследования содержания и организации геномов разных организмов.

Получение полных последовательностей геномов позволило пролить свет на степень различий между геномами разных живых организмов. Ниже в таблице представлены предварительные данные о сходстве геномов разных организмов с геномом человека. Сходство дано в процентах (отражает долю пар оснований, идентичных у двух сравниваемых видов).

Вид	Сходство
Человек	99,9 %
	100 %
Шимпанзе	98,4 %
	98,7 %
Бонобо, или карликовый шимпанзе	
Горилла	98,38 %
Мышь	98 %
	85 %
Собака	95 %
Банан	50 %
Нарцисс	35 %

Качественный скачок в этой области произошел благодаря японскому ученому СинъеЯманаке. Он смог вернуть взрослые клетки в состояние плюрипотентности, когда они могут превратиться в разные клетки организма. Их назвали индуцированными плюрипотентными клетками (iPS). «Мы еще не в полной мере осознаем революционность открытия, сделанного Синъей Яманакой в области стволовых клеток, это действительно новая эра», — говорили его коллеги и после вручения ему азиатской «нобелевки» ShawPrize в 2008 году, и после Нобелевской премии в 2012-м, и в конце февраля 2013-го.

Создание индуцированных плюрипотентных клеток журнал Science в конце прошлого года назвал одним из важнейших прорывов последнего десятилетия. В этом же списке еще один знаковый прорыв — технология редактирования генома. Генетические болезни обусловлены поломками в одном или нескольких генах. Ученые научились влезать в ядро клетки и с помощью технологии редактирования генома менять испорченный ген на здоровый. Поскольку здоровый ген должен работать в определенной ткани, наиболее перспективной исследователям представляется комби-

нация клеточного и генетического метода: у пациента берут, к примеру, клетки кожи, перепрограммируют их в iPS, затем в них редактируют геном, а после этого превращают скорректированные клетки, допустим, в нейроны и вводят их в определенную часть мозга.

Эволюция взглядов на терапию в иллюстративной форме представлена на рисунке ниже

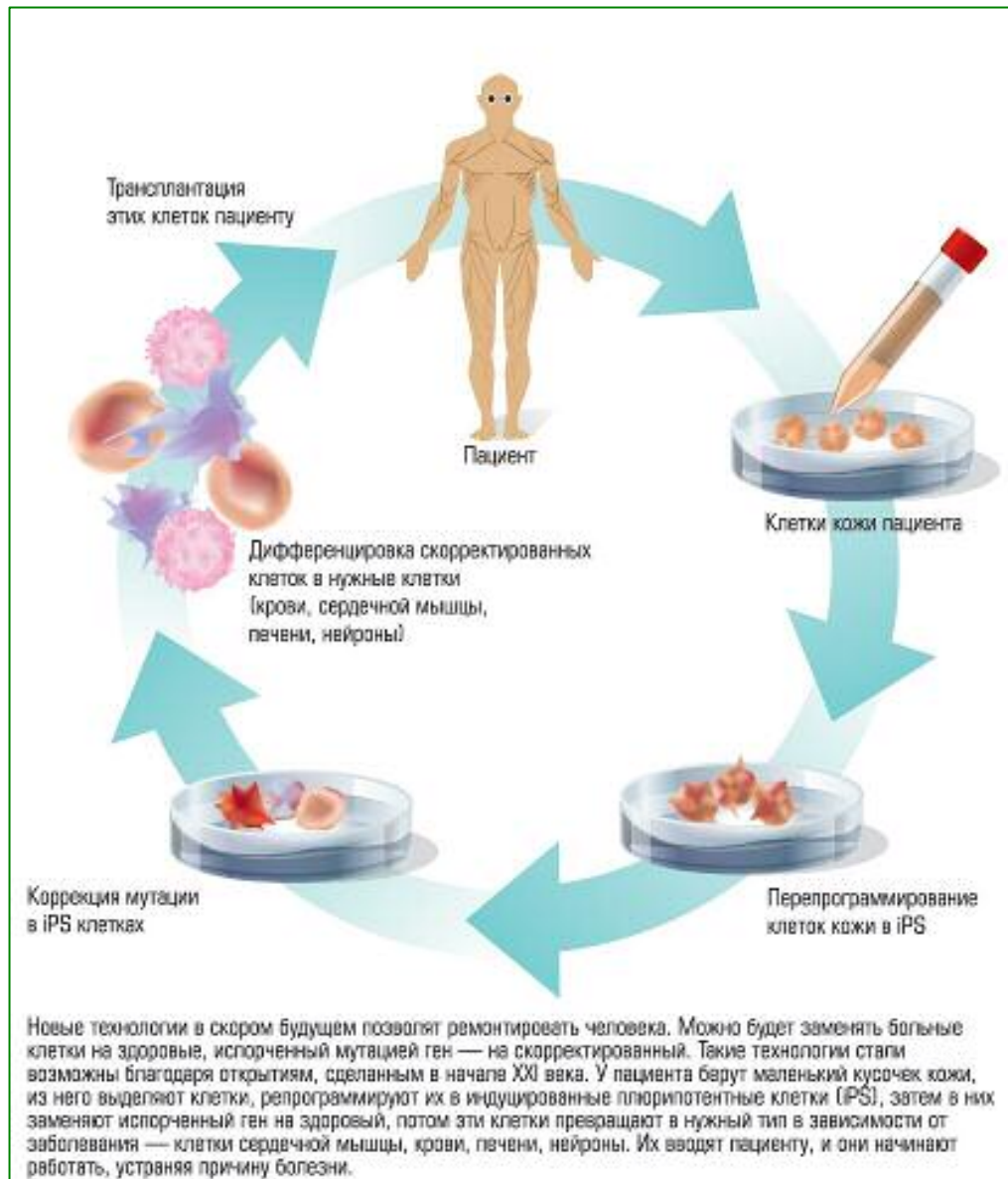


Рисунок 2 Генные технологии в терапии

Основные понятия геномики

Геном

С самого начала определимся, что мы здесь будем подразумевать под словом геном. Сам этот термин впервые был предложен в 1920 году немецким генетиком Г. Винклером. Тогда уже существовал другой научный термин — генотип, введенный в арсенал генетиков В. Иогансеном еще в 1909

году, под которым подразумевалась совокупность всех наследственных задатков данной конкретной клетки или данного конкретного организма. Впоследствии Иогансен сам с удивлением говорил, что его «словечко» неожиданно материализовалось в возникшей позднее хромосомной теории Т. Моргана. Но вот появился новый термин — геном. В отличие от генотипа этот термин должен был стать характеристикой целого вида организмов, а не конкретной особи. И это стало новым этапом в развитии генетики.

В биологическом словаре понятие геном определяется как совокупность генов, характерных для гаплоидного (одинарного) набора хромосом данного вида организмов. Такая формулировка звучит не совсем понятно для неспециалиста, а главное, она неточна в современном понимании этого слова. Основу генома составляет молекула дезоксирибонуклеиновой кислоты, хорошо известная в сокращенном виде как ДНК. Ведь все геномы (ДНК) содержат по крайней мере два вида информации: кодированная информация о структуре молекул-посредников (так называемых РНК) и белка (эта информация содержится в генах), а также инструкции, которые определяют время и место проявления этой информации при развитии и дальнейшей жизнедеятельности организма (эта информация в основном расположена в межгенных участках, хотя частично и в самих генах). Сами гены занимают очень небольшую часть генома, но при этом составляют его основу. Информация, записанная в генах, — это своего рода «инструкция» для изготовления белков, главных строительных кирпичиков нашего тела. «На плечах» генов лежит огромная ответственность за то, как будет выглядеть и работать каждая клетка и организм в целом. Они управляют нашей жизнью от момента зачатия до самого последнего вздоха, без них не функционирует ни один орган, не течет кровь, не бьется сердце, не работают печень и мозг.

Однако для полной характеристики генома недостаточно заложенной в нем информации о структуре белков. Нужны еще данные об элементах генетического аппарата, которые принимают участие в работе (экспрессии) генов, регулируют их проявление на разных этапах развития и в разных жизненных ситуациях.

Но даже и этого мало для полного определения генома. Ведь в геноме присутствуют также элементы, способствующие его самовоспроизведению (репликации), компактной упаковке ДНК в ядре и еще какие-то непонятные пока еще участки, иногда называемые «эгоистичными» (то есть как бы служащими только для самих себя). По всем этим причинам сегодня, когда речь идет о геноме, обычно имеют в виду всю совокупность последовательностей ДНК, представленных в хромосомах ядер клеток определенного вида организмов, включая, конечно, и гены. В этой книге мы будем подразумевать именно такое определение. Вместе с тем следует помнить, что в некоторых других структурах (органеллах) клетки также присутствует генетическая информация, необходимая для функционирования организмов. В частности, у всех животных организмов, в том числе и у человека, имеется еще и митохондриальный геном, то есть молекулы ДНК, присутствующие в таких внутриклеточных структурах, как митохондрии, и содержащие ряд так называемых митохондриальных генов. Митохондриальный геном человека очень небольшой по сравнению с ядерным геномом, расположенным в хромосомах, но, тем не менее, его вклад в клеточный метаболизм весьма существенен.

В отличие от многих других биологических наук генетика с момента своего возникновения стремилась быть точной наукой. И вся история генетики — это история создания и использования в эксперименте все более и более точных методов и подходов, что сближает ее с такими точными науками, как физика, химия и математика.

А начало всему было положено чешским монахом Грегором Менделем, который в 1865 году опубликовал свой фундаментальный труд с математическими расчетами, указывающими на существование неких абстрактных дискретных частиц, передающих наследственные свойства («частицы наследственности»), названных позднее генами. Эта феноменальная работа Менделя, осуществленная на горохе, не произвела особого эффекта на его современников и была забыта вплоть до 1900 года (научная мысль в то время еще не созрела для ее восприятия). Сам Мендель после неудачных попыток получить аналогичные результаты при скрещивании других растений прекратил

опыты и до конца жизни занимался садоводством, пчеловодством и метеорологическими наблюдениями.

Лишь спустя 35 лет произошло то, что и должно было случиться: законы Менделя были переоткрыты независимо и одновременно тремя разными исследователями (Г. де Фризом, Э. Чермаком и К. Корренсом), после чего и начала интенсивно развиваться наука, получившая позднее название генетика. С тех пор эта наука и рожденные на ее основе молекулярная генетика и геномика занимают лидирующее положение среди прочих наук о природе, став в итоге одними из основных, определяющих сегодняшний и завтрашний день развития человечества.

Вскоре после открытия основных законов генетики было установлено, что маленькие продолговатые тельца, наблюдаемые под микроскопом в ядрах клеток, которые были названы хромосомами, ведут себя именно так, как это ожидалось от «единиц наследственности» Менделя. Но уже тогда было ясно, что число генов должно быть больше, чем число хромосом. В 1910 году Томас Хант Морган начал изучать относительно простой и удобный для анализа генетический аппарат плодовой мушки дрозофилы, что привело в конечном итоге к созданию хромосомной теории наследственности. Согласно этой теории, существуют многочисленные гены, которые линейно расположены в хромосомах, и их последовательность в будущем может быть расшифрована.

Известно, что гены управляют развитием любого живого организма с момента его рождения и до смерти. Гены достаются нам от родителей, и от них в значительной мере зависят наши физические параметры, внешность, склонность к различным заболеваниям или, наоборот, своего рода иммунитет к ним. При этом следует обратить внимание на то, что такие черты, как характер, убеждения, привычки, поведение и даже способности также определяются в значительной мере генетически, хотя здесь существенную роль могут играть и социальные факторы, такие, как условия жизни, воспитание, образование, окружение.

Лишь в 40-е годы прошлого века была установлена материальная основа генов. Выяснилось, что ею служит одна из так называемых нуклеиновых кислот, а именно дезоксирибонуклеиновая кислота (сокращенно, ДНК). Само существование нуклеиновых кислот было обнаружено швейцарским биохимиком Ф. Мишером еще в 1868 году, то есть всего через три года после открытия Менделем своих законов (случайность или закономерность?). Тогда из спермы лосося Мишер выделил фосфорсодержащее вещество, происходящее из клеточных ядер, которое он назвал нуклеином (от слова нуклеус — ядро), а мы теперь его называем дезоксирибонуклеиновой кислотой. Примечательно, что два таких важных открытия, как обнаружение единицы наследственности и ее физического носителя, были сделаны почти одновременно. Однако, как и в случае с законами Г. Менделя, практически никто из исследователей в то время не смог оценить важность открытия Мишера. В дальнейшем существенный вклад в изучение нуклеиновых кислот внесли немецкий химик Альбрехт Кёссель и американский биохимик русского происхождения А. Ф. Левин. Первый установил, что в состав нуклеина входят четыре азотсодержащих вещества (их назвали азотистыми основаниями): аденин (А), гуанин (Г), цитозин (Ц) и тимин (Т) (за это в 1910 г. Кёссель получил Нобелевскую премию по физиологии и медицине). Затем А. Ф. Левин показал, что в состав нуклеина, кроме тетрады А, Г, Ц и Т, входит вдобавок к фосфорной кислоте еще и сахар дезоксирибоза, то есть «рибоза без кислорода». Так и возникло название дезоксирибонуклеиновая кислота (ДНК). Позднее в составе другой нуклеиновой кислоты — рибонуклеиновой кислоты (РНК) — была обнаружена рибоза вместо дезоксирибозы.

На этом этапе предложена одна из первых гипотез о структуре нуклеиновых кислот. Согласно этой гипотезе, нуклеиновые кислоты построены как линейная комбинация связанных друг с друга химической связью нуклеотидов. По мнению Левина, четыре разных нуклеотида, входящие в состав нуклеиновых кислот, связаны последовательно в стандартный тетра-нуклеотид, который многократно повторяется в структуре нуклеиновой кислоты. И многие исследователи приняли эту гипотезу на веру. Однако столь однообразная и монотонная последовательность не годилась на роль материальной структуры генов. По этой причине долгое время бытовало мнение, что ДНК выполняет какую-то чисто структурную функцию в хромосомах.

Серия открытий, которые привели к современному пониманию генетической важности ДНК и ее основополагающей роли в организации хромосом, началась в середине 20-х годов прошлого века, когда Л. Зильберт на протях и Ф. Гриффит на пневмококках описали опыты по серологической трансформации. В 1944 году американский биолог Освальд Теодор Эвери с соавторами в опытах с пневмококками показали, что с помощью чистого препарата ДНК могут быть специфически изменены их наследственные свойства. Однако даже эти безупречные результаты не убедили ученых полностью в том, что ДНК — это вещество наследственности. Они заставили ученых только усомниться в том, что ДНК играет чисто структурную роль в хромосомах. Продолжала господствовать теория белковой природы гена. И такая ситуация сохранялась вплоть до апреля 1953 года.

Обратимся к рассмотрению молекулы ДНК. Итак, ДНК состоит из 4-х азотистых оснований, а также сахара (дезоксирибозы) и фосфорной кислоты. Два азотистых основания (сокращенно называемых Ц и Т) относятся к классу так называемых пиримидиновых основания, а два других (А и Г) — к пуриновым основаниям. Такое разделение связано с особенностями их структур, которые показаны на рис. 3.

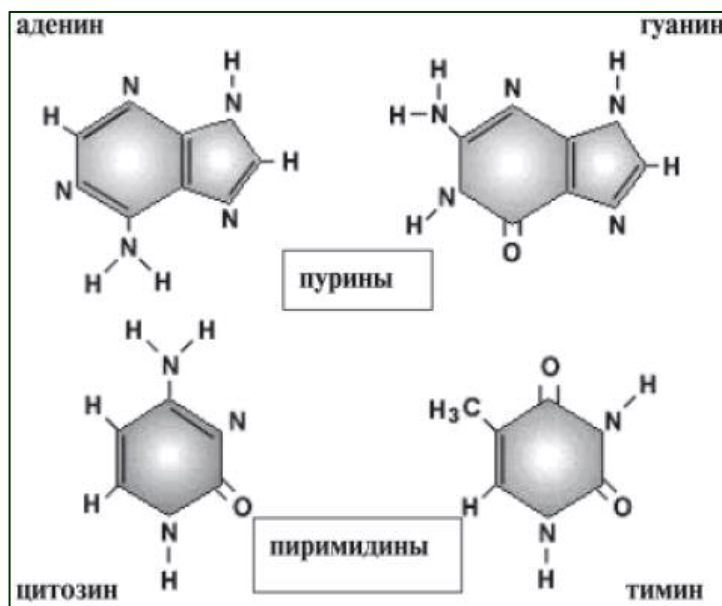


Рис. 3. Структура азотистых оснований (элементарных «букв»), из которых построена молекула ДНК

Отдельные основания связаны в цепочке ДНК сахаро-фосфатными связями. Эти связи изображены на рисунке (рис. 4).

Все это известно уже довольно давно. Но детальное устройство молекулы ДНК стало понятно лишь спустя почти 90 лет после знаменитых работ Менделя и открытия Мишера. 25 апреля 1953 г. в английском журнале «Nature» было опубликовано небольшое письмо молодых и тогда еще мало известных ученых Джеймса Уотсона и Френсиса Крика редактору журнала. Оно начиналось словами: «Мы хотели бы предложить свои соображения по поводу структуры соли ДНК. Эта структура имеет новые свойства, которые представляют большой биологический интерес». Статья содержала всего около 900 слов, но — и это не преувеличение — каждое из них оказалось на вес золота.

А началось все так. В 1951 году на симпозиуме в Неаполе американец Джеймс Уотсон встретился с англичанином Морисом Уилкинсом. Конечно же, они тогда не могли себе даже представить, что в результате этой встречи они станут нобелевскими лауреатами. В то время Уилкинс со своей коллегой Розалиндой Франклин проводили в Кембриджском университете рентгеноструктурный анализ ДНК и определили, что молекула ДНК представляет собой, скорее всего, спираль. После раз-

говора с Уилкинсом Уотсон «загорелся» и решил заняться исследованием структуры нуклеиновых кислот. Он перебрался в Кембридж, где познакомился с Френсисом Криком. Ученые решили совместными усилиями попытаться понять, как устроена ДНК. Работа началась не на пустом месте. Исследователи уже знали о существовании двух типов нуклеиновых кислот (ДНК и РНК), знали и то, из чего они состоят. В их распоряжении были фотографии рентгеноструктурного анализа, полученные Р. Франклином. Кроме того, Эрвин Чаргафф сформулировал к тому времени очень важное правило, согласно которому в ДНК число А всегда равно числу Т, а число Г равно числу Ц. А далее сработала «игра ума». Результатом этой «игры» и стала статья в журнале «Nature», в которой Дж. Уотсон и Ф. Крик описали созданную ими теоретическую модель строения молекулы ДНК. Согласно их «научной фантазии», основанной тем не менее на определенных твердо установленных фактах, молекула ДНК должна состоять из двух гигантских полимерных цепочек. Звенья каждого полимера состоят из нуклеотидов: углевода дезоксирибозы, остатка фосфорной кислоты и одного из 4 азотистых оснований (А, Г, Т или Ц).

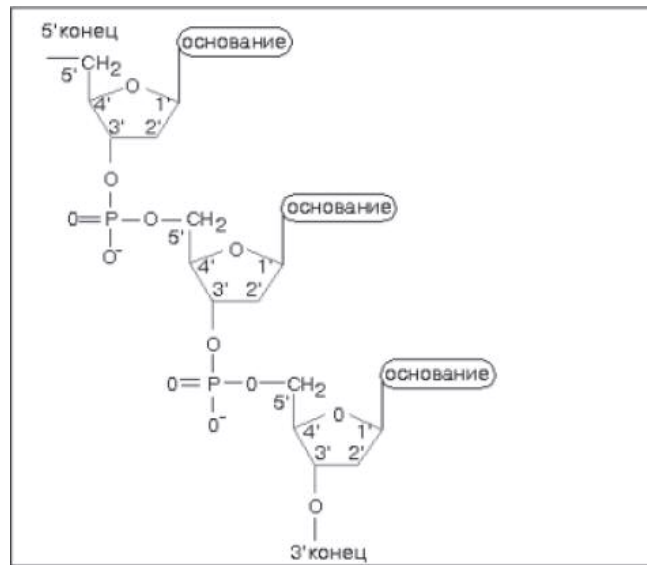


Рис. 4. Химическая структура цепи ДНК

Последовательность звеньев в цепочке может быть любой, но эта последовательность строго связана с последовательностью звеньев в другой (парной) полимерной цепочке: напротив А должно быть Т, напротив Т должно быть А, напротив Ц должно быть Г, а напротив Г должно быть Ц (правило комплементарности) (рис. 5).

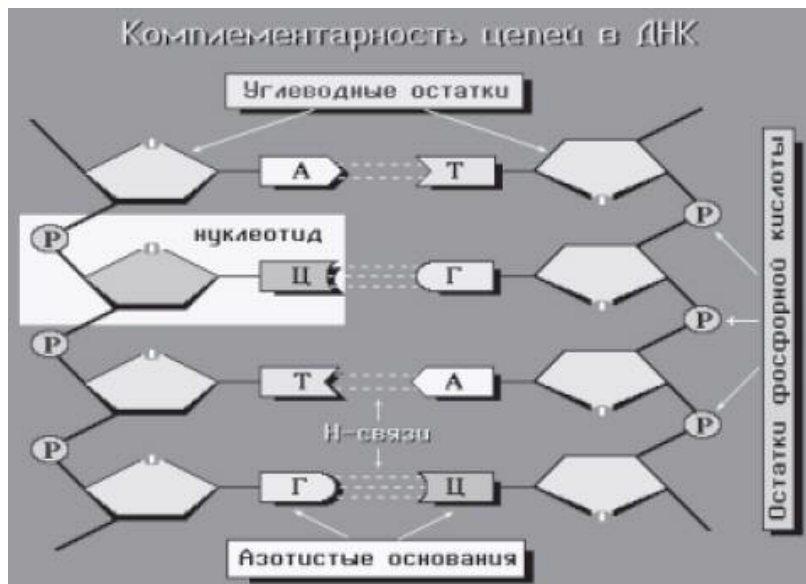


Рис. 3. Схема взаимодействия двух комплементарных цепей в молекуле ДНК

Две полимерные цепи закручены в правильную двойную спираль. Они удерживаются вместе посредством водородных связей между парами оснований (А–Т и Г–Ц) подобно ступенькам лестницы. По этой причине говорят, что две цепи ДНК комплементарны. Для природы это не удивительно. Известно множество примеров комплементарности. Комплементарны, например, древнекитайские символы «инь» и «янь», гнезда розетки и штырьки вилки.

Двойная спираль ДНК схематически изображена на рис. 4. Внешне она напоминает веревочную лестницу, завитую в правую спираль. Ступенями в этой лестнице являются пары нуклеотидов, а связывающие их «боковинки» состоят из сахаро-фосфатного остова.

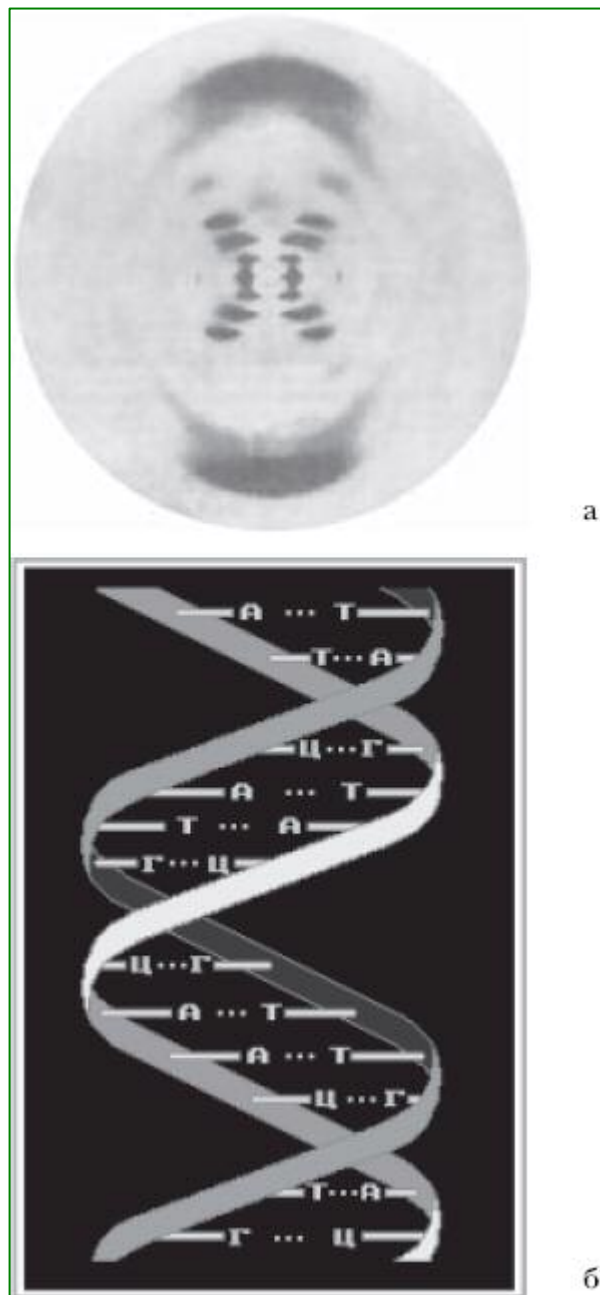


Рис. 4. Знаменитая двойная спираль ДНК а — Рентгенограмма ДНК, полученная Р. Франклин, которая помогла Уотсону и Крику найти ключ к двухспиральной структуре ДНК; б — Схематическое изображение двухспиральной молекулы ДНК

Так была открыта знаменитая «двойная спираль». Если последовательность звеньев (нуклеотидов) в ДНК рассматривать как ее первичную структуру, то двойная спираль — это уже вторичная структура ДНК. Предложенная Уотсоном и Криком модель «двойной спирали» изящно решала не только проблему кодирования информации, но и удвоения (репликации) гена.

В 1962 году Дж. Уотсон, Ф. Крик и Морис Уилкинс получили по достоинству за это достижение Нобелевскую премию. А ДНК была названа самой главной молекулой живой природы. Во всем этом, конечно же, сыграли свою роль точные сведения о строении ДНК, но не в меньшей мере и «провидческие» построения сложной пространственной структуры, что потребовало от исследователей не только логики, но и творческого воображения — качества, присущего художникам, писателям и поэтам. «Здесь, в Кембридже, произошло, быть может, самое выдающееся после книги Дарвина событие в биологии — Уотсон и Крик раскрыли структуру гена!» — писал в то время в Копенгаген Нильсу Бору его бывший ученик М. Дельбрюк. Известный испанский художник Сальвадор Дали после открытия двойной спирали сказал, что это для него явилось доказательством существования Бога, и изобразил ДНК на одной из своих картин.

Итак, интенсивный мозговой штурм, предпринятый учеными, завершился полным успехом! В историческом масштабе открытие структуры ДНК сопоставимо с открытием структуры атома. Если выяснение строения атома привело к появлению квантовой физики, то открытие структуры ДНК дало начало молекулярной биологии.

Каковыми же оказались главные физические параметры ДНК человека — этой главной его молекулы? Диаметр двойной спирали равен 2 нанометрам ($1 \text{ нм} = 10^{-9} \text{ м}$); расстояние между соседними парами оснований («ступеньками») составляет 0,34 нм; один поворот спирали состоит из 10 пар оснований. Последовательность пар нуклеотидов в ДНК нерегулярна, но сами пары уложены в молекуле как в кристалле. Это дало основание характеризовать молекулу ДНК как линейный аperiодический кристалл.

Число отдельных молекул ДНК в клетке равно числу хромосом. Длина такой молекулы в наибольшей по размеру хромосоме 1 человека составляет около 8 см. Подобных гигантских полимеров пока не выявлено ни в природе, ни среди искусственно синтезированных химических соединений.

У человека длина всех молекул ДНК, содержащихся во всех хромосомах одной клетке, составляет примерно 2 метра. Следовательно, длина молекул ДНК в миллиард раз больше их толщины. Так как организм взрослого человека состоит примерно из $5 \times 10^{13} - 10^{14}$ клеток, то общая длина всех молекул ДНК в организме равна 1011 км (это почти в тысячу раз больше расстояния от Земли до Солнца). Вот такая она, суммарная ДНК всего лишь одного человека!

Когда говорят о размере генома, то подразумевают общее содержание ДНК в единичном наборе хромосом ядра. Такой набор хромосом называют гаплоидным. Дело в том, что большинство клеток нашего организма содержит двойной (диплоидный) набор совершенно одинаковых хромосом (только у мужчин 2 половые хромосомы отличаются).

Измерения размера генома приводятся в дальтонах, парах нуклеотидов (п. н.) или пикограммах (пг).

Соотношение между этими единицами измерения следующие: $1 \text{ пг} = 10^{-9} \text{ мг} = 0,6 \times 10^{12} \text{ дальтон} = 0,9 \times 10^9 \text{ п. н.}$ (далее мы будем использовать в основном п. н.). В гаплоидном геноме человека содержится около 3,2 млрд. п. н., что равно 3,5 пг ДНК. Таким образом, в ядре одной клетки человека содержится около 7 пг ДНК. Если учесть, что средний вес клетки человека равен примерно 1000 пг, то легко рассчитать, что ДНК составляет менее 1% от веса клетки. И тем не менее, чтобы воспроизвести самым мелким шрифтом (как в телефонных справочниках) ту огромную информацию, которая содержится в молекулах ДНК одной нашей клетки, понадобилось бы тысяча книг по

1000 страниц в каждой! Вот таков полный размер генома человека — Энциклопедии, написанной четырьмя буквами.

Но не следует думать, что геном человека наибольший из всех существующих в природе. Например, у саламандры и лилии длина молекул ДНК, содержащихся в одной клетке, в тридцать раз больше, чем у человека.

Поскольку молекулы ДНК имеют гигантский размер, их можно выделить и увидеть даже в домашних условиях. Вот как описывается эта простая процедура в рекомендации для кружка «Юный генетик». Во-первых, надо взять любые ткани животных или растительных организмов (например, яблоко или кусок курицы). Затем надо нарезать ткань на кусочки и положить 100 г в обычный миксер. После добавления 1/8 чайной ложки соли и 200 мл холодной воды вся смесь взбивается на миксере в течение 15 секунд. Далее взбитая смесь процеживается через ситечко. В полученную мякоть надо добавить 1/6 от ее количества (это будет примерно 2 столовые ложки) моющего средства (для посуды, например) и хорошо размешать. Через 5–10 минут жидкость разливается по пробиркам или любым другим стеклянным емкостям, чтобы в каждой из них было заполнено не больше трети объема. Затем к ней добавляется по чуть-чуть либо сок, выжатый из ананаса, либо раствор, используемый для хранения контактных линз. Все содержимое встряхивается. Делать это надо весьма осторожно, так как если тряхи слишком сильно, то гигантские молекулы ДНК ломаются и после этого ничего нельзя будет увидеть глазами. Далее в пробирку медленно вливается равный объем этилового спирта, чтобы он образовал слой поверх смеси. Если после этого pokrutyть в пробирке стеклянной палочкой, на нее «намотается» вязкая и почти бесцветная масса, которая и представляет собой препарат ДНК.

Генетическая грамматика

После установления химического строения и пространственной структуры ДНК оставалось еще множество вопросов, основной из которых заключался в том, как же ДНК кодирует белки, то есть, что представляет из себя генетический код этой молекулы, какую «грамматику» она использует? На это в первую очередь и были направлены дальнейшие усилия исследователей.

Итак, установлено, что «буквами» в ДНКовом тексте служат нуклеотиды — элементарные звенья полимерной молекулы ДНК. В ДНК всего 4 нуклеотида (А, Т, Г, Ц). Следовательно, если сравнить каждый из этих нуклеотидов с отдельной буквой, то алфавит ДНКового текста содержит всего 4 «буквы». Как же из этих «букв» формируются «слова» и «предложения»?

Белковые молекулы всех существующих на земле организмов построены всего из 20 аминокислот. Сразу после создания модели ДНК стало ясно, что существует некий код, переводящий четырехбуквенный ДНКовый текст в двадцатибуквенный аминокислотный текст. Элементарные расчеты говорили о том, что число возможных сочетаний, в которых четыре нуклеотида могут быть по-разному расположены в «тексте», достигает астрономических значений. Так, молекула ДНК, состоящая, к примеру, всего из 100 пар нуклеотидов, может теоретически кодировать 4100 различных белковых «текстов». Какова же ситуация на самом деле?

Одним из первых в этом пытался разобраться русский физик Г. Гамов, эмигрировавший в то время в Америку. Наслушавшись многочисленных разговоров о ДНК и узнав, что она содержит — как и карты — всего четыре «масти», Гамов решил «разложить пасьянс» с целью понять устройство генетического кода. Ему сразу стало ясно, что код не может быть «двоичным», то есть одну аминокислоту в белке должна кодировать не двойка нуклеотидов — «букв», а как минимум тройка. Дело в том, что сочетание из 4 по 2 дает всего 16 комбинаций, а этого недостаточно для кодирования всех 20 аминокислот. Следовательно, рассуждал Гамов, код должен быть по крайней мере трехбуквенным, то есть каждую аминокислоту должна кодировать тройка «букв» в любых сочетаниях. На этом он и остановился, поскольку далее возникало множество вопросов. В частности, такой: число сочетаний из 4 по 3 равно 64, а аминокислот всего 20. Зачем же такая избыточность в трехбуквенном коде?

В то время уже существовал хорошо известный путь, который, в частности, был проделан в свое время французом Жаном Шампольоном при дешифровке иероглифов древнего Египта. В качестве основного подспорья для решения стоящей перед ним задачи он использовал базальтовую плиту, которую обнаружили во время военной компании Наполеона в Египет и которая получила название Розеттский камень. На плите одновременно присутствовали две надписи: одна была иероглифическая, а другая — сделанная греческими буквами на греческом языке. К счастью, и язык, и письмо древних греков были в то время уже хорошо известны ученым. В результате сравнение двух текстов Розеттского камня привело к расшифровке египетской иероглифики.

Этим путем и двинулись ученые при расшифровке генетического кода. Надо было сравнить два текста: текст, записанный в ДНК, с текстом, записанным в белке. Однако первоначально ученые не умели «читать» ДНК, а одного известного в то время белкового текста было недостаточно. Пришлось искусственно синтезировать разнообразные короткие фрагменты РНК и синтезировать на них в искусственных системах фрагменты белка. Весной 1961 года в Москве на Международном биохимическом конгрессе М. Ниренберг сообщил, что ему удалось «прочитать» первое «слово» в ДНКовом тексте. Это была тройка букв — ААА (в РНК, соответственно, УУУ), то есть три аденина, стоящие друг за другом, — которая кодирует аминокислоту фенилаланин в белке. Так было положено начало расшифровке генетического кода.

Такой путь в конечном итоге вскоре привел к полной расшифровке генетического кода. Подтвердилось предположение Гамова, что код триплетный: одной аминокислоте в белках соответствует последовательность из 3 нуклеотидов в ДНК и РНК. Такие кодирующие тройки нуклеотидов — «слова» — получили название кодонов.

Напомним, что еще Гамов Г.³ столкнулся с парадоксом: из четырех нуклеотидов может быть построено 64 разных кодонов, а для построения белков используется только 20 различных аминокислот. Решение этого парадокса оказалось в следующем. Большинство аминокислот может кодироваться несколькими кодонами. После выяснения этого обстоятельства генетический код назвали вырожденным.

В таблице 1 приведены кодоны, но не в самой ДНК, а в РНК-посреднике (матричной РНК, или мРНК), образующейся на ДНК, и соответствующие им аминокислоты в белках.

Кроме того, как видно из таблицы, реально для кодирования используются не все возможные кодоны. Три из этих «лишних» кодонов выполняют функцию стоп-сигналов, обеспечивая прекращение синтеза белковой цепи.

Если внимательно посмотреть на таблицу 1, то видно, что вырожденность генетического кода носит не совсем случайный характер. Хотя код триплетный, основную нагрузку несут первые два нуклеотида в каждом кодоне. Чаще всего в разных кодонах, кодирующих одну и ту же аминокислоту, отличается лишь третий нуклеотид.

Таблица 1. Генетический словарь. Указаны аминокислоты, встречающиеся в белках, и соответствующие им кодоны в комплементарной ДНК матричной РНК

Генетический код первоначально был расшифрован у таких простых организмов, как фаги и бактерии. В дальнейшем оказалось, что он универсален (за очень редким исключением) для геномов всех существующих ныне живых организмах (от бактерий до человека). Небольшие отличия, о которых мы поговорим далее, были выявлены при сравнении ядерного и митохондриального геномов.

³ Георгий Антонович Гамов (также известен как Джордж Гамов, англ. George Gamow) советский и американский физик-теоретик, астрофизик и популяризатор науки.

Он является автором первой количественной теории альфа-распада, одним из основоположников теории «горячей Вселенной» и одним из пионеров применения ядерной физики к вопросам эволюции звезд. Он впервые четко сформулировал проблему генетического кода. Широкую известность Гамову принесли его научно-популярные произведения, в которых живым и доступным языком рассказывается о современных научных представлениях.

Итак, как в привычном нам тексте книги, вся информация записана в ДНК последовательностью расположения четырех составляющих ее «букв» — нуклеотидов. Таким образом, ДНК-овый текст написан с помощью А, Т, Ц, Г-алфавита. При этом только текст одной из двух цепей ДНК обычно кодирующий, а другая цепь, как правило, не кодирующая. Хотя известно, что в каждом правиле есть исключения. Если читатель попробует написать этими четырьмя буквами какие-нибудь русские слова, то у него ничего не получится. «Словом» в ДНК-овом тексте, условно говоря, служит определенное сочетание трех нуклеотидов, которому соответствует конкретная аминокислота в белке, являющемся также полимером. Таким образом, в клетке четырьмя буквами записано два десятка «слов» (аминокислот — составных частей белков).

Таблица 1 Генетический словарь

Аланин	Аргинин	Аспарагин	Аспарагиновая кислота	Валин
ГЦУ, ГЦА, ГЦЦ, ГЦГ	ЦГУ, ЦГА, ЦГЦ, ЦГГ, АГА, АГГ	ГАУ, ГАЦ	ААУ, ААЦ	ГУУ, ГУЦ, ГУА, ГУГ
Гистидин	Глицин	Глутамин	Глутаминовая кислота	Изолейцин
ЦАУ, ЦАЦ	ГГУ, ГГА, ГГЦ, ГГГ	ГАА, ГАГ	ЦАА, ЦАГ	АУУ, АУА, АУЦ
Лейцин	Лизин	Метионин	Пролин	Серин
УУА, УУГ, ЦУЦ, ЦУГ, ЦУА, ЦУЦ	ААА, ААГ	АУГ	ЦЦГ, ЦЦЦ, ЦЦА, ЦЦУ	АГУ, АГЦ, УЦА, УЦГ, УЦУ, УЦЦ
Тирозин	Треонин	Фенилаланин	Триптофан	Цистеин
УАУ, УАЦ	АЦУ, АЦА, АЦГ, АЦЦ	УУУ, УУЦ	УГГ	УГУ, УГЦ
Нет аминокислоты				
УАА, УАГ, УГА				

И, наконец, как «предложение» в ДНК-овом тексте можно рассматривать полный набор триплетов, кодирующих определенный белок, то есть ген. Таким образом, генетический алфавит состоит всего из 4 букв, а генетический словарь из 20 слов. В этой связи вспомним, что даже словарь Эллочки-людоедки из романа И. Ильфа и Е. Петрова «Двенадцать стульев» состоял из 30 слов, а «Словарь языка произведений А. С. Пушкина» насчитывает примерно 20 тыс. слов.

Существует строгая закономерность: чем длиннее код (чем больше в нем знаков), тем короче тексты. Огромный по размерам код представляют собой, например, китайские иероглифы. В результате этого иероглифические тексты существенно более кратки по сравнению с другими системами письма, в том числе и нашей. Однако для создания генетического кода природа выбрала всего 4 «буквы». Такой код предполагает наличие длинных текстов, что и реализовалось природой в виде создания гигантских молекул ДНК. При написании полного «текста» генома человека потребовалось около 3,2 млрд. «букв».

Для сравнения: в священной книге Бытия, написанной на древнееврейском языке, содержится всего 78 100 букв.

Размножение ДНК (репликация)

Важно то, что структура ДНК, открытая Уотсоном и Криком, многое прояснила относительно разных механизмов функционирования этой молекулы в клетке. ДНК не только кодирует генетическую информацию, но и самовоспроизводится (удваивается) при каждом клеточном делении. И вскоре уже было экспериментально установлено, что одновременно с делением клетки ДНК снимает с самой себя точные копии в процессе удвоения, или репликации. Во время клеточного деления слабые связи между двумя цепями двойной спирали ДНК разрушаются, в результате чего нити разделяются. Затем на каждой из них строится вторая «дочерняя» (комплементарная) цепь ДНК. В результате этого молекула ДНК удваивается, как и клетка, и в обеих клетках оказывается по одной полной копии ДНК. Копии должны быть полностью идентичными, чтобы сохранить всю генетическую информацию.

Процесс репликации играет ключевую роль в сохранении одной и той же генетической информации в разных клетках, образующихся при их делении. Однако реальные механизмы репликации довольно сложны, и до настоящего времени еще не все тонкие детали этого процесса известны, особенно применительно к геномам высших животных организмов, включая человека. В общем виде этот процесс выглядит следующим образом. В каждой хромосоме ДНК удваивается не с начала до конца, а отдельными кусками (репликонами). Средний размер репликона составляет около 30 мкм. Тем самым в составе генома человека должно встречаться более 50 000 репликонов, участков ДНК, которые синтезируются в ядре как независимые единицы. И это имеет свой глубокий смысл. Если бы каждая из молекул ДНК удваивалась как один репликон от начала до конца молекулы, то при скорости синтеза 0,5 мкм в минуту (а она именно такова у человека) удвоение первой хромосомы, имеющей длину ДНК около 7 см, занимало бы 140 000 минут, или около трех месяцев. На самом деле благодаря полирепликонному строению молекул ДНК весь процесс занимает всего 7–12 часов. Отдельные относительно короткие репликоны соединяются друг с другом, обеспечивая этим процесс воспроизведения целой молекулы ДНК.

Перезапись генетического текста и перевод в белковый текст

Транскрипция и трансляция

В клетке ДНК служит в качестве матрицы, на которой первоначально происходит синтез разных РНК. Процесс перезаписи генетической информации из ДНК в РНК-овый текст получил название транскрипция. Этот процесс, как и репликация ДНК, осуществляется в ядрах клеток. Первоначально на генах, кодирующих белки, образуются РНК-предшественники, которые после ряда модификаций превращаются в так называемые матричные РНК (мРНК). Они-то непосредственно и служат матрицей для синтеза белков, то есть их кодируют. Установлено, что мРНК служат не только носителями ДНК-овой информации, но и переносчиками этой информации из ядра в цитоплазму клетки. Только там мРНК может играть роль матрицы для синтеза белковых молекул (этот процесс назван трансляцией). В цитоплазме на специфических «машинах» — рибосомах — осуществляется при трансляции мРНК синтез молекулы белка, т. е. происходит перевод информации с четырехбуквенного языка мРНК на двадцатипятибуквенный язык белка. Схематически этот процесс изображен на рис. 5.

В середине 60-х годов был сформулирован основной постулат (центральная догма) новой науки — молекулярной биологии, который первоначально выглядел следующим образом:
ДНК > РНК > белок.

Рибосома играет роль «машины», читающей генетическую информацию, записанную в мРНК. Это молекулярная машина, построенная по единой схеме у всех организмов (включая человека) с небольшими вариациями. Она состоит из двух рибонуклеопротеидных (состоящих из РНК и белков) субчастиц: малой и большой. На рибосоме происходит взаимодействие мРНК с транспортными РНК (тРНК), несущими по одной аминокислоте каждая. Белок синтезируется путем образования связей между последовательно доставляемыми тРНК аминокислотами. Этим процессом «руководит» сама рибосома.

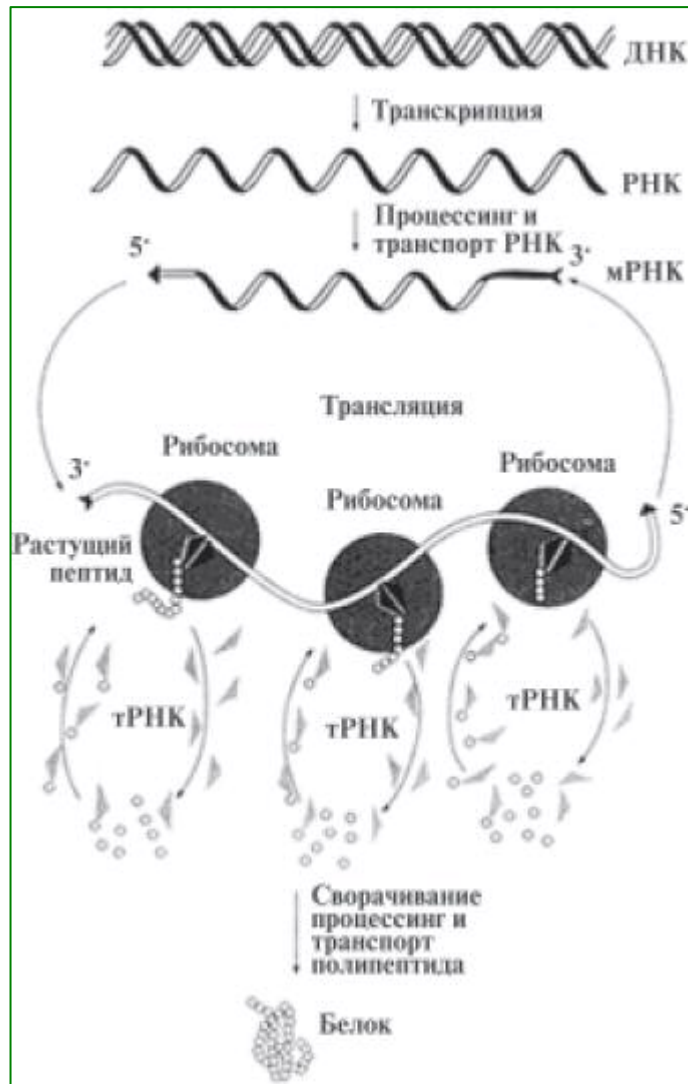


Рис. 5. Процесс трансляции мРНК на рибосомах с образованием белка

Позднее этот постулат был уточнен, и теперь он выглядит как



Сделанное дополнение вызвано обнаружением в 1970 году такого явления, как обратная транскрипция. Выяснилось, что у некоторых вирусов, генетический аппарат которых представлен РНК,

а не ДНК, как у других живых организмов, имеется специальный фермент, который позволяет осуществлять такой обратный процесс, как синтез ДНК на РНК. Этот фермент получил название обратная транскриптаза, или ревертаза.

Казалось бы, теперь уже все стало ясно и однозначно и более ничего не может измениться. Ряд открытий последних лет продолжает вносить уточнения в казавшуюся незыблемой центральную догму молекулярной биологии.

Во-первых, выяснилось, что центральная догма молекулярной биологии постулирует лишь путь передачи генетической информации от нуклеиновых кислот к белкам и, следовательно, к конкретным свойствам и признакам живого организма. Однако исследование механизмов реализации этого пути на протяжении нескольких десятилетий, следовавших за формулировкой центральной догмы, вскрыло гораздо более разнообразные функции РНК, помимо известных ранее.

Во-вторых, согласно первоначальной догме, все носители инфекционных болезней должны иметь генетический материал — ДНК или РНК. Оказалось, что и здесь есть исключения. В 1997 году Нобелевская премия была вручена Стэнли Прузинеру за открытие белковых инфекционных частиц, вызывающих такое заболевание, как, например, болезни «коровьего бешенства», или почесухи. Эти частицы были названы прионами. Проникая в клетку-хозяина, прионы «навязывают» свою болезнетворную конформацию (измененную пространственную структуру) нормальным белкам-аналогам, содержащимся в клетках. При этом ни РНК, ни ДНК никак не участвуют в развитии заболевания. Иными словами, при прионовых заболеваниях информация передается не от одной нуклеиновой кислоты к другой нуклеиновой кислоте, а от белка к белку, что в исходном варианте центральной догмы не предусматривалось.

Однако все эти дополнения в «центральную догму» молекулярной биологии не повлияли принципиально на ее общую закономерность. В центре всего и вся стоял и стоит ген. Сегодня он перестал быть чем-то таинственным, стал реальным химическим веществом, появилась возможность судить о нем так же, как и о других химических соединениях живых организмов, изучать с помощью доступных генетикам и биохимикам методов. Стало понятным, что такое генетический код и как реализуется в клетке та информация, которая записана в ДНКовом тексте. В результате этого в 1953 году родилась молекулярная генетика — раздел науки, который занялся детальным изучением процессов работы ДНК в клетке на молекулярном уровне.

Расшифровка химической и пространственной структуры ДНК — носителя генетической информации — оценена во всем мире как одно из наиболее выдающихся открытий XX века. Геном стоит в центре всех биологических проблем, всех свойств и способностей человека, всего разнообразия человека. Теперь это уже аксиома. Как говорил Козьма Прутков: «Многие люди подобны колбасам: чем их начиняют, то и носят в себе». Так вот, мы «начинены» ДНК, носим ее в себе, а она-то, главным образом, и определяет многое в нас.

В ядре клетки молекулы ДНК расположены в особых структурах, получивших название хромосомы. Их исследование началось еще свыше 100 лет назад с помощью обычного светового микроскопа. Уже к концу XIX века выяснилось кое-что о поведении хромосом в процессе деления клеток и высказывалась мысль об их участии в передаче наследственности.

Хромосомы становятся видимыми в микроскопе при делении клетки на определенной стадии клеточного цикла, называемой митозом. Хромосомы в этом состоянии представляют собой компактные палочковидные структуры разной длины с довольно постоянной толщиной, у большей части хромосом имеется перетяжка, которая делит хромосому на два плеча. В области перетяжки расположена важная для удвоения хромосом структура, называемая центромерой. При делении клетки в ходе митоза происходит удвоение числа хромосом, в результате которого обе вновь образующиеся клетки в конечном итоге обеспечиваются одним и тем же стандартным набором хромосом.

Лишь в 1956 г. впервые Ю. Тио и А. Леван описали хромосомный набор человека, определили количественный состав хромосом и дали их общую морфологическую характеристику. По сути дела эти работы и положили начало изучению структуры генома человека. У человека в каждой клетке тела содержится 46 хромосом, физические длины которых находятся в пределах от 1,5 до 10 мкм (рис. 6).

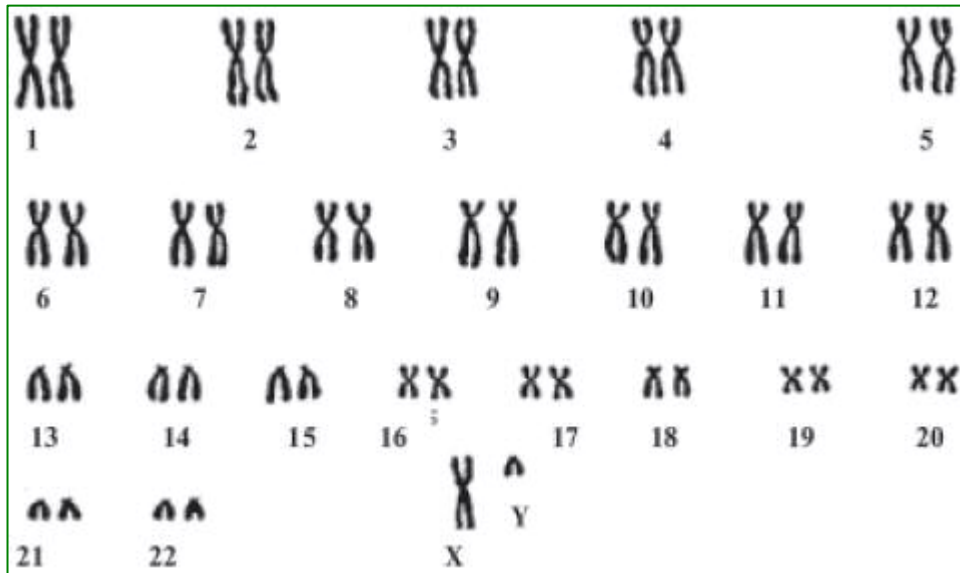


Рис. 6. Вид под микроскопом полного набора хромосом, содержащихся в ядре каждой отдельной клетки человека

Напомним, что набор хромосом во всех клетках человека (за исключением половых) называют диплоидным (двойным), поскольку каждая из хромосом представлена двумя копиями (всего 23 пары).

Каждая соматическая клетка человека (кроме красных кровяных клеток крови) содержит по 2 полных набора хромосом. В каждом единичном (гаплоидном) наборе присутствует **23 хромосомы** — **22 обычные хромосомы (аутосомы) и по одной половой хромосоме — X или Y.**

Таким образом, геном каждого конкретного человека состоит из 23 пар гигантских молекул ДНК, распределенных в разных хромосомах, а если говорить о геноме человека вообще (мужчин и женщин), то общее число таких молекул равно 24. Это первое базовое сведение, которое было получено о геноме человека при анализе хромосом.

Изучение строения (размера и формы) хромосом человека показало, что большинство из них по внешнему виду напоминают кегли, состоящие из двух толстых частей (хроматид) и тонкой перетяжки (центромеры) между ними. Сходство с кеглями, а не с гантелями заключается в том, что центромера чаще всего расположена не в центре хромосомы, а смещена к одному из ее концов. Размеры хромосом сильно варьируют, самая короткая хромосома примерно в десять раз меньше, чем самая длинная. Это второе принципиально важное сведение о структуре генома человека — составляющие его 24 молекулы ДНК имеют разный размер.

Если сравнивать число и размер хромосом у человека и у других видов организмов, то можно увидеть огромные отличия. Например, у коровы, размер генома которой примерно равен геному че-

ловека, имеется 60 пар хромосом. У шпорцевой лягушки содержится всего 18 хромосом, но даже самые маленькие из них больше, чем самые крупные хромосомы человека. У птиц, наоборот, число хромосом достигает 40 и более и все они очень небольшие по размерам. Таким образом, разнообразие хромосом в природе весьма велико.

С помощью световой микроскопии были определены размеры всех хромосом человека. Затем все неполовые хромосомы были пронумерованы по уменьшению размера — от 1 до 22. Половым хромосомам не присвоили номер, а назвали X и Y. Как показали более точные последующие исследования, хромосома 21 реально оказалась чуть меньше 22, однако нумерацию хромосом не изменили (чтобы не вносить путаницу). Различие в хромосомных наборах между мужчинами и женщинами состоит в том, что у женщин имеются две половые X-хромосомы (т. е. хромосомы во всех 23-х парах одинаковы), а у мужчин пару с X-хромосомой образует мужская половая хромосома — Y. Каждую хромосому можно рассматривать как отдельный том большого двадцатичетырехтомного собрания сочинений под названием Энциклопедия человека.

Половые клетки человека, в отличие от клеток тела взрослого организма (соматических клеток), содержат не 2 набора томов ДНК-ового текста, а всего лишь один. Перед зачатием каждая отдельная хромосома (отдельный том в Энциклопедии человека) сперматозоида отца и яйцеклетки матери состоят из смешанных в разном сочетании различных глав ДНК-ового текста их родителей. Любая из хромосом, полученная нами от отца, образовалась в его семенниках незадолго до того, как мы были зачаты. Ранее, за всю историю человечества, точно такая хромосома никогда не существовала. Она была сформирована в процессе случайного перемешивания, происходящего при делении, постепенно образуясь из объединяющихся друг с другом участков хромосом предков со стороны отца. Также обстоит дело и с хромосомами яйцеклеток, за исключением того, что они формируются в организме нашей матери задолго до нашего рождения (почти сразу после рождения самой матери).

В зиготе, образующейся в результате слияния сперматозоида и яйцеклетки, материнские и отцовские гены смешиваются и перетасовываются в разных сочетаниях. Это происходит в результате того, что хромосомы не остаются неизменными в поколениях — они вступают во взаимодействие со своей случайно встреченной парой, обмениваясь с ней материалом. Такой постоянно идущий процесс получил название рекомбинации. И следующему поколению часто достается уже гибридная хромосома — часть от бабушки и часть от дедушки. Далее в ряду поколений пути генов постоянно пересекаются и расходятся. В результате слияния уникальной яйцеклетки с уникальным сперматозоидом и возникает уникальный во всех отношениях геном. И в этом смысле все мы уникальны. Каждый человеческий индивид хранит уникальную генетическую информацию, состоящую из случайной комбинации разных вариантов генов.

Отдельный ген можно рассматривать как единицу, продолжающую существовать в ряду многочисленных поколений. И в этом смысле ген бессмертен!

Существует даже такая оригинальная точка зрения, что не сами люди, а их гены правят миром, а каждый конкретный живой организм служит лишь временным прибежищем для них. Эта не бесспорная мысль принадлежит Ричарду Докинзу, автору книги «Эгоистичный ген». По его мнению, гены практически бессмертны в отличие от живых организмов, в которых они существуют. Некоторым генам десятки и даже сотни миллионов лет. Гены, пользуясь терминологией Докинза, делают все возможное, чтобы выжить. Приспосабливаются к жаре и холоду, выбирая себе местечко получше, мигрируют с помощью человека и вступают в новые комбинации. Человек оказался довольно непоседливым хозяином. За тысячи лет он сильно исколесил мир, распространяя свое присутствие, влияние и свою начинку — гены.

Такая точка зрения далеко не бесспорна, и из дальнейшего изложения нам станет понятно, что гены — это в первую очередь не эгоисты, а трудоголики. Имеются гены — «сторожа» генома, гены — «дворники», гены — «повара» и гены — «домоправители». Обеспечивая свое существование, они обеспечивают и существование нас.

Сразу после зачатия будущий человек представляет собой всего одну клетку (зиготу), наделенную одной исходной ДНКовой библиотекой, содержащей 46 томов. Среди 46 томов всегда 23 получены от отца, а другие 23 — от матери. Тексты 23 отцовских и 23 материнских томов хотя и очень сходны в целом, тем не менее отличаются в деталях. Например, в отцовском томе № 18 на странице 253 существует предложение-предписание (в виде гена), в котором сказано, что глаза у ребенка должны быть карими, а в этом же материнском томе на той же странице тоже написано о цвете глаз, но согласно этому тексту цвет должен быть голубыми. Первое указание более строгое (доминирующее), чем второе, и в результате у ребенка глаза будут иметь карий цвет. Ген, который диктует свои права, называют доминирующим, а тот, который уступает свои права, — рецессивным. Голубой цвет глаз имеют только те люди, у которых и в материнском, и в отцовском тексте содержатся рецессивные гены, в которых есть указание на голубоглазость. Затем зигота делится на две клетки, каждая из них вновь делится и так до появления миллиардов клеток. Схематически процесс деления клеток изображен на рис. 7.



Рис. 7. Основные стадии клеточного цикла, приводящего к делению клетки

При каждом делении клетки содержащиеся в библиотеках тома ДНКового текста точно копируются, причем практически без ошибок. Организм взрослого человека состоит в среднем из 10¹⁴ клеток. Например, в головном мозге и печени насчитывается примерно по 10 млрд. клеток, в иммунной системе — 300 млрд. клеток. В течение всей жизни человека в его организме происходит около 10¹⁶ клеточных делений. Клеточный состав многих органов за 70 лет жизни обновляется несколько раз. И каждая из этих клеток содержит одни и те же 46 томов ДНКового текста. В конце 60-х годов XX века был осуществлен важный прорыв в исследовании хромосом. Обусловлен он был всего лишь тем, что для их окраски стали использовать специальное контрастное вещество — акрихин-иприт, а затем и другие сходные с ним соединения. Такая окраска позволила выявить внутри хромосом большое число разных субструктур, которые не обнаруживались под микроскопом без окрашивания. После окрашивания хромосом специфическим красителем Гимза-

Романовского они выглядят как зебры: вдоль всей длины видны поперечные светлые и темные полосы, имеющие окраску разной интенсивности.

Эти полосы получили название хромосомных G-сегментов или полос (рис. 8). Картина сегментации сильно отличается у разных хромосом, но расположение хромосомных сегментов постоянно у каждой из хромосом во всех типах клеток человека.

Природа полос, выявляемых при окраске, до конца еще не ясна. Сейчас установлено только, что участки хромосом, соответствующие темным полосам (названные R-полосами), реплицируются раньше, чем светлые участки (названные G-полосами). Таким образом, полосатость хромосом скорее всего все же имеет некий до конца еще не понятый смысл.

Окрашивание хромосом очень облегчило их идентификацию, а в дальнейшем способствовало определению расположения на них генов (картированию генов).

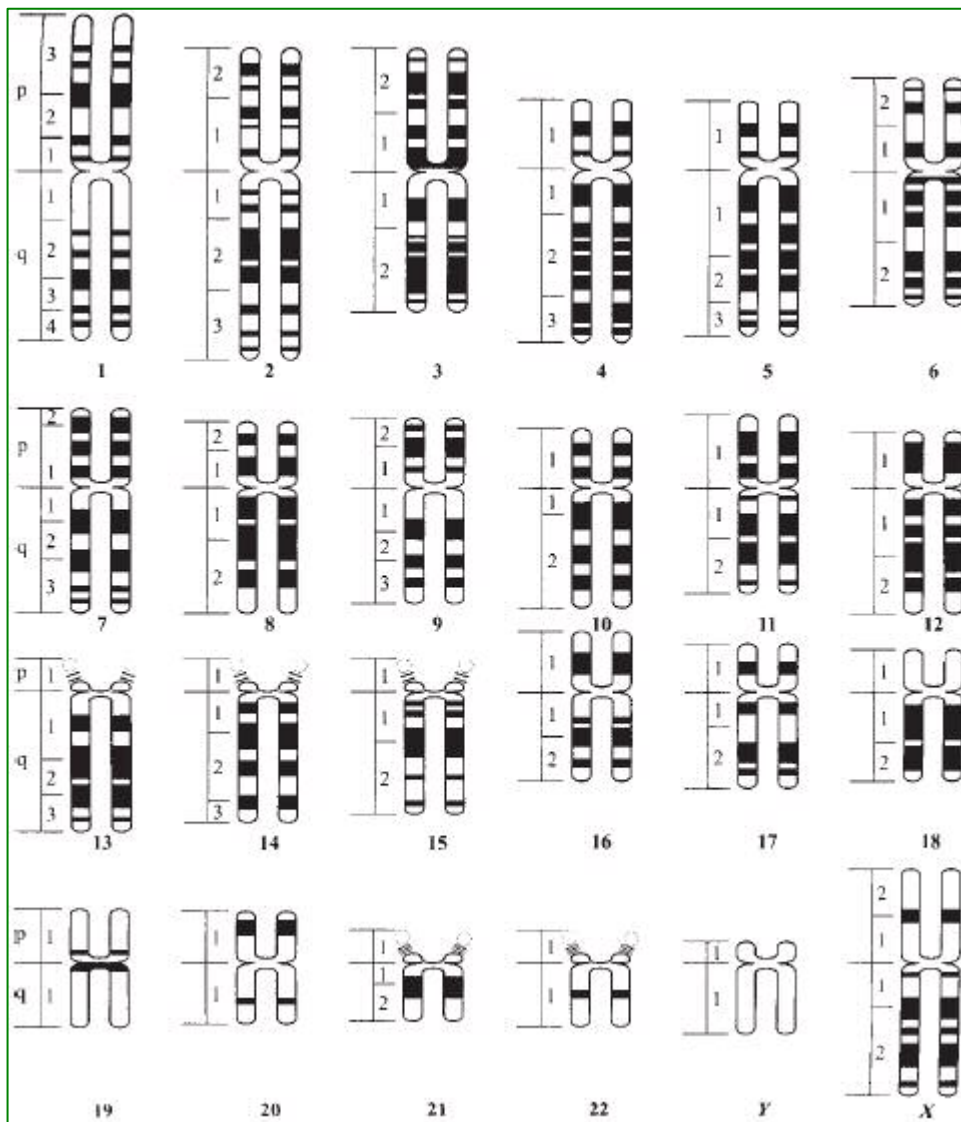


Рис. 8. Специфические хромосомные G-сегменты, выявляемые при окраске хромосом человека, и система их обозначения согласно решению международной конференции в Париже в 1971 году. Цифрами под хромосомами указаны их номера. X и Y — половые хромосомы, p — короткое плечо, q — длинное плечо хромосом

Хотя детальные процессы, происходящие при окрашивании, еще не до конца ясны, очевидно, что картина окраски зависит от такого параметра, как увеличенное или уменьшенное содержание в отдельных полосах хромосом AT или GC-пар. И это еще одно общее сведение о геноме — он не однороден, в нем есть районы, обогащенные определенными парами нуклеотидов.

Это, в частности, может быть связано с повторяемостью некоторых типов нуклеотидных последовательностей ДНК в определенных районах.

Дифференциальная окраска хромосом нашла широкое применение для выявления и идентификации небольших индивидуальных изменений генома конкретного человека (полиморфизма), в частности, приводящих к различным патологиям. Примером этому может служить обнаружение так называемой филадельфийской хромосомы, встречающейся у больных с хроническим миелоидным лейкозом. С помощью окраски хромосом установлено, что у пациентов с этим заболеванием определенный фрагмент исчезает на хромосоме 21 и появляется на конце длинного плеча хромосомы 9 (перенос фрагмента или транслокация, сокращенно t). Генетики обозначают такое событие как t(9; 21). Таким образом, хромосомный анализ свидетельствует о том, что разные молекулы ДНК могут обмениваться между собой отдельными участками, в результате чего в геноме образуются «гибриды», состоящие из молекул ДНК разных хромосом. Анализ уже изученных свойств хромосом позволил сформировать представление о полиморфизме генома человека.

Для выяснения локализации отдельных генов на хромосомах (то есть картирования генов) используют целый арсенал специальных зачастую весьма сложных по замыслу и исполнению методов. Один из основных — молекулярная гибридизация (образование гибрида) гена или его фрагмента с фиксированными на твердой подложке препаратами хромосом, выделенными из клеток в чистом виде (это называют гибридизацией *in situ*). Суть метода гибридизации *in situ* заключается во взаимодействии (гибридизации) между денатурированными (расплетенными) нитями ДНК в хромосомах и комплементарными нуклеотидными последовательностями добавленных к препарату хромосом, индивидуальных одонитевых ДНК или РНК (их называют зондами). При наличии комплементарности между одной из нитей хромосомной ДНК и зондом между ними образуются довольно стабильные молекулярные гибриды. Зонды маркируют предварительно с помощью разных меток (радиоактивных, флуоресцентных или др.). Места образования гибридов на хромосомах выявляют по положению этих меток на препаратах хромосом.

Так, еще до появления методов генной инженерии и секвенирования ДНК выяснили, например, расположение в геноме человека генов, кодирующих большие и малые рибосомные РНК (рРНК). Гены первых оказались локализованными в пяти разных хромосомах человека (13, 14, 15, 21 и 22), тогда как основная масса генов малой рРНК (5S РНК) сконцентрирована в одном месте на длинном плече хромосомы 1.

Гены, расположенные на одной хромосоме, определяют как сцепленные (связанные) гены. Если гены расположены на разных хромосомах, они наследуются независимо (независимая сегрегация). Когда же гены находятся на одной и той же хромосоме (т. е. сцеплены), они неспособны к независимой сегрегации. Изредка в половых клетках могут происходить различные изменения хромосом в результате рекомбинационных процессов между гомологичными хромосомами. Один из таких процессов получил название кроссинговера. Из-за кроссинговера сцепление между генами одной группы никогда не бывает полным. Чем ближе расположены друг к другу сцепленные гены, тем

меньше вероятность изменения расположения таких генов у детей по сравнению с родителями. Измерение частоты рекомбинаций (кроссинговера) используется для установления линейного порядка генов на хромосоме внутри группы сцепления. Таким образом, при картировании хромосом первоначально устанавливаются, находятся ли данные гены в одной и той же хромосоме, без уточнения, в какой именно. После того, как хотя бы один из генов данной группы сцепления локализуется в определенной хромосоме (например, с помощью гибридизации *in situ*), становится ясным, что все другие гены этой группы сцепления находятся в той же самой хромосоме.

Первым примером связи генов с определенными хромосомами может служить обнаружение сцепления определенных наследуемых признаков с половыми хромосомами. Чтобы доказать локализацию гена в мужской половой Y-хромосоме, достаточно показать, что данный признак всегда встречается только у мужчин и никогда не обнаруживается у женщин. Группа сцепления женской X-хромосомы однозначно характеризуется отсутствием наследуемых признаков, передающихся от отца к сыну, и наследованием признаков матери.

Особенно важным для изучения генома человека на первых этапах его исследования стал метод, называемый гибридизацией соматических клеток. При смешивании соматических (неполовых) клеток человека с клетками других видов животных (чаще всего для этой цели использовали клетки мышей или китайских хомячков) в присутствии определенных агентов может происходить слияние их ядер (гибридизация). При размножении таких гибридных клеток происходят потери некоторых хромосом. По счастливой для экспериментаторов случайности в гибридных клетках человек-мышь происходит потеря большей части хромосом человека. Далее отбираются гибриды, в которых остается только какая-нибудь одна человеческая хромосома. Исследования таких гибридов позволили связать некоторые биохимические признаки, свойственные клеткам человека, с определенными хромосомами человека. Постепенно благодаря использованию селективных сред научились добиваться сохранения или потери отдельных хромосом человека, несущих определенные гены. Схема отбора, хотя и не очень проста на первый взгляд, довольно хорошо показала себя в эксперименте. Так, придумали специальную селективную среду, на которой могут выживать только те клетки, в которых синтезируется фермент тимидинкиназа. Если для гибридизации с клетками человека взять в качестве партнера мутантные клетки мыши, не синтезирующие тимидинкиназу, то будут выживать только те гибриды, которые содержат хромосомы человека с геном тимидинкиназы. Таким путем впервые удалось установить локализацию гена тимидинкиназы на хромосоме 17 человека.

Несмотря на то, что изучение генома человека на уровне хромосом дало ряд важных его характеристик, они были самыми общими и дали относительно мало для полного понимания устройства и функционирования генетического аппарата человеческих клеток.

Возникает вопрос как двухметровая молекула уместается в микроскопическом ядре? Как уже говорилось, молекулы ДНК, содержащиеся во всех хромосомах одной клетки человека, имеют общую длину около 2 метров. Это в среднем в миллион раз больше, чем диаметр клеточного ядра, составляющий порядка микрометра. Как же уместаются столь гигантские молекулы в таком маленьком ядре? Изучение структуры хромосом позволило ответить на этот вопрос. Оказывается, в ядре осуществляется «насилованная» упаковка молекул ДНК. Это достигается с помощью специальных механизмов, обеспечивающих изгибание двойной спирали ДНК. Существует несколько уровней «компактизации» ДНК в клетке (рис. 9).

Некоторые из особенностей упаковки ДНК изучены хорошо, а о некоторых существуют пока лишь приблизительные представления. Первый уровень компактизации заключается в накручивании нити ДНК, как нитки на катушку, на специальный комплекс ядерных белков (гистонов).

Нить ДНК делает около двух оборотов вокруг одного комплекса, а затем снова около двух оборотов вокруг второго комплекса и т. д. В результате образуется структура, напоминающая бусы. От-

дельные бусинки в этой структуре получили название нуклеосомы. На одной нуклеосоме размещается около 200 пар нуклеотидов ДНК. Между нуклеосомами остается фрагмент ДНК размером до 60 пар оснований, называемый линкером. Этот уровень укладки позволяет уменьшить линейные размеры ДНК в 6–7 раз.

На втором уровне компактизации «бусы» укладываются в спираль, состоящую из шести нуклеосом на виток. При этом линейные размеры ДНК уменьшаются в сумме до 1 мм, т. е. в 25–30 раз. Третий уровень компактизации молекул ДНК изучен еще плохо. Скорее всего, это петельная укладка фибрилл — образование петельных доменов, которые под углом отходят от основной оси хромосомы (уплотнение в 680 раз). Их можно видеть в обычный световой микроскоп.

На последнем уровне компактизации ДНК происходит ее уплотнение примерно в 10000 раз.



Рис. 9. Стадии упаковки ДНК в хромосомах (от двойной спирали до целых хромосом)

Анализ суммарной ДНК — новые сведения о структуре генома человека

На первом этапе непосредственного исследования структуры генома человека, когда еще не существовала методология генной инженерии, для изучения ДНК применяли традиционные физико-химические методы. В этих опытах использовали суммарные препараты ДНК, целиком выделенные из ядер клеток человека.

Пожалуй, первые сведения о молекулярной структуре генома человека были получены в результате центрифугирования в пробирке растворов ДНК в хлористом цезии при довольно высоких скоростях. В процессе вращения соли цезия создают тонкие слои раствора с различной плотностью (градиент плотности) вдоль пробирки и молекулы ДНК перемещаются в этом градиенте, пока не достигнут такой области, где плотность солевого раствора будет точно такой же, как их собственная. А плотность ДНК сильно зависит от содержания АТ и ГЦ-пар нуклеотидов, т. е., как

говорят, от нуклеотидного состава. Оказалось, что основная масса ДНК человека после центрифугирования располагается преимущественно в одной зоне градиента (это соответствует среднему содержанию ГЦ-пар в геноме человека, равному 42%). Однако наряду с этим неожиданно обнаруживались и небольшие (минорные) полосы, в которых также содержались молекулы ДНК, но с иной плотностью и, следовательно, с иным содержанием нуклеотидных пар. Такие минорные, или дополнительные фракции ДНК получили название «сателлитных». Такое имя дали этим фракциям не случайно. В то время как раз был запущен первый советский спутник (лат. *satellit*is — спутник).

Вскоре после того, как была установлена двухспиральная структура ДНК, обнаружили, что при сильном нагревании ДНК две ее цепи расходятся (расплавляются) и ДНК из двунитевой превращается в одонитевую. Это приводит к нарушению ее естественной структуры, что получило отражение в названии данного процесса — денатурация ДНК. Однако при охлаждении денатурированной ДНК комплементарные цепи находят друг друга и соединяются строго так, как они располагались в исходной неденатурированной молекуле, по типу застёжки — «молнии». Этот процесс получил название ренатурации или реассоциации. Сразу же после обнаружения этого явления оно было использовано экспериментаторами в целях изучения структуры генома.

Немного совсем простой математики для тех самых любопытных. Процесс реассоциации ДНК во многом сходен с обычной химической реакцией второго порядка и, по этой причине, может быть описан довольно простой формулой:

$$C_t / C_0 = 1 / (1 + k_2 (C_0 t)),$$

где C_0 и C_t — концентрации одонитевых ДНК соответственно в начальный (нулевой) момент времени и в момент времени t после начала реакции реассоциации, k_2 — константа скорости реакции второго порядка.

Для удобства изображения процесса реассоциации строят график, в котором по оси ординат откладывают долю реассоциированных молекул, а по оси абсцисс — величину $C_0 t$. При таком изображении кривые кинетики реассоциации ДНК простейших организмов (вирусов и бактерий) имеют S-образный вид. Это соответствует кинетике химической реакции второго порядка. Если в эксперименте наблюдают отклонения от этой кривой, то это должно указывать на гетерогенность цепей ДНК по скорости взаимодействия друг с другом: одни реагирует быстрее, другие, соответственно, медленнее. Когда ДНК человека порезали на куски небольшого размера и измерили кинетику их реассоциации, то обнаружилось, что кривая, описывающая этот процесс, далека от стандартной (рис. 12).

Это явилось результатом того, что в ДНК человека имеются нуклеотидные последовательности, реассоциирующие с разной скоростью, а суммарная кривая реакции, наблюдаемая в опыте, отражает совокупность множества независимых реакций второго порядка. Когда были проведены эти эксперименты, уже существовал математический аппарат, позволяющий, хотя и очень грубо, вычленять из сложной суммарной кривой отдельные относительно однородные кинетические компоненты. Различия в скоростях реассоциации разных компонентов ДНК человека были связаны с разной представленностью в ДНК отдельных нуклеотидных последовательностей. Участки, которые присутствуют в геноме всего один раз, назвали уникальными. Если в геноме определенная нуклеотидная последовательность не уникальна, а представлена неким числом одинаковых копий (т. е. повторяется, отсюда и название — повторяющаяся), то такая последовательность, естественно, по чисто химическим законам, будет в растворе находить комплементарную цепь и взаимодействовать с ней (реассоциировать) значительно быстрее первой.

В результате такого анализа нашли, что в геноме человека около 75% участков ДНК представлены 1 копией (уникальные) на гаплоидный геном (естественно, в ядре, являющемся диплоидным, имеется 2 копии каждой уникальной нуклеотидной последовательности). Остальную часть генома со-

ставляют повторяющиеся последовательности (повторы), среди которых 10% представлены очень быстро реассоциирующими повторами (10^4 и более копий на геном) и около 15% — умеренными повторами. Заметим сразу, что в дальнейшем эти оценки были существенно пересмотрены. Однако некоторые из этих результатов остались в силе до сих пор.

ДНК разбивают на небольшие фрагменты, денатурируют путем нагревания, а затем при охлаждении разошедшиеся цепи ДНК вновь соединяются. Чем чаще в смеси встречаются те или иные последовательности, тем быстрее они находят друг друга в растворе и реассоциируют. По этой кинетике определяли общее содержание повторяющихся и неповторяющихся (уникальных) нуклеотидных последовательностей в геноме человека

Первоначальные анализы показали, что среди быстро взаимодействующих друг с другом при реассоциации фрагментов ДНК присутствует некоторое количество таких, кинетика реассоциации которых отличается от реакции второго порядка. Причина этому оказалось в том, что эти участки представляют собой так называемые обращенные повторы, или палиндромы — взаимокплементарные последовательности, расположенные не на разных, а на одной нити ДНК.

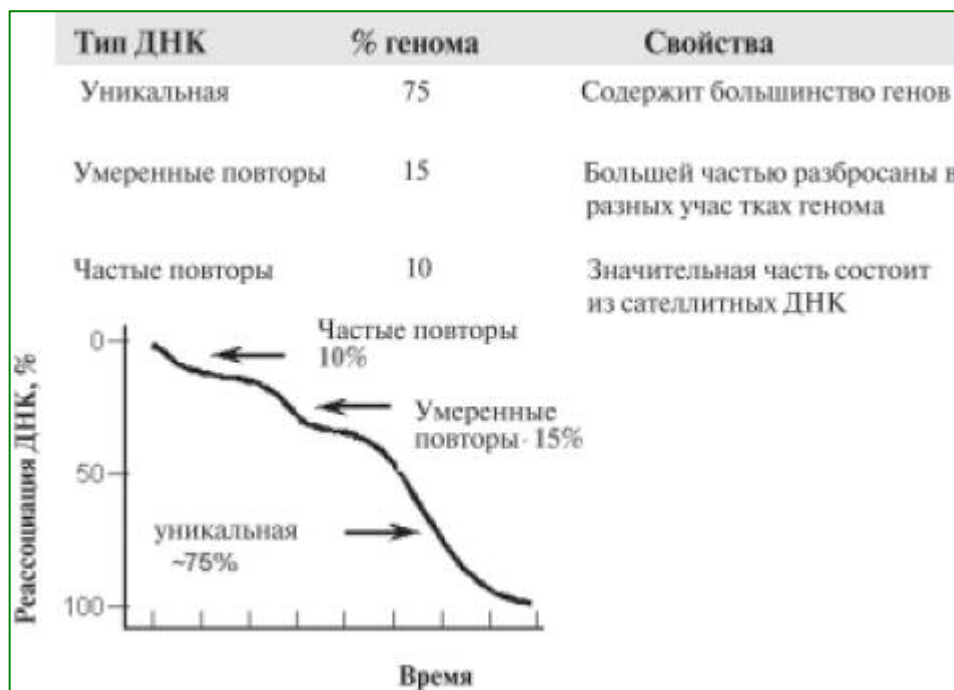
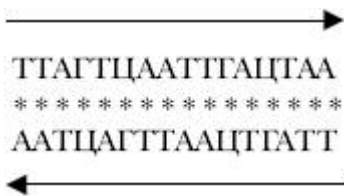


Рис. 12. Кинетика восстановления двунитовых молекул из искусственно разделенных комплементарных цепей ДНК человека (реассоциация).

Таким образом, в ДНКовом тексте присутствуют «предложения» — палиндромы («перевертыши»), одинаково читаемые слева направо и справа налево. Перевертыши хорошо известны из литературы — это предложения, которые читаются одинаково слева направо и справа налево без учета знаков препинания и интервалов между словами. В качестве примера приведем один из таких перевертышей:

УЖРЕДКОРУКОЮОКУРОКДЕРЖУ.

В ДНК перевертышами называют отрезки двойной антипараллельной спирали, которые имеют одинаковую нуклеотидную последовательность при чтении по обеим цепям в одинаковом направлении. Это выглядит как, например, в ниже приведенном случае:



Здесь стрелками показано направление цепей ДНК, а звездочками — водородные связи, образуемые между парами нуклеотидов. Общее число таких «перевертышей» в геноме человека оценено в интервале от 10^5 до 10^6 . При этом они относительно равномерно распределены по ДНК.

Имеются в геноме человека и нуклеотидные последовательности, которые на всем своем протяжении построены из одной единственной «буквы». Если в одной цепи ДНК эта буква А, то в другой цепи, соответственно, будет буква Т. Такие участки названы гомополимерными. В случае приведенного выше примера последовательность нуклеотидов записывается, как поли(А) — поли(Т). Таким образом, выяснилось, что в геноме человека (а параллельно это изучали и в геномах других организмов) имеются все варианты нуклеотидных последовательностей, состоящих из 4 «букв», которые только можно себе мысленно представить.

Используя различные модификации метода реассоциации ДНК, установили также, что в большей части генома человека повторяющиеся и уникальные нуклеотидные последовательности перемежаются друг с другом, а средние длины перемежающихся повторяющихся и уникальных фрагментов ДНК составляют соответственно 300 и 2000 п. н.

Приведенные выше данные были, конечно же, усредненными и поверхностно отражали общую картину устройства генома человека. Тем не менее они послужили хорошей основой для дальнейших более углубленных исследований. Важно, что одновременно такие же работы проводились и с использованием других эукариотических организмов. Многие моменты оказались сходными у разных высших организмов и растений. Так постепенно начали вырисовываться в общих чертах основные принципы организации генома человека.

Анализ индивидуальных элементов ДНК (генная инженерия)

Продолжение идей конструирования ДНК привело к тому что в результате в 1970-х годах были разработаны принципиально новые методы работы с ДНК. Ее научились резать на фрагменты, сшивать их друг с другом, а затем переносить в клетки и целые организмы. На практике это сводится к созданию в пробирке (т. е. *in vitro*) новых молекул ДНК, состоящих из фрагментов разного происхождения (так называемых рекомбинантных ДНК), которые в естественных условиях не могут возникнуть благодаря установленным и тщательно охраняемым Природой межвидовым барьерам. Датой рождения нового революционизирующего направления, получившего название генной инженерии, принято считать 1972 год, когда в США была создана первая рекомбинантная молекула ДНК.

Становление генной инженерии происходило очень непросто. Уже в самом начале (в 1973 году) участники Гордоновской конференции направили письмо президенту Национальной Академии наук США, в котором высказывали серьезные опасения о возможной опасности рекомбинантных ДНК для здоровья человека и окружающей среды. Прозвучал даже призыв к мораторию на некоторые виды генно-инженерных исследований. Два года спустя состоялась конференция в Асиломаре (США), где также шла речь об опасности генно-инженерных манипуляций и их ограничении. Один из основных исследователей структуры ДНК Э. Чаргафф, в частности, вопрошал: «Имеем ли мы право посягать необратимым образом на эволюционную мудрость миллионов лет только для того, чтобы удовлетворить амбицию и любопытство нескольких ученых?» Озабоченность ученых объяснялась теоретической возможностью создания в пробирке новых генетических структур, способных вызывать необычные формы инфекционных заболеваний человека, или созданием организмов, неблагоприятно воздействующих на окружающую среду. Однако как тогда, так и по

прошествии трех десятков лет не существует никаких подтверждений тому, что рекомбинантные ДНК представляют опасность для человека. Современное состояние молекулярной генетики и как вершина ее развития — секвенирование генома человека — свидетельствуют о том, что генная инженерия полностью оправдала большие надежды ученых и поддерживающей ее общественности. Сейчас уже практически нет спора в отношении важности и необходимости использования генной инженерии для решения как научных, так и практических задач, стоящих перед человечеством.

Генная инженерия как принципиально новая технология возникла в значительной мере на базе достижений, которых ученые добились при изучении молекулярных процессов, которые происходят в бактериальной клетке. Было сделано несколько важных открытий, прежде чем исследователи научились легко и направленно манипулировать с ДНК, а не просто смешивать фрагменты ДНК с бактериями и отбирать среди них случайно возникающие варианты (рекомбинанты). Бактерии имеют высокоэффективный механизм защиты от чужеродной ДНК. Исследование этой проблемы привело к обнаружению в бактериях специальных ферментов, получившие название рестриктаз. Первые из них получены еще в 1970 г. Сейчас уже выделено несколько сот разнообразных рестриктаз из клеток разных видов бактерий. Каждый из таких ферментов «узнает» строго определенные короткие (обычно 4–6 п. н.) нуклеотидные последовательности ДНК, которые чаще всего представляют собой палиндромы, и разрезает ДНК на фрагменты по этим участкам. В результате этих свойств рестриктазы стали незаменимым инструментом в исследовании генома, своеобразным скальпелем для ДНК.

Большой удачей для экспериментаторов было то, что в большинстве случаев при нарезании рестриктазами ДНК образующиеся фрагменты содержат «липкие» концы, т. е. способны в дальнейшем соединяться друг с другом. Для восстановления химической связи таких слипшихся фрагментов стали использовать еще один новый класс ферментов — лигазы. Так появились первые инструменты, позволяющие целенаправленно манипулировать с молекулами ДНК *in vitro*. Следующий важный этап — размножение сконструированных рекомбинантных ДНК, чтобы можно было с ними работать дальше. Для этой цели стали использовать специальные ДНК — векторы, которые представляют собой или бактериофаги, или внехромосомные кольцевые молекулы — плазмиды, которые обладают способностью размножаться в бактериальных клетках. Если любой фрагмент ДНК внести в состав ДНК бактериофагов или плазмид, то при переносе в бактериальные клетки он будет там размножаться вместе с вектором. Наконец, были разработаны специальные подходы для отбора бактериальных клеток, в которых содержатся необходимые рекомбинантные ДНК. Все эти достижения позволили проводить клонирование ДНК, т. е. выделять из сложной смеси молекул определенные участки ДНК и размножать их.

Клонирование фрагментов ДНК дало в руки исследователей принципиально новый подход к выделению из сложной смеси ДНК-фрагментов конкретных индивидуальных последовательностей. При этом они могли содержать осмысленные тексты (гены) или не содержать таковых. Традиционная биохимия не могла обеспечить подобную процедуру в силу того, что различные фрагменты ДНК химически очень сходны. Но генной инженерии это оказалось вполне под силу. Схематически процесс клонирования ДНК изображен на рис. 11.

После разрезания ДНК на фрагменты с помощью рестриктазы все они соединяются ковалентно с вектором (вектор разрезан той же рестриктазой) в результате «слипания» концов молекул и последующей обработки «гибридов» лигазой. Так в пробирке создается искусственная или рекомбинантная ДНК. Затем осуществляют перенос множества образовавшихся рекомбинантных молекул в бактериальные клетки — трансформацию. Особенность процесса трансформации заключается в том, что в одну бактериальную клетку попадает лишь одна рекомбинантная молекула ДНК. После этого, благодаря отбору, размножаются только бактерии с рекомбинантными ДНК, содержащими требуемые нуклеотидные последовательности.

Для изучения структуры генома необходимо не только клонировать все его участки в виде рекомбинантных молекул, но, главное, расшифровать в них последовательность нуклеотидов, т. е. «прочитать» ДНК-текст.

Важным моментом во всей этой генно-инженерной «кухне» было создание принципиально новой технологии размножения индивидуальных фрагментов ДНК *in vitro*. Этот удивительно простой метод получения фрагментов ДНК в неограниченном количестве копий был придуман при необычных обстоятельствах — ночью, во время автомобильной поездки в горах Калифорнии. Автор этой идеи нобелевский лауреат Кэри Б. Мюллис так вспоминает момент озарения. «Иногда удачная идея приходит в голову совершенно неожиданно. Со мной, например, это случилось в одну из апрельских ночей в 1983 г., когда я, сидя за рулем автомобиля, пробирался по освещенной луной горной дороге в секвойные леса Северной Калифорнии. Мысль моя случайно натолкнулась на процесс, благодаря которому можно получать копии генов в неограниченных количествах. Теперь он называется «полимеразной цепной реакцией». За создание новой технологии ее автор получил Нобелевскую премию.

Полимеразная цепная реакция (сокращенно ПЦР) дает возможность в течение дня из одной молекулы ДНК получать 100 миллиардов идентичных по структуре молекул. Эта реакция довольно проста в исполнении: нужны лишь ДНК, пробирка, несколько реагентов и источник нагревания и охлаждения. Препарат ДНК, который необходимо копировать, может быть получен и из кусочка ткани, и из капли засохшей крови, и из засушенной мумии, и даже из тела мамонта, пролежавшего несколько тысяч лет в вечной мерзлоте.



Рис. 11. Схема процесса клонирования ДНК с использованием в качестве вектора бактериальной плазмидной ДНК

Сама процедура ПЦР заключается в следующем. Прежде всего, ДНК денатурируют, нагревая ее до 98°C. При этом две цепи ДНК, относительно слабо связанные между собой, разделяются друг от друга. Далее в ход идут так называемые затравки (праймеры) — короткие однонитевые фрагменты ДНК, которые комплементарны краям того участка генома, который собираются размножить (амплифицировать). Эти затравки гибридизуются с комплементарными участками однонитевых молекул ДНК, после чего специальный фермент (ДНК-полимераза) удлиняет затравки, строя молекулу ДНК по матрице (цикл 1). Затем продукты реакции вновь денатурируют и процедуру повторяют необходимое число раз. За 20–30 циклов из одной молекулы можно получить миллионы копий ДНК. Этот процесс подобен цепной ядерной реакции, но осуществляется в пробирке с помощью специального фермента — термостабильной ДНК-полимеразы. По этой причине он и получил название полимеразной цепной реакции.

Определение полной последовательности нуклеотидов в ДНК человека

Длительное время искали эффективные методы, позволяющие определять последовательность нуклеотидов в длинных молекулах ДНК (такое определение нуклеотидной последовательности получило название секвенирование, от англ. sequence — последовательность). Только в середине 70-х годов секвенирование протяженных участков ДНК стало возможным. Связано это было с изобретением принципиально новых подходов. В 1976 г. А. Максомом и У. Гилбертом был разработан метод прямого секвенирования, основанный на химической дегградации ДНК.

В данном случае осуществляется специфическая химическая фрагментация длинной цепи ДНК (полинуклеотида), радиоактивно меченной с одного конца. Затем препарат меченой ДНК разделяют на четыре порции и каждую из них обрабатывают реагентом, модифицирующим одно или два из четырех оснований, содержащихся в ней. После разделения в специальном геле меченых фрагментов по размерам (с помощью электрофореза) на рентгеновских пленках смотрят, что при этом происходит с нуклеотидной последовательностью, и на основании этого делают вывод о порядке расположения нуклеотидов друг за другом в каждом фрагменте ДНК. Однако метод оказался довольно сложным.

Английский ученый Фред Сэнгер предложил иной способ расшифровки структуры ДНК, который в конечном итоге после разных модификаций стал основным. Согласно методу Сэнгера молекулу ДНК с помощью специальной обработки ферментами не только расщепляют на фрагменты, но и «расплетают» ее двойную спираль на две нити. Потом по каждому из полученных обрывков «нитей» с помощью специальных затравок восстанавливают недостающую вторую нить ДНК. При этом синтез второй нити осуществляется не полностью, а с помощью специального приема обеспечивается прекращение синтеза строго на одном из 4 нуклеотидов.

В результате этого получался набор цепей ДНК с разной длиной — «лесенка». Эта «лесенка» становится видимой глазом после разделения всех фрагментов по размерам за счет того, что концы всех фрагментов содержат специфическую — или радиоактивную, или флуоресцентную метку. Следует отметить, что в разработку этого поистине революционизирующего метода существенный вклад был сделан и двумя российскими учеными: профессором С. К. Василенко и академиком Е. Д. Свердловым.

За разработку методов секвенирования Гилберт и Сэнгер получили Нобелевскую премию. Интересно, что для Сэнгера эта была уже вторая премия, первую он получил за расшифровку аминокислотной последовательности первого белка — инсулина.

Все автоматы-секвенаторы, которые появились позднее, были построены по принципу метода Сэнгера, поскольку он оказался более удобным для автоматизации процесса и компьютерной регистрации результатов. В настоящее время ряд крупных фирм занимается производством таких автоматов, а также стандартных наборов реактивов для анализа ДНК. Секвенирование стало достаточно рутинной лаборантской работой. А метод Максама–Гилберта имеет сейчас скорее историческое, чем практическое значение.

Во всех случаях основное правило заключается в том, чтобы в препарате ДНК, который собираются секвенировать, все фрагменты имели, по крайней мере, один одинаковый конец, т. е. одну точку отсчета. В этот фиксированный конец вводят радиоактивную (или флуоресцентную) метку. После этого проводят дальнейшие специфические расщепления молекул с помощью ферментов или химических агентов по строго определенным азотистым основаниям. Условия расщепления при этом таковы, что на каждый меченый фрагмент в среднем приходится один разрыв. В результате после расщепления образуется набор фрагментов разной длины, каждый из которых начинается в точке отсчета, а заканчивается на или перед определенным типом нуклеотида.

Полученные фрагменты разделяют по размерам с помощью электрофореза в гелях. Маленькие фрагменты движутся быстрее более крупных, и в результате такой процедуры все фрагменты выстраиваются друг за другом в зависимости от своей длины. Результат такого деления изображен схематически на рис. 12.

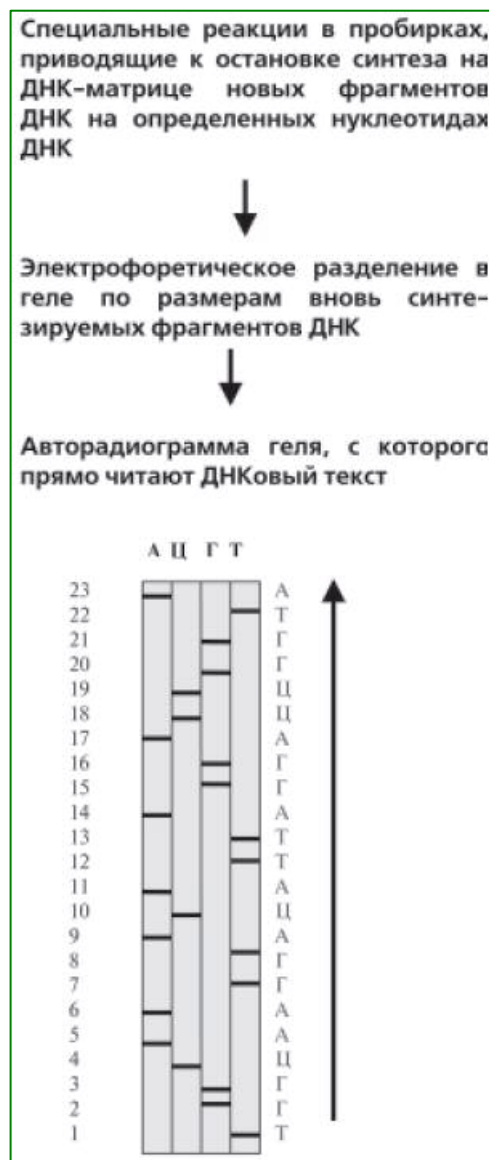


Рис. 12. Схема, поясняющая процесс подготовки и прочтения ДНКового текста

Так осуществляют секвенирование небольших фрагментов ДНК (не более 1000 нуклеотидов). Но геномы даже самых примитивных организмов имеют размер свыше 500 000 п. н. Как же их секвенируют? Для этого разработано несколько стратегий. Не будем вдаваться в подробности, но укажем на некоторые основные из них.

Секвенирование простых геномов

Для определения нуклеотидной последовательности (т. е. первичной структуры) конкретного района ДНК в первую очередь необходимо упростить ее, что достигается путем разрезания ее на относительно короткие фрагменты. Сделать это можно, например, с помощью специальных «скальпелей» для ДНК — ферментов рестриктаз, о которых уже шла речь выше.

При секвенировании простейших организмов, у которых геном относительно невелик, обычно используют процедуру, называемую условно «сверху вниз». Вся ДНК «разрезают» на кусочки с помощью уже упоминавшихся выше ферментов рестриктаз, затем секвенируют эти кусочки по отдельности, а после «склеивают» из них полный геном. «Склеивание» оригинала осуществляется за счет того, что нуклеотидные последовательности разных кусочков перекрываются друг с другом, т. е. одинаковы по концам. Эта методология получила название «дробовика». Суть данной процедуры

отражена на рис. 13.

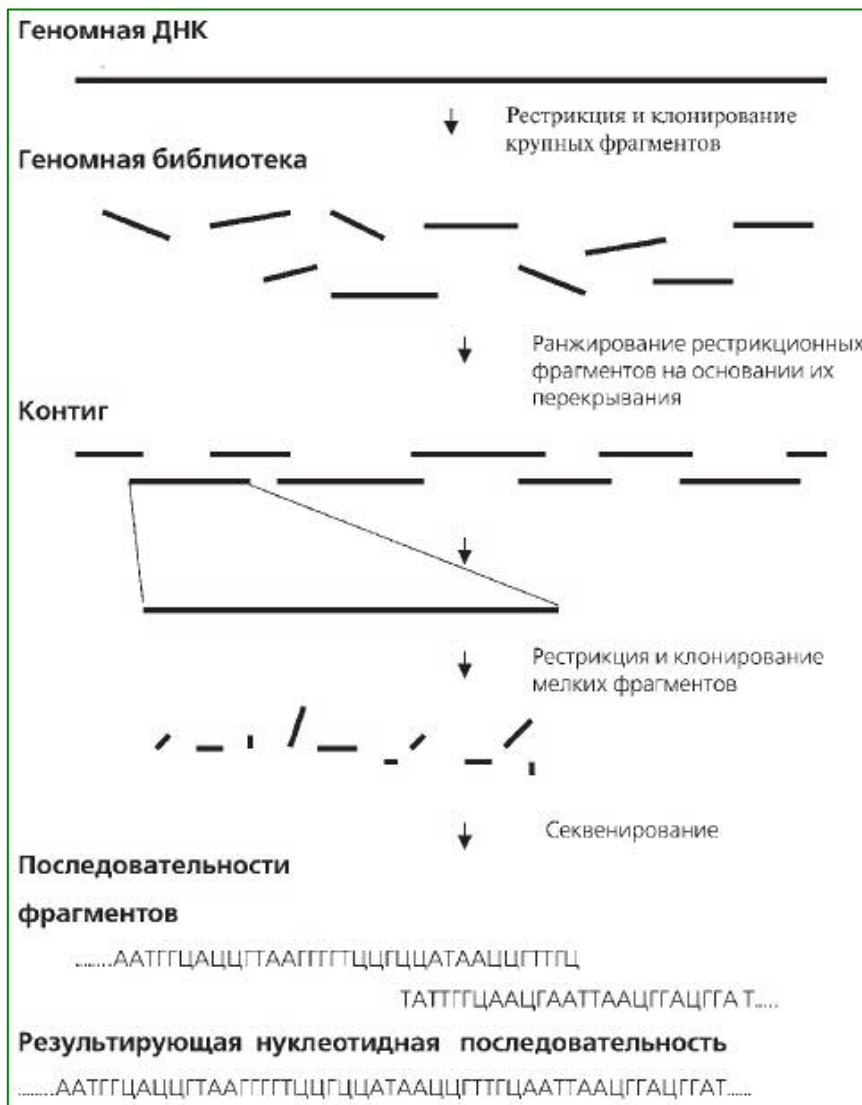


Рис. 13. Схема стратегии «дробовика», используемая для секвенирования больших молекул ДНК

Однако в случае такого очень сложного генома, как геном человека, начали с другого, а именно с определения положения известных генов и других генетических маркеров на отдельных хромосомах, то есть с генетического картирования генома. Подобную задачу генетическими экспериментами на более простых объектах пытался решить еще Т. Морган, который за свои работы получил Нобелевскую премию в 1933 году. Теперь появились более эффективные методы. Один из них, называемый методом «радиационных гибридов», заключается в следующем. Клетки человека, растущие вне организма в питательной среде, облучают рентгеном, что приводит к гибели клеток в результате разрыва хромосомной ДНК на куски, содержащие от 2,5 до 25 млн. п. н.

Но прежде, чем убитые облучением клетки распадутся, их сливают с клетками хомяка, в результате чего в разные клетки хомячка попадают разные наборы фрагментов ДНК человека. Затем «гибридные» клетки размножают в культуре, при этом в них наряду с собственной ДНК реплицируются и фрагменты чужеродной ДНК. Затем определяют состав известных генов и других генетических маркеров в каждой клеточной линии и, статистически обработав полученные данные, выводят наиболее вероятное их взаимное расположение в хромосомах. В качестве генетических маркеров использовали как гены, так и фрагменты ДНК с неизвестной функцией. Для картирования хромосом важным свойством маркеров являлся их полиморфизм, т. е. существование разных форм среди индивидуумов.

Так были построены первые генетические карты генома человека, на которых первоначально были отмечены различные генетические маркеры, отстоящие друг от друга на расстоянии не более 2 миллионов нуклеотидных пар (млн. н. п.).

Затем были составлены физические карты хромосом: первоначально с разрешением 0,1 млн. н. п., а затем 0,001 млн. н. п. Для этой цели на первом этапе применяли методы окрашивания хромосом и гибридизации с хромосомами *in situ*. И лишь позднее использовали рестриктазы. С помощью этих удивительно точно работающих ферментов «дробили» ДНК по строго определенным участкам на миллионы перекрывающихся между собой по нуклеотидной последовательности фрагментов, «разбирали» каждый из них по отдельности, после чего из них «склеивали» оригинал. «Склеивание» проводили на основании перекрывания фрагментов по нуклеотидной последовательности. Так постепенно шли все выше и выше.

Поэтому данная стратегия получила название «снизу вверх». Со всей очевидностью можно сказать, что решалась грандиозная по масштабам и сложности задача. И решить ее ученые смогли только с помощью суперкомпьютеров. В результате были созданы физические карты разных областей ДНК и целых хромосом, состоящие из последовательно перекрывающихся друг с другом фрагментов. Набор таких «родственных» фрагментов получил название контиг (рис. 16).

А далее наступила очередь секвенирования отдельных рестрикционных фрагментов. Это привело в конечном итоге к построению секвенсовых карт, на которых степень разрешения была доведена до своего максимального значения. Если 20 лет назад расшифровка нуклеотидной последовательности ДНК длиной 1000 п. н. считалась большим научным достижением (за это можно было сразу

получить степень доктора наук), то уже к 1990 году секвенирование ДНК стало массовой технологией. А сейчас квалифицированный лаборант проделывает такую работу всего за несколько часов. После появления эффективных методов секвенирования ДНК и нескольких стратегий использования этого метода стал стремительно нарастать вал расшифрованных нуклеотидных последовательностей, в первую очередь таких простых организмов, как вирусы, а также отдельных клонированных фрагментов ДНК различных видов высших организмов. Так, еще в конце 1970-х годов была расшифрована структура первого живого объекта — вируса бактерий — бактериофага, обозначаемого как ?? 174, имеющего длину 5386 п. н. Затем последовала очередь других.

В конце 80-х гг. были начаты крупные международные проекты, целью которых было полное секвенирование геномов бактерий, грибов, дрожжей, дрозофилы, мыши, пшеницы, человека и др. Первоначально была определена первичная структура ДНК микроорганизмов с размерами генома до 20 млн. п. н., к концу 1998 г. их число составило уже 18 (см. табл. 2). Наименьший размер генома у свободно живущих организмов (например, у бактерии *Mycoplasma genitalium*) составляет лишь около 600 000 п. н. У этой бактерии содержится всего 500 генов, причем 150 из них (если их удалять по одиночке) никак не влияют на её жизнеспособность. Так возникло предположение, что элементарная «машина жизни» теоретически должна работать при наличии всего лишь 350 генов. Первые объекты для секвенирования были выбраны неслучайно. Большинство из этих микроорганизмов (архебактерии, спирохеты, хламидобактерии, кишечная палочка, возбудители пневмоний, сифилиса, гемофилии, метанобразующие бактерии, микоплазмы, риккетсии, цианобактерии) способны вызывать различные патологии у человека. В настоящее время многие из этих проектов уже завершены; исследовано свыше 800 полных клеточных геномов микоплазм, кишечной палочки, возбудителей ряда болезней человека, а также пекарских дрожжей, маленького червя-нематоды, дрозофилы и весьма интересного в практическом плане растения арабидопсиса. Весьма вероятно, что истинное число секвенированных к настоящему времени геномов гораздо больше, потому что многие фармацевтические фирмы засекречивают свои результаты, не публикуя их в открытой печати.

Таблица 2. Некоторые микроорганизмы, геномы которых полностью секвенированы к настоящему времени

Организм (латинское видовое название)	Размер генома (млн. п. н.)	Количество генов
<i>Archaeoglobus fulgidus</i>	2,18	2436
<i>Aquifex aeolicus</i>	1,55	1512
<i>Bacillus subtilis</i>	4,20	4100
<i>Borrelia burgdorferi</i>	1,44	853
<i>Chlamydia trachomatis</i>	1,06	895
<i>Escherichia coli K-12</i>	4,60	4288
<i>Haemophilus influenzae</i>	1,83	1740
<i>Helicobacter pylori 199</i>	1,64	1495
<i>Helicobacter pylori 26695</i>	1,67	1590
<i>Methanobacterium thermoautotrophicum</i>	1,75	1855
<i>Methanococcus jannaschi</i>	1,66	1738
<i>Mycobacterium tuberculosis</i>	4,40	4033
<i>Mycoplasma genitalium</i>	0,58	468
<i>Mycoplasma pneumoniae</i>	0,82	716
<i>Pyrococcus horikoshii</i>	1,74	2061
<i>Rickettsia prowazekii</i>	1,12	834
<i>Synechocystis sp.</i>	3,57	3168
<i>Treponema pallidum</i>	1,14	1041

Напомним еще раз, как все это начиналось и происходило. В 1988 году был создан проект «Геном человека» Национального института здоровья США. Главой работ по полному секвенированию генома человека стал нобелевский лауреат Джеймс Уотсон. Одновременно в России с этой же идеей выступил академик А. А. Баев. Государственная программа «Геном человека» была принята в этом же году в Советском Союзе. Позднее к программе подключились другие страны, и широкомасштабные координированные исследования стали проводиться под эгидой международной организации Human Genom Organisation (HUGO). На работу по секвенированию генома человека, осуществленную этим мировым консорциумом, было потрачено в сумме более 3 миллиардов долларов. В России членами этой организации являются сегодня около 80 российских ученых (председатель Российского совета по геному человека — академик Л. Л. Киселев). Из 24-х хромосом генома человека ученые России в основном концентрировались на 3-й, 13-й и 19-й хромосомах. При этом непосредственно секвенированием ДНК они занимались очень мало, основное внимание было обращено на структурно-функциональные исследования генома. В материальном отношении наш вклад был настолько скромным, что сейчас при общих подсчетах его вообще не учитывают. Первоначально огромный объем предстоящей работы очень пугал многих исследователей. И действительно ситуация на первых этапах не была слишком обнадеживающей. В первые 2 года работы проекта «Геном человека» скорости секвенирования были еще очень низкими. Для полного завершения работы такими темпами потребовалось бы около 100 лет. Это удручало одних исследователей, но не останавливало других. Последним стало очевидно, что решение данной задачи невыполнимо традиционными методами и необходим поиск новых подходов и технологий расшифровки нуклеотидных последовательностей фрагментов ДНК, а также создание новой вычислительной техники и оригинальных компьютерных программ. Все это, конечно, было невыполнимо в рамках отдельно взятого государства. По этой причине к проекту были привлечены около 20 различных стран, организован международный банк данных, куда поставлялась вся информация, полученная в процессе решения данной задачи. В результате технических новаций постепенно возрастала производительность аппаратуры, стали использовать промышленные роботы, многие процессы были автоматизированы. И вскоре скорость секвенирования достигла 10 млн. нуклеотидных пар в сутки! Теперь уже выполнение проекта стало реальностью. За 1995 г. длина участков ДНК человека с установленной последовательностью нуклеотидов увеличилась почти в 10 раз, но это составляло еще менее 0,001% от всего генома человека. К началу 1998 г. было секвенировано всего около 3% генома.

Правила образования информационных пакетов

Анализ генома человека позволил на конец 2002 года обнаружить порядка 30 000–40 000 генов, кодирующих белки. Эти гены сильно отличаются друг от друга по размерам. Подсчитано, что средняя длина гена у человека составляет около 27 тыс. п. н. Такой усредненный ген содержит 9 экзонов (средний размер каждого около 150 п. н.) и 8 интронов (средний размер каждого чуть меньше 3400 п.н.). Но это лишь средние значения. Самые короткие гены содержат всего два десятка букв-нуклеотидов, например, гены эндорфинов — полипептидов, вызывающих ощущение удовольствия. В то же время самый длинный ген, кодирующий один из белков мышц, — миодистрофин содержит 2,4 млн. п. н.

Кроме того, гены не равномерно распределены между разными хромосомами. При средней плотности один ген на 100 т. п. н. генома их содержание в хромосоме 19 составляет около 2 на 100 т. п.н., а в хромосомах 2, 13 и Y-хромосоме — от 0,15 до 0,7. Если сравнить плотность генов с плотностью расселения людей, то Y-хромосома напоминает нашу Сибирь, а хромосома 19 — Европейскую часть России. Для сравнения: в геноме бактерий содержится свыше 1000 генов на 1,0 млн. п. н., у дрожжей около 450 генов на 1,0 млн. п. н., а у червя *C. elegans* — около 200. Следовательно,

плотность расположения генов на единицу длины генома заметно падает по мере эволюционного усложнения организмов.

После секвенирования генома человека и ряда других организмов был проведен детальный компьютерный анализ нуклеотидных последовательностей. Подсчитано, например, сколько в ДНК может идти подряд букв А, или как часто Г встречается после Ц. И тут выяснилось, что в генах эти сочетания подчиняются определенным правилам, тогда как в промежутках между генами, там, где ничего существенного в ДНК не записано, частота сочетаний разных нуклеотидов близка к случайной. Где много генов, там много букв Г и Ц, а где генов нет, там много букв А и Т. Здесь опять можно провести параллель с обычным текстом. По правилам грамматики русского языка, которые мы учили в школе, «жи, ши — пиши через и», в словах после букв ж и ш буква ы не встречается. Нечто подобное характерно и для генетических текстов — в кодирующих участках некоторые сочетания нуклеотидов практически не встречаются, а распределение других сочетаний сильно отличается от случайного. Это был один из первых, но далеко не единственный результат анализа ДНК.

Как в письменном языке пунктуация (точки, запятые, тире и т.д.) используется для членения и графической организации текста, так и в ДНКовом тексте существуют всевозможные знаки-сигналы, обозначающие границы отдельных функциональных элементов. Без таких сигналов генетическая информация не может быть правильно считана с носителя. Классический пример из лингвистики — предложение «Убить нельзя помиловать». Если не поставить запятую, то остается непонятным, то ли убить, то ли помиловать. Так и в гене. Знаки препинания, в отличие от генетического кода, не так универсальны в живом мире. Часть из них у низших организмов (прокариот) существенно отличаются по сравнению с высшими (эукариотами), включая человека.

Вот как выглядит, например, одна из страниц ДНКового текста, содержащаяся в томе (хромосоме) 11 человека (рис. 14).

```

1  ГЦГЦЦЦЦТТА ТГТАЩЦАЩА ААААТЦТАТТ ТЦААААААГ ТТЦЦЦЦТААГ ААТАГАТТА
61  ТЦААГТТААГ ТААААГТЦА АТАГЦЦТТТТ ААТТТААТТТ ТТААТГТТТТ ТАТЦАТЦЦТТ
121  ТГЦААТААТА АААЦАГТААЦ ТТТАТАЦТТТ ТТААТТТААТ ГТАТАГААТА ГАГАТАТАЩА
181  ТАГТАТАТТТ АААТАГАТАЦ АЦАГТТГТАТА ТТТРАТТААА АТАТАААТГГГ АГАТТЦААТЦ
241  АГАААААААТ ТТЦТАААААГ ГЦЦЦГГГГТТ ААААГАГГАА ГГАААЩААТА АТГААААААА
301  ТТТТТТГАГА ААААЩАГЦТТ ААААЩЦААТТ ТАААГАТТТТ АТАААГАААГ ЦАААААГАГА
361  АГТАГААААТ ААЩАЩАГТТТ ЦАТТТТГААА АТГАААЩА ГТАТТГТЦЩЦ ТАТТТААГЦ
421  ТАГТЦАЩААА ГЦААГТТЦТТ ЦАГАГААЩЦТ ГГАГЦЦТААГ ГТТТАГТЦЦЦ АЦЦАТТТЦА
481  АЦЦАГТЦТАГ ЦАГЦАТЦГЦ ААЦАТЦТАЩА АТГГЦЦТТГА ЦЦТТТГЦТТТ АЦТТТТГЦЦЦ
541  ЦТЦЦТТГТГЦ ТЦАГЦГЦАА ГЦААГЦТГЦ ЦТТТТТГЦТ ГТРАЦТГЦЦ ТЦАААЩЦАЦ
601  АГЦЦТТГТГА ГЦАГГАГТАЦ ЦТГГАТГЦЦ ЦТГЦАЩАГА ГГАГГАГААТ ЦТЦЦТТТЦЦ
661  ТЦЦТГЦТТГА АГГАЩАЩАЩА ТГАЦТТТГГА ТТЦЦЦЦААГГ АГГАГТТТТГ ЦААЩАГТЦЦ
721  ЦАААААГЦТТ АААЩАТЦЦЦ ТГЦЦЦЦАТ ГАГАТГАТЦЦ АГЦАГАТЦТТ ЦААТЦТЦТЦ
781  АГЦАЩААААГ АЦТЦАТЦГЦ ТГТТТТГГАТ ГАГАЩЦЦЦЦ ТАГАЩАААТТ ЦТАЩАЦТГАА
841  ЦТЦТАЩЦАГЦ АГЦТГААТГА ЦЦТТГААГЦЦ ТГТТГГАТАЦ АГТГТТТТТ ГТГАЩАГАГ
901  АЦЦЦЦЦЦТГА ТГААГГАГГА ЦТЦАТТЦТТ ГЦТТТГАГГА ААТАЦТЦЦА ААГААТЦАЦТ
961  ЦТЦТАЩЦТГА ААГАГАГАА АТАЩАГЦЦЦТ ТГТЦЦЦТТТ АГТТТТЦАГ АГЦАГАААТЦ
1021  АГТАГАТЦТТ ТТЦТТТГТЦ ААЩАААЦТТ ЦААГАААГТТ ГААГААГТАА ГГААТГАААА
1081  ЦТГТТЦААЦ АТГААААТГА ТТТЦАТТГА ТЦГТАТГЦЦ АЦЦЦАЩЦТТ ТТТАТГАТЦТ
1141  ГЦЦАТТТЦАА АГАЦЦАТТТ ТТЦТЦТААГ АЩАТГАЩАЦ ГАТТТАААТЦ ТТТЦАААТГ
1201  ТТТТТАГГАГ ТАТГААЩАА ЦАТТГТАТЦ АГЦЦТТААГ ГЦАЦГАРЦЦ ЦТТАЩАГАТГ
1261  АЦЦАТЦЦТГА ЦТГАТЦАТТ АЦЦАТТТГА АТАТТТТГАА ААТАТТАТТТ АТТТААЦТАГ
1321  ТТАТААААЩА АЦТГАТТТТТ ГТЦАТАТГА ТГЦАТТГЦЦ АЩЦТТГЦАЦ АГТТТТТААТ
1381  ГТААТААААТ ГТГТЦЦТТТ ТАТТТТГТАА АТТТАТТТТ ГТТГТТЦАТ ГГААЦТТТТГ
1441  ЦТАТГГААЦТ ТТГТЦАТТТ ТТТАТЦТТТ ААААТГАААТ ТЦАААГЦЦА АТТТГЦААЦ
1501  ЦТГАТТАЩАГ ААТААЦТТТТ АЦАЦТЦАТТ ТГТЦАТЦАА ТАТТАТАТТЦ ААГАТАТААГ
1561  ТАААААТААА ЦТТТЦТГТАА АЦАААГТТТ АТТТТТАЦТ ЦААГАТААЩА ГТТГГААЦЦТ
1621  ААЩАААТАЩА АТЦЦГЦЦЦЦ ЦТТТТТТАТТ ТГАТТТТТТТ АТГААААААА ЦТАААААТТГ
    
```

Рис. 14 Последовательность нуклеотидов (фрагмент) ДНК человека

Ген интерферона выделен курсивом. Подчеркнут участок, кодирующий белок, состоящий из 188 аминокислот. Жирным шрифтом выделены знаки пунктуации в этом гене, обеспечивающие его правильную транскрипцию с образованием мРНК и последующую правильную трансляцию этой мРНК.

Глядя на этот рисунок, даже специалисту сразу трудно что-то понять. Однако простая компьютерная программа быстро позволяет найти здесь участок, кодирующий белок интерферон. Это осуществляется в результате того, что в генах имеются специальные сигналы-ярлычки, которые можно рассматривать как знаки препинания в ДНКовом тексте. И компьютер способен их различить. Рассмотрим главные из них.

Гены построены из кусков

Успехи в изучении генома человека, а еще ранее — проведенный анализ отдельных генов других высших организмов, привели к ряду принципиально новых и совершенно неожиданных для генетиков результатов в отношении устройства генов, особенно тех, которые кодируют белки. Со времени появления центральной догмы молекулярной биологии думали, что непрерывная линейная последовательность нуклеотидов в ДНК с помощью триплетов нуклеотидов кодирует непрерывную последовательность нуклеотидов в мРНК, а последняя как посредник передает эту информацию на рибосомы, которые, считывая ее, осуществляют синтез линейных молекул белков. Это соответствовало ситуации, которую первоначально наблюдали при изучении генов у бактерий. Однако у высших организмов, включая человека, (и даже у некоторых низших) картина оказалась более сложной.

Когда сравнивали мРНК и соответствующий белок, то действительно обнаруживали непрерывность их нуклеотидных и аминокислотных последовательностей. То же наблюдали и при сопоставлении последовательностей аминокислот в белках и нуклеотидов в геномах вирусов. Однако с появлением возможности секвенировать протяженные участки ДНК (т. е. непосредственно генов) как у некоторых вирусов, так и у высших организмов получили совершенно неожиданные результаты. Внутри нуклеотидных последовательностей ДНК, кодирующих белок, были обнаружены какие-то иные, ничего не кодирующие участки.

Оказалось, что многие гены «расчленены» на отдельные куски. Одни из этих кусков, как и положено, кодируют белок (их назвали экзонами), а другие никакие белки кодировать не способны (их назвали интронами) и расположены между экзонами. Таким образом, в гене кодирующие и не кодирующие участки перемежаются друг с другом (рис. 15). Так устроено большинство белок-кодирующих генов. Хотя имеются и исключения. В этих случаях говорят, что ген устроен из одного экзона и интронов не содержит.



Рис. 15. Схематическое изображение экзон-интронного строения некоторых генов человека.

Эзоны — прямоугольники, интроны — линии между прямоугольниками. Темные прямоугольники — участки генов, кодирующие белок, белые прямоугольники в первом и последнем экзонах — специальные участки гена, кодирующие мРНК, но не кодирующие белок

Такое разорванное (экзон-интронное) строение оказалось характерным для подавляющего большинства (хотя и не для всех) генов человека. Вместе с тем у большинства простейших организмов интроны не обнаружены.

Размеры интронов часто значительно превышают размеры экзонов, что существенно увеличивает общую длину нуклеотидных последовательностей ДНК, образующих ген. При этом в геноме человека интроны в целом существенно длиннее, чем у других организмов.

Каким же образом разорванный ДНКовый текст реализуется в клетках в виде неразорванного РНКового, а затем белкового?

1. Выяснилось, что первоначально при транскрипции гена синтезируется большой РНК-предшественник, копия ДНКового текста (с экзонами и интронами вместе).
2. Далее в результате работы специфических ферментов происходит нарезание РНК-предшественника на куски. Те из них, которые ничего не кодируют (интроны), отбрасываются, а кодирующие куски (экзоны) соединяются между собой, обычно в том же порядке, как они располагались исходно в ДНК.
3. В результате формируется зрелая молекула мРНК. Этот сложный процесс и получил название сплайсинга (рис. 16).

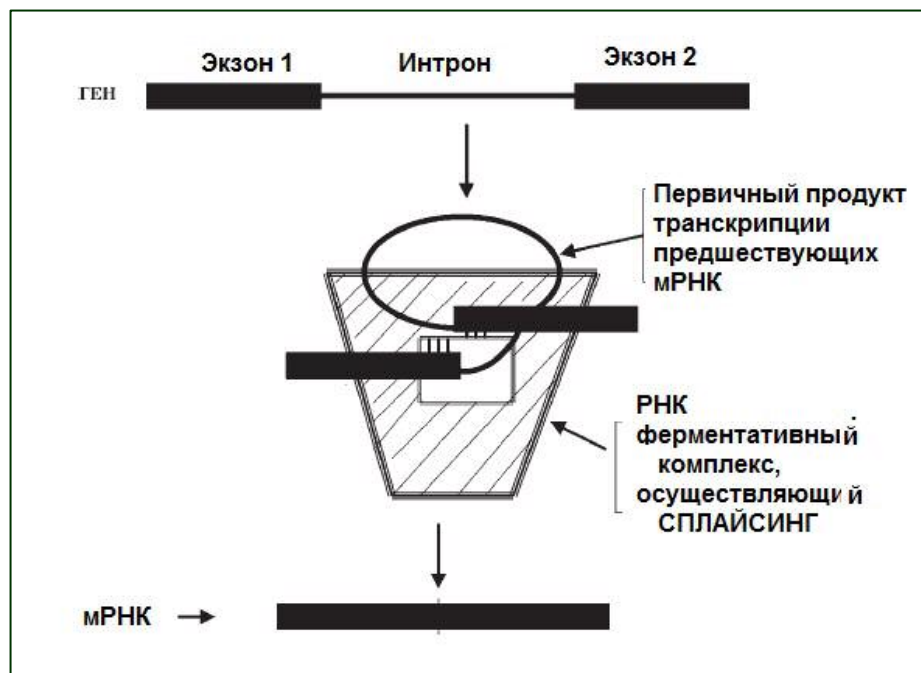


Рис. 16. Образование молекул мРНК на РНК-предшественнике в результате сплайсинга.

Из РНК-предшественника вырезаются фрагменты, синтезированные с интронных последовательностей гена, а участки, синтезированные на экзонах, соединяются друг с другом, что в конечном итоге приводит к формированию зрелой функционально активной мРНК

В интронах содержатся особые сигналы, которые узнаются специальными ферментами, осуществляющими сплайсинг. Так, в 98% случаев интроны начинаются с динуклеотида ГТ, а заканчиваются динуклеотидом АГ. Они-то и служат главными сигналами для правильного осуществления сплайсинга.

Во всем этом и состоит основное ноу-хау в организации генов человека и других эукариотических генов. «Лоскутное» устройство большинства генов — чрезвычайно важное эволюционное приобретение высших организмов. За счет сплайсинга в РНК может происходить соединение не только между соседними экзонными нуклеотидными последовательностями, но и между другими, отстоящими порой в гене на значительном расстоянии. Это называют альтернативным сплайсингом (рис. 17). В геноме человека альтернативный сплайсинг характерен для более трети генов. Этому можно привести следующую аналогию: из слова «администрация» за счет удаления отдельных букв и слогов можно образовать множество совершенно разных по смыслу слов: ад, министр, амнистия, астра, нация.

Существование множества альтернативных вариантов сплайсинга в конечном итоге обеспечивает появление более чем одного белка при экспрессии одного единственного гена. В частности, на гене белка тропонина человека, содержащем 18 экзонов, за счет альтернативного сплайсинга может образовываться 64 различных продукта! Сейчас подсчитано, что в среднем один ген у человека способен кодировать около 3-х разных белков. Альтернативный сплайсинг зависит от множества внутриклеточных и внешних факторов, в том числе, как показано недавно, даже от вирусного воздействия на клетку.

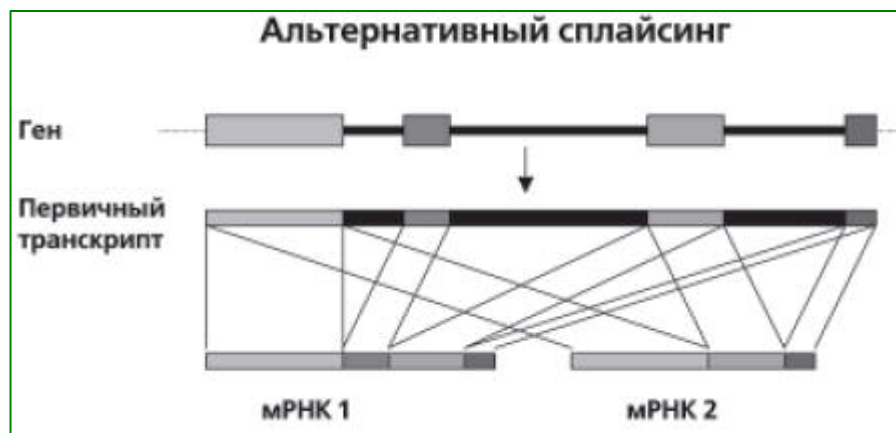


Рис. 17. Схематическое изображение альтернативного сплайсинга. Прямоугольниками разного оттенка обозначены экзоны, черными линиями — интроны. Показаны лишь две из множества возможных мРНК, образующихся на одном гене

Экзон-интронное строение генов дало принципиально новую возможность для эволюции генов — комбинировать отдельные элементы при формировании генов. В результате этого в эволюции возникали новые гены, кодирующие разные белки за счет разного сочетания экзонов одного и того же гена. То есть из уже готовых «слов» составляются совершенно разные «предложения». И такой подход, используемый природой, оказался весьма продуктивным.

Следует отметить, что и в зрелой мРНК, которая образуется уже после удаления из первичного транскрипта интронных нуклеотидных последовательностей, также не все участки кодируют бе-

лок. Однако, в отличие от интронов, такие последовательности нуклеотидов расположены в начале и в конце молекулы (см. рис. 18). Их размеры также существенно различаются.

Так, в разных мРНК человека некодирующая область, расположенная в начале молекулы, имеет длину от 2800 до 18 нуклеотидов (в среднем 200 нуклеотидов), а та, которая расположена в конце, составляет от 8500 до 20 нуклеотидов (в среднем 1000 нуклеотидов). Эти некодирующие белок области гена очень важны для нормального функционирования мРНК, ее трансляции, стабильности и локализации в клетке. Как уже говорилось, на одном гене за счет альтернативного сплайсинга может образовываться несколько мРНК. Следует подчеркнуть, что и на одной уже сформированной мРНК могут образовываться разные белки. Это происходит за счет наличия в 50% мРНК человека не одного, а двух АУГ-кодонов, являющихся старт-сигналами для синтеза белка на мРНК. За счет этого, в разных ситуациях клетка может начать синтез белка с разных мест мРНК, в результате чего образуются разные по длине, а порой и по своим свойствам, белки. Опять же пример из лингвистики. Если читать текст не с первой, а со второй буквы, смысл зачастую совершенно меняется (сравните: удочка и дочка).

В связи со всем этим понятно, что классическая формула генетиков «один ген — один белок», предложенная лауреатами Нобелевской премии Джоржем Бидлом и Эдуардом Татумом, оказывается в реальности справедливой лишь для небольшой группы генов, а многие гены (возможно, большинство) кодируют семейства родственных, но существенно отличающихся белков, то есть нередко действует принцип «один ген — много белков». За счет этого в геноме осуществляется принцип экономии генетического материала (компактная запись) и достигается большой набор генных продуктов — белков — на относительно небольшом числе генов.

В таблице 3 приведена общая характеристика белок-кодирующих генов, выявленных в результате секвенирования генома человека.

Таблица 3. Усредненные данные о структуре белок-кодирующих генов генома человека

Характеристика генов	Средние значения
Число экзонов	8,8
Размер интронов	3365 п. н.
Длина кодирующей последовательности	1340 п. н.
Размер	27000 п. н.

В реальности размеры разных генов, а также их экзонов и интронов сильно отличаются. И основной вклад в эту вариабельность вносят интроны. Уже говорилось, что максимальный по размерам ген дистрофина занимает в геноме человека 2,4 млн. п. н. Подавляющая часть занята именно интронами (их средний размер составляет несколько десятков тысяч п. н.). Наибольшую кодирующую последовательность в геноме человека имеет ген по имени титин. Его размер равен около 81000 п. н. Этот ген — чемпион и по числу интронов (178 штук!), и по длине единичного экзона (17106 п. н.).

Важно отметить, что разбивка генов на отдельные экзоны неслучайна. Индивидуальный экзон часто кодирует не просто какую-то часть белковой молекулы, а определенный фрагмент белка, выполняющий в целом белке отдельную функцию. Такой отдельный структурно-функциональный фрагмент, входящий в состав целого белка, называют доменом.

Перекрытие генных текстов

Принцип компактности записи информации в ДНК проявляется еще в одном. Важный факт, обнаруженный первоначально при анализе последовательностей генов у вирусов, заключается в том, что в одном и том же участке ДНК иногда может быть записана информация о двух совершенно разных РНК (и, соответственно, о совершенно разных белках). Такое случается и в геноме высших

организмов, включая человека. Как же это осуществляется? Вспомним, что записанный в ДНК текст явно не разбит на слова (буквы — нуклеотиды — следуют в молекуле ДНК одна за другой). Но слова в этом тексте есть, просто знаки пунктуации и разрывы между словами скрыты от глаза и узнаются клеткой после перекодировки информации из ДНК в РНК. Для большей наглядности приведем широко известный пример такого текста:

НАПОЛЕОНКОСИЛТРАВУПОЛЯКИПЕЛИСОЛОВЬЯМИ.

В зависимости от разбивки этого текста на слова получается два предложения с совершенно разным значением:

НА ПОЛЕ ОН КОСИЛ ТРАВУ ПОЛЯ КИПЕЛИ
СОЛОВЬЯМИ

НАПОЛЕОН КОСИЛ ТРАВУ ПОЛЯКИ ПЕЛИ
СОЛОВЬЯМИ.

Теоретически на двух цепях ДНК с учетом того, что код триплетный, можно записать 6 разных текстов: три на одной цепи и три на другой. Это может быть достигнуто за счет сдвига точки начала считывания информации, что называют сдвигом рамки считывания (начало считывания возможно с любого из 3-х разных рядом лежащих нуклеотидов) (см. рис. 18). Такое явление получило название перекрывания генов. Впервые оно было обнаружено у вирусов. И здесь было понятно, зачем это потребовалось. Вирусные геномы обычно очень маленькие. Перекрывание генов диктовалось необходимостью как можно компактнее разместить информацию на относительно коротких молекулах ДНК. Большой геном человека теоретически позволяет обойтись без такого перекрывания. Однако и в геноме человека, хотя и относительно редко, перекрывание генов также имеется.

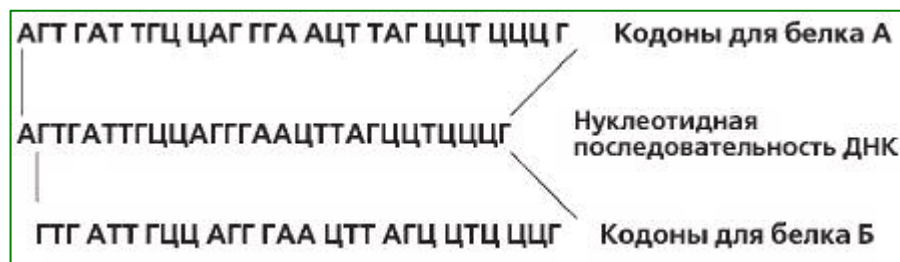


Рис. 21. Схематическое изображение перекрывания двух ДНК-текстов, записанных на одном участке ДНК.

Белок А образуется на мРНК, которая транслируется с одних кодонов, а белок Б образуется на мРНК, читаемой (транслируемой) со сдвигом в один нуклеотид. В результате кодоны (они разделены пробелами) совершенно разные, и, соответственно, при трансляции мРНК образуются совершенно разные белки

Каков может быть смысл в сохранении этого феномена в геноме человека? Пока еще не все ясно, но кое-какие моменты уже проясняются. Например, в ряде случаев обнаружили считывание РНК с разных цепей одного участка ДНК. Это приводит к образованию таких РНК, которые в силу комплементарности могут взаимодействовать друг с другом в клетке, образуя в результате двунитевые РНК. А в дву-нитевом виде мРНК не способна к такому важному процессу, как трансляция. Таков реальный механизм специфической регуляции (инактивации) экспрессии генов человека, число которых, по последним оценкам, составляет ~ 1600.

Избирательное кодирование

Не все гены кодируют белки. Прежде всего, следует отметить, что кроме генов, кодирующих белки, в геноме имеются еще гены, на которых синтезируется РНК, которые не являются мРНК (то есть, не кодируют белок), но выполняют ряд самостоятельных важных функций в клетках. В результате получается ситуация, что традиционное привычное определение гена надо расширять, включив в него гены, кодирующие белки, и гены, не кодирующие белки, но кодирующие функционально значимые РНК.

Давно уже известно, что кроме мРНК на ДНК синтезируются разнообразные вспомогательные РНК, которые сами не транслируются (т. е. они не кодируют белки), но участвуют в разных клеточных процессах. В первую очередь это РНК «домашнего хозяйства» — рибосомные РНК (рРНК), транспортные РНК (тРНК) и др., которые участвуют непосредственно в синтезе белка на мРНК. Эти РНК кодируются участками, которые также называются генами, и составляют основную массу РНК в клетках. рРНК входят в состав рибосом, являясь важным компонентом их структурной организации. У человека присутствует два основных типа рРНК размером около 1900 нуклеотидов и более 5000 нуклеотидов в малой и большой субчастицах соответственно.

Еще одни некодирующие белок РНК — тРНК, которые обеспечивают аппарат трансляции, подтаскивая к рибосомам различные аминокислоты, вступающие в реакцию друг с другом. В геноме человека гены, кодирующие рРНК и тРНК, представлены многочисленными копиями. Так, в секвенированной ДНК человека обнаружено около 500 генов, кодирующих тРНК. Генов для рРНК в геноме человека выявлено около 200 копий, которые расположены на 5-ти разных хромосомах (13, 14, 15, 21 и 22).

Изучение свойств РНК привело к тому, что представление об исключительности белков в катализе биохимических реакций пришел конец. Выяснилось, что в природе имеются виды РНК, которые, подобно белкам, обладают высокоспецифической каталитической активностью.

Очень важным стало обнаружение в геноме человека множества других генов, также производящих РНК, но не способных кодировать белок. Постепенно выяснилось, что некоторые из таких не кодирующих белки РНК принимают участие в важнейших процессах, происходящих в клетке: регуляции транскрипции ДНК, сплайсинга и трансляции мРНК, модулировании функций белков и их пространственного распределения в клетке. По этой причине их называли риборегуляторами. И примеров таких риборегуляторов уже сейчас можно привести немало. Так, установлено, что не кодирующий белок участок гена H19 имеет отношение к ряду процессов, протекающих в клетках, и, в частности, к их злокачественному перерождению.

Другой РНК-кодирующий ген контролирует работу белок-кодирующего гена HFE, вовлеченного в метаболизм железа и связанного с наследственным заболеванием хемохроматозом. В последнем случае РНК-продукт кодируется тем же самым геном, который он регулирует, но его образование осуществляется на другой нити ДНК. В результате так называемый антисмысловый (комплементарный) РНК-продукт способен взаимодействовать с мРНК, образуя гибриды, неспособные транслироваться в рибосомах с образованием белка. Еще один интересный ген — ген РНК-активатора стероидного рецептора. Он обеспечивает активность стероидных рецепторов за счет образования комплекса с этим белком.

В клетках человека, как и у других организмов, выявлены короткие двунитевые РНК (микроРНК), отдельные из которых, по-видимому, могут участвовать в процессе регуляции экспрессии генов через механизм, названный РНК-интерференцией.

Этот механизм впервые был обнаружен в 1998 году у низших организмов. Различные микроРНК в клетках червя *C. elegans* оказались в состоянии «приглушать» работу строго определенных генов

путем воздействия на процесс синтеза кодируемых ими белков. Вполне вероятно, и у человека микроРНК играют подобную роль.

Таким образом, мы видим, что продукты генов, кодирующих только РНК, вмешиваются в различные клеточные процессы, используя при этом совершенно разные механизмы. Гены рибосомных регуляторов составляют, по-видимому, заметную часть генома человека. Примерные оценки говорят о величине на порядок больше, чем доля белок-кодирующих нуклеотидных последовательностей. Сам факт существования таких генов, неспособных кодировать белок, но реально проявляющих себя в производстве функционирующих в клетках РНК, ставит большой вопрос перед исследователями генома. И, в первую очередь, что следует после этого считать собственно геном?

Генная матрешка

Иногда обнаруживаются варианты, когда внутри одного гена целиком содержится другой, меньший по размерам ген. Этакая своеобразная «матрешка», построенная из генов. Такая организация генов весьма редка. Так, в хромосоме 22 имеется лишь 2 таких случая.

Чаще всего белок-кодирующий ген располагается в интроне другого белок-кодирующего гена. Но встречаются и другие варианты. В качестве примера можно привести ситуацию, имеющую место для митохондриального гена одной из рибосомных РНК. Ген, кодирующий эту рРНК, обеспечивает ею рибосомы митохондрий в качестве структурного компонента (т. е. не кодирует белок). Однако вместе с тем небольшой участок, расположенный внутри этого гена, кодирует короткий белок (полипептид), получивший название гуманин (от англ. human — человек), который принимает участие в процессе программированной клеточной гибели. То есть в РНК-кодирующем гене может содержаться белок-кодирующий ген.

Другой вариант — уже упоминавшийся выше ген H19. Здесь, наоборот, ген, кодирующий белок, содержит внутри своей кодирующей части другой более короткий ген, кодирующий только РНК, которая принимает участие в регуляции работы этого гена.

Генные семейства

Как у людей имеются семьи, так и у генов нередко существуют семейства. Семейством генов называют набор из двух или более генов, чьи экзоны родственны между собой, т. е. похожи (хотя и не идентичны) по нуклеотидной последовательности. В геноме человека присутствует около полутора тысяч таких семейств генов. Причем только около сотни из них специфичны для человека и других позвоночных животных, тогда как основная масса генных семейств имеется и у человека, и у червя.

Разные члены одного семейства генов возникали в эволюции из одного гена-предшественника (условно — от одного прапрапрадеда, как и в человеческой популяции). Процесс, в результате которого появлялись «копии», получил название дупликации, т. е. удвоения. Такие дупликации в некоторых случаях происходят не однократно, а многократно. Сейчас стало ясно, что дупликации играют очень важную роль в эволюции. Подсчитано, что в геноме человека в сумме дублировано около 3,6% нуклеотидных последовательностей размером в 1000 и более п. н.

Разные копии одного семейства генов могут располагаться в геноме рядом и следовать друг за другом (это называют тандемной дупликацией). Рассмотрим несколько примеров. Так, гены, кодирующие рибосомную РНК (рРНК) и белки гистоны, расположены в геноме человека в виде кластеров, построенных из одних и тех же последовательностей, следующих друг за другом (рис. 19).

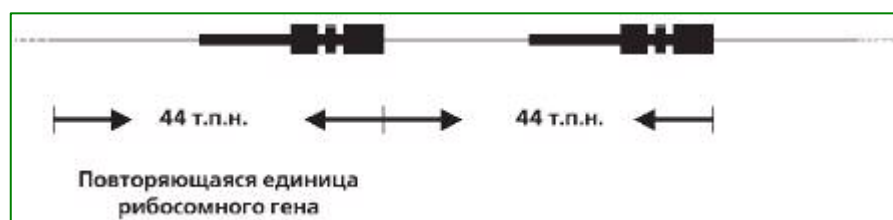


Рис. 19. Строение участка генома человека, состоящего из множества генов, кодирующих рибосомные РНК. Утолщенные линии — транскрибируемые участки

Такие «гроздьи» одинаковых генов обычно требуются организму для того, чтобы нарабатывать большие количества определенного продукта. Действительно, рРНК составляет по массе основную часть всей клеточной РНК, входя в состав таких важных «машин» клетки, как рибосомы. Огромная масса гистонов нужна клетке для упаковки гигантских молекул ДНК и плотной укладки их в хромосомы. Во всех таких случаях говорят о большой «дозе» гена в геноме, которая в конечном итоге и обеспечивает клетку большим количеством продукта.

Однако в большинстве других случаев возникающие в результате дупликации «копии» генов постепенно претерпевают изменения (мутации), что приводит к возникновению определенных различий между ними, хотя формальное родство при этом сохраняется. Такой процесс был назван дивергенцией, т. е. расхождением. Дивергенцию обычно выражает в процентах различий между двумя родственными последовательностями нуклеотидов в ДНК или последовательностями аминокислот в белках.

Однако даже после того, как разные «копии» генов одного семейства становятся не совсем похожими друг на друга, они чаще всего выполняют сходную функцию. Они остаются очень дружной семейкой! Понятие «схожести» формулируют биологи, а с точки зрения математики здесь есть о чем подумать !

Но экспрессируются разные члены одного такого семейства обычно в разное время или в разных типах клеток. Так, в геноме человека обнаружено тридцать генов, кодирующих родственные факторы роста клеток кожи (фибробластов). А вот у низших организмов число таких генов существенно меньше (у дрозофилы и червя их всего 2).

Другой пример — гены, кодирующие кератины — белки наружного слоя кожи и ее производных (волосы, ногти). Их у человека 111 штук! Но и это не предел. Геном человека, как ни удивительно, имеет около 1 000 копий генов-рецепторов обоняния.

Дивергировавшие в ходе эволюции человека гены семейства иногда оказываются разбросанными в разных местах на одной или даже разных хромосомах. При этом ген может быть удвоен целиком, а может происходить удвоение небольшого набора его экзонов или даже всего лишь одного из его экзонов.

В семействах генов встречаются иногда такие члены, которые не утруждают себя никакой работой. Нарушение их способности функционировать в клетке обычно связано с накоплением опечаток в их ДНКовом тексте. Такие гены-тунеядцы получили название «псевдогенов».

Перед их названием ставят греческую букву ?. Не совсем ясно, зачем геному нужны такие гены, почему он сохранил их в эволюции, не избавился от них. Но факт остается фактом. По имеющимся оценкам, в геноме человека имеется около 20000 таких, подобных вымершим реликтам, псевдогенов. В частности, в огромной семействе генов-рецепторов обоняния, состоящем из примерно 1000 генов, около 60% являются псевдогенами. Считается, что массивная потеря функциональных генов произошла за последние 10 млн. лет, и связано это со снижением роли обоняния у человека по сравнению с другими позвоночными организмами. Здесь, по-видимому, действует принцип от-

мирования из-за неупотребления. Установлено, что существует, по крайней мере, два механизма, приводящих к образованию псевдогенов.

Один из них заключается в удвоении уже существующего гена (дупликация), а потом дополнительная копия накапливает мутации, которые препятствуют его работе. Другой механизм связан с процессом сплайсинга и обратной транскрипции. На сформированной мРНК синтезируется ее ДНК-копия, которая в дальнейшем встраивается в геном. Такие псевдогены не содержат ни интроны, ни промоторные участки. Кроме того, в них со временем накапливаются многочисленные мутации, не позволяющие этим генам обеспечивать синтез нормального белка.

Все вышесказанное можно проиллюстрировать на примере семейства генов, кодирующих белки крови человека (рис. 20). У человека имеется два семейства глобиновых генов (альфа и бета). Эти семейства расположены на разных хромосомах. Каждое семейство состоит из нескольких членов, имеются здесь и псевдогены. Нормальные гены глобиновых семейств экспрессируются в клетках крови на разных стадиях развития человека. Одни у эмбрионов (?), другие у плода (G?, A?, ?1), а третьи — у взрослого организма (?2, ?, ?). Знаки вопроса вставлены с целью указать на отсутствие достаточно четких сведений. Имеется лишь приблизительное представление о роли этих псевдогенов.

Функция у всех глобиновых генов одинакова — синтез глобинов, участвующих в переносе кислорода клетками крови. И место их работы одно — клетки крови эритроциты. Отличаются они лишь тем, что включаются в разное время, в результате чего выполняют одну и ту же роль, но на разных стадиях развития организма.



Рис. 20. Строение двух локусов генома, состоящих из кластеров родственных глобиновых генов человека.

Экспрессирующиеся гены указаны стрелками (направление транскрипции), молчащие гены (псевдогены) — квадратиками

Организация кластеров глобиновых генов у человека совершенно одинакова с организацией этих генов у гориллы и павиана. Это указывает на то, что такая организация полностью сформировалась еще от 20 до 40 млн. лет назад и с тех пор не изменялась. Анализ мутаций в межгенных областях позволил сделать вывод об эволюции не отдельных генов, собранных в такой кластер, а кластера как единого целого супергена.

По-иному ведут себя многочисленные гены семейства актинов, которые кодируют сократительные белки клетки. Эти гены экспрессируются почти все время, и их экспрессия происходит как в мы-

шечных клетках, так и во многих других немышечных. При этом функция актинов везде одинакова.

Генные конструкции - текстовые конструкции

Экзон-интронная организация генов способствует еще одному механизму возникновения новых генов — созданию новых ДНКовых текстов из фрагментов старых. Это осуществляется в результате «тасования» уже предсуществующих экзонов, что может приводить к появлению нового их сочетания и новых белков с новыми функциями. Чаще всего это происходит за счет процесса присоединения новых экзонов к старым. Примером этому может служить случай, представленный на рис. 21.

В процессе эволюции из общего предшественника образовались гены дрозофилы *trx* и ген человека *alr*. Ген человека *hrx* возник в результате присоединения к гену *trx* четырех новых экзонов. Это приводит к появлению в белке, кодируемом этим геном, четырех новых функциональных доменов, что существенно меняет его свойства по сравнению с предшественником. Такая ситуация вообще-то характерна для всех высших организмов, но у человека она встречается в 2–5 раз чаще, чем, например, у дрозофилы или червя.

После получения всех этих новых многочисленных сведений об особенностях организации и эволюции генов у человека (сходным образом обстоит дело и у других высших организмов) стало ясно, что дать строгое определение того, что же такое собой представляет ген, довольно сложно. Но даже если бы такое определение сегодня имелось, его, как и всякое другое, не следовало бы считать незыблемым. Не зря говорил Спиноза: «Всякое определение есть ограничение». Однако в целом базовое представление о гене осталось в силе. Ген — это физическая (определенный участок ДНК) и функциональная (кодирует белок или РНК) единица наследственности.

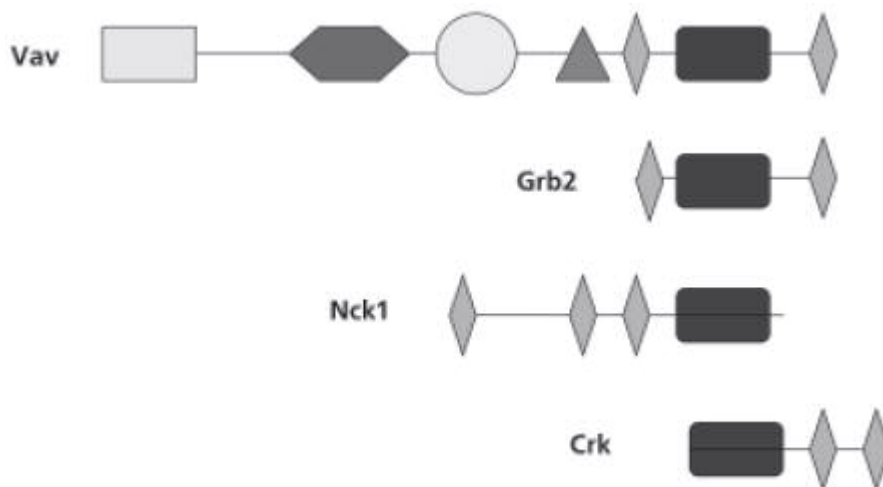


Рис. 21. Пример формирования новых генов за счет различного сочетания уже имеющихся экзонов.

Изображены четыре разных белка, состоящих из сходных блоков-доменов (обозначены разными символами), которые кодируются разными экзонами

Количество генов у человека

Это наиболее интересный вопрос, ради которого собственно и затевалось полное секвенирование генома человека. После получения основной информации о структуре генома человека в первую очередь были произведены различные анализы по поиску генов и определению их числа. Однако задача оказалась не простой. Это может показаться странным для читателя, но однозначного ответа на поставленный вопрос до сих пор нет.

Сколько же всего генов в ДНК человека? Еще несколько лет назад полагали, что их около 100 тыс., затем решили, что не более 80 тыс. В конце 1998 г. пришли к выводу, что в геноме человека не более 50–60 тыс. генов и на их долю приходится около 3% общей длины ДНК.

Последние подсчеты общего числа генов в геноме человека проводили несколько международных команд ученых. Уже упоминавшаяся компания «Celera» провела собственные исследования, результаты которого изложены в журнале «Science» в 2001 году. По ее оценкам общее число генов в геноме человека составляет от 26383 до 39114. Средний размер гена оценивается равным примерно 3000 п. н. Если принять, что число генов у человека порядка 30 тысяч и на каждый ген приходится примерно 3 тыс. п. н., то нетрудно подсчитать, что в кодировании белков принимает участие менее 1,5% хромосомной ДНК. Таким образом, генетические инструкции по формированию человеческой личности занимают меньше 3 сантиметров на двухметровой молекуле ДНК. Удивляет и малое количество генов, несущих эти инструкции, — их всего в пять раз больше, чем, например, у такого на наш взгляд совершенно примитивного организма, как муха дрозофила.

Вторая команда исследователей из Национального института геномных исследований США во главе с Френсисом Коллинзом, подсчитав число генов у человека независимым способом и на основе своих данных, получила сходный результат — около 32000 генов содержится в геноме каждой клетки человека.

Разнобой в окончательные оценки пока вносят два других коллектива ученых. Доктор Вильям Хезелтайн (руководитель фирмы «Хьюмэн Геном Сайенс») продолжает настаивать, что в их банке содержится приватизированная информация на 120 тыс. генов. Этой информацией он не собирается пока делиться с мировой общественностью. Фирма вложила деньги в патенты и собирается заработать на полученной информации, поскольку она относится к генам широко распространенных болезней человека. Фирма «Инсайт» сообщила о том, что имеет в настоящее время каталог, состоящий из 140 тысяч идентифицированных ей генов человека, и также настаивает на этом количестве общего числа генов у человека.

Очевидно, что наспех приватизированная генетическая информация будет еще тщательно анализироваться и проверяться в ближайшие годы, пока точное число генов станет окончательно «канонизировано». Дело в том, что устройство генов весьма многообразно и до конца еще не поняты все возможные варианты. Вот мы прочитали последовательность нуклеотидов ДНК. Определено, что она способна кодировать белок. Но один ли? Выше уже говорилось о том, как транскрипция и последующие модификации РНК, а затем трансляция и модификации полипептидов, способны обеспечить огромное многообразие белков, кодируемых одним участком ДНК. И понять это исходя только из нуклеотидной последовательности ДНК очень часто просто невозможно. Тем не менее структура генома представляет собой единственную базу для осмысления данных, получаемых такими новыми направлениями, рожденными геномикой, как транскриптомика (исследует совокупность РНК-транскриптов организма), протеомика (исследует совокупность белков организма), метаболомика (исследует обмен веществ — метаболизм — в организме). Эти направления призваны дополнить лежащий в основе структурной геномики метод геномного секвенирования, дать возможность выйти за пределы его разрешающей способности.

Выше уже также говорилось об альтернативном сплайсинге. Сейчас хорошо известно, что за счет этого процесса с одних и тех же генов могут считываться разные белки, которые затем взаимодействуют друг с другом, образуя неповторимую смесь, как из основных цветов в живописи — желтого, красного и голубого можно получить мириады оттенков. Такой сплайсинг характерен не менее чем для половины генов человека. Считается, что в среднем с одного гена человека за счет альтернативного сплайсинга может образовываться три разных пептида. Но некоторые гены имеют до 10 альтернативно сплайсируемых экзонов, что позволяет теоретически получать более 1000 различных вариантов белков всего лишь на одном гене. В реальности число разных белков, кодируемых одним геном, достигает 10. Кроме того, существуют еще и альтернативные промоторы, альтернативные кодоны инициации трансляции, редактирование РНК (превращение Ц в У или А в аналог Г — инозин). Все вышесказанное пока еще невозможно учесть при оценке общего числа генов у человека.

Но и это не все. Кроме генов, кодирующих белки, имеются еще гены, конечным продуктом которых являются РНК. Вспомним об упоминавшихся выше генах-риборегуляторах — они не кодируют белки, но производят функционирующую в клетках РНК. Так что скорее всего окончательная оценка числа генов у человека будет сделана еще нескоро.

На сегодняшний день ученым известны функции всего лишь около восьми-десяти тысяч из них. А детальные сведения о механизмах их регуляции еще более скудны. Тем не менее, приведенные выше данные о строении и функционировании генов человека свидетельствуют о том, что у человека, царствующего в природе, в отличие от других существующих на нашей планете организмов, очень высока сложность протеома — полного набора функциональных белков в клетке, которая обеспечивается не просто за счет крупного размера генома или большого числа генов, а благодаря всевозможным инновациям, связанным с функционированием генов и формированием белков: большее число доменов-модулей, более высокая комбинаторика (перемешивание) этих модулей в белках, активное использование альтернативного сплайсинга и многое другое, о чем мы поговорим дальше.

Информационная оценка биосинтеза

Теперь стало возможным оценить, что РНК синтезируется лишь максимум на 25–28% нуклеотидных последовательностей генома человека. А на кодирование белков используется вообще всего лишь чуть более 1% генома. Остальные транскрибируемые участки ДНК — это интроны белок-кодирующих генов и гены для РНК, не кодирующих белки. Таким образом, большая часть генома (свыше 70%) не имеет, по-видимому, прямого отношения ни к каким генам и инертна в плане транскрипции. Как это ни странно на первый взгляд, но факт остается фактом. Большая часть генома не кодирует ни белки, ни какую-нибудь РНК вообще, т. е. не относится к генам или каким-то другим, функционирующим на уровне транскрипции (РНК), нуклеотидным последовательностям.

В ДНК, на которой не образуется никакая РНК и которая, таким образом, для кодирования информации совершенно нейтральна, записаны длинные тексты, которые сегодня представляются для ученых полной абракадаброй, смысл и происхождение которой совершенно не ясен. Зачем нужны клетке эти «бессмысленные» и никак, на первый взгляд, не работающие участки ДНК, пока никто не знает, хотя априори ясно, что ничего лишнего в природе не должно быть. При полном секвенировании генома человека, естественно, была определена последовательность не только генов, но и этих пока ничего не обозначающих для ученых участков. Но из чего-то они все-таки состоят. Когда провели их детальный анализ, то выяснилось, что здесь присутствует скопище всех мыслимых и немыслимых вариантов расположения букв, бессмысленных слов и предложений, различных повторов одного и того же «текста», палиндромов и так далее и так далее. Рассмотрим подробнее хотя бы основные из них.

ДНКовый текст человека перенасыщен повторами

Как уже было показано в ранних экспериментах, описанных выше, в ДНК человека содержатся многочисленные элементы, которые не кодируют никакие белки, но при этом многократно встречаются в ДНКовом тексте.



Рис. 22. Основные варианты расположения повторяющихся последовательностей в геноме человека. Повторы обозначены стрелками, неповторяющиеся участки — серыми линиями

На рис. 22 приведены некоторые основные варианты расположения повторов в геноме человека, а далее мы подробнее остановимся на описании отдельных из них.

В «догеномную» эру считалось, что такие повторы составляют около 25% генома человека. Секвенирование ДНК показало, что их существенно больше: повторяющиеся последовательности в сумме занимают около 50% ДНКового текста. А общее число таких повторов в геноме человека равно приблизительно 5 миллионам. Для сравнения: у дрозофилы доля повторов в геноме составляет менее 5%, у червя — 6–7%, а у других млекопитающих — чуть меньше, чем у человека. Повторы напоминают бессмысленные фразы, которые с упорным постоянством в разных вариациях повторяются в тексте, но от этого смысл в них все равно не возникает.

Перевертыши (обращенные повторы)

В ДНКовых текстах человека, как, впрочем, и в ДНК других организмов, присутствуют уже упоминавшиеся обращенные повторы, перевертыши или, как их еще часто называют, палиндромы. Уже говорилось о том, что перевертыши-палиндромы хорошо известны из литературы — это предложения, которые читаются одинаково слева направо и справа налево без учета знаков препинания и интервалов между словами. Считается, что палиндромы известны еще со времен Древней Греции, т. е. свыше двух тысяч лет. На Руси палиндромы когда-то назывались «рачьими стихами». В качестве примера приведем две державинские строки:

Я разуму уму заря, Я иду с мечем судия
Рислинг сгнил сир

Особое свойство таких участков ДНК заключается в том, что при определенных состояниях ДНК они способны образовывать структуры, называемые крестами. На рис. 23 дан пример одного из таких возможных крестов.

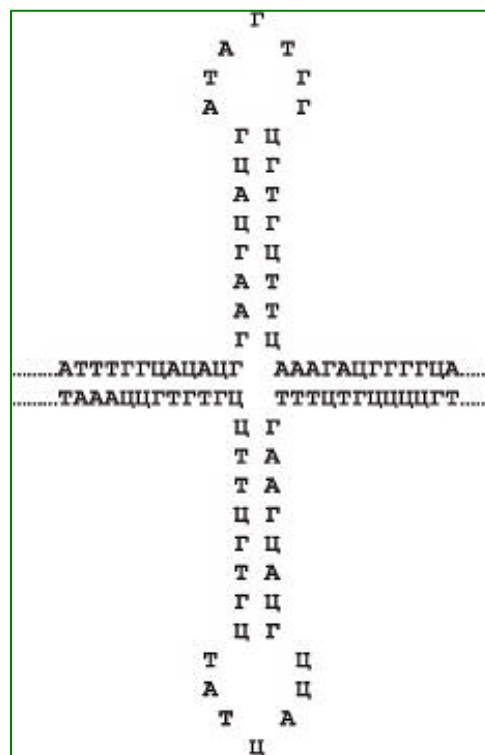


Рис. 23. Крест, способный образовываться последовательностью ДНК, состоящей из обращенных повторов

Важно понимать, что такие представления о пространственной структуре возникают при хорошем знании предмета. Введение в рассмотрение визуального образа значительно помогает «усмотреть» некоторые стереотипы. В реальной ситуации в клетке такие структуры действительно могут формироваться. Но зачем это нужно? Окончательного ответа на этот вопрос нет. Одно из предположений сводится к тому, что кресты могут быть местами связывания с какими-то специфическими белками, обеспечивающими регуляцию или процессов транскрипции, или репликации ДНК. С другой стороны, при транскрипции таких структур в РНК должны формироваться «шпильки», а такие структуры в РНК часто являются значимыми (то есть выполняют определенные функции — или способствуют транскрипции, или, наоборот, препятствуют ей).

Простые тандемные повторы (сателлиты)

Как и у других организмов, в геноме человека присутствуют короткие повторяющиеся нуклеотидные последовательности, которые расположены вплотную друг за другом (тандемно). Их часто называют сателлитными ДНК. Мы уже говорили о сателлитных ДНК ранее. Это те участки ДНК, которые при центрифугировании в градиенте солей цезия осаждаются отдельно вне области основной массы ДНК. В разных типах сателлитной ДНК различные копии представлены неодинаково — их численность варьируется от сотен до миллионов. Если записать их структуру текстом, то это выглядит, как, например, предложение из одного и того же слова: столстолстолстолстолстолстолстол. Ясно, что такие предложения не могут кодировать никакую серьезную информацию. Тем не менее выяснилось, что с сателлитными нуклеотидными последовательностями в некоторых случаях могут специфически взаимодействовать отдельные белки. А это уже некоторое довольно серьезное указание на возможное участие таких повторов в каких-то клеточных процессах.

К настоящему времени в геноме человека обнаружено по крайней мере шесть видов сателлитных повторов. Это в первую очередь три типа так называемых классических сателлитов. Сателлит 1 состоит из элементарной повторяющейся последовательности нуклеотидов длиной 42 п. н. (располагается на семи разных хромосомах); сателлиты 2 и 3 состоят из элементарных повторяющихся единиц длиной 5 п. н. (первый из них обнаружен на четырех, а второй — на двенадцати разных хромосомах).

Кроме того, во всех хромосомах человека присутствуют альфа-сателлитные повторы. Они имеют элементарную повторяющуюся нуклеотидную последовательность длиной 171 п. н. Существуют еще бета-сателлитные повторы с элементарной повторяющейся единицей длиной 68 п. н., которые встречаются лишь на девяти хромосомах. Еще реже представлены в геноме человека гамма-сателлитные повторы, имеющие элементарную повторяющуюся единицу длиной 220 п. н. Такие повторы присутствуют только на двух хромосомах человека (8 и X).

Распределение всех сателлитных ДНК вдоль хромосом не выглядит случайным. Так, некоторые из них «гнездятся» в районах, прилежащих к концам хромосом, и в области центромеры (перетяжка на хромосомах, которая не всегда расположена в самом центре, см. рис. 7). В частности, центромеры всех хромосом человека, размер которых варьирует от 0,1 до 4 млн. п.н., состоят преимущественно из альфа-сателлитных ДНК.

Наряду с типичными сателлитными ДНК, в геноме человека имеются нуклеотидные последовательности, получившие название мини- и микросателлитов, которые разбросаны по большинству участков всех хромосом и в сумме составляют около 3% генома человека. В среднем один из таких сателлитов содержится в каждых 2 тыс. п. н. человеческого генома. Минисателлиты представляют собой повторяющиеся нуклеотидные последовательности размером от 11 до 500 п. н. Кластеры таких повторов занимают от 0,1 до 20 тысяч п. н. в человеческом геноме, а общее число таких районов во всех хромосомах человека составляет несколько тысяч. Еще меньше по размерам микросателлиты. Их повторяющаяся нуклеотидная последовательность состоит из 10 и меньше п. н. Микросателлиты занимают в геноме человека обычно небольшие участки (до 150 п. н.). Когда же в тексте повторяется всего одна «буква» из четырех, такие структуры называют гомополимерными нуклеотидными последовательностями. Примером гомополимерной последовательности нуклеотидов в геноме человека может служить последовательность

AAAAAAAAAAAAAAAAAAAAAAAAA

на одной нити и, соответственно, TTTTTTTTTTTTTTTTTTTTTTTT на другой.

Микросателлиты, состоящие из двух пар нуклеотидов (динуклеотидные повторы), занимают около 0,5% генома человека. Наиболее частыми динуклеотидами являются АЦ и АТ (50 и 35% всех динуклеотидов, соответственно). А вот динуклеотидный повтор, состоящий из пары ГЦ, в ДНКовом тексте человека занимает всего лишь 0,1%. Наиболее частые из более «сложных» микросателлитов (тринуклеотидных повторов) являются ААТ (33% от всех тринуклеотидных повторов) и ААЦ (21%). Триплет ЦГГ, который довольно редок в геноме, присутствует в гене FMR-1, определяющем синдром ломкости X-хромосомы.

У нормальных людей число копий ЦГГ в гене варьирует от 6 до 54. А вот у больных индивидумов такие триплеты «размножаются» и их число в гене может достигать 200–1300 копий. Подробнее о связи между сателлитными повторами и различными заболеваниями человека мы еще поговорим далее. Но остается загадкой: зачем такие нуклеотидные последовательности вообще нужны геному.

Выше приведенный пример и множество других случаев указывают на то, что микро- и минисателлиты весьма изменчивы, т. е., как говорят, полиморфны. Их распределение и число в геноме каждого конкретного человека столь специфично, что может использоваться в качестве аналога отпечатков пальцев. Как только это было обнаружено, этим фактом сразу воспользовались. В настоящее время ненужная вроде бы фракция генома человека (а на самом деле возможно и нужная, но мы просто о ней мало что знаем) стала востребована человеком для своих практических целей. На основании вариабельности сателлитных ДНК была разработана диагностика родственных связей между людьми, склонности человека к различным заболеваниям и др. (см. об этом подробнее далее в разделе о геномной дактилоскопии).

И, наконец, еще один вид сателлитной ДНК был обнаружен при секвенировании генома человека, который получил название мегасателлитов. Эти тандемные повторы имеют размер от 2,5 до 5 тыс. п. н. и повторяются от 50 до 400 раз. Мегасателлиты присутствуют лишь в небольшом числе хромосом (в 4, 8, 19 и X).

Диспергированные повторы

В ДНКовом тексте генома человека содержится множество более длинных повторяющихся нуклеотидных последовательностей, которые в отличие от сателлитных повторов рассеяны (диспергированы) по всему геному. Эти повторы называют вездесущими, так как они присутствуют во всех районах всех хромосом, а иногда обнаруживаются даже и внутри генов. В сумме около половины ДНКового текста генома человека представлено такими повторами.

Исследователи разделяют эти разбросанные по геному (диспергированные) повторы на четыре основных типа. Наиболее представленными в геноме являются длинные диспергированные повторы (сокращенно ДДП). Их длина достигает 6 тыс. п. н., а повторяются они в геноме человека

примерно $8,5 \times 10^5$ раз (в сумме ДДП составляют около 21% генома). Далее — это короткие диспергированные повторы (сокращенно КДП) длиной 100–400 п. н. В ДНК человека содержится примерно $1,5 \times 10^6$ копий КДП (это составляет около 13% генома). Имеются еще повторяющиеся элементы, окруженные с двух сторон дополнительными повторами (последние называют длинными концевыми повторами). Их назвали ретротранспозонами. (Как появилось такое сложное название, будет ясно из дальнейшего изложения). Длина этих повторов варьирует от 1,5 до 11 тыс. п. н., а общее число копий в ДНК человека составляет около $4,5 \times 10^5$, (8% генома). Наконец, в геноме человека присутствуют повторы, названные ДНК-транспозонами. Длины разных ДНК-транспозонов существенно отличаются (от 80 до 3000 п. н.). В геноме их меньше, чем других повторов (всего 3%), а общее число составляет 3×10^5 .

Более подробно о структуре и характерных особенностях этих типов повторов мы поговорим далее в связи с одной чрезвычайно важной их особенностью — способностью перемещаться в геноме. Здесь же отметим, что имеется большое различие в распределении всех этих диспергированных повторов вдоль молекул ДНК. Хотя повторы имеются во всех участках хромосом, некоторые места почему-то сильно заселены ими, а другие, наоборот, практически их не содержат. Так, область размером 525 тыс. п.н. в хромосоме 11 на 89% состоит из таких повторяющихся элементов. А вот в хромосоме 2 имеется участок ДНК размером 100 тыс. п.н., в котором расположены кластеры генов некоего семейства НОХ, ответственных за процессы развития при эмбриогенезе, и который содержит всего около 2% рассеянных повторов.

Однозначного объяснения этому нет, можно строить только разные предположения. Например, в случае с блоком генов НОХ отсутствие здесь повторов, возможно, связано с очень сложной координированной регуляцией в этой области, а наличие бесполезных повторов могло бы нарушить такую регуляцию.

Диспергированные повторы умеют «перемещаться»

В соответствии с законами классической генетики ожидалось, что в геноме все должно быть четко размечено, все гены-предложения должны располагаться в строго определенных местах текста. И большинство элементов генома действительно подчиняется этому общему правилу. Но из общего правила всегда есть исключения. Оказалось, что некоторые участки ДНК могут «путешествовать», меняя свое место, вытесняя друг друга. Подавляющее большинство генов никогда не покидают родную хромосому. Но в отличие от них определенные специфические нуклеотидные последовательности ДНК изредка с корнем вырываются из генома, выпрыгивают из одной хромосомы и приземляются в случайном месте на другой. При этом они могут влезть в середину гена, вызывая хаос, а могут примкнуть с края, слегка видоизменяя его регуляцию.

Впервые данное явление обнаружила у растений американская исследовательница Барбара Мак-Клинток еще в конце 30-х годов. Но очень долгое время ей никто не верил, исходя из общего принципа: «этого не может быть, поскольку не может быть никогда».

Генетикам трудно было даже представить, чтобы гены прыгали с места на место, «ломая» наследственную информацию. Маститые ученые, если речь заходила о Мак-Клинток, с иронией приводили «прыгающие гены», «скачущие гены» как пример научной торопливости, некритического отношения к собственным результатам, стремления прославиться. Но прошло время и российские ученые (академик Г. П. Георгиев и член-корреспондент В. А. Гвоздев) обнаружили перемещения некоторых участков и в ДНК животного организма (дрозофилы). В конечном итоге научная общественность воздала Мак-Клинток должное. Она дождалась своего звездного часа и получила за свое открытие под конец своей жизни Нобелевскую премию (1983 г.).

Таким образом, было установлено, что отдельные участки генома высших организмов, подобно некоторым индивидуумам в нашем обществе, могут не хотеть жить по общим правилам. Их мало интересует постоянная прописка, и при определенной ситуации они берут и просто «переезжают» в любое заинтересовавшее их по каким-то причинам место или, как говорят, «перепрыгивают».

вают» в иные, видимо, более удобные для них районы ДНК. Другими словами, стало ясно, что геном не так уж и стабилен, законсервирован, как он казался многим исследователям до этого.

У человека возможность перестроек в геноме впервые была продемонстрирована при изучении генов иммуноглобулинов, связанных непосредственно с иммунной системой организма. Оказалось, что в процессе индивидуального развития в лимфоцитах (белые кровяные шарики) происходит перетасовка отдельных участков этих генов, в результате чего и формируются тысячи и даже миллионы новых вариантов иммуноглобулинов-антител, убивающих многочисленные чужеродные вмешательства различных живых агентов в человеческий организм. Это позволяет за счет сравнительно небольшого набора исходных генов обеспечивать образование огромного разнообразия всесторонне направленных иммуноглобулинов.

Другие хорошо известные гены никогда не «прыгают». Такая ситуация оказалась уникальной именно для иммуноглобулиновых генов. Основное перемещение в геноме человека осуществляют не гены, а элементы совершенно другого типа, получившие название транспозонов. К ним относят различные элементы, входящие во все четыре выше перечисленных основных типа диспергированных повторов генома.

Транспозоны способны к разным инновациям в геноме. Перемещаясь, они могут изменять регуляцию других генов и даже принимать участие в создании новых генов. Так, после полного секвенирования генома человека в нем удалось обнаружить около пятидесяти генов, произошедших из транспозонов. Несколько сотен генов в геноме человека используют длинные концевые повторы транспозонов в качестве терминаторов при транскрипции.

Теперь рассмотрим по отдельности структурные особенности разных семейств этих «прыгающих» время от времени элементов генома человека.

Семейство повторов по имени LINE

Как среди населения земли имеются многочисленные нации и меньшинства, так и в геноме одни повторы редки, а другие представлены большим числом копий. Основной представитель длинных диспергированных повторов генома человека — семейство повторов, получившее имя LINE (сокращенно от англ. long interspersed elements). Отличительная особенность этих повторов заключается в том, что они способны кодировать два белка-фермента, которые и принимают участие в перемещении их самих, а заодно они способны обеспечивать «перепрыгивание» и коротких диспергированных повторов (КДП).

Это осуществляется главным образом за счет уже упоминавшегося фермента ревертазы. РНК, образующиеся после транскрипции LINE-повторов, превращаются в ДНК-копии, которые и встраиваются в новые места генома. В процессе синтеза ДНК-копии на РНК иногда происходит преждевременная остановка этого процесса, поэтому в геноме человека наряду с «полноразмерными» ДДП присутствуют элементы укороченной длины. Ферменты, кодируемые LINE-повторами, участвуют и в процессе появления в геноме некоторых псевдогенов, т. е. генов, которые вообще не способны функционировать. Но, как оказалось, не только их.

Сейчас уже известно с десятков нормально функционирующих генов, которые возникли за счет обратной транскрипции мРНК. Такие гены, конечно же, не содержат интронов (их называют безинтронными).

Итак, в геноме человека источником активной ревертазы являются, по-видимому, ретротранспозоны L1, число копий которых достигает 100 тыс. Однако реально число активных перемещающихся копий составляет всего 30–60 тыс., тогда как остальные настолько повреждены, что не транскрибируются и, следовательно, уже не могут перемещаться.

Семейство повторов по имени Alu

Основную массу коротких диспергированных повторов (КДП) составляют так называемые Alu-повторы, которые занимают в ДНК-тексте генома человека почти в 10 раз больше места, чем все последовательности, кодирующие белки. Свое название они получили по имени рестрик-

тазы, для которой у них имеется определенный сайт. В сумме в геноме насчитывают около одного миллиона Alu-повторов. Такое их огромное количество не поддается нашему пониманию.

Для чего они нужны? Ведь это относительно короткие повторы (около 300 п. н.), которые не способны кодировать никакие белки. Более того, разные члены семейства Alu-повторов не полностью похожи друг на друга, хотя родство между ними прослеживается однозначно (в среднем последовательности двух любых Alu-повторов сходны на 85%). Считается, что этот вид повторов появился в ходе эволюции у приматов (и только у них) свыше 65 млн. лет назад.

Попав в геном человека, Alu чудовищно размножились и причудливо расселились по разным хромосомам, то скапливаясь в некоторых местах, то перемежаясь с генами или даже влезая в некоторые из них. В частности, Alu-повторы нередко присутствуют в тех районах генов, которые транскрибируются при образовании мРНК, но не транслируются при синтезе белка. При этом, как число, так и расположение Alu-повторов в геноме может заметно отличаться у разных индивидуумов.

Большинство из диспергированных повторяющихся элементов ДНКового текста обладают способностью к перемещению в геноме. Свидетельства этому «перепрыгиванию» были получены, в частности, при сравнительных исследованиях структуры отдельных районов генома у человека и других близкородственных приматов. Более того, и сейчас продолжается время от времени регистрация перемещения отдельных повторов. Так, установлено, что Alu-повторы изредка встречаются у некоторых людей в тех местах генома, где у большинства других их нет. Один из таких ярких примеров — обнаружение внедрения Alu-повтора в ген фактора свертываемости крови VIII при такой патологии, как гемофилия А.

Итак, геном — это не застывшая ДНК, а динамическая структура, чем-то напоминающая атом. Последний имеет не только стабильное ядро, но и целый рой перемещающихся вокруг его элементов (в частности, электронов).

Стоит сразу оговориться, что перемещение транспозонов происходит крайне редко. Тем не менее его значимость для генома человека весьма велика и определяется, по-видимому, тем, что этот процесс продолжается безостановочно на протяжении многих тысячелетий. Можно думать, что в результате этого подвижные или мобильные элементы способны в ходе эволюции заметно «тасовать» текст Энциклопедии человека подобно тому, как игроки тасуют колоду карт.

Перечень подвижных элементов генома не ограничивается различными описанными выше диспергированными повторами. К таковым относятся и некоторые элементы, доставшиеся нам в ходе эволюции от таких простейших организмов, как вирусы и бактерии.

Вирусы — составная часть генома человека

Незримая борьба между человеком и такими «внутриклеточными паразитами», как вирусы, происходила на протяжении тысячелетий. Только после обнаружения вирусов Д. И. Ивановским в 1892 году человечество наконец-то распознало своего невидимого врага, который, проникая в клетку человека, использует все ее возможности в своих корыстных целях, тем самым нарушает нормальный метаболизм клетки и вызывает различные тяжелые патологии⁴.

Ряд вирусов обладает способностью внедряться в геном человека и по сути становится как бы его собственными генами. В первую очередь это относится к так называемым ретровирусам. Они так были названы по своему образу жизни. Исходно геном этих вирусов представляет собой РНК. Но, попав в клетку, вирус на своей РНК с помощью обратной транскриптазы строит ДНК-копию (см. основную догму молекулярной биологии).

⁴ Вирус (лат. *virus* — «яд») — неклеточный инфекционный агент, который может воспроизводиться только внутри живых клеток.

После этого ДНК-копия вируса встраивается в геном клетки, что служит обязательным условием для жизненного цикла ретровируса. Встраиваемая в геном клеток человека ДНК-копия вируса была названа провирусом. Затем на провирусе синтезируются вирусные РНК, на базе которых образуются новые вирусные частицы. Таков жизненный цикл обычного ретровируса. Так ведет себя, например, хорошо известный ретровирус, получивший название вируса иммунодефицита человека (ВИЧ), когда он инфицирует клетки крови.

Казалось, что все это есть проблема отдельных инфицированных клеток. И действительно, вирусы инфицируют в основном соматические клетки. Затем вирусы и провирусы погибают вместе с теми клетками, в которых они похозяйничали. Другими словами вирусы сами подготавливают и свою собственную смерть. Однако это не всегда так происходит. Очень редко в эволюции человека происходило инфицирование вирусами клеток зародышевого пути, образовывались провирусы, но организм выживал и внедренный провирус становился наследуемым элементом генома человека. Так в клетке появляется «лжепрограмма» (провирус), которая изменяет геном гораздо сильнее, чем это возможно при «нормальной» эволюционной изменчивости.

Когда секвенировали геном человека и многих других млекопитающих, оказалось, что в их составе содержится очень большое число повторяющихся элементов, имеющих сходство с инфекционными вирусами. Повторяющиеся элементы, способные кодировать 2–3 белка и окруженные с двух сторон еще одними особыми повторами — названными длинными концевыми повторами (ДКП), — были отнесены к семейству, получившему название ретротранспозонов.

У человека они составляют довольно существенную часть — около 8% генома. Такие элементы называют часто эндогенными ретровирусами, в отличие от типичных ретровирусов, существующих в природе вне организмов (их называют экзогенными ретровирусами). Структура типичного эндогенного ретровируса (провируса) изображена на рис. 24 на цветной вклейке. Она напоминает структуру отдельных современных экзогенных ретровирусов (на рисунке вверху), хотя часто существенно отличается от них. Все-таки за миллионы лет древним провирусам трудно было остаться неизменными. Большая часть этих элементов представляет собой дефектные вирусные последовательности или даже отдельные их короткие фрагменты. Происходило это главным образом за счет накопления точечных мутаций и различных делеций. В частности, у многих эндогенных ретровирусов отсутствует ген *env*. Концевые участки ретровирусов, названные длинными концевыми повторами (ДКП), нередко существуют в геноме человека сами по себе, оторвавшись от своего эндогенного ретровируса в ходе эволюции и собственного перемещения по геному.

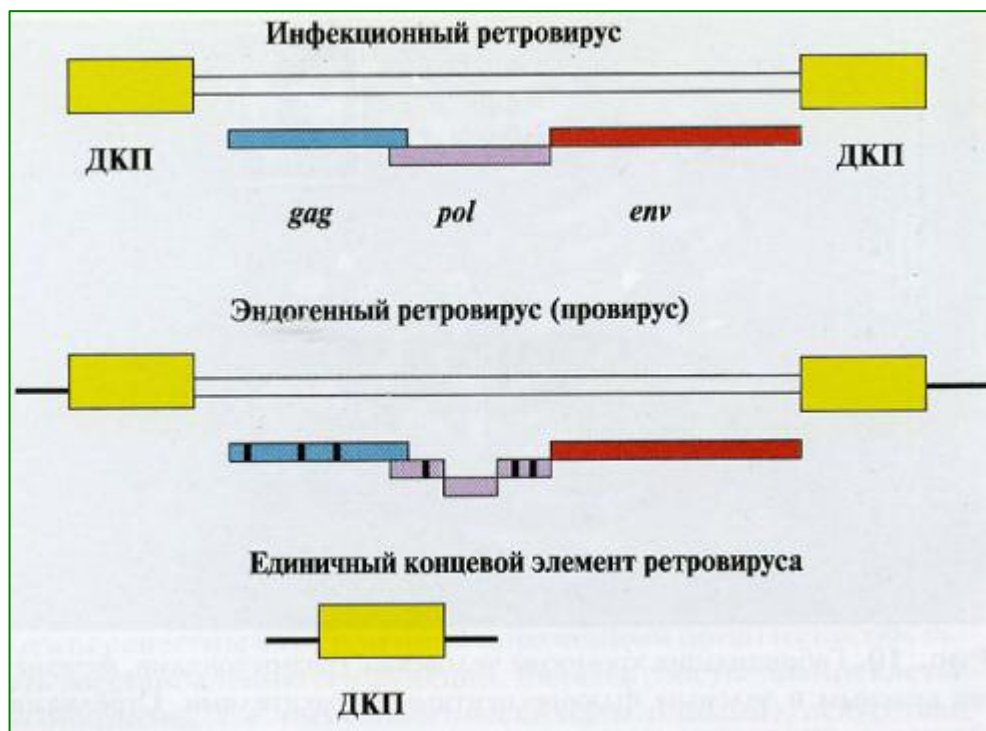


Рис. 24. Типичные структуры инфекционных и эндогенных ретровирусов в геноме человека.

ДКП — длинный концевой повтор (обозначен желтым цветом). Типичные гены ретровирусов — gag, pol, env. Черные горизонтальные линии — последовательности генома человека. Черные поперечные полосы в генах — мутации.

Мутации, накопившиеся в ходе эволюции в эндогенных ретровирусах, не позволяют им образовывать новые инфекционные вирусные частицы, как это происходит обычно в случае типичных экзогенных вирусов (например, вируса иммунодефицита человека, вызывающего такое страшное заболевание, как СПИД).

Согласно проведенным оценкам, такие провирусы появились в геноме человека от 10 до 50 млн. лет назад в результате инфицирования зародышевых клеток наших предшественников и с тех пор передаются по наследству, как и все другие собственные элементы генома. Так «чужие» молекулы ДНК стали частью нашего генома. В сумме их насчитывается несколько десятков тысяч, они состоят из более 200 различных групп и подгрупп. Ретротранспозоны генома человека часто рассматривают как «ископаемые останки» древних ретровирусов.

Появление провирусов в геноме человека сродни процессу, который в лингвистике называют лексическим заимствованием. Во всяком языке имеются слова «свои» и «чужие» — те, которые пришли из других языков. Заимствование слов считается естественным ходом событий, необходимым для языкового развития. Возможно, и появление провирусов повлияло на ход становления человека как вида. По крайней мере известно, что число провирусов в геноме человека существенно больше, чем в геноме обезьян.

Итак, эндогенные провирусы являются «следами» от вирусов, которые инфицировали наших предков миллионы лет назад. Поселившиеся в геноме наших предков, они со временем потеряли способность образовывать новые вирусные частицы. Большинство таких «реликтовых» ретровирусов «молчит» (не функционирует). Однако сейчас уже установлено, что при определенных воздействиях на геном они могут «заговорить» и этим нарушить нормальный метаболизм в клетке. Таким образом, потенциально некоторые из них и по прошествии тысячелетий продолжают представлять потенциальную опасность для человека.

Выяснилось, что отдельные члены семейства ретровирусов сохранили способность синтезировать РНК, а также обладают склонностью к рекомбинациям и перемещению в геноме. Так, анализ эндогенных провирусов генома человека позволяет заключить, что отдельные из них являются копиями, появившимися в результате так называемой ретропозиции. Ретропозиция — это процесс появления новых копий в геноме не за счет нового инфицирования вирусом, а в результате транскрипции уже имеющегося в геноме провируса. Сначала образуется вирусная РНК-копия, затем на ней с помощью обратной транскрипции образуется ДНК-копия, и уже она внедряется в новое место в геноме. Другие элементы эндогенных провирусов, обнаруживаемые в геномах современных людей, возникли в результате рекомбинаций, таких, как, например, рекомбинации между концами провирусов, что привело к исчезновению внутренних участков провирусов и образованию коротких ретровирусподобных элементов. Таким образом, провирусы эндогенных ретровирусов — не просто окаменелые остатки прошлых эпох, а элементы генома человека, которые определяют присущую ему нестабильность (об этом подробнее поговорим дальше).

В геноме человека эндогенные ретровирусы встречаются во многих местах, но их распределение по хромосомам не совсем равномерное. Имеются участки, обогащенные ими, и в то же время некоторые области хромосом человека не содержат подобные элементы. Иногда отмечают, что местонахождение эндогенных ретровирусов и их производных коррелирует с распределением генов в геноме. Часто они располагаются в регуляторных областях генов. И в этом ученые усматривают особый смысл: регуляторные элементы эндогенных ретровирусов могут вмешиваться в регуляцию работы обычных генов человека.

Академик Е. Д. Свердлов считает, что вирусы могли сыграть решающую роль в «очеловечивании» обезьяны. Возможно, что в процессе эволюции молекулы ДНК человека включили в свой состав уже готовые фрагменты генетического материала вирусов с одной лишь целью — облегчить конструирование собственных генов для кодирования новых признаков. Включение генетического материала вирусов в геном человека служит еще одним подтверждением универсального характера молекул ДНК, имеющихся у всех живых существ — от самых простейших организмов вплоть до человека.

Еще одна группа спонсоров человеческого генома (ДНК-транспозоны). В результате секвенирования ДНК человека был раскрыт еще один секрет генома. Кроме вирусов, проникших в наш геном извне и со временем сильно размножившихся там, заметную часть генома человека составляют и гены, пришедшие к нам от других наших постоянных симбионтов — бактерий. В геноме человека содержится множество так называемых ДНК-транспозонов. Они в сумме составляют около 3% генома человека и представлены в нем примерно 300 000 копиями, которые в свою очередь подразделяются на 7 разных классов. Все они по своей структуре очень напоминают ранее обнаруженные перемещающиеся элементы у бактерий — бактериальные транспозоны. Сходство выражается в первую очередь тем, что ДНК-транспозоны человека, как и бактериальные транспозоны, способны кодировать специальный фермент транспозазу, который и обеспечивает их подвижность.

Кроме бактериальных транспозонов, в геноме человека содержится свыше 220 генов, доставшихся нам «в наследство» непосредственно от бактерий, обитающих у нас в кишечнике. Половина из этих генов широко распространена в разных видах бактерий, но среди эукариот встречается только у позвоночных. Теперь стало окончательно ясно, что бактерии, также как и вирусы, внесли свой вклад в формирование генома современного человека. Все это, по-видимому, есть результат длительного сожительства человека с бактериями, когда при некоторых случайных событиях гены бактерий попали в геном человека и там закрепились. Мы давно знаем, что в нашем кишечнике живет много разных видов бактерий, которые не только для нас не вредны, но, наоборот, полезны, так как они служат поставщиками ряда биологически активных продуктов (в частности, некоторых витаминов). Перенос генов «по горизонтали», т. е. от одного вида организма в другой, в принципе был известен давно. В частности, он активно реализуется между разными видами бактерий.

Но то, что он столь масштабно выражен и между геномами таких эволюционно далеких организмов, как человек и бактерии, безусловно, стало для исследователей полной неожиданностью. Интересно, что гены, доставшиеся человеку от бактерий, содержат, как и многие другие, интроны. Последних нет у бактерий. Следовательно, в процессе эволюции после переноса бактериальных генов в геном человека произошло еще одно событие: в состав «бактериальных» генов, ставших генами человека, внедрили дополнительные нуклеотидные последовательности, которые теперь являются интронами.

Гены, доставшиеся человеку от бактерий, не случайно сохранились в геноме. Многие из них играют важную роль в метаболизме ксенобиотиков (чужеродных для человеческого организма веществ) и в ответе клеток на стресс. Можно отметить один из таких генов — ген MAO, кодирующий белок, на который направлено действие разных психофизиологических лекарств. Следовательно, приобретение по крайней мере некоторых из бактериальных генов дало человеку определенные селективные преимущества.

На этой почве существуют разные спекуляции. Одна из них изложена З. Ситчиным в своей книге «Генезис пересмотренный и космический код». Согласно его рассуждениям, некие пришельцы из космоса по имени аннуаки прибыли на Землю приблизительно 450 тысяч лет назад. Они высадились на Землю в поисках полезных ископаемых, и им понадобилась рабочая сила. Рабочих решили не привозить издалека, а слегка усовершенствовать существовавшего тогда на Земле гоминоида, добавив ему некоторые гены более «продвинутых» аннуаков. И космические генные инженеры внесли в геном гоминоида те гены, которые мы теперь называем бактериальными.

ми. А бактерии потом приобрели их от нас. Сформулировано, конечно, красиво, однако весьма напоминает фантастический роман.

В любом случае, геном человека представляет собой некое собрание генов, которые достались человеку от их предков, и генов, которые явно ведут свое происхождение от неродственных организмов — вирусов и бактерий. То есть можно сказать, что наш геном — это книга, в которой отдельные «предложения» заимствованы из других независимо написанных источников.

Нуклеотидные последовательности — «близнецы»

В результате секвенирования генома человека в нем, наряду с типичными повторяющимися элементами, были обнаружены протяженные нуклеотидные последовательности, которые представлены всего в двух-пяти копиях. Чаще всего эти копии очень похожи друг на друга, как близнецы (сходство по последовательности достигает 99%). Общее количество таких элементов в геноме человека составляет порядка 5%. Это значительно больше, чем в других секвенированных к настоящему времени геномах, таких, например, как червь и дрозофила. Появление в геноме ДНК-подобных текстов-близнецов связывают с процессом, названным сегментальной дубликацией, то есть случайным удвоением отдельных сегментов генома. Размер таких «близняшек» составляет от 1 до 200 тыс. п. н. Последовательности-близнецы иногда располагаются на одной и той же хромосоме, а иногда присутствуют на совершенно разных хромосомах. Например, геномный сегмент X-хромосомы размером 9,5 тыс. п. н., известный как участок, ответственный за развитие аденолейкодистрофии, присутствует практически в неизменном виде на хромосомах 2, 10, 16 и 22. А в хромосоме 17 имеется пять копий одного сегмента размером 200 тыс. п. н., которые разделены между собой довольно длинными последовательностями. При анализе разных хромосом выяснилось, что дублицированные последовательности распределены по ним не равномерно. Больше всего их имеется в Y-хромосоме, а менее всего в хромосоме 7. Неравномерно они распределены и вдоль хромосом, концентрируясь в районах центромер и теломер. Предполагается, что появление последовательностей-близнецов может быть связано с эволюционным процессом и приводит к увеличению разнообразия белков за счет «тасования» уже существующих экзонов.

В таблице 4 даны суммарные данные о содержании разных видов последовательностей в геноме человека.

Таблица 4. Нуклеотидные последовательности, входящие в состав генома человека

Тип последовательности	Содержание, %
Экзоны генов	-1
Интроны генов	25
Транспозоны	45
Большие дубликации	5
Простые повторы (микросателлиты)	3
Другие межгенные последовательности	-20

«Эгоистичная» ДНК

Когда полностью секвенировали довольно крупный геном круглого червя *C. elegans*, то обнаружили, что в нем 27% нуклеотидных последовательностей кодируют структурные белки, 26% принимают участие только в начальной стадии кодирования и в дальнейшем от синтеза белка отстраняются (интроны), а оставшиеся 47% приходятся на повторы и межгенные участки, которые в кодировании белков не участвуют и в настоящее время представляются нам «лишним» материа-

лом. При сравнении этих данных с геномами животных, стоящих на различных ступенях эволюционной лестницы, выяснилось, что чем сложнее устроен организм, тем меньше в его геноме доля кодирующих участков и тем больше в нем доля непонятной для нас информации. И больше всего ее в геноме человека. На первый взгляд это выглядело парадоксально.

Так в чем же дело? Этот вопрос возник еще в конце 70-х годов прошлого века. В 1980 году лауреатом Нобелевской премии Дж. Уотсоном и рядом других исследователей было впервые предложено считать участки, не способные кодировать белки, и при этом повторяющиеся в геноме, никчемными, лишними. Появился даже специальный термин «selfish DNA», что переводится как «эгоистичная ДНК». По мнению Уотсона, эгоистичная ДНК существует сама по себе и сама для себя. Накопление избыточной некодирующей белки ДНК — не что иное, как «эгоистическое» самоумножение селективно нейтральных молекул. В попытках объяснить наблюдаемое явление ученые пренебрежительно назвали «лишнюю» «эгоистичную» ДНК строительным «мусором», издержкой эволюции, которая служит платой за совершенство остальной ее части. В одном из последних изданий Oxford English Dictionary приводится такое значение слова эгоистичный: «Применительно к гену или генетическому материалу: обладающий тенденцией к воспроизведению или распространению, несмотря на отсутствие фенотипического (т. е. проявляемого внешне) эффекта».

После секвенирования генома человека, когда было установлено, что и в ДНК человека подавляющая часть последовательностей представляют собой участки, не кодирующие никакие белки или функционально значимые РНК (рРНК, тРНК и др.), стало ясно, что в отличие от трудягенов основная масса ДНК действительно не работает, т. е. не выдает никакой информации для производства белков или РНК. И это при любом расчете более 70% всей ДНК! Что-то слишком много для эгоистичной ДНК или «мусора». Такого количества геномного «мусора» нет ни у бактерий, ни у дрожжей, ни у мух.

Что это значит? Или наш геном в отличие от геномов более примитивных организмов ленился выгрести накапливающийся «мусор», или это совсем не мусор, а некое ценное эволюционное приобретение, которое дало когда-то человеку определенные преимущества, и поэтому геном от такого «мусора» не спешит избавиться.

На сегодняшний день мы еще многого не знаем об этой огромной части генома человека, пока условно называемой «эгоистичной». Однако уже сейчас ясно, что определение «эгоистичности» или «бессмысленности» для большинства участков генома на самом деле нельзя признать полностью справедливым. Если мы сегодня не в состоянии понять, для чего нужны какие-то участки ДНК, это еще не значит, что они пренебрежительно могут быть названы мусором. У бактерий «бессмысленных» участков вообще нет. У дрожжей почти нет. Но по мере повышения уровня организации живого организма накапливается все больше и больше вроде бы ничего не кодирующей ДНК. Трудно согласиться с предположением, что это результат некоей неосознанной деятельности эволюции. Можно предположить, например, что эти нуклеотидные последовательности ДНК являются резервуаром эволюции, складом «запчастей». Если с каким-либо геном что-то не в порядке, клетка использует фрагменты некодирующей ДНК для «ремонта» поврежденного. Сегодня ученые вынуждены отказываться от многих своих первоначальных, довольно ограниченных представлений. Время идет, и сейчас уже по крайней мере у некоторых из этих участков обнаружены определенные специфические функции.

И каковы же эти функции у ДНК, считавшейся много лет «эгоистичной»? Давно уже известно, что повторами в геноме человека являются представленные сотнями копий гистоновые гены, имеются тысячи копий генов, кодирующих рибосомные РНК, транспортные РНК и др. В дальнейшем выяснилось, что существуют и повторы, расположенные в концевых некодирующих белки участках мРНК, которые принимают, по-видимому, участие в таком важном процессе, как трансляция этих молекул. Оценено, что у человека такие элементы встречаются в 12% начальных некодирующих участков мРНК и в 36% некодирующих участков, расположенных в конце молекулы, что существенно выше, чем у других организмов. Повторы, расположенные в участках, кодирующих мРНК, несут определенную функциональную нагрузку. Они нередко участвуют в правильном и своевременном считывании информации при образовании белка на мРНК. Однако в

сумме доля таких значимых или потенциально значимых повторяющихся участков генома человека весьма невелика.

Среди транспозонов, которые также были прежде отнесены к «эгоистичной» ДНК, постепенно обнаруживаются различные неизвестные ранее регуляторные последовательности, которые управляют работой ряда генов. Некоторые из транспозонов выступают в роли энхансеров (усилители транскрипции), сайленсеров (ослабители транскрипции), терминаторов транскрипции и прочих красиво звучащих элементов генома. Более того, выяснилось, что не менее полусотни генов генома человека имеют свое происхождение из транспозонов. Наконец, следует отметить, что определенные некодирующие белки участки ДНК обеспечивают процесс ее репликации, другие — упаковку в хромосомы, третьи — прикрепление к определенным структурам в ядре и т. д.

Можно привести ряд конкретных примеров. Так, Alu — повторы, о которых шла речь выше, обычно рассматривали как типичные представители «эгоистичной» ДНК, поскольку они имеют высокую повторяемость и при этом ничего не кодируют. Тем не менее и им может быть найдено применение в геноме, например, в регуляции работы генов. Похоже, в эволюции геномной ДНК действует некий «принцип слоненка Киплинга». Хобот у слоненка возник из-за его желания узнать, что ест крокодил на обед. Слоненок вначале огорчился своему носу, но не растерялся и нашел ему разные полезные применения. Так и повторяющиеся последовательности генома. Они возникают и меняются по своим внутренним молекулярно-генетическим законам, но для их вариаций потом может найтись полезная функция в геноме. Сейчас выяснилось, что способность Alu-повторов к перемещению может играть существенную роль в эволюции генома. Например, описаны случаи, когда, внедряясь в экзон или интрон какого-нибудь гена, такой повтор менял регуляцию его экспрессии. Alu-повтор способен также модулировать процесс репликации ДНК, регулировать сплайсинг и некоторые другие генетические процессы. Имеются данные, что эти повторы участвуют в подавлении трансляции РНК при клеточном ответе на стрессовые воздействия. Различные описанные в литературе функции, выполняемые Alu-повторами, помогают понять и объяснить причину их длительной «фиксации» в человеческом геноме. Скорее всего действительно многие из Alu сидят без дела до поры до времени, но отдельные представители не просто живут и размножаются, а и иногда работают.

Не находятся совсем без дела и длинные повторяющиеся последовательности. Так, показано, что они могут участвовать в процессе «тасования экзонов» (подробно тасованию карт), т. е. перемещения экзонов из одних генов в другие и формирования за счет этого новых генов.

Наконец, «бессмысленные» участки ДНК иногда могут осмысленно работать в хромосоме, например, когда они защищают свой вид от вторжения чужеродной ДНК. Выяснилось, что пустые концевые участки хромосом, как и область центромер (первичные места спаривания родительских парных хромосом), важны для сохранения вида: они определяют строгое распознавание макро-рельефа хромосомы как органеллы клетки (а не микро-рельефа — как молекулы ДНК) одного вида по принципу «ключ-замок». В результате этого сперматозоиды обезьяны не оплодотворяют яйцеклетки человека, т. к. хромосомы двух видов не распознают друг друга.

Таким образом, согласно современным оценкам, в участках «эгоистичной» ДНК скрыта определенная эволюционная информация. В результате перестройки некодирующих белки участков, что вообще-то не влияет на структуру самих генов, может происходить изменение в наборе формирующихся в конкретной клетке мРНК за счет изменения в их процессинге (альтернативный сплайсинг), использование альтернативных промоторов или даже формирование новых генов.

И последний пример функциональной значимости «эгоистичной» ДНК, который приведем здесь, — данные о строении концов хромосом.

Концы молекул ДНК — объект повышенного внимания

Устройство концов хромосом давно интересовало исследователей. Теперь стало ясно, что они устроены по особому. И эту некодирующую белки ДНК никак нельзя назвать «эгоистичной». Концевой участок ДНК, называемый теломерным, представляет собой область размером 3–20 тыс.

п. н., состоящую из tandemно расположенных повторов ТТАГГГ (рис. 28). С этими участками образуют комплекс различные специфические белки, и такой комплекс получил название теломера. Самый-самый (в данном случае левый) кончик хромосомы человека выглядит так:

ГГГАТТГГГАТТГГГАТТГГГАТТГГГАТТГГГАТТ...
ЦЦЦТААЦЦЦТААЦЦЦТААЦЦЦТААЦЦЦТАА...

То есть каждый том Энциклопедии человека начинается с буквы Г, и первые 6 букв — это однострочная ДНК, а вовсе не двунитевая, какова она на всем дальнейшем протяжении. В английском языке с буквы Г (G) начинается слово Бог (God), и некоторые исследователи склонны видеть в этом некий глубокий смысл. Конечно к этим суждениям можно относиться с определенной долей скепсиса. При определенном желании, конечно же, можно найти много всяких аналогий, но эта — безусловно случайная.

За короткими повторами располагаются более длинные повторяющиеся последовательности, которые назвали повторами, ассоциированными с теломерой. Далее следуют уникальные для каждой хромосомы последовательности, которые на практике используются для различения разных концов разных хромосом. Теломеры защищают основную часть ДНК от действия ферментов, которые в отсутствие теломер разрушали бы молекулы ДНК. Кроме того, теломеры способствуют правильной рекомбинации хромосом и их прикреплению к оболочке ядра.

Концевые участки хромосом не статичны. При каждом делении клеток происходит их укорочение в среднем на 50 п. н. Все известные ДНК-полимеразы, которые участвуют в процессе репликации ДНК, неспособны полностью реплицировать концы молекул. Поэтому каждая следующая копия короче предыдущей. Как если бы машинистка, перепечатывая книгу, всегда ленилась бы перепечатать последнюю страничку.

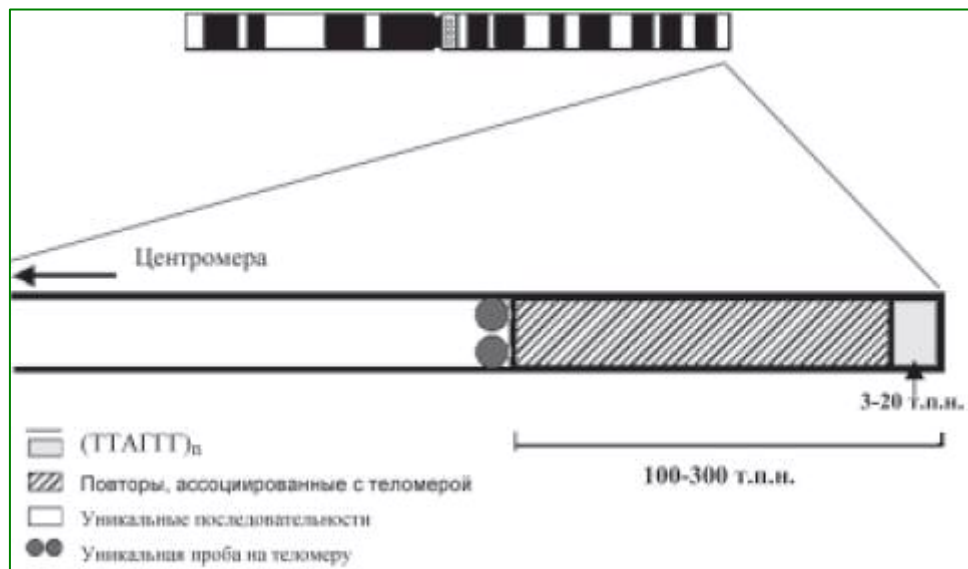


Рис. 25. Структура концов ДНК в хромосомах — теломер

Чтобы не происходила потеря генетического материала при делении клеток, у эмбрионов концы молекул ДНК наращиваются перед каждой репликацией короткими повторяющимися последовательностями. В этом участвует специальный фермент — теломераза. Однако в клетках взрослого организма этот фермент перестает работать, и в результате хромосомы постоянно укорачиваются, съеживаются, как шагреновая кожа, при каждом цикле репликации ДНК. Так происходит до определенного предела, после чего деление клеток становится, по-видимому, невозможным и наступает их гибель.

«Белые пятна» генома

На фоне всеобщей эйфории, связанной с секвенированием генома человека, мало внимания было обращено на то, что необходимо еще некоторое время для получения действительно полной последовательности нуклеотидов генома. На современной карте генома человека все еще зияет множество «белых пятен». В сумме они составляли на начало 2002 года всего около 3–5%, что вообще-то сопоставимо с долей генома, занимаемой всеми генами. Причина этого заключается в первую очередь в существовании серьезных технических трудностей при секвенировании и картировании отдельных специфических участков ДНК.

Так, некоторые нуклеотидные последовательности обладают нестандартной структурой, что усложняет процесс их расшифровки стандартными методами. В частности, это относится к таким районам хромосом, как теломеры — концевые участки, которые имеют необычную структуру. Их порой трудно выделять из общей массы нуклеотидных последовательностей в силу того, что они расположены на концах молекул ДНК. Другие участки трудно идентифицировать в связи с тем, что они часто повторяются и их пока не удастся «вписать» в общий текст, поскольку не ясно, где их истинное место.

Эти участки порой так похожи друг на друга, что нелегко понять, где какой из них точно расположен («Фигаро здесь, Фигаро там»). Такая ситуация характерна в основном для центромерных областей хромосом, содержащих тандемные кластеры альфа-сателлитных последовательностей, и районов расположения микро- и минисателлитов. Имеются проблемы и с секвенированием коротких «плечей» некоторых хромосом.

Однако если сравнить эти проблемы с начальными этапами работы, то они теперь сродни «семечкам», большинство из которых исследователи безусловно «отщелкают» в самое ближайшее время, возможно даже еще до выхода в свет этой книги.

Опечатки в тексте из ДНК

Каждый человек уникален, следовательно, уникальным должен быть и его геном. По этой причине между любыми двумя людьми, если они не однояйцовые близнецы, всегда имеются значительные индивидуальные различия (их называют различиями в фенотипе). Они проявляются во многом, в частности, во внешности, в физических возможностях, в отношении предрасположенности к развитию определенных заболеваний, в реакции организма на воздействие факторов окружающей среды или лекарственных средств, в умственных способностях и поведении.

Однако, как оказалось, из свыше 3,2 миллиардов генетических букв, из которых состоит ДНК человека, 99,9 процента одни и те же у всех людей. Выходит, что всего от одной десятой процента генома зависит, кто мы есть, — высокие или низкие, красивые или не очень, большие или здоровые, сильные или слабые, умные или глупые, оптимисты или пессимисты, добрые или, наоборот, жестокие и так далее.

Различия между нуклеотидными последовательностями ДНК разных индивидуумов связаны главным образом с тем, что за время длительной человеческой эволюции в нашем геноме накопилось довольно много случайных, вызванных разными причинами, изменений. Эти «опечатки» в ДНКовом тексте, которые каждому из нас достаются от двух родителей в случайном сочетании и наследуются, называют мутациями.

Очень трудно издать книгу (особенно если она большая по объему) без опечаток, хотя в процессе подготовки их тщательно отслеживают и автор, и редактор, и корректор. Известна старая история о споре двух издателей, когда один из них утверждал, что сумеет выпустить книгу без единой опечатки. И, действительно, при самой тщательной проверке в тексте этой книги опечаток не нашли. Но издатель проиграл спор: опечатка была в самом названии книги! И книга, которую сейчас читаете вы, наверняка будет содержать опечатки (хотелось бы только, чтобы их число было минимальным и чтобы их не было в названии).

Мутации в ДНК — естественные наследуемые изменения — опечатки в ДНКовом тексте. Каждый раз, когда клетка в очередной раз делится, необходимо выполнить грандиозную задачу — копировать свыше 3 млрд. букв. В клетке этим занимается целая армия специальных ферментов,

которые выполняют свою гигантскую работу в среднем за один рабочий день (7 часов). При этом неправильные буквы первоначально возникают в 10000 участках нового ДНКового текста, а потом эти же ферменты, выступая уже в роли «корректоров», исправляют его. В конечном итоге при копировании ДНК остается только 1 ошибка на 10 млн. п. н. Из этого легко понять, что чем больше раз копируется (реплицируется) ДНК, тем больше в ней накапливается ошибок.

Изменения в структуре ДНК имеют разнообразный характер. В первую очередь это «точечные мутации». Такие мутации представляют собой опечатки в ДНКовом тексте, когда одна из 4 букв выпадает из текста или заменяется на какую-нибудь другую из них. В сумме эти опечатки в масштабах одного генома, как правило, невелики. (Уже говорилось, что мы в среднем отличаемся друг от друга приблизительно в одной позиции нуклеотидов из каждой тысячи). Это явление называют полиморфизмом генома. Отличие между мутациями и полиморфизмом достаточно условно. Когда один из вариантов нуклеотидных последовательностей определенного участка ДНК является более чем у 1% людей в популяции, то такую вариабельность чаще всего называют полиморфизмом, если же менее 1% — просто мутацией.

Мутации могут быть простейшими заменами одного нуклеотида на другой, а могут представлять собой более сложные изменения ДНКового текста, в виде вставки в готовый текст других «букв» и текстов или, наоборот, потери ряда «букв», «слов» или целых «предложений». Изредка отдельные куски ДНКового текста даже переворачиваются. Все это и есть «опечатки». Однако опечатка опечатке рознь. Чаще всего мутации не сказываются заметным образом на фенотипе человека. Это и понятно. В абракадабре, которую из себя представляет свыше 70% ДНКового текста человека, появление 1–2 и даже порой сотни «опечаток» может ровным счетом ничего не менять. Но это не всегда так. Некоторые из них очень сильно влияют на человека.

Этому есть аналогия в лингвистике. Так, существуют слова-паронимы, которые сходны по строению, имеют общий корень, но смысл которых сильно отличается при разнице всего в одной букве. Пример: нового посетителя представляют гостям, но бывают нерадивые хозяева, которые предоставляют гостей самим себе.

Известно и множество случаев из жизни, которые свидетельствуют о важном значении некоторых на первый взгляд незначительных «опечаток». Яркий примером может служить судьба марок бывшей британской колонии — острова Маврикий. В давние времена на этом острове решили вдруг срочно отпечатать свои марки. На этом настояла жена губернатора, которая организовывала пышный прием, а вот марок для приглашений на острове не оказалось. Срочность заказа привела к тому, что гравер в спешке, да еще и спьяну перепутал надпись на марке и вместо «Почтовый сбор оплачен» написал «Почтовая контора». Теперь эти марки стоят у коллекционеров бесценных денег. Такова цена полиграфической опечатки.

Сходная ситуация имеется и в геноме человека. Некоторые «опечатки» в ДНКовом тексте, особенно расположенные в генах или непосредственно рядом с ними, весьма существенны для человека. В качестве примера можно привести широко известный случай одной из мутаций (делеция части ДНКового текста) в человеческом гене CCR5. Укороченный вариант этого гена обеспечивает невосприимчивость людей к вирусу иммунодефицита человека. В результате этого чума XX века — СПИД — таким людям не страшна. В Европе такой вариант довольно широко распространен и достигает в некоторых регионах 9%. А вот в Африке он обнаружен менее чем у 0,1% жителей. Хотя известны различные виды генетического полиморфизма, однако полиморфизм единичных нуклеотидов является самым распространенным типом вариабельности нуклеотидных последовательностей ДНК. Более того, он оказался наиболее перспективным и для чисто практического применения.

Сейчас подсчитано, что в геноме человека приблизительно 3 млн. п.н. должны быть вариабельными. Такие вариабельные (полиморфные) участки были названы снипсами (по-английски snps — от Single Nucleotide Polymorphisms — единичные (сингулярные) нуклеотидные полиморфизмы). Термином «снипсы» обозначаются «точечные» отличия (полиморфизмы) в нуклеотидных последовательностях ДНК различных индивидуумов, вызванные заменой единичного химического звена ДНК-нуклеотида. Постараемся запомнить этот непривычный пока еще термин — далее это нам очень пригодится.

Уже в процессе секвенирования генома человека в 1999 г. был создан специальный международный консорциум для идентификации и картирования только снипсов — «The SNP Consortium Ltd». Целью нового многосотмиллионного исследовательского проекта стало создание сводной карты, показывающей расположение в геноме человека этих «генетических маркеров». В настоящее время установлены 2,2 млн. снипсов (напомним, что общее их число считается равным 3 млн.). Поскольку гены могут содержать от тысячи до миллиона нуклеотидов, то практически в каждом гене или рядом с ним можно обнаружить один или несколько снипсов. Сейчас выявлено 33,5 тыс. снипсов, расположенных в экзонах, что примерно соответствует общему числу генов у человека. Следовательно, большая часть снипсов (около 99%) расположена в тех участках молекулы ДНК, которые не кодируют синтез белков. По этой причине, как правило, снипсы не имеют существенного биологического эффекта. Хотя это и не всегда так.

Изначально было ясно, что создание карты снипсов — маркеров человеческого генома может стать весьма полезным для проведения различных генетических исследований в медицине и фармакологии, даст ключ к пониманию уникальности личности, роли наследственности в интеллектуальных способностях и чертах характера. И снипсы стали использовать, в первую очередь, в качестве генетических маркеров, поскольку пока нет возможности секвенировать целиком геном каждого индивидуума. Нет сомнения, что в обозримом будущем на этой основе станет возможным создание генетического паспорта каждого человека.

Когда такие опечатки в ДНКовом тексте, как снипсы, не меняют сильно сам текст, их называют нейтральными или «молчащими» мутациями. Однако имеются опечатки, которые так искажают «слова», что их уже невозможно прочесть. Их называют значащими, «говорящими» мутациями. Такие значащие мутации часто приводят к возникновению разнообразных наследственных болезней. Поскольку снипсы достаточно стабильно передаются от поколения к поколению, их можно использовать в качестве маркеров в генетических исследованиях по выявлению связи генов с теми или иными заболеваниями, по пониманию особенностей метаболизма лекарственных веществ у разных пациентов.

Так, например, сравнивая «профиль» снипсов в группах больных и здоровых лиц, можно установить, какие из них наблюдаются чаще у пациентов с данным заболеванием. Последующий анализ этих и соседних со снипсами участков ДНК дает возможность идентифицировать гены, которые обуславливают предрасположенность к заболеванию. Таким образом, применение этого метода позволяет ускорить поиск генов, ответственных за развитие заболеваний, сделав его более направленным. В частности, ученые компании «Glaxo Wellcome» уже использовали методику картирования снипсов для выявления генов, принимающих участие в развитии таких распространенных заболеваний, как сахарный диабет II типа, псориаз и мигрень.

Такой же подход можно использовать для выявления у пациентов особенностей метаболизма, распределения или выведения лекарственных средств, то есть для исследований по фармакогеномике (подробнее об этом речь пойдет далее в специальном разделе). Врачи смогут использовать данные генетических исследований для обоснования индивидуального подхода при назначении лекарственных средств, что значительно снизит риск развития побочных реакций. Хотя перспективность такого подхода очевидна, подробная карта снипсов-маркеров пока еще не составлена. Но работа в этом направлении идет очень активно. Составление таких карт человеческого генома сделает в конечном итоге технически возможным проведение быстрого и относительно недорогого генотипирования большого количества пациентов.

Поскольку снипсов в геноме человека довольно много, то было предложено изучать их не только по отдельности, но и группами, в которые объединяются снипсы, расположенные относительно рядом в геноме. Такие наследуемые вместе группы снипсов получили название гаплотипов. Проведенный анализ показал, что число гаплотипов для каждого гена сильно варьирует и может составлять от 2 до 53 (в среднем, по разным оценкам, — от 4 до 14). Было проведено сравнение таких гаплотипов у 21 белого человека, 20 афроамериканцев, 20 человек азиатского происхождения, 18 латиноамериканцев и 3 американских индейцев. В результате не удалось обнаружить ва-

риантов гаплотипов, характерных для той или иной этнической группы. Вместе с тем некоторые варианты чаще встречались у людей, которые родом из одного географического региона.

По оценке ученых компании Genaissance Pharmaceuticals в среднем существует 14 вариаций каждого человеческого гена. Ученые проанализировали 313 генов у большого числа американцев, относящихся к 4 расам. Если в геноме человека присутствует примерно 35 тысяч человеческих генов, следовательно, имеется от 400 до 500 тысяч их вариаций. Обнаружение столь большой амплитуды генетических изменений может потребовать пересмотра традиционных представлений о геноме человека. Сейчас планируется создание каталога гаплотипов каждого человеческого гена. Для этого потребуется проанализировать ДНК у тысяч представителей разных континентов нашей планеты.

Помимо высокой плотности, снипсы имеют очень низкий уровень мутаций на поколение (~10–8), что делает их удобными маркерами для анализа молекулярной эволюции (далее мы поговорим о «молекулярных часах», которые эффективно используются генетиками в геногеографии и эволюционной геномике).

Другой вид полиморфизма генома человека связан с повышенной нестабильностью таких уже упоминавшихся повторов генома, как мини- и микросателлиты. Эта нестабильность проявляется в разном числе повторяющихся элементов в отдельных локусах у разных индивидуумов и приводит к сильному полиморфизму. Уровни нестабильности таких тандемно расположенных повторов значительно отличаются в разных локусах. Особое внимание привлекают к себе гипермутабельные сателлиты. В частности, нестабильность ди- и тринуклеотидных минисателлитов обуславливает так называемые динамические мутации, при которых увеличение числа копий повторяющихся элементов (экспансия) выше некоего критического уровня приводит к опухолевому росту клеток и развитию некоторых наследственных нейромышечных заболеваний (миотоническая дистрофия, хорей Хантингтона, болезнь Кеннеди и др.).

В результате экспансии число повторяющихся элементов в одном локусе может достигать нескольких тысяч. Например, число тринуклеотидных повторов ЦТГ в гене миотонинпротеинкиназы в норме не превышает 27, а при развитии миотонической дистрофии это число увеличивается обычно в тысячу и более раз. Причины такой нестабильности пока до конца не ясны. Как наиболее вероятный механизм этого явления — «проскальзывание» в ходе репликации ДНК. Сильная изменчивость мини- и микросателлитов нашла практическое применение и в так называемой геномной дактилоскопии. Это отдельная тема.

Молекулярные часы

Достаточно твердо установлено, что большинство из происходящих точечных мутаций в геноме не представляются ни полезными, ни вредными для человека, то есть являются нейтральными. Скачки скорее всего не свойственны эволюции. Частоты возникновения точечных мутаций невысоки (порядка 10–8 на генерацию) и относительно постоянны. Расчеты показывают, что при размере диплоидного генома человека около 7×10^9 п. н. в нем происходит в среднем 175 новых мутаций на генерацию (25 лет).

Накопление мутаций можно сравнить с перетеканием песчинок в песочных часах, которые служат мерой времени. Но на практике ученые применили не песочные, а «молекулярные часы». Поскольку большинство мутаций никак не сказываются на самом организме, они могут спокойно сохраняться в эволюции на протяжении длительного времени. Именно за эти факты и ухватились ученые, когда создавали новый подход, названный «молекулярными часами».

На первом этапе была проведена «юстировка» этих часов. Для этой цели были использованы данные о скорости изменения ДНК разных видов организмов, для которых время расхождения в эволюции было надежно установлено по палеонтологическим и археологическим данным. И только после этого началось «чтение» разных эволюционных глав Энциклопедии человека.

«Молекулярные часы» довольно строго показывают, как часто за миллион лет (в среднем) в ДНК того или иного гена или генома в целом происходят мутации. На этой основе по различиям в

ДНК можно судить о том, как давно два разных вида организмов, время возникновения которых не известно, были одним видом, когда произошла их дивергенция, то есть расхождение на две разные ветви эволюции.

Таким же путем можно сравнивать человеческие популяции или даже ДНК отдельных людей и судить об общности их происхождения или родственных связях. В частности, по «молекулярным часам» удалось оценить более точно время отделения человека от обезьян. Согласно молекулярной оценке, это произошло примерно 5 млн. лет назад. Это внесло существенную поправку в данные палеонтологов, которые долгое время полагали, что разделение человека и обезьян произошло около 25 млн. лет назад.

Массовый анализ митДНК из разных рас показал, что одни митДНК отличаются от других числом замен одних нуклеотидов другими, то есть числом мутаций. Были определены такие показатели, как количество индивидуальных мутаций, их расположение и тип. Эти получившие широкую известность данные выявили общность происхождения всех ныне живущих людей по женской линии. Если читатель не прекратит знакомство с книгой в этом месте, то более подробный рассказ об этих интереснейших исследованиях он найдет далее.

Метод «молекулярных часов», применяемый к ДНКовым текстам, очень похож по своему смыслу на метод глоттохронологии⁵, используемый в лингвистике при установлении родства разных языков. Это специальный статистический метод определения возраста родственных языков, т. е. давности их разделения, по количеству слов, имеющих в этих языках одинаковое происхождение. Ведь язык человека, как и геном человека, все время меняется. Если племя или народ, говорящий на одном языке, по какой-то причине разделится на две части, которые потеряют связь между собой, то язык, как и геном, каждого из этих двух новых племен или народов будет меняться по-своему. Чем больше времени пройдет после разделения двух народов, тем меньше общего сохранится в их языках и геномах, но они по-прежнему останутся родственными.

По степени сходства можно судить о том, когда произошло разделение. Лингвисты установили, что за 1000 лет в так называемом базовом словаре (он включает те слова, которые есть в любом языке, — «дом», «земля», «небо», названия частей тела и т. д.) сохраняется 86% слов, т. е. каждый из языков двух народов, обособившихся 1000 лет назад, имеет 86% общих слов с предковым языком. В результате, друг с другом эти языки имеют 74% (86% от 86%) общих слов. При сопоставлении эволюционного древа популяций человека с лингвистическим анализом выяснилось, что в большинстве случаев языки генетически родственных популяций принадлежат к одной лингвистической группе. Вывод ученых гласит: чем раньше разделились две популяции, тем дольше они эволюционировали независимо и тем больше накопилось замен, как в их ДНК, так и в их языках. Конечно, языки напрямую не зависят от генов, и корреляции генетического и лингвистического родства определяются лишь историческими обстоятельствами. Но для нас важно, что здесь одно исследование довольно часто подтверждает другое.

Хромосомы - отдельные части целого

Организм человека поддерживается 24 информационными ресурсами. Фактически это 24 отдельных тома ЭНЦИКЛОПЕДИИ — хромосом, которые исторически были пронумерованы согласно цитологически измеренному размеру: наибольший размер у хромосомы 1 (том № 1) и наименьший — у хромосомы 22 (том № 22). Хотя со временем выяснилось, что хромосома 22 содер-

⁵ Глоттохронология (от др.-греч «время» и λόγος «слово», «учение») — гипотетический метод сравнительно-исторического языкознания для предположительного определения времени разделения родственных языков, основанный на гипотезе, что скорость изменения базового словаря языка остается примерно одинаковой. Эта гипотеза предложена Моррисом Сводешом как попытка аналогии с радиоуглеродным методом измерения возраста органических веществ. В лингвистике предлагается оценивать «лексический полураспад». Этим методом определяется период времени, за который два или более языка разошлись от общего праязыка, путем подсчёта количества заменённых слов в каждом языке. Затем вычисляется приблизительное время появления этих языков. Глоттохронология является приложением лексикостатистики,

жит больше ДНК, чем хромосома 21, порядок нумерации томов не стали менять, чтобы не вносить путаницу. Отдельно идут две половые хромосомы: X и Y (условно их можно назвать томами № 23 и № 24). Разные тома Энциклопедии сильно отличаются друг от друга как по числу страниц (размер ДНКового текста в хромосомах человека варьирует от 50 до 280 млн. п. н.), так и по важности того текста, который в них записан (в первую очередь по числу и свойствам генов). Тем не менее, только совокупность всех хромосом обеспечивает клетки полной информацией, позволяющей человеку нормально развиваться и жить.

В Энциклопедии женщин в норме содержится лишь 23 тома из 24-х, и все они представлены в соматических клетках двумя экземплярами. У мужчин же в клетках содержится вся Энциклопедия человека (все 24 тома-хромосомы), но два из них (хромосомы X и Y) существуют в единичных экземплярах. Это весьма строгое правило. Но, как и любое другое, оно не без исключений. В огромной человеческой популяции встречаются разные варианты. Известно множество хромосомных аномалий, и все они приводят к тяжелым хромосомным болезням человека. Их можно разделить на две группы: вызванные геномными мутациями, т. е. изменением числа хромосом (полиплоидии, анеуплоидии) при сохранении структуры последних, и обусловленные хромосомными мутациями, т. е. изменением структуры хромосомы (транслокации, делеции, инверсии).

В отсутствие любой из пар томов-хромосом в Энциклопедии человека жизнь конкретного индивидуума становится невозможной. Причина понятна — исчезает целостный план-эскиз, необходимый для формирования и/или существования организма. При потере по каким-либо причинам только одной из пары хромосом состояние человека сильно отличается от нормы. Так, например, изредка у людей встречается синдром Шерешевского—Тёрнера, который обусловлен моносомией по X-хромосоме. Этот синдром сопровождается недоразвитием половых органов, низким ростом и рядом врожденных пороков развития. Частичная моносомия хромосомы 5 (потеря части короткого плеча этой хромосомы) приводит к синдрому кошачьего крика. У детей с этой аномалией отмечается необычный плач, что обусловлено изменением гортани, а также черепа и лица. Здесь можно провести такую аналогию. Если в типографии при плохом издании многотомного произведения произойдет потеря лишь одного из томов или только части текста в одной из глав, это сделает издание неполноценным и существенно обеднит его для читателя.

С другой стороны, трудно себе представить, что кто-то, покупая многотомное издание, приобретет какой-то том в двойном количестве. Иногда при издании книг в результате типографических ошибок значительная часть текста повторяется несколько раз, что, безусловно, плохо характеризует издательство и вызывает раздражение у читателя. Такое же «раздражение» возникает и в клетках организма при появлении в нем изредка в силу тех или других причин большего числа томов-хромосом, чем в норме.

Вроде бы ничего страшного не произошло — просто определенная информация дублировалась. Но и это чаще всего приводит к серьезным патологиям. Наиболее известная из них — болезнь Дауна, возникающая у людей, содержащих три хромосомы 21 (трисомия) вместо двух. Клинический синдром Клайнфельтера обусловлен трисомией XXУ.

В таблице 5 приведены общие сведения о размерах молекул ДНК, содержащихся в отдельных хромосомах, суммарном числе генов в каждой из хромосом и количестве генов, которые к настоящему времени ассоциированы с разнообразными патологиями человека, а также о числе снипсов, о которых мы уже упоминали выше.

Представленные в таблице 5 данные на сегодняшний день отражают хотя и довольно точную, но пока еще некую приближенную к абсолютной ситуацию (до сих пор имеются небольшие отличия при сравнении разных источников и разных баз данных). Предстоит еще полностью закончить секвенирование генома человека, затем необходимо найти и изучить функции всех генов и других элементов генома. Поэтому пока еще нельзя дать полную и окончательную характеристику всех томов Энциклопедии человека. Но тем не менее общий план уже ясен, уже выявлены некоторые хромосомо-специфические моменты, особенности устройства отдельных хромосом.

Одной из целей исследования генома человека являлось построение точной и подробной карты всех этих хромосом. Генетическая карта представляет собой схему, описывающую порядок

расположения на хромосоме генов и других генетических элементов-маркеров, а также расстояние между ними. Карты генома, как и географические карты, можно строить в разном масштабе. Максимально возможный уровень разрешения, достигаемый на генетической карте, — один нуклеотид, т. е. одна буква молекулы ДНК. В таком варианте генетическая карта не менее точна, чем, скажем, химическая формула, алгоритм или чертеж. Но для изображения таких карт генома человека требуется несколько сотен тысяч книг.

На рис. 26 изображены карты всех хромосом человека в уменьшенном масштабе, позволяющем наглядно отразить ряд основных параметров, характерных для хромосом: распределение снипсов, распределение установленных и гипотетических генов, расположение повторов, а также количество ГЦ-пар нуклеотидов в разных участках. Если эти карты рассматривать по отдельности, то может сложиться впечатление, что они очень похожи. Но при детальном сравнении выясняется, что на самом деле они существенно отличаются. «Узоры» расположения разных элементов на представленных картах весьма индивидуальны для разных хромосом, их можно сравнить с папиллярными узорами пальцев, которые строго специфичны для каждого человека.

Таблица 5. Общая характеристика хромосом —

Из базы данных Ensembl <http://www.ensembl.org> (октябрь 2002 г.)

Хромосома	Содержание ДНК, млн. п. н.	Минимальное оценочное количество генов	Число генов, ассоциированных с болезнями	Число снипсов, тыс.
1	263	2237	157	180
2	255	1558	103	163
3	214	1213	93	109
4	203	900	67	111
5	194	1042	82	113
6	183	1195	90	128
7	171	1132	79	124
8	155	780	53	99
9	145	899	66	131
10	144	895	66	110
11	144	1239	135	92
12	143	1089	92	77
13	114	384	33	66
14	109	737	54	59
15	106	745	49	68
16	98	996	68	72
17	92	1206	98	59
18	85	338	30	59
19	67	1437	70	39
20	63	593	36	48
21	50	204	23	55
22	56	469	40	47
X	164	717	208	65
Y	59	82	3	27
Сумма	3277	22087	1795	2101

Теперь рассмотрим по отдельности краткие аннотации содержанию ДНК-овых текстов человека, расположенных в 24 томах-хромосомах. Начнем по порядку, т. е. с самой большой хромосомы под номером 1.

Хромосома 1

Эта наибольшая по физическим размерам и по содержанию ДНК хромосома человека. На ее долю приходится почти 10% генома человека. Поскольку хромосома 1 самая большая, то вполне естественно, что в ней содержится больше всего генов и больше всего снипсов, то есть полиморфных участков. По последним оценкам число генов в этой хромосоме приближается к 3000. Гены и снипсы расположены неравномерно вдоль хромосомы, они практически отсутствуют в области центромеры.

подавляющая часть снипсов концентрируется в трех участках хромосомы 1. В этой хромосоме имеется множество (порядка 160) генов, связанных с разнообразными заболеваниями человека. Многие из этих болезней едва ли знакомы читателю, но для тех, кто с ними столкнулся в жизни, они являются как правило очень большой бедой. Перечислим лишь некоторые из них, связанные с генами хромосомы 1: болезнь Альцгеймера, нейропатия–2а, болезнь Гоше, синдром Чедиака—Хигаси, рак протоков молочной железы, кардиомиопатия, катаракта, эктодермальная дисплазия, гипотирозидизм, острая лимфобластная лейкемия, нейробластома, рак простаты, атеросклероз.

Хромосома 2

Это вторая по размерам хромосома. Наибольшая плотность снипсов имеется в районе центромеры, а вот повторы здесь практически отсутствуют. На единицу длины в ней содержится заметно меньше генов, чем в хромосоме 1 и ряде других хромосом. Тем не менее, число заболеваний, связанных с мутациями в генах этой хромосомы, достаточно большое. Вот некоторые из них: синдром Альстрема, голопрозенцефалия 2, дистрофия большеберцовой мышцы, аутосомно-рецессивная глухота–9, дистрофия мышц конечности 2b, рак прямой кишки, цистинурия, диабет, фиброматозис, гипотирозидизм, ожирение, болезнь Паркинсона, тромбофилия, синдром Валленберга, синдром Гилберта и др.

Хромосома 3

Это еще одна довольно большая хромосома. В отличие от хромосомы 2 у нее в области центромеры содержится мало как снипсов, так и повторов. Наибольшее количество снипсов расположено ближе к концам этой хромосомы, а наибольшее число генов — на коротком плече. Гены, содержащиеся в этой хромосоме, связаны с более чем 90 различными заболеваниями, среди которых можно отметить такие, как кардиомиопатия, рак прямой кишки, колоректальный рак, диабет, гемолитическая анемия, гипокальцемия, миелоидный лейкоз, В-клеточная лимфома, миотоническая дистония, карцинома почки, шизофрения.

Хромосома 4

Гены, повторы и снипсы распределены в хромосоме 4 довольно равномерно (за исключением района центромеры, где все они представлены малым количеством). Подсчитано, что общее число генов здесь меньше, чем в среднем на единицу длины генома. Среди заболеваний, ассоциированных с мутациями в генах этой хромосомы, можно отметить такие, как дисфибриногенемия, болезнь Хантингтона, синдром Эллиса—Ван-Кревельда, гипохондроплазия, гипогонадизм, маннозидозис, болезнь Паркинсона, фенилкетонурия, синдром Ригера, острый иммунный дефицит, синдром Вольфрама, склонность к алкоголизму и др.

Хромосома 5

Большинство генов этой хромосомы сконцентрировано в двух областях длинного плеча и одном районе короткого ближе к его концу. Имеются два района, расположенных вокруг центромеры, обогащенные снипсами. С генами хромосомы 5 связан ряд тяжелых заболеваний: мегало-

пластическая анемия, колоректальный рак, капиллярная гемангиома, дистрофия роговицы, ауто-сомно-доминантная глухота, синдром Гарднера, болезнь Гиршпрунга, кетоацидоз, острая промиелоцитарная лейкемия, мышечная дистрофия, миелодиспластический синдром, астма и др.

Хромосома 6

Плотность и генов и снипсов наибольшая в нескольких районах на коротком плече этой хромосомы, а вот повторы распределены вдоль хромосомы довольно равномерно (их мало только в области центромеры). С генами хромосомы 6 связан ряд патологий человека: диабет, спиноцеребральная атрофия, гемолитическая анемия, аргининемия, дислексия, гепатоцеллюлярная карцинома, лейкемия, множественная миелома, болезнь Паркинсона, тромбофилия, чувствительность к туберкулезу и др.

Хромосома 7

Плотность снипсов наибольшая в прицентромерной области длинного плеча этой хромосомы. А вот гены расположены довольно равномерно вдоль хромосомы, за исключением одного участка в середине длинного плеча, где содержится наибольшее их количество. Среди заболеваний, ассоциированных с генами хромосомы 7, можно отметить такие, как хронический гранулематоз, рак прямой кишки, кистозный фиброз, ауто-сомно-доминантная глухота, вялая кожа, эритремия, гемолитическая анемия, карликовость, семейный гиперинсулинизм, врожденная миотония, остеопороз, панкреатит, трипсиногеновая недостаточность, болезнь коронарной артерии и др.

Хромосома 8

Большинство снипсов в этой хромосоме сконцентрировано на конце короткого плеча, а на конце длинного плеча имеется область, сильно обогащенная генами. Число генов, ассоциированных с заболеваниями, в хромосоме 8 относительно небольшое. Среди них имеются гены, мутации в которых приводят к таким заболеваниям, как хондросаркома, эпилепсия, гипотиреозидизм, синдром Пфаффера, восприимчивость к атеросклерозу, синдром Вернера, лимфома Беркитта, сфероцитоз и ряд других.

Хромосома 9

Здесь и снипсы, и повторы, и гены распределены очень неравномерно вдоль хромосомы. Кроме того, хромосома 9 обогащена снипсами по сравнению с другими хромосомами (при расчете их числа на единицу длины). При этом наибольшее их число сконцентрировано в прицентромерной области длинного плеча. Одновременно эта же область, наоборот, обеднена повторами. Целый ряд заболеваний ассоциирован с мутациями в генах хромосомы 9: альбинизм, карцинома базальных клеток, недостаточность карнитинацетилтрансферазы, торзионная дистония, эпителиома, гиперглицинемия, ангиомиолипома, галактоземия, хроническая миелоидная лейкемия, меланома, порфирия, стоматоцитоз, синдром Валкера—Варбурга.

Хромосома 10

Эта хромосома является средней по числу содержащихся в ней генов, повторяющихся участков и снипсов на единицу длины, но распределение их по хромосоме далеко не равномерное: несколько участков на длинном плече сильно обогащены генами и снипсами. Среди заболеваний, ассоциированных с этой хромосомой, можно отметить такие, как кардиомиопатия, почечная гиперплазия, катаракта, болезнь Гиршпрунга, синдром Джексона—Вейсса, острая Т-клеточная лейкемия, болезнь Гоше, глиобластома, гиперфенилаланинемия, эндокринная неоплазия, наследственная нейропатия, аденокарцинома простаты, шизэнцефалия, орнитинаминотрансферазная недостаточность.

Хромосома 11

На конце короткого плеча и в прицентромерном районе длинного плеча этой хромосомы имеет место концентрация генов. Содержание снипсов повышено лишь в районе конца короткого

плеча, а вдоль хромосомы оно относительно одинаковое. От общего числа генов этой хромосомы доля генов, с которыми ассоциированы различные патологии, здесь одна из наибольших. Среди заболеваний, определяемых генными мутациями в хромосоме 11, имеются такие, как альбинизм, амилоидоз, злокачественная анемия, не-Ходжкинская лимфома, рак груди, рак мочевого пузыря, рак простаты, глухота, эритремия, гиперпротромбинемия, острый комбинированный иммунодефицит, мужское бесплодие, множественная миелома, аномалия Петерса, серповидноклеточная анемия, талассемия, тромбоцитопения, пигментная ксеродерма, адренкортикальная карцинома, наследственная ангиодемия, гиперпроинсулинемия, параганглиома, паратироидный аденоматоз, остеопороз и др.

Хромосома 12

Эта хромосома является средней по большинству параметров. Гены распределены в ней весьма неравномерно. С ними ассоциирован ряд заболеваний: адренолейкодистрофия, амилоидоз, злокачественная неходжкинская лимфома, рак прямой кишки, эмфизема, энурез, задержка роста, гемолитическая анемия, кератодерма, липома, множественный липоматоз, наследственная миопатия, остеоартроз, фенилкетонурия, аденома слюнных желез, синдром Вагнера, болезнь фон Виллебранда, гистидинемия и др.

Хромосома 13

Короткое плечо этой хромосомы пока плохо секвенировано. Имеется концентрация снипсов в районе центромеры на длинном плече. Хромосома 13 относительно других хромосом обеднена генами (на 1 млн. букв в ней в среднем приходится всего около 5 генов). Наибольшее их количество содержится на обоих концах длинного плеча. Тем не менее, ряд тяжелых заболеваний ассоциирован с мутациями в генах этой хромосомы: рак мочевого пузыря, глухота, недостаточность факторов свертываемости крови VII и X, мышечная дистрофия, панкреатический агенезис, рак поджелудочной железы, пинеалома, ретинобластома, болезнь Вилсона, пигментная ксеродерма и др.

Хромосома 14

По картине распределения генов, повторяющихся последовательностей и снипсов эта хромосома весьма напоминает хромосому 13. Ее короткое плечо также долгое время (до января 2003 г.) не было секвенировано. Но по числу генов различия между двумя этими хромосомами значительны (в хромосоме 14 их примерно в два раза больше на единицу длины). В хромосоме 14 содержатся гены, важные для работы иммунной системы. С мутациями в ее генах (их около 60) связан ряд серьезных заболеваний: ранняя форма болезни Альцгеймера, разные формы потери слуха, кардиомиопатия, дистония, эллиптоцитоз, фибродисплазия, гипертироидизм, сфероцитоз, фенилкетонурия, тироидная аденома, температурочувствительный апоптоз, агаммаглобинемия.

Хромосома 15

Эта хромосома тоже пока секвенирована не полностью. В ней имеется несколько областей, обогащенных снипсами и генами, которые расположены в разных участках длинного плеча. Микроделеции в определенном районе хромосомы 15 приводят к возникновению синдрома Эйнджелмена и синдрома Прадера—Вилли. Выявлен большой спектр заболеваний, ассоциированных с мутациями в генах этой хромосомы: альбинизм, синдром Барттера, синдром Блюма, гинекомастия, острая промиелоцитарная лейкемия, гипомеланоз, мышечная дистрофия, шизофрения, тирозинемия, эпилепсия и ряд др.

Хромосома 16

В этой хромосоме содержание генов на единицу длины выше среднего по геному. Большинство генов, как и большинство снипсов, сконцентрировано в 5–6 областях хромосомы. С генами хромосомы 16 связан ряд серьезных генетических заболеваний, таких, как рак груди, рак желудка, синдром Гительмана, хронический грануломатоз, эритроцитоз, болезнь рыбьих глаз, гепа-

тический гликогенез, альфа-талассемия, поликистозная болезнь почек, миелоидная лейкемия, карцинома яичника, спастическая параплегия, тирозиемия, мукополисахаридоз, синдром Лиддла, болезнь Моркио и др.

Хромосома 17

Здесь также имеет место довольно высокое содержание генов по сравнению с другими хромосомами (при пересчете на единицу длины). Выявленные гены, так же как и повторы и снипсы, распределены вдоль хромосомы неравномерно. Большинство снипсов содержится в районах длинного и короткого плечей, прилежащих к центромере. Ряд заболеваний ассоциирован с мутациями в генах этой хромосомы: спорадический рак груди, рак прямой кишки, диабет, буллезный эпидермолиз, гемолитическая анемия, эссенциальная гипертензия, рак языка, острая лейкемия, мышечная дистрофия, миостенический синдром, миелопероксидазная недостаточность, наследственная миотония, нейробластома, рак яичника, синдром Ватсона и др.

Хромосома 18

Как общее число генов, так и число генов, мутации в которых связаны с патологиями человека, в ней относительно невелико. В области центромеры практически отсутствуют гены, повторы и снипсы. Можно отметить концентрацию снипсов в прицентромерной области короткого плеча этой хромосомы. Среди заболеваний, ассоциированных с этой хромосомой, имеются амилоидоз, рак прямой кишки, буллезный эпидермолиз, рак поджелудочной железы, В-клеточная лейкемия/лимфома, эритропоэтическая протопорфирия и др.

Хромосома 19

ДНК, входящая в эту хромосому, имеет размер около 67 млн. п. н. Ее особенностью является то, что она наиболее богата ГЦ-парами нуклеотидов по сравнению с другими хромосомами человека. В ней также содержится большее, чем теоретически можно было рассчитать, число генов и подобных им участков (экспрессирующихся последовательностей, так называемых ESTs), что находится в соответствии с высоким ГЦ-содержанием ДНК этой хромосомы (на 1 млн. п.н. приходится 23 гена при среднем значении по геному около 10). Еще одно отличие этой хромосомы от других заключается в том, что в ней имеются последовательности, которые гомологичны последовательностям, расположенным на шестнадцати других хромосомах человека. Кроме того, для хромосомы 19 характерна наибольшая плотность расположения минисателлитов. Среди заболеваний, ассоциированных с этой хромосомой, имеются такие серьезные патологии, как гипертрофическая кардиомиопатия, рак прямой кишки, церебральная артериопатия, сахарный идиопатический диабет, гиперлиппротеинемия, наследственный гипотирозидизм, острая лимфобластная лейкемия, миотоническая дистрофия, эритроцитоз, атеросклероз коронарной артерии и многие др.

Хромосома 20

Хромосома 20 стала третьей по времени полностью секвенированной хромосомой человека. По размеру эта хромосома составляет всего около двух процентов генетического кода генома человека. Гены, повторы и снипсы распределены вдоль хромосомы весьма неравномерно. В настоящее время выявлено, что в хромосоме 20 наряду с 600–700 генами, расположенными в нескольких основных областях, содержится множество (около 170) псевдогенов. Гены этой хромосомы несут информацию о ряде заболеваний человека, начиная от ожирения и экземы и заканчивая слабоумием и катарактой. С мутациями в генах хромосомы 20 связаны такие генетические заболевания, как болезнь Крейтцфельда—Якоба, тяжелые нарушения иммунной системы, болезни сердца, астма, скелетная дисплазия, гепатопленомегалия, диабет и др.

Хромосома 21

Эта хромосома является самой маленькой по размерам и информационной емкости (на ее долю приходится не более 1,5% от всего генома человека). Но секвенирована она была только вслед за хромосомой 22. Число генов в хромосоме 21 относительно невелико. При размере около

50 млн. пар нуклеотидов в ней обнаружено пока только около 200 генов. Для сравнения, 22-ая хромосома при чуть большем размере (56 млн. пар нуклеотидов) содержит почти 500 генов. Иногда говорят, что 21-ая хромосома представляет из себя «геномную пустыню». В ней, например, есть участок длиной в 7 млн. пар нуклеотидов (это больше полного генома бактерии *Escherichia coli*), содержащий только один ген. Имеющиеся гены концентрируются главным образом на конце длинного плеча, а снипсы — в области центромеры.

Хромосома 21 хорошо известна тем, что при наличии в клетках трех ее копий возникает болезнь Дауна. Предполагается, что из-за малого числа генов в этой хромосоме люди, больные синдромом Дауна, доживают до взрослого состояния (тогда как дети с тремя копиями других хромосом, кроме половых хромосом X и Y, после рождения живут всего несколько дней).

Мутации в генах хромосомы 21 способны вызывать такие серьезные болезни, как голопрозэнцефалию, синдром Ушера и некоторые формы злокачественных опухолей.

Хромосома 22

ДНК этой хромосомы была секвенирована первой (декабрь 1999 г.), поэтому она и описана более полно. В хромосоме 22 остались нерасшифрованными всего несколько участков (менее 3% длины ДНК). Она содержит около 500 генов и 134 псевдогена. Все эти генные последовательности вместе с внутригенными некодирующими районами (интронами) составляют всего 13 млн. п. н. Среди генов, присутствующей в хромосоме 22, имеются такие известные, как гены иммуноглобулинов, генные семейства, кодирующие различные ферменты, аполипопротеины и кристаллины. Подсчитано, что средний размер гена этой хромосомы — 19,2 тыс. п. н., хотя самый маленький состоит лишь из 1 тыс. п. н., а самый большой — из 583 тыс. п. н. Несколько генов содержат только один экзон, но, с другой стороны, имеется ген, состоящий аж из 54 экзонов.

Величины экзонов варьируют от 8 до 7,6 тыс. п. н. Установлено, что два известных гена локализованы в протяженных интронах двух других генов («генные матрешки»). Что касается псевдогенов, на их долю приходится всего 204 тыс. п. н., причем большинство — это процессированные псевдогены, т. е. у них отсутствуют интроны. Установлено, что псевдогены хромосомы 22 в основном относятся к семействам иммуноглобулинов, кристаллинов, цитохромов и др. Обнаружен только один кластер, состоящий из 26 псевдогенов, рядом с центромерной областью хромосомы. Математическая обработка полученных данных показала, что еще определены не все гены и псевдогены хромосомы 22. Различные повторяющиеся последовательности, которые не кодируют белки, составляют 41,9% этой хромосомы. Алу-повторы формируют блоки в районе центромеры и почти в центре длинного плеча.

Для хромосомы 22 в настоящее время установлены функции примерно половины генов. Около 160 генов, расположенных на хромосоме 22, показывают значительную гомологию с генами мыши. Несмотря на свои небольшие размеры и малое число генов, ее патология установлена при некоторых генетических и онкологических заболеваниях. Сейчас известно 27 заболеваний, вызванных нарушениями в разных генах хромосомы 22 человека: от рака (миелоидная лейкемия) до предрасположенности к шизофрении и паркинсонизму, а также серьезных аномалий сердца и нервной системы. Известны многочисленные геномные и хромосомные мутации, связанные с этой хромосомой. Так, трисомия (три вместо двух копий) хромосомы 22 вызывает синдром «кошачьего глаза» (колобома радужной оболочки), атрезию ануса, некоторые пороки развития и умственную отсталость. Трисомия хромосомы 22 — вторая по значению причина выкидышей у беременных.

Выпадение (делеция) одного из районов длинного плеча (22q11.2) приводит либо к синдрому Ди Джорджи, который сопровождается аплазией тимуса, пороками сердца и аномалиями развития, несовместимыми с жизнью, либо, если делеция меньших размеров, к велокардиофациальному синдрому с характерными пороками сердца и крупных сосудов. При лейкозах и лимфомах также выявлены трисомии и моносомии, обмены участками (транслокации) различных хромосом и хромосомы 22. Самый известный пример — филадельфийская хромосома, образованная в результате транслокации между хромосомами 9 и 22. В солидных опухолях также достаточно часто обнаруживают различные транслокации с вовлечением хромосомы 22.

Хромосома X

Это женская половая хромосома. Наличие двух хромосом X определяет женский пол. Пара для хромосомы X у мужчин — омертвевшая и короткая Y-хромосома. У женщин в одной из 2 хромосом X происходит инактивация всех тех генов, которые не имеют пары на хромосоме Y. В ходе эмбрионального развития эта хромосома на всех нужных участках покрывается особым инактивирующим белком Xist или сжимается «пеленкой» белка-гистона так туго, что не может работать.

Генов в хромосоме X немного, и они сосредоточены, главным образом, в двух областях: прицентромерном районе короткого плеча и на конце длинного плеча. Снимсы концентрируются в середине длинного плеча хромосомы. Укажем лишь на небольшую часть заболеваний, ассоциированных с генами хромосомы X: мышечная дистрофия Дюшенна, сидеробластическая анемия, рак груди, рак простаты, кардиомиопатия, хориоидеремия, дискератоз, эпилепсия, глицеролкиназная недостаточность, гонадотропиновая недостаточность, гемофилия B, ихтиоз, синдром Барта, мукополисахаридоз 2.

Y-хромосома

Эта маленькая половая хромосома определяет мужской пол у человека. Содержащиеся в ней последовательности рассматривают как очень «юные». Скорости мутаций в этой хромосоме в 4 раза выше, чем в хромосоме X. Гены, повторы и снимсы выявлены лишь на левом конце этой хромосомы (правый пока не полностью секвенирован). Здесь содержится большое число палиндромов. По этой причине появилось даже выражение: «Y-хромосома — зал зеркал». А вот генов в ней немного. По последним уточненным данным их максимальное число не достигает и 100. На 1 млн. букв-нуклеотидов в среднем приходится всего около 5 генов. Основная роль тех генов, которые имеются, заключается в контроле дифференцировки пола, формировании яичек и процесса сперматогенеза.

В частности, основной ген «самцовости», названный SRY (Sex-determining Region of Y), кодирует белок, который включает в работу многие гены других хромосом и тем самым вызывает каскад биохимических реакций (конечный результат — образование яичек). Ген SRY практически одинаков у всех людей, но в десять раз более отличен у людей и обезьян, чем все другие их гены. Иными словами, на сегодняшний день это самый консервативный ген внутри вида и самый динамичный между видами. Отмечены случаи, когда в клетках имеется не одна, а две и даже 3 копии хромосомы Y. Характерными признаками такой хромосомной патологии являются асоциальное поведение и различные психологические нарушения, характерные для 35% больных. Совсем немного генов в хромосоме Y ассоциировано с болезнями человека. Основные из них — гонадный дисгенез и синдром клеток Сертоли.

Митохондриальный геном

Когда сейчас громогласно заявляют о полном секвенировании генома человека, то, как правило, подразумевают ядерный геном. На этом фоне как-то забывается, что в клетках имеются молекулы ДНК, расположенные не только в хромосомах, но и в таких уже упоминавшихся специфических внутриклеточных структурах, как митохондрии. И это тоже геном человека, но он называется митохондриальным, а ДНК — митохондриальной (сокращенно митДНК). МитДНК теперь называют иногда хромосомой 25 или M-хромосомой.

Эта ДНК была секвенирована еще в 1981 году уже упоминавшимся Ф. Сенгером, что тоже было в свое время сенсацией, которая, однако, имела резонанс несопоставимо меньший, чем секвенирование ядерного генома. Что же представляет собой эта 25-ая хромосома человека?

В клетке человека насчитывается от 100 до 1000 митохондрий, в каждой из которых содержится от 2 до 10 молекул кольцевой митДНК длиной 16569 п. н. Таким образом, размер митохондриального генома примерно в 200 000 раз меньше ядерного. Интересно, что размер митДНК у человека — один из наименьших среди высших организмов (эукариот). Например, у дрожжей митДНК состоит из 78520 п. н.

Человеческая митДНК содержит 37 генов, кодирующих 13 белковых цепей, 22 тРНК и 2 рибосомных РНК (рРНК) (рис. 27). Белковые цепи входят в состав белков, которые участвуют в основном в важнейшем внутриклеточном процессе, называемом окислительным фосфорилированием, который обеспечивает клетку энергией. За счет окислительного фосфорилирования в митохондриях осуществляется производство более 90% специальных молекул АТФ, являющихся основой энергетики клетки.

Стрелками указано направление транскрипции генов. Сокращения: ND1—ND6, ND4L — гены субъединиц НАД-Н-дегидрогеназного комплекса; COI—COIII — гены субъединиц цитохром-с-оксидазы; ATP6, ATP8 — гены субъединиц АТФ-синтетазы; Cyt b — ген цитохрома b

Всего же в процессе окислительного фосфорилирования задействовано 87 генов, но все недостающие 74 кодируются не митохондриальным, а ядерным геномом. Интересно, что в ядерном геноме обнаруживаются участки, подобные митДНК. Предполагается, что в процессе эволюции и при различных патологиях имела место миграция части митДНК в ядерный геном.

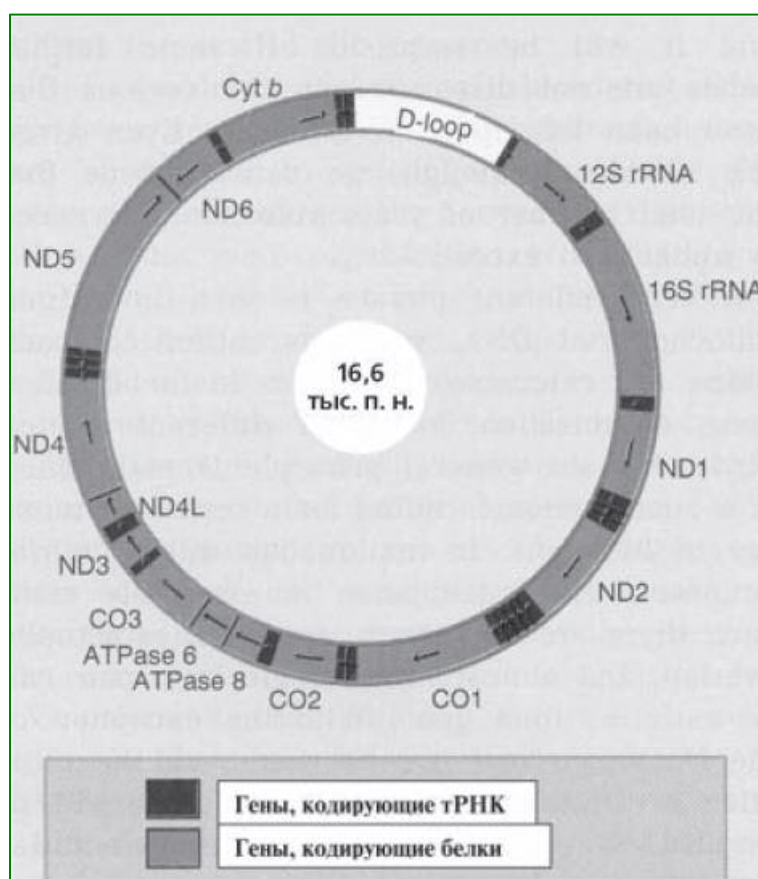


Рис. 27. Структура митохондриального генома человека (митДНК). В митДНК содержится 22 гена, кодирующих тРНК, 2 рибосомных гена (16S и 12S рРНК) и 13 белок-кодирующих генов.

Важно, что устройство митохондриального генома существенно отличается от ядерного. В первую очередь, для митДНК характерно очень компактное расположение генов, как и в геноме бактерий. В отличие от ядерного генома митохондриальные гены соседствуют друг с другом и между ними практически отсутствуют межгенные промежутки. В ряде случаев они даже перекрываются на один нуклеотид: последний нуклеотид одного гена оказывается первым в следующем за ним. То есть гены набиты в митохондриальной ДНК, как сельди в бочке. Кроме того, большинство митохондриальных генов не содержит такие характерные для ядерных генов структуры, как ин-

троны. Но это еще не все отличия. Выяснилось, в частности, что митДНК не подвержена такой модификации, как метилирование, которая характерна для ядерной ДНК.

Однако особенное удивление исследователей вызвал генетический код, используемый в митДНК. Хотя генетический код универсален (за очень небольшим исключением) во всем живом мире, в митохондриях используется некий его необычный вариант. Большинство кодонов в митохондриальных генах сходны с теми, которые имеются в ядерной ДНК, но наряду с этим имеются и принципиальные отличия. В митДНК человека изменили свой смысл четыре кодона. Терминирующими стали кодоны АГА и АГГ. Кодон УГА, являющийся в ядерной ДНК терминирующим, в митДНК не только не вызывает остановки трансляции, а кодирует аминокислоту триптофан. Аминокислоту метионин кодируют не один кодон АУГ, а еще кодон АУА, который в ядерном геноме кодирует аминокислоту изолейцин.

МитДНК ответственна в клетке за синтез всего лишь нескольких митохондриальных белков. Но эти белки очень важны для клетки, поскольку участвуют в осуществлении одного из важнейших процессов — обеспечении клетки энергией. Таким образом, митДНК — весьма ценное приложение к Энциклопедии человека. Белки, кодируемые непосредственно генами митДНК, синтезируются тут же в митохондриях. Для этой цели используется собственная РНК-полимераза и собственный аппарат белкового синтеза. Причина ясна — генетический код митохондрий особый, нужна и особая система биосинтеза.

Далеко не все белки, которые нужны для автономного существования митохондрий, кодируются митохондриальным геномом и синтезируются здесь же. Для этого их геном слишком мал. Большая часть митохондриальных белков и отдельных субъединиц этих белков кодируется основным, т. е. ядерным геномом и синтезируется в цитоплазме клеток. Затем они транспортируются в митохондрии, где взаимодействуют со специфическими белками, кодируемыми митДНК. Таким образом, между ядерным и митохондриальным геномом существует тесная взаимосвязь, они дополняют друг друга.

Почему в эволюции клетки случилось так, что очень небольшая часть ДНК содержится не в хромосомах ядра, а отдельно внутри митохондрий? В чем необходимость или преимущество такого распределения генетического материала, пока неизвестно. Для объяснения этого удивительного факта было придумано много гипотез.

Одну из первых еще в далеком 1890 году высказал Р. Альтман⁶. Однако она и на сегодняшний день сохранила актуальность. Согласно этой точке зрения, митохондрии появились в клетках высших организмов не в ходе внутриклеточного развития и дифференцировки, а в результате естественного симбиоза высших организмов с низшими аэробными организмами. Такое объяснение предполагает, что митохондриальный генетический код более древен, чем код, используемый в ядерной ДНК у современных организмов.

Но наряду с этим высказывалась и иная точка зрения, которая пока в равной мере имеет право на существование. Согласно последней, после перехода большинства генов из митДНК в ядерную ДНК в аппарате, обеспечивающем синтез белка в митохондриях, произошли какие-то мутации. Для того, чтобы процесс трансляции не нарушался, потребовались специальные мутации в генах митДНК, которые бы «компенсировали» нарушения и позволили бы измененному аппарату белкового синтеза осуществлять свою работу. Если исходить из этого предположения, тогда митохондриальный код следует рассматривать не как более древний, а, наоборот, скорее как более молодой.

В любом случае язык митДНК — это в определенном смысле «жаргон». Зачем он нужен митохондриям? Здесь можно провести параллель с жаргонами определенных социальных или профессиональных групп. Жаргон используется ими для сокрытия своих намерений и действий от посторонних, избежания чужого вмешательства в их дела. Возможно, и митДНК, благодаря исполь-

⁶ В теории Альтмана мы видим концепцию, аналогичную клеточной теории, согласно которой сама клетка является элементом, построенным из еще гораздо более мелких живых структур. По гранулярной теории Рихарда Альтмана (1852—1900), которая рассматривает клетку, состоящей из значительного числа зернышек-гранул.

зованию видоизмененного кода — жаргона, — изолируется от белок-синтезирующего аппарата клетки, специализируясь на выполнении одной, но очень важной для клетки функции — производстве энергии.

Замечено, что митохондриальный геном более раним, чем ядерный геном. В результате в нем часто происходят различного рода мутации (точковые мутации, небольшие потери ДНК — делеции и, наоборот, вставки — инсерции). Сейчас уже установлены многочисленные болезни человека, связанные с изменениями в митДНК. Патологические мутации обнаружены почти во всех митохондриальных генах. При этом отмечают огромное разнообразие клинических признаков, обусловленных одним и тем же молекулярным повреждением. Обнаружена взаимосвязь некоторых мутаций и изменений в экспрессии генов митДНК с возникновением рака. В частности, неоднократно отмечалось при раке груди и лимфомах усиление транскрипции гена, кодирующего одну из цепей белкового комплекса, участвующего в снабжении клеток энергией (субъединицы II цитохром-с-оксидазы). Некоторые, к счастью, редкие, тяжелые наследственные болезни человека также обусловлены мутациями в отдельных генах митДНК. В России сейчас существует специальная программа диагностики и профилактики митохондриальных болезней.

Еще один удивительный факт, связанный с митДНК, касается ее наследования. Оказалось, что митДНК передается из поколения в поколение принципиально иначе, чем хромосомная ДНК. Организм человека развивается из оплодотворенной яйцеклетки, которая содержит хромосомы обоих родителей. При оплодотворении в яйцо проникает сперматозоид с набором отцовских хромосом, но практически без отцовских митохондрий и, следовательно, без какой-либо отцовской митДНК. Только яйцеклетка предоставляет зародышу свою митДНК. Это ведет к важным последствиям: митДНК передается только по женской линии. Мы все получаем митДНК только от своей матери, а она еще раньше от своей, и так далее в ряду только женских поколений. Сыновья в отличие от дочерей не передают свою митДНК — цепочка оборвется.

Таким путем ДНК образуют клоны — наследственные линии, которые могут только разветвляться (если у женщины родилось несколько дочерей), но в отличие от хромосомной ДНК не могут объединяться в одном организме и создавать новые генетические комбинации. По этой причине было интересно провести сравнение митДНК у представителей различных человеческих этнических популяций, то есть рас и народностей. Такого рода сравнения были начаты еще в конце 80-х годов прошлого века и продолжаются до сих пор. Подробнее мы еще поговорим об этом далее.

Таким образом, такие базовые процессы в клетке, как транскрипция, трансляция, репликация и репарация митДНК, в значительной мере зависят от ядерного генома, но пока не до конца ясно, как эти два генома интегрированы друг с другом. Изучение механизмов межгеномного взаимодействия может стать полезным во многих отношениях, в частности для понимания интегральной картины различных патологий человека, включая и злокачественное перерождение клеток.

Функциональная геномика

Хорошо известно выражение, что от слова «халва» во рту слаще не становится. Так же обстоит дело и с нашим геномом. В нем есть много разных слов и предложений, но клетке от этого ни сладко, ни горько, пока геном не начинает работать (по научному, функционировать). Только когда в клетке образуются белки, тогда она почувствует, «сладко» ей (если ген работает правильно) или «горько» (при наличии, например, в гене «плохой» мутации).

То, что в природе (да и не только в ней) всякая конкретная структура тесно связана с выполняемой ею определенной функцией, — это аксиома. Но раскрыть и увидеть эту тонкую взаимосвязь порой весьма не просто. Особенно если это касается такой сложнейшей структуры, как геном человека. От относительной статики генома, когда изучалась структура и последовательность нуклеотидов ДНК, теперь ученым предстоит перейти к сверхсложной динамике клеточного

метаболизма, зависящего от этой молекулы жизни. Это есть предмет новейших направлений в молекулярной генетике — функциональной геномики и протеомики.

В геноме человека имеются гены, которые функционируют во всех типах клеток. Такие гены, как говорят, заведуют «домашним хозяйством». Считается, что домоуправителей среди генов около 20%. Но вот остальные 80% ведут себя чрезвычайно специфически. Клетка от клетки сильно отличается по своему виду и функции. И это в первую очередь связано с разными наборами работающих в них генов и белков.

Многие гены, функционирующие в мозгу, не работают ни в печени, ни в селезенке, ни в каком-либо другом органе или ткани. И, наоборот, в печени или селезенке функционирует ряд генов, которые нигде больше не работают. Одни из них работают только на определенных этапах развития организма, обеспечивая формирование нашего облика (хорошего или плохого — это уже другое дело). Другие гены — специализированные — работают только в очень ограниченном наборе клеток.

Хорошо известный пример последних — глобиновые гены. Они функционируют только в клетках крови, поскольку их основная функция — обеспечивать перенос кислорода к другим клеткам и тканям.

Другой пример. Наша способность воспринимать запахи связана с обонятельными рецепторами. Всего найдено около тысячи генов, кодирующих такие белки-рецепторы. Но в каждой клетке специализированного органа — обонятельной луковицы — работает только один ген из этой тысячи. Происходит высочайшая специализация. Как это связано с нашим восприятием запаха, пока совсем не ясно. И вдруг обнаружилось (ученых это сильно удивило), что некоторые из этих генов работают кроме обонятельной луковицы еще только в одном типе клеток — в сперматозоидах.

Итак, можно сказать, что каждая клетка предпочитает «читать» лишь определенные предложения на избранных страницах Энциклопедии человека, в результате чего в ней появляется строго определенный набор молекул мРНК и белков. Другие страницы этой книги для конкретной клетки неинтересны или, может быть, написаны так, что клетка просто не может их прочесть (условно говоря, текст других предложений и страниц слишком мелок для нее, а у клетки нет необходимого для их прочтения «увеличительного стекла»).

Важно понимать наличие фактов: разные возможности для чтения геномного текста существуют в разных типах тканей и на разных этапах развития организма. На этот процесс могут влиять также многочисленные внешние или внутренние факторы, которым постоянно подвержен в природе организм. Но в любом случае клетки нашего мозга, например, не могут вырабатывать желудочный сок ни при каких обстоятельствах. Такая чрезвычайно сложная ситуация существует в миллиардах клеток человека, и все эти ситуации, не поддающиеся пока даже приблизительной оценке в отношении их числа, еще предстоит исследовать.

Путь от обнаружения генов до выяснения их функций весьма непрост. Для полного понимания жизненных процессов, протекающих в клетках и в организме в целом, для выяснения природы различных заболеваний человека необходимо тщательное изучение молекулярных механизмов, осуществляющих тонкую регуляцию работы генов. Пока, увы, и этот вопрос еще далек от своего полного решения.

Теперь перед исследователями в области функциональной геномики встала значительно более серьезная задача: определить механизмы работы генома человека и функции многочисленных вновь выявленных генов и их продуктов. Одни гены будут кодировать белки, которые окажутся специфическими ферментами (катализаторами биохимических реакций), другие будут представлять собой банальный строительный материал, третьи будут идентифицированы как гормоны, факторы роста, рецепторы факторов роста, четвертые — как передатчики нервных импульсов и так далее.

И порой у одного и того же белка может оказаться несколько функций. Каждому белку, в конечном итоге, должно быть найдено его истинное место в сложнейшем устройстве и функцио-

нировании человеческого организма. То есть, условно говоря, должна быть создана система, подобная таблице Менделеева. А еще нужно выявить все возможные взаимные влияния динамично работающих генов и их продуктов друг на друга. Огромная сложность заключается в том, что функционирование любого гена осуществляется на фоне работы множества других генов, участвующих в выполнении определенной функции организма, то есть в составе так называемой генной сети. Такие генные сети включают в себя обычно от нескольких десятков до многих сотен координированно функционирующих генов. Например, генная сеть системы кроветворения включает в себя не менее 500–600 генов.

Даже неполный список генов, входящих в генную сеть, контролирующую цикл клеточного деления, состоит из более чем сотни генов и т.д.

Постепенно становится понятным, что гены, как и люди, могут выполнять в своей жизни порой совершенно разные функции, оставаясь при этом самими собой. То есть, они иногда могут участвовать немножко в одном, в иной раз немножко в другом, а при случае немножко в третьем. И благодаря этому общее количество генов совсем не обязательно должно совпадать с количеством функций, присущих человеку.

Иногда совокупность генов сравнивают с большим музыкальным оркестром. В зависимости от дирижера и исполнителей одна и та же партитура может звучать так, что вызывает у слушателей или восторг, или глубокое разочарование. Каким образом каждый музыкант — отдельный ген — играет в этом «оркестре», как осуществляется слаженное «концертирование» многих тысяч разных исполнителей? Какую роль в основном «дирижировании» играет сложнейшее взаимодействие между различными генами, какое влияние оказывают постоянные воздействия на гены многочисленных факторов внутренней и внешней среды, таких, как радиация, температура, пища, солнечный свет, бактерии, вирусы, лекарственные вещества, внутриклеточные метаболиты? Не менее важен и вопрос о «дирижерских палочках» — сигнальных путях и молекулярных механизмах, контролирующих программы работы генов, так как дисгармония этих процессов сильно отражается на слаженной игре «оркестра» в целом.

Таким образом, когда мы говорим о групповой ответственности генов за те или иные функции, следует помнить, что внутри определенной группы всегда есть такие, кто собственно и выполняет всю работу, другие ее контролируют, а кто-то обеспечивает чисто вспомогательные функции. Речь идет сейчас о том, чтобы определить иерархию внутри таких сложных групп взаимозависимых и взаимосвязанных между собой генов.

Все эти вопросы еще предстоит решить в будущем. И все это, безусловно, гораздо более сложная и трудоемкая задача, чем уже проделанная гигантская работа по секвенированию генома человека.

Однако и на этом сложности предстоящей работы не заканчиваются. В изучении регуляции экспрессии генов сейчас все большее и большее значение придается химическим изменениям, происходящим иногда в определенных участках генома. Наиболее известное из них — уже упоминавшееся метилирование — это присоединение в ДНК к букве Ц специальной химической группы, состоящей из одного атома углерода и трех атомов водорода (метильная группа). Такая модификация существенно сказывается на работе гена, без изменения его текста. Метилирование, а также другие способы, которыми природа изменяет активность генов, меняя лежащие рядом с ними регулировочные участки, стали предметом изучения новой «постгеномной» дисциплины — эпигеномики (от греческого «эпи» — рядом, около).

Явление немутационной эпигенетической изменчивости, открытое в конце 50-х годов при изучении генетики простейших и генетики соматических клеток, тогда не вписывалось в хромосомную теорию наследственности, «вызывая тень Ламарка» (последний утверждал, что признаки, приобретенные организмами в процессе своей жизни, могут передаваться по наследству, что отвергается современной генетикой). В наше время, когда понята молекулярная природа этого явления, эпигеномика стала еще одним из важных направлений биологического поиска. Об этом говорит хотя бы тот факт, что еще до завершения проекта «Геном человека» произошло объединение усилий большой группы ведущих европейских научных центров с целью создания Европейского

эпигенетического консорциума. Его основная задача — выявление сотен тысяч участков генома, подвергающихся метилированию, и анализ различных вариаций.

Каковы же магистральные пути развития функциональной геномики человека в настоящее время? Разработано огромное множество подходов для решения данной задачи. Обо всех рассказать здесь просто невозможно, поэтому остановимся вкратце лишь на некоторых основных из них.

Микрочипы — новый рубеж в исследовании экспрессии генома

В последние годы появились принципиально новые экспериментальные технологии, позволяющие одновременно следить за динамикой экспрессии сотен, а порой и тысяч генов в ходе функционирования и развития клеток. И здесь в первую очередь следует отметить технологию так называемых микрочипов.

В чем же причина проявления столь жгучего интереса к биочипам?

История биочипов началась еще в конце 80-х годов, когда исследователи сразу трех стран (Югославии, США и СССР) опубликовали первые результаты работ по оценкам возможности применения биочипов для анализа ДНК. В нашей стране ведущая роль в разработке данных технологий принадлежит Институту молекулярной биологии им. В. А. Энгельгардта РАН

Спустя 10 лет в США было объявлено о начале проекта по выпуску на рынок первых коммерческих продуктов, использующих биочипы. В настоящее время эта технология очень широко используется в самых разных целях большим числом фирм. С ее помощью осуществляется широкомасштабная автоматизация процесса обнаружения множества различных генных мутаций, одновременный анализ работы десятков тысяч генов в здоровых и поврежденных тканях, диагностика генетических заболеваний и многое другое.

Современный биологический микрочип имеет весьма небольшой размер (чаще всего с ноготь) и представляет собой набор нанесенных и закрепленных (иммобилизованных) на предметном стекле специальных проб в количестве от нескольких десятков до 10000 и более.

В качестве таких проб могут служить очень короткие нуклеотидные последовательности ДНК (олигонуклеотиды), протяженные фрагменты ДНК, ферменты, антигены, антитела и т.д. В зависимости от поставленной задачи сложность микрочипа (количество иммобилизованных проб и плотность их расположения) может варьировать в широких пределах. Пробы расположены на чипе в определенной последовательности.

Чип гибридизуют (о гибридизации мы уже говорили) с изучаемым образцом, содержащим сложную смесь меченых тем или иным способом РНК или ДНК, а далее анализируют полученный результат с помощью специальных приборов — сканеров. По получаемой картине можно судить о том, что именно содержится в анализируемом образце. Таким образом, обеспечивается проверка на соответствие известных и неизвестных образцов ДНК и автоматизируется процесс их идентификации.

Иными словами, биочип⁷ — это новое изделие, которое выступает в роли биологического эквивалента микрочипа компьютера, за исключением того, что, вместо выполнения миллионов математических действий, использующих электричество, здесь происходит расшифровка работы тысяч генов за счет процесса гибридизации.

Биологические микрочипы

⁷ ДНК-микрочип (англ. DNA microarray) — это технология, используемая в молекулярной биологии и медицине. Современный ДНК-микрочип состоит из тысяч дезоксиолигонуклеотидов (зондов, или проб), сгруппированных в виде микроскопических точек и закреплённых на твёрдой подложке. Каждая точка содержит несколько пикомолей ДНК с определённой нуклеотидной последовательностью. Олигонуклеотиды ДНК-микрочипа могут быть короткими участками генов или других функциональных элементов ДНК и используются для гибридизации с кДНК или мРНК (кРНК). Гибридизация зонда и мишени регистрируется и количественно характеризуется при помощи флюоресценции или хемилюминесценции, что позволяет определять относительное количество нуклеиновой кислоты с заданной последовательностью в образце.

В последнее время активно развиваются ДНК-технологии, которые позволяют не только определять признак, но и одновременно проводить дифференциальный сиквенс, т.е. определение точечных мутаций или полиморфизма в известных участках генома. Данные технологии имеют значительные преимущества перед традиционными молекулярно-биологическими методами, т.к. они позволяют миниатюризировать исследуемый образец и анализатор, что значительно снижает стоимость анализа и время его проведения, а также одновременно определять различные параметры исследуемого образца, причем без потери чувствительности амплификационных методов. Главным преимуществом методов основанных на использовании чипов с олигонуклеотидами всех возможных нуклеотидных последовательностей данной длины, является их универсальность. Наличие на чипе олигонуклеотида любой последовательности делает возможным анализ любой исследуемой последовательности. В основе применения микрочипов лежит принцип быстрого определения взаимодействий тех или иных лигандов со множеством различных зондов одновременно. Собственно биологические микрочипы представляют собой ту или иную твердую подложку, на которой нанесены или определенные фрагменты нуклеиновых кислот, или белки, или углеводы, или какие-либо иные молекулы-зонды, способные быть узнаваемыми или проявлять биологическую активность. Количество различных зондов на подложке может достигать сотен тысяч, причем чипы каждого типа строго идентичны и при существующих технологиях могут быть реплицированы в сотнях тысяч и миллионах копий нанесенных на подложку.

ДНК-микрочипы

Существуют белковые, ДНК, углеводные, тканевые чипы. Особого внимания заслуживают ДНК-чипы. Они представляют собой уникальный аналитический инструмент, позволяющий определять наличие в анализируемом образце (как правило, биологического происхождения) заданных последовательностей ДНК (т.н. гибридизационный анализ). Проведение анализа с помощью ДНК-чипов обходится в несколько раз дешевле, чем при использовании альтернативных технологий (электрофорез, ПЦР в реальном времени) и допускает, при наличии детектора несложной конструкции, работу вне лаборатории.

Впервые ДНК-чипы были использованы в исследованиях в конце 80-х годов прошлого века. В основе этого теперь уже широко распространенного метода, позволяющего одновременно анализировать экспрессию множества генов, лежит принцип узнавания мРНК-овых или кДНК-овых мишеней посредством их гибридизации с иммобилизованными на микрочипе одноцепочечными фрагментами ДНК.

ДНК-чип – это твердая подложка, на которой иммобилизованы (как правило, ковалентно) одностранные фрагменты ДНК разной длины: короткие – 15-25 нуклеотидов, длинные – 25-60 нуклеотидов и кДНК фрагменты – от 100 до 3000 нуклеотидов. В качестве материала подложки используют стекло, кремний, различные полимеры, гидрогели (например, на основе полиакриламида) и даже золото.

Гибридизация-основа технологии

Основа всех современных ДНК-технологий – гибридизация. В результате гибридизации молекулы нуклеиновых кислот способны формировать устойчивые двухцепочечные структуры благодаря связям между элементами молекул - нуклеотидами. Нуклеотид аденин (А) комплементарен тимину (Т), гуанин (Г) - цитозину (Ц). В результате одноцепочечная последовательность нуклеотидов АТГЦ будет образовывать стабильную ассоциацию, двухцепочечную структуру, с одноцепочечной молекулой ДНК состава ТАЦГ.

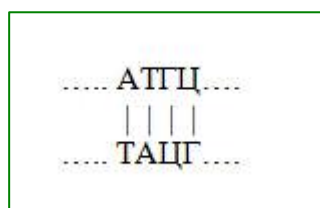


Рисунок 28 Схема рабочего процесса на биочипе

Подобная комплементарность приводит к «слипанию» двух молекул ДНК, одна из которых может быть неподвижно закреплена на подложке, и являться элементом ДНК-чипа. Чем больше содержится в образце молекул комплементарных элементам чипа, тем больше их будет связываться с чипом, и тем выше будет интенсивность сигнала,

воспринимаемого от данного элемента. На рис. 28 показан принцип действия ячейки ДНК или олигонуклеотидного биочипа, основанный на комплементарных взаимодействиях основания аденина (А) с тимином (Т) и гуанина (G) с цитозином (С) в двух нитях ДНК.

Если последовательность оснований в одной нити ДНК (или олигонуклеотида) полностью комплементарна последовательности другой нити, то образуется стабильная совершенная двухнитчатая спираль - дуплекс. Однако присутствие в дуплексе даже одной неправильной пары, например G-G, предотвращает образование дуплекса. Если иммобилизовать в одном из элементов микрочипа специфическую одноцепочечную ДНК или, положим, 20-мерный олигонуклеотид (пробу), то при добавлении к микрочипу меченных флуоресцентными красителями фрагментов ДНК, например генома человека, будет происходить их высокоспецифичное взаимодействие. Заданный олигонуклеотидный элемент биочипа специфически свяжет только одну комплементарную последовательность из $4^{20}=1.09 \times 10^{12}$ всех возможных последовательностей этой длины в ДНК. В результате флуоресцентное свечение наблюдается только на этом комплементарном элементе биочипа. Таким образом, один элемент биочипа производит одну выборку примерно из триллиона возможных вариантов, в отличие от элемента электронного чипа, где происходит двойная выборка: ДА или НЕТ.

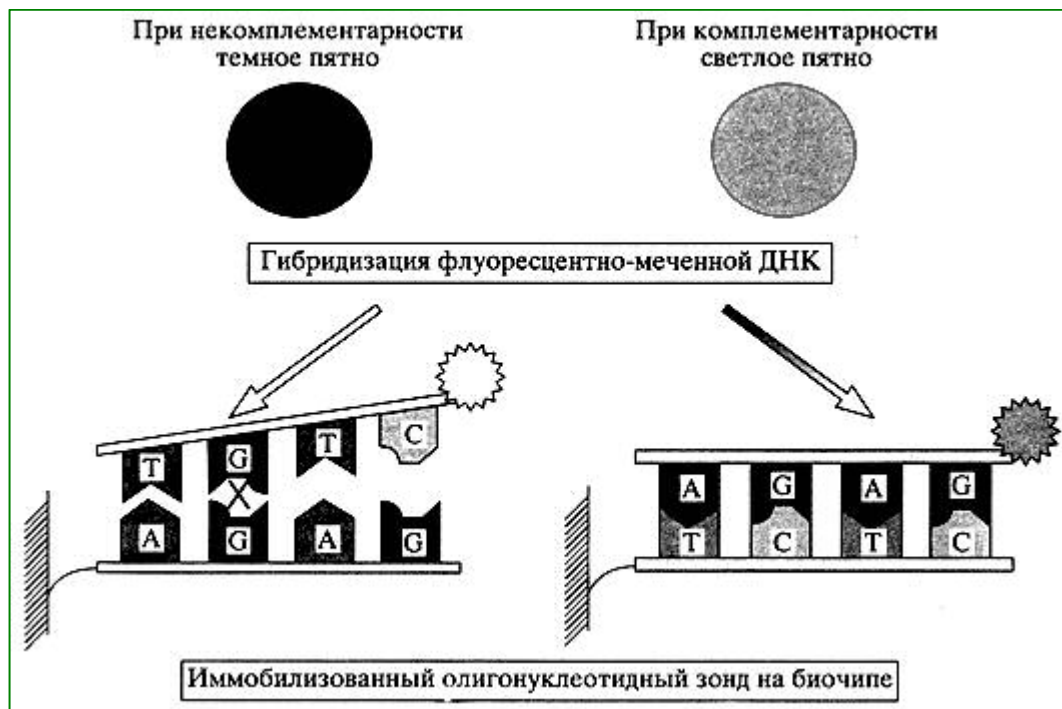


Рис. 29. Схема образования двойной спирали ДНК на биочипе.

Олигонуклеотид фиксирован на одном из элементов биочипа и избирательно связывает из многих флуоресцентно меченых фрагментов ДНК только комплементарный. В результате только этот элемент начинает светиться. Это происходит благодаря высоко-специфичным взаимодействиям комплементарных пар нуклеотидов А с Т и G с С. Присутствие некомплементарной пары, например G-G, предотвращает взаимодействие и оставляет элемент микрочипа темным.

Используемые для определения параметров гибридизации устройства позволяют регистрировать не только конечный результат, но и кинетику ассоциации и диссоциации комплементарных цепей. Эти технологии, делая возможным многопараметрический анализ образцов, могут предоставлять большое количество информации. Результаты гибридизации зависят от длины ДНК - пробы, химического состава меченой ДНК - мишени, температуры, при которой проводится гибридизация, состава гибридизационной смеси, типа флуоресцентной метки. Здесь необходимо отметить,

что в ДНК-чипах в основном используется пассивная гибридизация, т.е. взаимодействие ДНК-мишени с иммобилизованной пробой является вероятностным процессом и зависит от различных условий.

Применение ДНК-чипов

Оценка состояния и идентификация всех генов исследуемого организма является одной из важнейших задач, поставленных перед разработчиками ДНК-чипов. Решение этой задачи может быть реализовано в иммобилизации всех генов организма на биологический чип, что позволит комплексно оценить состояние генов и генома в целом. Биогенетические базы данных, содержащие всю информацию (систематизированную) по генам и геномам различных организмов, представляют исследователям огромные возможности в дизайне ДНК-чипов.

К основным причинам широкого распространения биочиповых исследований относят высокую чувствительность, специфичность и воспроизводимость, простоту процедуры выполнения, возможность одновременного анализа множества параметров и относительно невысокую стоимость работ. Эти же причины заставляют рассматривать биочипы как перспективный инструмент в различных областях народного хозяйства.

Подводя некоторые итоги нужно отметить, что микрочипы являются эффективным подходом для одновременной идентификации от десятков до тысяч генов и их структурного анализа, для выявления специфичных нуклеотидных последовательностей и нуклеотидных вариаций в их структуре. Однако, когда гены присутствуют в геноме в количестве одной или нескольких копий, с чем постоянно и приходится сталкиваться в клинической практике, требуется их предварительная амплификация. Наиболее эффективным методом амплификации ДНК является полимеразная цепная реакция, в процессе которой происходит экспоненциальное увеличение количества молекул ДНК от нескольких до миллионов и более копий, а основное достоинство такого вида ПЦР⁸ как Real Time, позволяют давать даже количественную оценку исследуемой матрицы. Это имеет важное значение для решения задач в области развития фундаментальных и интегральных наук, а так же оптимизации условий методов диагностики.

Итак, два метода, ставшие уже традиционными для некоторых областей науки и прикладных технологий, наряду со своими недостатками имеют совершенно уникальные достоинства.

RealTime ПЦР:

- дает возможность оценивать количество исходной матрицы;
- не требует дополнительных трудоемких этапов работы;
- отсутствие стадии электрофореза позволяет минимизировать риск контаминации и таким образом уменьшить число ложноположительных результатов;
- использование математических методов анализа позволяет проводить автоматическую интерпретацию полученных результатов и снимает проблему субъективной оценки электрофореграмм;
- обеспечивает менее строгие требования к организации ПЦР-лаборатории и автоматической регистрации и интерпретации результатов;
- позволяет экономить время.

Биологические микрочипы:

- позволяют миниатюризировать образец и анализатор;
- экономит время и стоимость анализа;
- позволяет одновременно определять несколько параметров исследуемого образца;
- обладает высокой чувствительностью амплификационных методов, специфичностью и воспроизводимостью;
- обеспечивает простоту процедуры выполнения работы.

⁸ Полимеразная цепная реакция (ПЦР) — экспериментальный метод молекулярной биологии, позволяющий добиться значительного увеличения малых концентраций определённых фрагментов нуклеиновой кислоты (ДНК) в биологическом материале (пробе).

Возможно, что объединение этих методов за счет перевода полимеразной цепной реакции в формат микрочипов позволит создать диагностическую систему нового поколения, которую будут характеризовать следующие качества: более высокая чувствительность и, главным образом, специфичность определения нуклеиновых кислот, высокая производительность при низкой себестоимости выполнения анализа, общее сокращение количества манипуляций в пределах каждого этапа анализа.

В настоящее время исследования генной экспрессии часто проводятся с использованием высокоэффективных микрочипов ряда фирм. У фирмы Affymetrix они представляют собой маленькие пластинки, в микролунки которых нанесены короткие (20–25 нуклеотидов) пробы, комплементарные последовательностям исследуемых генов. При этом каждому гену соответствует несколько (10–20) таких проб. Это существенно повышает точность и воспроизводимость количественного анализа генной экспрессии. Плотность размещения проб, а следовательно количество генов, анализируемых на одном чипе, могут быть столь велики, что один такой чип способен охватить все гены человека. С таким микрочипом проводят гибридизацию меченых копий суммарной РНК клетки (кДНК). Благодаря комплементарности цепей, в каждой лунке гибридизуется отдельная фракция кДНК. Затем автоматически микрометодами определяется спектр генов, изменяющих свою экспрессию в одном типе клеток по сравнению с другим. В результате можно расписать функциональное состояния огромного числа генов на каждый конкретный момент времени в каждом определенном типе клеток.

Моделирование — важный подход к пониманию функции генов

Естественно, что ученые не всегда могут проводить испытания на живом человеческом организме. Но из этого положения есть определенный выход. Сравнения геномов простых и сложных организмов указывает на существенный консерватизм функций генов. Этот консерватизм не только демонстрирует тот факт, что жизнь на земле основана на общих принципах, но и предоставляет экспериментаторам инструмент, позволяющий, исследуя функции генов у простых организмов, делать заключения о функции генов у более сложных, включая человека. По этой причине уже давно в экспериментах стали широко использовать различные животные организмы, говоря в этих случаях о моделировании процессов, происходящих у человека. Для моделирования используются различные животные, начиная от нематоды и кончая обезьянами.

Модель — это конечно же образец, некое подобие оригинала (свирель не может заменить соловья), но она зачастую много говорит ученым о самом оригинале. При этом модель помогает решить главную проблему — понять человека, не затрагивая в экспериментах его самого. Напрашивается такая, хотя и несколько отдаленная, литературная аналогия. Это все равно, как сравнивать достоверную летопись об определенном событии и легенду или сказку, где это также отражено.

Одним из наиболее широко используемых для моделирования организмов служит мышь. С ней мы часто боремся в быту, ругаем ее, травим, а вот в науке она остается одним из основных объектов исследования. Дело не только в том, что мышь — весьма доступный и дешевый объект, а главным образом в том (как это, может быть, не совсем приятно читателю), что устройство генома и сами гены у мыши довольно сходны с человеческими.

В последние годы ученые научились даже получать мышей с генами человека. Для этой цели в качестве мишени используют обычно ее оплодотворенную яйцеклетку. Чаще всего ген вводят с помощью микропипетки в ядро этой клетки. При удачном стечении обстоятельств (обычно такое стечение происходит в 5–10% случаях) ген встраивается в геном мыши и после этого становится таким же, как и собственные мышьиные гены. В результате, когда из прооперированной яйцеклетки вырастает потомство, оно содержит новый, ранее не имевшийся у них ген — трансген. Такие животные получили название трансгенных. Например, когда мышам ввели ген гормона роста человека, они увеличили размер своего тела почти в два раза (рис. 30).

Когда же им вводили человеческие «больные» онкогены, то у таких трансгенных мышей часто развивались опухоли того же типа, что и у человека. Подобного рода эксперименты были проведены уже с десятками сотен различных генов. Изучая биологические эффекты разнообразных (порой совершенно не изученных) генов у мышей, ученые довольно уверенно делают выводы об их вероятной функции у человека.

Более того, в последние годы были созданы специальные весьма изощренные и методически сложные генетические приемы, которые позволяют направленно изменять работу определенных, уже имеющихся мышинных генов в тех или других органах и на тех или иных этапах развития. В частности, найдены молекулярные подходы, которые позволяют полностью выключать работу строго определенных генов (это называют нокаутированием генов). Мыши с такими, находящимися «в нокауте» генами, дают возможность как выяснять роль для жизнедеятельности уже известных генов, так и идентифицировать новые гены, важные для различных аспектов жизни человека.



Рис. 30. Фотография нормальной мыши

Протеомика

От гена к белку. Сам по себе ген — лишь определенная последовательность нуклеотидов. Его основная задача — обеспечить производство на свет строго определенного белка (в крайнем случае РНК). Существует даже выражение: «гены — поваренная книга, испещренная тысячами рецептов; белки — угощение, выставленное на стол». В постгеномную эру появились новые возможности не только для функциональной геномики, но и для исследования самих белков — основных кирпичиков живого.

Белки (или, как еще их называют, протеины) известны нам уже около 200 лет. В начале XIX столетия химики выбрали имя «протеины» для этих веществ от греческого слова «proteios», означающего «первый» или «важнейший». В русском языке эти вещества чаще называются белками. И вот теперь, на базе геномики возникло новое направление исследований — протеомика.

Протеомика — это изучение всей совокупности белков клеток и их взаимодействия в целом организме. Ученые в области протеомики изучают «производство» белков, их структуру и состав, различные модификации после синтеза, функции и метаболизм. Все эти исследования имеют одну общую цель: идентифицировать все белки, работающие в каждом типе клеток в каждый определенный момент их жизненного цикла, и понять в совокупности их сложный метаболизм.

Масштаб предстоящей работы представляется огромным. Ведь во взрослом организме человека в 10¹⁴ его различных клеток функционирует несколько десятков тысяч генов. Но гены кодируют лишь белки, а не их сборку в работающую «машину». Белковый «текст» — это трехмерный мир, намного превышающий возможности ДНКового текста. Как пишет американский биолог

Роберт Поллак, «гены — это линейный текст, а белки — трехмерная скульптура». В белковом тексте, условно говоря, «буква» — это аминокислота, корень «слова» — последовательность аминокислот в единичном полипептиде. Свертываясь в трехмерную структуру, единичный полипептид формирует по сути дела полное «слово», которое уже может быть и простым «предложением». Вдобавок, в живой клетке форма белков может динамически меняться, что превращает их в подобие более сложной, «кинетической» скульптуры. Далее возникают сложнейшие «предложения» — комплексы, состоящие из десятков и даже сотен белков.

Таким образом, разнообразие белковых вариантов у нас значительно больше, нежели генов. По некоторым оценкам, белковых вариантов и их всевозможных комплексов в наших клетках может набраться аж миллион! Связано это с тем, что, как говорилось выше, с одного гена за счет альтернативного сплайсинга возможно получение нескольких разных РНК-копий, которые способны кодировать разные белковые цепи (полипептиды). Последние, в свою очередь, могут по-разному взаимодействовать друг с другом, образуя комплексы, состоящие из 2–4 и более полипептидов. При этом формируются многочисленные белковые комплексы, обладающие порой разными функциями.

Например, ген по имени *Ikarus* за счет альтернативного сплайсинга способен обеспечивать образование 6-ти различных полипептидов. А далее из них в клетке может сформироваться около 20 разнообразных белковых комплексов, состоящих или из одинаковых, или разных полипептидов. Таким образом, наш организм содержит гигантское число вариантов белковых молекул. Развивая эту мысль можно предложить разработку базы ЗНАНИЙ (БЗ), где будут отображены все известные варианты реализации белковых комплексов. Фактически, по своему замыслу такая БЗ проста. Одна запись - общего плана, создает ссылки к менее информативным записям. Получаем вложенную систему знаний.

Отсутствие такой БЗ существенно осложняет на сегодняшний день возможность создания «каталога белков». Кроме белков, в нашем организме содержится большое количество других молекул: липидов, углеводов, разнообразных низкомолекулярных органических соединений. Но все-таки основу основ составляют именно белки. Они выполняют в клетках весьма разнообразные функции: одни служат строительными материалами клетки (структурные белки), другие выполняют функции ферментов, катализируя многочисленные клеточные химические реакции, третьи являются регуляторами процессов репликации, транскрипции и трансляции.

Существуют еще транспортные белки, участвующие в транспортировке метаболитов в клетке, белки-гормоны, хранилищные белки, рецепторные, сократительные и др. Белки в клетке постоянно синтезируются и постоянно разрушаются (деградируют). Кроме того, в клетке они распределены между разными, относительно изолированными областями, окруженными мембранами, которые называются органеллами.

Важно также отметить, что в настоящий момент анализировать белки существенно труднее, чем гены. Если процесс севенирования ДНК уже автоматизирован так, что «с ней справилась бы любая обезьяна», как едко заметил нобелевский лауреат Джеймс Уотсон, то методы анализа белков оказались гораздо более сложными. В далеком 1962 г. вместе с Уотсоном и Криком в Стокгольме были приглашены из Кембриджа Джон Кэндриу и Макс Перутц. Они были тогда удостоены Нобелевской премии по химии за впервые осуществленную расшифровку трехмерной структуры двух белков — миоглобина и гемоглобина, — ответственных за перенос кислорода в мышцах и эритроцитах соответственно.

Даже в начале 1990-х годов расшифровка структуры каждого нового белка представляла значительные трудности. Каждый анализ занимал порой до десятка лет. Сейчас методы анализа белков значительно усовершенствованы. Однако все еще перспективы создания «каталога протеинов» намного более отдаленны, чем перспективы создания упомянутых выше каталогов генов, снипсов или вариаций метилирования ДНК. Это другая актуальная проблема геномики.

Хотя и от самого по себе синтеза белка еще далеко до создания целого организма, но без белков этот процесс никогда не пойдет. Белки не только составляют большую часть клетки как структурные элементы, они осуществляют также постоянный тонкий контроль за всеми химиче-

скими процессами, происходящими внутри клетки, избирательно включая и выключая их в строго определенные сроки и в строго определенных местах. Это теперь понятно. Но не понятно до конца другое: как именно происходит в конечном итоге развитие целого организма из одной единственной клетки — зиготы. Белки, конечно же, активно участвуют в этом процессе как основные компоненты.

Работа белков, как и работа генов, тоже зависит от многих факторов, прежде всего влияющих на их пространственную структуру. Более того, многие белки со временем могут меняться (модифицироваться), и здесь уже гены не оказывают никакого влияния. Происходит это путем присоединения к белковым молекулам особых побочных групп — фосфатидов, сахаридов или ненасыщенных углеродных цепочек. Все эти события — как и образование пространственной структуры белка — никак не отмечены в генах. Гены содержат лишь общий план активной жизни клеток, а сама жизнь определяется в конечном итоге белками.

Гены во всех клетках всегда одни и те же, но вот белки, наоборот, постоянно меняются в зависимости от стадии развития клетки, возраста организма, состояния клетки и множества других обстоятельств.

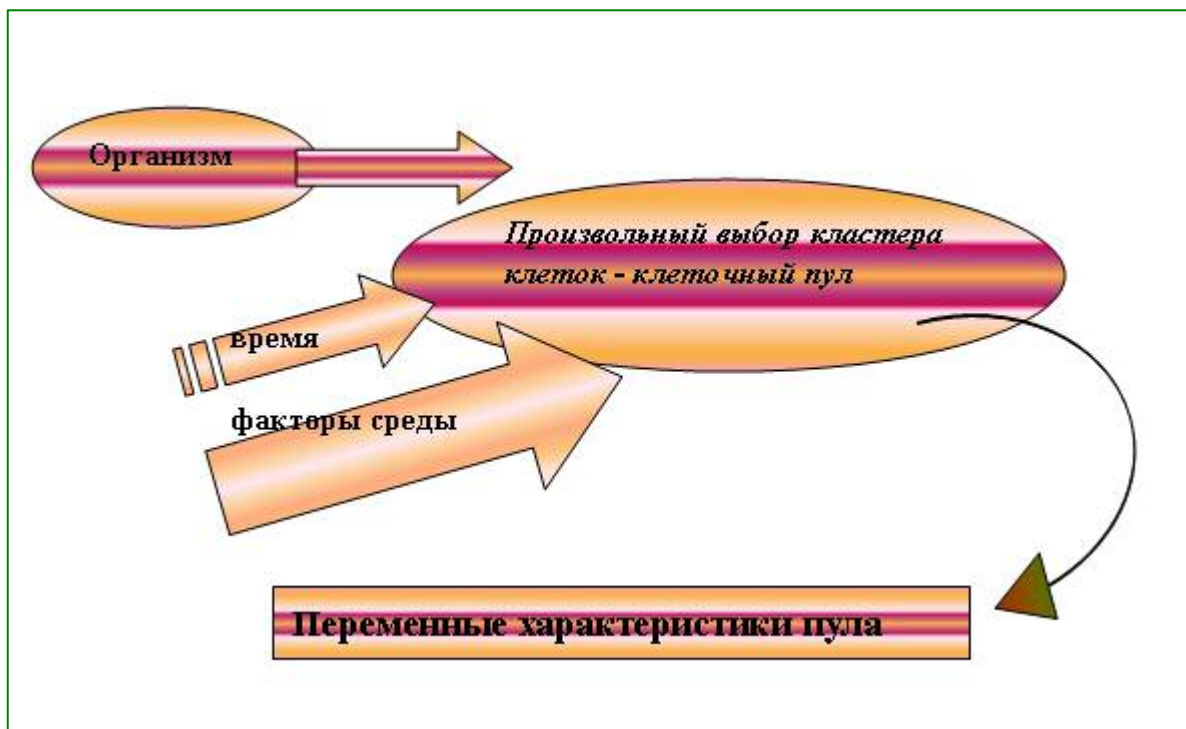


Рисунок 31 Схема постоянной смены событий в клеточном пуле

Белки позволяют себе импровизацию типа джазовой, т.е. коллективной, когда вместо одних белков их функцию выполняют другие. Это очень важно, поскольку все ситуации, с которыми сталкивается организм в процессе своей жизни, заранее предусмотреть нельзя. Известно и другое — один и тот же белок может иметь отличающийся эффект в разных типах клеток и на разных стадиях развития. Все это имеет аналогии с обычным литературным текстом. В лингвистике для описания подобных случаев имеются специальные термины.

Так, термин **синонимия** обозначает предложения, несущие один смысл, но состоящие из разных слов (например, «Мамонтов погубило внезапное похолодание» и «Резкое понижение температуры на планете привело к гибели мамонтов»). Под термином омонимия подразумевают смысловое различие внешне совпадающих предложений (например, предложение «Мы проехали остановку» может обозначать две разные ситуации: мы проехали расстояние, равное перегону между двумя остановками, или мы миновали то место, где должны были выйти).

Одна из важнейших целей, стоящих перед учеными, занятыми анализом протеома, заключается в поиске характерных изменений в белках, присущих различным видам заболеваний. Действительно, нередко удается проследить взаимосвязь между изменениями белков и разными болезненными состояниями.

Секвенирование генома человека дало в руки исследователей важную информацию для борьбы с наследственными недугами. Однако не все болезни передаются нам по наследству. Многие никак не связаны с «родовым проклятием». Выявить эти болезни можно, лишь узнав, как изменился состав белков внутри патологической клетки. По этой перемене можно заранее заметить и патологические процессы, начавшиеся в организме.

Чтобы увидеть белковый состав клетки, ученые прибегают к так называемому двумерному электрофорезу. Сначала белки сортируют по их заряду, а затем они попадают в гель, выполняющий роль сита, в котором белки разделяют по их размерам. Таким образом, удается разделить до десяти тысяч белков и маркировать их. На этой основе можно составить что-то вроде паспорта данной клетки, где примерно указан состав белков. Если человек заболевает, «паспорт» изменится. Сравнивая наборы белков больной и здоровой клеток, удастся оценить течение болезни и процессы, ей сопутствующие. Все это позволяет значительно ускорить поиск белков-мишеней и осуществлять разработку лекарственных средств, чтобы как можно быстрее обеспечить больных новыми эффективными лекарствами. Сегодня более 95% всех фармакологических средств на рынке нацелены на воздействие на белки-мишени. Протеомика порой может помочь идентифицировать и оценить новые целевые белки гораздо эффективнее, чем геномика. Это позволяет резко ускорить разработку новых диагностик и терапевтических средств.

Можно привести такой пример. Известно, что к повышению артериального давления прямое отношение имеет так называемый ангиотензин-конвертирующий фермент (АСЕ). Ангиотензин, образующийся под действием АСЕ, воздействует на стенки артерии, что и ведет к гипертонии. Уже относительно давно были найдены блокаторы АСЕ, которые стали продаваться в качестве лекарств от повышенного давления. Однако все эти лекарственные средства оказались малоэффективными.

В результате анализа генома человека был выявлен наиболее распространенный и эффективный вариант фермента. Затем была определена структура белкового продукта, после чего подобраны химические вещества, активно связывающиеся с ним. В результате был найден новый эффективный препарат против артериального давления, причем за вдвое меньшее время

Окончательных результатов этого нового грандиозного исследования, которое можно назвать «Протеом человека», конечно же, в ближайшие месяцы ожидать не приходится. Понадобится, видимо, много десятков лет, если не все новое столетие. Лишь после того, как человечество получит в свое распоряжение всю эту информацию, можно будет говорить об окончательном завершении программы «Геном человека».

Но и это еще не конец предстоящих исследований. Не завершение процесса создания Энциклопедии человека. Не вызывает сомнения, что гены и кодируемые ими белки — продукт эволюции. Вместе с тем очевидно, что когда гены еще вообще не существовали, в неорганическом мире, у кристаллов и минералов, имелись определенные формы, подобные тем, что есть сейчас у вирусов, белков и даже клеток. До сих пор мы еще мало знаем о тех механизмах, которые побуждают группы молекул объединяться вместе, формируя те или иные клетки и ткани. Тем не менее, следует отметить, что в последние годы возникло некое новое представление о процессе самосборки молекул, которое может оказаться важным для нашего понимания процессов формирования и вообще функционирования человека.

Возможно, во всем многообразии природных систем существует некий единый фундаментальный способ построения разнообразных структур. Его иногда называют «тенсегрити» (при переводе слова «tensegrity» с английского, получится что-то вроде «напряженности стойкости», что довольно коряво и малопонятно звучит по-русски). Термин «тенсегрити» означает, что система стабильна за счет баланса в ее структуре сил сжатия-растяжения. Это общий принцип, который, вероятно, работает и в организме человека.

Считается, что все 206 костей человека противостоят силе тяжести, что дополнительно к мышцам, сухожилиям и связкам позволяет ему удерживаться в вертикальном положении. Таким образом, по тенсегрити, основные структуры внутри нас — кости, которые служат прочными распорками, а также мускулы, сухожилия и связки, являющиеся упругими элементами. Согласно этому представлению, законы тенсегрити действуют как на молекулярном уровне (им подчиняются наши белки и другие важные для клетки молекулы), так и на клеточном и даже организменном уровнях (при формировании тканей и органов, при морфогенезе). И все эти «сопроматные» механизмы, действующие *in vivo* наряду с генетическими, нам еще предстоит изучить.

Медицинская геномика

Медицинская геномика - первая ступень понимания создания терапевтических процедур. Лечение того что возможно состоится или уже произошло следует проводить с учетом имеющихся ресурсов организма. Это в первую очередь относится к знаниям о генной предрасположенности организма к нарушениям.

Многие годы важную роль в решении всех этих «исключений из правила» играло специальное направление в медицине — молекулярная генетика. Основоположниками медицинской генетики в России были С. Н. Давиденков, занимавшийся изучением роли наследственности в болезнях нервной системы, и С. Г. Левит, организовавший в Москве и возглавивший первый в России (и один из первых в мире) Медико-генетический институт. Это направление сыграло важную роль на начальных этапах, когда мало что было известно о структуре генома человека.

Первоначально в медицине было принято считать, что любое заболевание представляет собой результат дисбаланса биохимических реакций, возникающих в результате тех или иных молекулярных поломок. В связи с этим на первом этапе нужно всего лишь определить «молекулярную мишень», а затем с помощью лекарства исправить существующее нарушение. Такой подход в целом оправдал себя при разработке лекарственной терапии инфекционных заболеваний, где в качестве мишени выступали различные бактерии и вирусы. Но при других типах заболеваний, таких, как, например, атеросклероз, рак, преждевременное старение, заболевания нервной и сердечно-сосудистой систем, старая идеология «забуксовала», потому что у этих заболеваний оказалось слишком много разных причин, включая и генетические. Стало понятным, что надежда на одну единственную «волшебную пулю», направленную на одиночную мишень, здесь не может помочь — во многих случаях «мишеней» в клетках очень много, к тому же они еще и подвижны.

Только революция, которая произошла в медицине в последние годы — как результат развития молекулярной генетики и полного секвенирования генома человека, позволила перейти от в известной мере иррационального знания к рациональному. Сегодня на основе геномики возникли такие направления, как медицинская геномика и биомедицина, которые тесно переплетаются с функциональной геномикой человека.

Опытные *врачи говорят, что нет двух одинаковых болезней и двух одинаковых пациентов.* И не менее чем на 50% эти различия связаны с особенностями структуры генома. Если в будущем удастся проводить сравнение генов больных людей с генами идеального, «чистого» гомо сапиенс, это существенно поможет поиску пути к эффективному лечению.

Геном — это в первую очередь реальный текст, опечатка в котором, допущенная Природой, может сделать человека или уродом, или инвалидом, или даже не позволит ему родиться вообще. В результате реализации проекта «Геном человека» появилась твердая уверенность, что уже в недалеком будущем большинство болезней можно будет лечить на генетическом уровне. А потому медицинскую геномику считают чуть ли не самым перспективным направлением в развитии современной медицины. В конце прошлого столетия с этим согласились даже политики, когда правительства Америки, Англии, России и ряда других стран стали инициаторами грандиозного научного проекта под названием «Геном человека».

Полное секвенирование генома человека стало важнейшим этапом для всей мировой медицины. На основе успехов в молекулярной генетике и определения полной последовательности ге-

нома человека произошла «генетизация» медицины, что привело к возникновению медицинской геномики, или молекулярной медицины. Если раньше считали, что генетика имеет отношение только к наследственным заболеваниям (а они составляют всего около 2% от общего числа заболеваний человека), то сейчас стало совершенно очевидным, что все болезни в той или иной мере от генов (а не от нервов, как это широко было принято считать, хотя и они, конечно, тоже играют свою роль). Один из руководителей геномного проекта Френсис Коллинз высказался на этот счет весьма категорично: «Нет болезней, за исключением некоторых случаев травм, которые не имели бы наследственного вклада. Я не знаю ни одного примера».

При наследственных болезнях наблюдаются опечатки в ДНКовом тексте, которые приводят к дефекту гена. В случае других заболеваний нарушения обычно затрагивают не структуру гена, а скорее регуляцию его экспрессии.

Как теперь ясно — очень многое. В первую очередь то, что теперь стало возможным диагностировать большинство (а в перспективе практически все) известных наследственных заболеваний на любой стадии развития человека, в том числе и до рождения (пренатальная диагностика). ДНКовый текст позволяет обнаружить предрасположенность индивидуума к тем или иным заболеваниям, включая и заболевания, обусловленные внешними причинами (химическими факторами среды, вирусами, питанием и др.). ДНК стала базой для генотипирования, идентификации личности. Далее, детальные данные об индивидуальных особенностях генома пациента служат основой для фармакогеномики — новой науки, изучающей, как генетика влияет на вариабельность ответа пациентов на лекарства. Рассмотрим теперь все эти вопросы по порядку.

Любая болезнь, для которой не существует эффективной терапии, представляет собой тайну, которую не так легко раскрыть. Ученым-медикам, прежде чем изготовить эффективное лекарство или найти иной способ лечения пациента, нужно узнать первопричину заболевания, а не его внешнее проявление. Ключ к этому — ген или гены, так или иначе вовлеченные в патологию. Не так давно в науке появилась новая профессия — «охотник за генами». По аналогии с популярными в конце прошлого века «охотниками за микробами» — микробиологами — век спустя «охотниками за генами» называют молекулярных биологов, которые выслеживают, находят и секвенируют (расшифровывают) человеческие гены, и в первую очередь гены серьезных наследственных болезней.

Один из наиболее многообещающих способов сделать это — проведение тщательного сравнения ДНКовых текстов у людей, связанных родственными узами, одна часть которых подвержена какому-то характерному заболеванию, а другая — нет. Конечная цель такого поиска — определить у носителей болезни ответственный за нее ген с нарушенным генетическим кодом. Уже говорилось, что геномы всех нас с вами очень похожи друг на друга. Если и имеются различия, то они весьма незначительные. Однако в них-то и заключается проблема. По этой причине медицинская геномика существенное внимание уделяет генетическому полиморфизму. Ранее уже шла речь о полиморфизме генома человека и об использовании этого явления для решения различного рода теоретических проблем. Но для медиков полиморфизм нашего генома имеет не столько теоретическое, сколько сугубо практическое значение.

Достоверно установлено, что многочисленные наследственные заболевания связаны с мутациями в генах, то есть с изменениями в самой структуре белков. Иногда достаточно мутации лишь в одном гене (такие патологии называют моногенными наследственными болезнями), иногда заболевание связано с наличием мутаций одновременно в нескольких генах (мультигенные наследственные болезнями). Медикам, конечно же, проще иметь дело с первыми, хотя и здесь могут быть проблемы с лечением.

Наряду с эти, выяснилось, например, что белки содержат специальные сигнальные последовательности, которые можно сравнить с адресами на почтовых отправлениях или с ярлычками на багаже, где написано место, куда предполагается его доставить. Фактически такие сигнальные последовательности представляют из себя цепочки аминокислот, которые или прикреплены к белковой молекуле в виде хвоста, или, иногда, лежат непосредственно внутри нее самой. Очень важно, что множество наследственных заболеваний человека вызывается нарушениями именно в этих

сигналах и в транспортных механизмах, с ними связанных. Таким образом, человеческие патологии связаны не только с нарушением структуры генов и кодируемых ими белков, но и с транспортом (перемещением) функционально активных белков в клетке.

Полиморфизм генома чаще всего для человеческого организма нейтрален. Тем не менее, изменения в целом ряде генов (и в их регуляторных участках) при определенных условиях могут способствовать, а иногда, наоборот, препятствовать проявлению отдельных заболеваний. Если обнаруживают гены, определенный полиморфизм в которых ассоциируется с тем или иным заболеванием, то их называют «генами предрасположенности» и они становятся объектом пристального внимания медиков. Первоначально определяют, к какому классу относятся эти гены. Принято считать, что среди генов предрасположенности к заболеваниям есть «гены детоксикации», которые реагируют на факторы внешней среды, и «гены-триггеры», связанные с запуском патологического процесса под действием внутренних факторов организма.

В поиске «больных» генов весьма полезными оказались уже упоминавшиеся снипсы. На них сейчас возлагают большие надежды. Изучение «снипсов» показало, что иногда замена одного единственного нуклеотида может вызвать у данного человека повышенную склонность к развитию того или иного заболевания. Это происходит в тех случаях, когда, например, точковая мутация возникает вблизи гена, на участке ДНК, регулирующем его работу. Изучение «точечных» различий между индивидуумами может открыть путь к «индивидуализированной» медицине. Большая роль здесь отводится новому направлению — фармакогеномике, о котором подробнее мы поговорим в следующем разделе.

Секвенирование генома человека и анализ снипсов существенно изменили стратегию обнаружения и клонирования генов, ответственных за наследственные заболевания. Итак, имеется некая наследственная патология. Понятно, что за нее отвечает какой-то ген или группа генов. Что требуется, чтобы обнаружить такой ген? Сначала надо грубо определить локализацию этого неизвестного гена (его называют геном-кандидатом) на одной из 23 хромосом. Начинается масштабный поиск каких-либо изменений в ДНК разных больных с этой патологией. Обычно сперва в результате такого анализа на определенной хромосоме выявляется довольно большой район, в котором могут располагаться десятки или даже сотни разных генов. Для поиска конкретного гена-кандидата заболевания необходимо сузить исследуемый участок ДНК. На это уходило раньше несколько лет. Однако после расшифровки генома ситуация упростилась. Теперь почти вся нуклеотидная последовательность генома человека известна и помещена в базы данных, с которыми можно связаться в любой момент по Интернету.

Если известна хромосомная локализация какого-либо заболевания, достаточно запросить информацию о нуклеотидной последовательности этого района, провести компьютерный поиск генов, содержащихся в нем, а затем выбрать один или несколько наиболее подходящих и протестировать их на мутации у ряда пациентов с данной патологией. Ген, в котором таковые будут обнаружены у большинства больных и не обнаружатся у здоровых людей, и есть искомый ген заболевания.

Подробные генетические и физические карты генома человека и особенно последовавшее за этим полное секвенирование ДНК позволили в последние годы успешно клонировать многие гены, ответственные за различные наследственные заболевания. Как уже говорилось, важная роль в этих достижениях принадлежит и сравнительному анализу геномов разных индивидуумов, принадлежащих к одному семейству, а также геномов нормальных и «больных» клеток.

Важными материалами для медицинской геномики служат относительно изолированные популяции людей и наличие подробных данных об их родственных связях. Так, в отличие от многих других стран, Исландия уникальна не только своим климатом, но и своими жителями. В течение многих веков здесь строго документировали все формировавшиеся брачные отношения, и, что очень важно, эти сведения сохранили в целостности и сохранности до наших дней. Это дало в руки исследователей огромный материал, охватывающий около 270 тыс. ныне живущих людей и 330 тыс. их ранее живших предков.

В результате появилась возможность создавать родословную для каждого из исландцев на глубину до 10 веков. Сейчас там проводят массовый анализ геномов жителей и исследуют взаимо-

связь наиболее распространенных заболеваний с различными генетическими маркерами. И это уже начинает приносить первые плоды. Во-первых, удалось обнаружить связь между развитием панических расстройств психики и аномалиями строения хромосомы 9. Во-вторых, исследователи выявили ген, мутация в котором резко увеличивает вероятность развития ишемического инсульта (эта мутация получила название STRK 1 от англ. «stroke» — инсульт).

В принципе существует два основных подхода к поиску гена-кандидата, отвечающего за патологию: от белка к гену и от генома к гену. Последний путь оказался наиболее эффективным. Часто для поиска «больного» гена используют полиморфизм микросателлитных маркеров, расположение которых на хромосомах хорошо известно. Сначала выявляют «сцепление» между заболеванием и определенной формой того или иного микросателлита, затем по расположению микросателлита определяют область, которая может иметь отношение к развитию заболевания. В этой области и ищут ген-кандидат болезни. Новая стратегия получила название «позиционного клонирования гена-кандидата».

Так, в частности, поступил российский генетик Е. И. Рогаев совместно со своими зарубежными коллегами для обнаружения одного из генов, ответственного за болезнь Альцгеймера, который определяет прогрессирующее старческое слабоумие. Это заболевание начинается с незначительных на первый взгляд нарушений памяти (на имена и близкие по времени события), но уже в течение двух—трех лет приводит к полной деградации личности и мучительной смерти. Болезнь Альцгеймера как причина смертности стоит на четвертом месте в развитых странах. Еще несколько лет назад врачи не могли предвидеть развитие болезни Альцгеймера у своих пациентов, а порой даже неправильно ее диагностировали. И вот к этой проблеме подключились «охотники за генами». Они проанализировали ДНК, выделенные из крови больных и их здоровых родственников, и определили, что у всех больных есть некий особый общий участок хромосомы 14.

Схема дальнейшего поиска гена-кандидата изображена на рис. 32. На следующем этапе определили, с какими маркерами (микросателлитами) сцеплен искомый ген (на рисунке указано 8 таких микросателлитов). На следующем этапе в библиотеке генов человека были найдены большие (около 1 млн. п. н.) фрагменты хромосомы 14, соответствующие исследуемому участку ДНК. С их помощью был более точно установлен участок хромосомы, несущий «больной» ген (на рисунке он обозначен двунаправленной стрелкой). Половина дела была сделана. Далее определили, какие гены здесь имеются. Их оказалось несколько десятков (отмечены светло-серыми прямоугольниками). И вот конечный результат: лишь в одном из множества генов были найдены мутации, которые встречаются только при болезни Альцгеймера (отмечен темным прямоугольником).

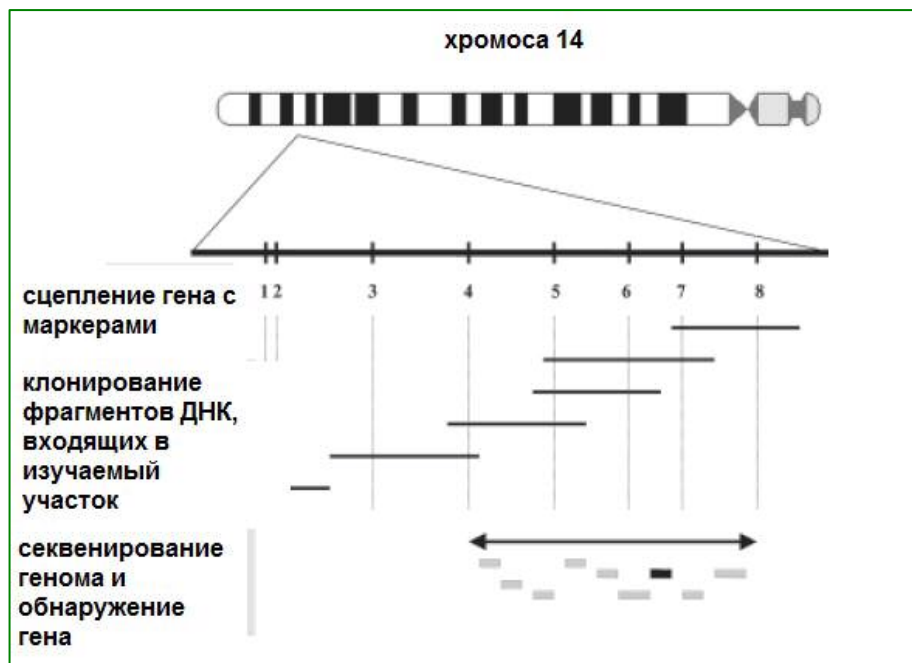


Рис. 32. Схема процедуры поиска гена-кандидата, ответственного за болезнь Альцгеймера.

Фрагмент с геном пресенилина указан внизу темным цветом. Так был найден ген, ответственный за изучаемую патологию, который назвали пресенилин 1. Позднее еще один похожий ген (ген пресенилина 2) был найден на хромосоме 1. Люди с мутациями генов пресенилинов заболевают к 30–60 годам. При этом никакие условия среды, к сожалению, не могут предотвратить или остановить развитие болезни, ведущей в конечном итоге к летальному исходу.

А первый успех на пути использования полиморфизма для определения гена, ответственного за такое заболевание человека, как болезнь Хантингтона, пришел еще в 1983 году. Тогда был локализован на коротком плече хромосомы 4 ген, ответственный за это заболевание. Затем последовали другие успешные работы. В частности, геномный скрининг аллелей предрасположенности в семьях с накоплением диабета типа 1 позволил выявить в хромосоме 6 полиморфный вариант гена, тесно связанного с развитием этого заболевания.

В настоящее время более-менее детально картировано около 1000 генов, связанных с различными болезнями человека. А всего в Энциклопедии человека описано свыше 7,5 тыс. участков, существенные изменения в которых потенциально являются причинами тех или иных наших болезней. Но со всем этим придется еще долго разбираться. Сейчас установлено, что продукты большинства (около 30%) из хорошо изученных генов, вовлеченных в патогенез, обладают энзиматической активностью. Еще примерно 14% составляют гены, продукты которых модулируют функции различных белков. Функции продуктов остальных генов многочисленны: факторы транскрипции и их рецепторы, белки, входящие в матрикс, трансмембранные транспортеры и др. Уже известны гены, мутантные формы которых приводят к сердечно-сосудистым заболеваниям, атеросклерозу, раку и другим заболеваниям.

Удалось уже выделить группу генов, которые отвечают за предрасположенность к наркомании и алкоголизму, за различные психические заболевания человека. В частности, обнаружены молекулярные механизмы свыше десятка неврологических заболеваний, получивших общее название болезней триплетных повторов. Суть генетических изменений в этих случаях заключается в размножении (это называют экспансией, захватом) микросателлитов — повторяющихся триплетов — внутри или на концах отдельных генов. Так, давно был известен синдром ломкости X-хромосомы, который проявляется в виде разрывов в длинном плече этой хромосомы. Результат этого — нарушение нормального функционирования центральной нервной системы. «Охотники за генами» обнаружили «больной» ген, который определяет эту патологию (ген FMR1). Оказалось, что в первом экзоне гена FMR1 имеется участок, состоящий из повторяющихся тринуклеотидов ЦТГ. У здоровых людей число таких повторов варьирует от 6 до 20, а вот у больных число копий резко увеличено (до 230), в результате чего этот ген перестает работать. Другие заболевания, связанные с увеличением (экспансией) тринуклеотидного повтора, — болезнь Хантингтона (здесь триплет в «больном» гене другой — ЦАГ), миотоническая дистрофия (триплет ЦТГ), болезнь Кеннеди (снова триплет ЦАГ). В сумме умеренная экспансия триплетов ЦАГ (они кодируют аминокислоту глутамин) обнаружена при 8 нейродегенеративных заболеваниях. В этом случае экспансия глутамина вызывает, по-видимому, приобретение белком токсических свойств.

Анализ «больных» генов показал, что часто разные мутации в одном гене приводят к одному и тому же типу патологии. Например, в уже упоминавшемся гене пресенилина обнаружено множество различных мутаций, и все они ассоциированы с болезнью Альцгеймера. Синдром Холта—Орама, характеризующийся скелетными и сердечными аномалиями, обусловлен мутациями в так называемом гомеобоксном гене, расположенном на 12-ой хромосоме. Продукт этого гена скорее всего функционирует как фактор транскрипции в сердце и конечностях. В гене обнаружено 8 мутаций, большая часть которых связана с возникновением преждевременного стоп-кодона. При муковисцидозе — болезни одного гена — проявление патологии весьма разнообразно. Выяснилось, что разные формы этого заболевания связаны чуть ли не с 600 различными мутациями в одном гене.

С другой стороны, в отдельных случаях разные мутации в одном и том же гене могут приводить к развитию разных патологий. Например, четыре вида заболеваний связывают с мутациями в одном гене тирозинкиназы (RET). Одно из них — болезнь Гиршпрунга — представляет собой сложное наследственное заболевание. Кроме того, с другими мутациями в гене RET связаны такие болезни, как множественные эндокринные неоплазии двух разных типов.

Кроме болезнетворных генов обнаружены еще некоторые гены, имеющие прямое отношение к здоровью человека. Выяснилось, что существуют гены, обуславливающие предрасположенность к развитию профессиональных заболеваний на вредных производствах. Так, на асбестовых производствах одни люди болеют и умирают от асбестоза, а другие устойчивы к нему. В будущем возможно создание специальной генетической службы, которая будет давать рекомендации по поводу возможной профессиональной деятельности с точки зрения предрасположенности к определенным профессиональным заболеваниям.

Оказалось, что предрасположенность к алкоголизму или наркомании тоже может иметь генетическую основу. Открыто уже семь генов, повреждения которых связаны с возникновением зависимости от химических веществ. Из тканей больных алкоголизмом был выделен мутантный ген, который приводит к дефектам клеточных рецепторов дофамина — вещества, играющего ключевую роль в работе центров удовольствия мозга. Недостаток дофамина или дефекты его рецепторов напрямую связаны с развитием алкоголизма. В хромосоме 4 обнаружен ген, мутации которого приводят к развитию раннего алкоголизма. Уже в раннем детстве это отражается на поведении ребенка в виде повышенной подвижности и заметного дефицита внимания.

Например, недавно выяснилось, что более двух третей всех полинезийцев наделены особой аллелью гена алкогольдегидрогеназы 2, которая оберегает их от алкогольной зависимости. Большинство людей с таким геном признают, что попытки выпить пива, водки или вина неизбежно оставляют после себя очень неприятное чувство и стойкое отвращение к спиртным напиткам.

В настоящее время также интенсивно изучается проблема зависимости психики человека, его способностей и талантов от его генов. Главная задача будущих исследований — это изучение однонуклеотидных вариаций (снимпов) в ДНК и выявление различий между людьми на генетическом уровне. Это позволит создавать генные портреты людей и, как следствие, эффективнее лечить болезни, оценивать способности и возможности каждого человека, выявлять различия между популяциями, оценивать степень приспособленности конкретного человека к той или иной экологической обстановке и т. д. Подробнее об этом мы расскажем далее в специальной главе.

Сейчас уже выявлено свыше двух сотен генов детоксикации, полиморфизм которых тесно взаимосвязан с существенно измененной функциональной активностью и отражается в результате в предрасположенности индивидуума к различным заболеваниям под действием ксенобиотиков — чужеродных для организма веществ, таких, как пестициды, препараты бытовой химии, лекарственные средства и т. п. Рассмотрим несколько характерных примеров. Функционально неполноценный вариант гена детоксикации, называемого геном N-ацетилтрансферазы 2, влияет на возникновение рака молочной железы, и этот эффект напрямую связывают с курением.

С определенным видом полиморфизма другого гена (ген GSTM1), участвующего в детоксикации ароматических углеводородов, связан у курильщиков повышенный риск развития такого злокачественного заболевания, как карцинома.

Интересно, что мутации генов не всегда приводят к негативным последствиям — они иногда могут быть даже полезными для конкретного индивидуума. Приведем один пример. Некоторые спортсмены используют в качестве допинга гормон эритропоэтин, который стимулирует размножение клеток крови. За это их строго наказывают. А вот у двукратного призера зимних Олимпийских игр 1964 года в Иннсбруке финна Еро Мянтиранта имела редкая мутация в гене рецепторе этого гормона (выяснилось это значительно позже). Эта мутация, доставшаяся спортсмену от деда, и внесла существенный вклад в его победы. Дело в том, что благодаря мутации у него образовывалось в крови красных кровяных клеток на 50% больше, чем в норме, что приводило к существенному повышению обмена кислорода и сильно увеличивало выносливость организма к перегрузкам.

Конечно, золотые медали — это хорошо, но, скажем, устойчивость к вирусу иммунодефицита не менее важна для человека. Ведь в период до 2020 года от СПИДа в мире может погибнуть до 70 миллионов человек! Эти пугающие цифры приводятся в отчете, представленном в 2002 году на ежегодном заседании Экономической и социальной комиссии ООН.

В Уганде и Танзании инфицированность вирусом иммунодефицита среди проституток доходит до 60–80%, однако некоторые из них не только не умирают от СПИДа, но и рожают здоровых детей. Было предположено, что имеется мутация (или мутации), защищающая человека от СПИДа. Люди с такой мутацией могут быть инфицированы вирусом иммунодефицита, но не заболевают СПИДом. В конечном итоге был идентифицирован ген, получивший имя CCR–5, с которым все это связано. За этой непонятной для непросвещенного читателя аббревиатурой скрывается ген, определенное изменение в котором (выпадение небольшого участка ДНК) придает человеку невосприимчивость к вирусу иммунодефицита человека. Большинство людей подвержено этой страшной вирусной инфекции, а вот индивидуумам с такой мутацией она не страшна. В настоящее время создана карта, примерно отражающая распределение этой мутации в гене CCR–5 в Европе. Особенно часто (до 15% населения) она встречается среди финно-угорской группы населения. Обнаружение такого мутантного гена может привести в конечном итоге к созданию надежного способа борьбы с одним из самых страшных заболеваний нашего века.

Полиморфизм в генах-триггерах (генах-переключателях) часто связан с тяжелыми мультифакторными заболеваниями. В частности, отдельные варианты гена ApoE ассоциированы с развитием таких патологий, как атеросклероз и болезнь Альцгеймера. В качестве генов-триггеров рассматривают также многочисленные онкогены и антионкогены, мутации в которых изменяют метаболизм клетки в целом и приводят к ее злокачественному перерождению.

Зная все это, молекулярная медицина имеет возможность предсказать развитие тех или других патологий у конкретного человека. За этим следует ряд превентивных мер, а в отдельных случаях и направленное лекарственное воздействие.

Многочисленные важные для медицины результаты получают на животных как на моделях. Для этого работу мутантного гена изучают с использованием специально полученных культур клеток и трансгенных животных. Смоделировав на животных болезнь человека, их можно затем использовать для подбора лекарств, влияющих на разрушительное действие мутантных генов, например препаратов, которые останавливают гибель клеток.

Иногда идут другим путем. У лабораторных животных с помощью «генного нокаута» полностью выключают тот или иной ген-кандидат, чтобы лучше понять его функцию. Так, итальянские ученые обнаружили ген, играющий ключевую роль в развитии кокаиновой наркотической зависимости. У мышей без гена mGluR5 (этот ген был выключен — нокаутирован), в отличие от нормальных мышей, пристрастие к кокаину не возникало. Нормальные же мыши быстро привыкали к наркотику. С помощью специального приспособления они вводили себе дозу кокаина по 25 раз за два часа. Эти эксперименты открывают новое направление в изучении причин наркомании у людей и поиска путей борьбы с этим тяжелым недугом.

Когда установлена связь заболевания с одним из генов, можно начинать поиск методов лечения. Выяснение механизмов развития того или иного заболевания помогает разрабатывать новые препараты, которые действуют на причину болезни, а не на симптомы. В конечном итоге должны быть разработаны способы и средства лечения большинства болезней человека. На это, скорее всего, потребуются десятилетия, может быть, даже сменится нынешнее поколение людей, однако любой успех на этом пути, без сомнений, стоит затраченных средств, усилий и времени. Благодаря молекулярной медицине в будущем, вероятно, будут созданы новые лекарства, гораздо более избирательные и эффективные, чем ныне существующие, поскольку они будут целенаправленно действовать на строго определенные генные и белковые мишени.

По мнению главы американской программы исследований генома человека Френсиса Коллинза, уже в первой половине XXI века лечение самых разных недугов будет основано на использовании синтетических генных продуктов, которые станут изменять работу заболевших клеток и органов в нужном направлении.

Теперь, когда молекулярная природа многих наследственных патологий установлена, их можно обнаружить даже на внутриутробной стадии. Это позволяет прерывать беременность или заблаговременно начать лечение ребенка. Например, в результате таких мер в некоторых странах за последние годы доля детей, больных талассемией (болезнь крови), снизилась более чем в 20 раз. Моногенные наследственные заболевания легче определяются, часто первые клинические симптомы проявляются на сравнительно поздних стадиях. Это дает возможность диагностировать их до появления первых признаков патологии. Речь идет о таких заболеваниях, как рак молочной железы, хорея Хантингтона, рак толстой кишки, миотоническая дистрофия и др.

Прогресс в секвенировании генома человека ускорил понимание и молекулярной природы сложных генетических заболеваний. Со строго определенными генами связано относительно немного заболеваний. Подавляющее же их большинство, в том числе такие «главные убийцы», как сердечные, возникает при участии многих генов и белков, с одной стороны, и под влиянием окружающей среды — с другой. Все это сильно усложняет ситуацию. Для выхода из этого положения используют несколько методов. Один из них, весьма трудоемкий, заключается в изучении большого числа семей с определенным типом заболевания. Так, американец Р. Лифтон провел со своими коллегами обширное исследование многочисленных семей с синдромом повышенного давления. Это позволило обнаружить у них специфические мутации примерно в десятке разных генов. Многие из этих генов кодируют почечные белки, которые участвуют в транспорте солей и других метаболитов.

Другой путь — изучение соответствующих мутаций у больных мышей, которые имеют значительное генетическое родство с человеком. Однако и здесь пока не все так просто.

Весьма полезной базой для решения этих вопросов может служить большая клиническая информация, разбросанная сейчас по разным институтам и медицинским центрам. Для формирования такой информационной системы сейчас уже создаются огромные международные базы данных по сложным генетическим заболеваниям. Например, программа LURIC (The Luwigshafen Risk and Cardiovascular Health Study) поставила своей целью идентифицировать новые факторы риска генетической природы и окружающей среды для болезней сердечно-сосудистой системы. Далее с помощью функциональной геномики будет осуществляться поиск новых маркеров и генов, вовлеченных в эти сложные заболевания.

К сожалению, наши новые знания пока еще слабо отражаются в медицинской практике на процессе лечения. И это понятно — мы находимся только на первом этапе пути к полному пониманию функций человеческого генома. Тем не менее, в самое ближайшее время наиболее важный практический выход наших знаний о геноме человека ожидается в генной диагностике болезней, лечении и профилактике наследственных и врожденных заболеваний и пороков.

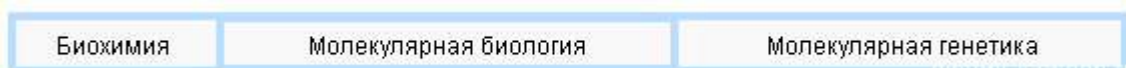
Уже давно бытует хорошо известное мнение, что все болезни от нервов. Это утверждение, как и многие другие категоричные заявления, конечно же, однобоко. Из вышеизложенного очевидно, что в основе большинства патологий лежат не нервы, а человеческий геном. Большая часть заболеваний человека имеет генетический компонент. Роль генетических нарушений может быть ведущей или второстепенной, но она есть почти всегда. «Психологический фактор» действительно играет порой существенную роль как в развитии заболевания, так и при его лечении. Но это уже проблема не генетиков, а психологов.

2. Заключительные понятия по дисциплине

2.1 Клеточный уровень организации живой природы

2.1 Включает в себя предыдущий — молекулярный уровень организации.

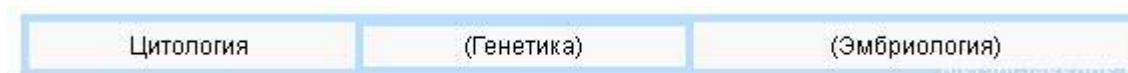
Науки, которые изучают живые организмы именно на этом уровне:



На этом уровне уже появляется термин «клетка» как «мельчайшая неделимая биологическая система»

- Обмен веществ и энергии данной клетки (разный в зависимости от того, к какому царству принадлежит организм);
- Органеллы клетки;
- Жизненные циклы — зарождение, рост и развитие и деление клеток

2.2 Науки, изучающие клеточный уровень организации:



Генетика и эмбриология изучают этот уровень, но это не основной объект изучения.

2.3. Тканевый уровень организации:

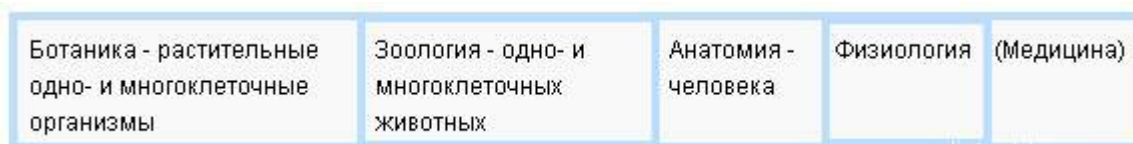
Включает в себя 2 предыдущих уровня — молекулярный и клеточный. Этот уровень можно назвать «многоклеточным» — ведь ткань представляет собой совокупность клеток со сходным строением и выполняющих одинаковые функции.

Наука — Гистология

2.4. Органный (ударение на первый слог) уровень организации жизни

У одноклеточных органы — это органеллы — есть общие органеллы — характерные для всех эукариотических или прокариотических клеток, есть отличающиеся. У многоклеточных организмов клетки общего строения и функций объединены в ткани, а те, соответственно, в органы, которые, в свою очередь, объединены в системы и должны слаженно взаимодействовать между собой.

Тканевый и органный уровни организации — изучают науки:



2.5. Организменный уровень

Включает в себя все предыдущие уровни: молекулярный, клеточный, тканевый уровни и органный.

На этом уровне идет деление Живой природы на царства — животных, растений и грибов. Характеристики этого уровня:

- Обмен веществ (как на уровне организма, так и на клеточном уровне тоже)
- Строение (морфология) организма
- Питание (обмен веществ и энергии)

- Гомеостаз
 - Размножение
 - Взаимодействие между организмами (конкуренция, симбиоз и т.д.)
- Взаимодействие с окружающей средой

Науки:

Анатомия	Генетика	Морфология	Физиология
----------	----------	------------	------------

Размеры объекта	Объект изучения	Уровень	Уровень организации
0,1 мм (100мкм) и более	Организм, органы	Организменный, органный	Анатомический
100—10 мкм	Ткани	Тканевый	Гистологический (светооптический)
20—0,2 мкм (200 нм)	Клетки (эукариотические и прокариотические)	Клеточный	Цитологический
200—1 нм	Клеточные компоненты	Субклеточный	Ультраструктурный (электронно-микроскопический)
Менее 1 нм	Молекулы	Макромолекулярный	Физико-химический

2.6. Популяционно-видовой уровень организации жизни

Включает молекулярный, клеточный, тканевый уровни, органный и организменный.

Если несколько организмов схожи морфологически (проще говоря, одинаково устроены), и имеют одинаковый генотип, то они образуют один вид или популяцию.

Основные процессы на этом уровне:

- взаимодействие организмов между собой (конкуренция или размножение)
- микроэволюция (изменение организма под действием внешних условий)

Науки, изучающие этот уровень:

Генетика	Эволюция	Экология
----------	----------	----------

2.7. Биогеоценотический уровень организации жизни

На этом уровне уже учитывается почти все:

- Пищевое взаимодействие организмов между собой - пищевые цепи и сети
- Меж- и внутривидовое взаимодействие организмов — конкуренция и размножение

- Влияние окружающей среды на организмы и, соответственно, влияние организмов на среду их обитания

Наука, изучающая этот уровень — Экология

Последний уровень — высший!

2.8 . Биосферный уровень организации живой природы

Представлен процессами:

- Взаимодействие как живых, так и неживых компонентов природы
- Биогеоценозы
- Влияние человека — «антропогенные факторы»
- Круговорот веществ в природе

2.2. Типы мутаций

Мутация — устойчивое (то есть передающееся по наследству) преобразование генотипа, происходящее под влиянием внешней или внутренней среды. Процесс возникновения мутаций получил название мутагенеза.

К мутациям способны абсолютно все формы жизни на Земле — мутируют как вирусы, так и высокоразвитые организмы. Изменение организма идет на уровне ДНК, поэтому не всегда может быть заметна фенотипически.



Соматический тип мутаций — изменения происходят в соматических (т.е. неполовых) клетках — клетках тела.

Яркий пример — появление на ветках черной смородины красных ягод или гетерохромия — явление разного цвета глаз у человека.

Они не передаются по наследству

Генеративный тип мутаций — это изменения, происходящие в генеративных, т.е., половых клетках (гаметах). Соответственно, этот тип мутаций передается по наследству.

Доминантные мутации появляются уже в первом поколении, а рецессивные — только во втором и последующих поколениях.

типы мутаций:

Генные мутации — мутации, приводящие к изменению одного гена

Это может быть выпадение одного или нескольких нуклеотидов (в части С ЕГЭ по биологии такие задания встречаются довольно часто). Кажется, изменение не существенное, но оно влечет за собой серьезные последствия. Т.к. белок строится при «считывании» молекулы РНК, а она, в свою очередь, строится на «базе» ДНК, то такое изменение ведет к изменению белкового состава организма.



Наиболее частой причиной генных мутаций является сбой при удвоении (репликации) ДНК. Яркие примеры проявления этого типа мутации — серповидно-клеточная анемия (смертельное заболевание), альбинизм и дальтонизм.

Наиболее вероятная мутация генов происходит при скрещивании тесно связанных организмов, которые унаследовали мутантный ген у общего предка. По этой причине вероятность возникновения мутации повышается у детей, чьи родители являются родственниками.

Геномные мутации — изменение количества хромосом кариотипа организма.

- изменение всего на одну хромосому — 47 хромосом вместо 46 — синдром Дауна;
- полиплоидия — увеличение хромосомного набора в несколько раз (триплоидия, тетраплоидия и т.д.)- у животных не бывает, у растений — довольно часто. Почему у растений такое изменение встречается чаще? Потому что у растений возможно самоопыление — одна клетка организма с таким нарушением оплодотворяется клеткой с точно таким же нарушением.

Причина такого типа мутации — неправильное расхождение хромосом в процессе мейоза.

Хромосомные мутации – изменение строения хромосом

Это может быть изменение или даже выпадение целого участка хромосомы, изменение последовательности, перенос негомологичного участка с одной хромосомы на другую и т.д. Из-за такой перестройки может меняться функции генов. Зачастую это приводит к меньшей жизнеспособности организма.

Мутации происходят постоянно. Как уже было сказано, не всегда они видны внешне (не всегда проявляются фенотипически). Чем более заметен внешний эффект изменения, тем более вредной считается мутация.

Мутация — основа изменчивости, а она, как известно, двигатель эволюции.

Организм не может знать, какие мутации будут полезны в следующем поколении. Нет и не может быть механизма, который бы обеспечивал направленное появление полезных для организма мутаций.

Литература

1. Ливитин В. Удивительная генетика Энас-книга 2013 г.
2. Курчанов Н Генетика человека с основами общей генетики СпецЛит 2009 г.
3. Терри А. Браун Геномы Институт компьютерных исследований 2009