

УДК 575.571.27

РОЛЬ США В РАЗРАБОТКЕ МЕЖДУНАРОДНОГО ПРОЕКТА «ГЕНОМ ЧЕЛОВЕКА»

© 2011 г. **Л.П. Жиганова***
Институт США и Канады РАН, Москва

В статье описывается роль США в разработке и выполнении международного проекта «Геном человека». Анализируется участие государственного и частного секторов в осуществлении проекта. Приводится прогноз геномных исследований.

Ключевые слова: *геном, секвенирование, геномика, хромосомы, гены.*

26 июня 2000 г. человечество шагнуло в новую эру. Сообщение о событии исключительной важности облетело мир и попало практически во все газеты: получен первый вариант расшифровки генома человека. В Восточном зале Белого дома в Вашингтоне прошла торжественная церемония, посвящённая успешному завершению работ. Рядом с президентом США Клинтоном стояли американский учёный Фрэнсис Коллинз, руководитель международного некоммерческого проекта «Геном человека», и Крейг Вентер, глава конкурирующей частной компании «Селера Дж. Геномикс», проводившей аналогичные исследования. Премьер-министр Великобритании Тони Блэр был на связи по спутниковому каналу, и торжества проходили одновременно во многих частях света.

Президент Клинтон начал свою речь со сравнения составленной карты генома человека и той карты новых земель, которую почти 200 лет назад в этом самом зале развернул знаменитый путешественник Мэриуэзер Льюис перед президентом Томасом Джефферсоном. «Без сомнения, – сказал Клинтон, – это самая важная и самая дивная карта, какую когда-либо составляло человечество». Но более всего общее внимание привлекла та часть его речи, где он перешёл от научного значения проекта к его духовному аспекту. «Сегодня, – произнёс он, – мы изучаем язык, посредством которого Бог создал жизнь. И мы испытываем ещё большее благоговение перед сложностью и дивной красотой драгоценнейшего и священнейшего из его даров» [3].

* ЖИГАНОВА Лариса Петровна – кандидат биологических наук, старший научный сотрудник Центра аграрных проблем ИСКРАН. E-mail: Larissa-Zhiganova@rambler.ru

В ответном выступлении Фрэнсис Коллинз откликнулся так: «Сегодня счастливый день для всего мира. Смирением и благоговением наполняет меня сознание того, что мы впервые сумели заглянуть в инструкцию, по которой сотворены и которая до сих пор была известна одному лишь Богу» [3].

Геном – (нем. *Genom*, англ. *genom(e)*) гаплоидный (одинарный) хромосомный набор, совокупность генов, локализованных в одиночном наборе хромосом данного организма. Термин предложен в 1920 г. немецким биологом Г. Винклером. Под геномом принято понимать совокупность генов, сосредоточенных в хромосомах без учёта наследственных детерминант, связанных со структурами цитоплазмы. В гаметах, половых клетках диплоидных организмов, а также в клетках гаплоидных организмов содержится один геном; в соматических клетках диплоидных организмов – два генома. С увеличением пloidности клеток растёт число геномов. При оплодотворении происходит объединение геномов отцовских и материнских гамет. Как правило, геномы, полученные от отцовской и материнской гамет, гомологичны. Под абсолютной гомологией двух геномов понимают совпадение линейного расположения генов в каждой хромосоме. Наличие такого совпадения обеспечивает возможность нормальной конъюгации хромосом в мейозе. Изменение числа хромосом называют геномными мутациями. Организм, у которого несколько раз повторяется один и тот же геном, называется автополиплоидом. Организм, в котором объединены разные геномы, называется аллополиплоидом. Для понимания структуры и функционирования геномов большое значение имело установление строения молекул нуклеиновых кислот – ДНК и РНК, а также механизмов их синтеза – репликации, транскрипции и трансляции [1, т. 6].

Немного из истории молекулярной биологии

Журнал «Нейчер» 25 апреля 1953 г. опубликовал небольшое письмо молодых и никому тогда не известных Ф. Крика и Дж. Уотсона редактору журнала, которое 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 900 слов, и каждое из них было на вес золота. «Ершистая молодёжь» посмела выступить против нобелевского лауреата Лайнуса Полинга, американского физика и химика, автора знаменитой альфа-спирали белков, общественного деятеля. Он преподавал и вёл исследовательскую работу в Калифорнийском технологическом институте, с 1969 г. – профессор Стэнфордского университета. Председатель Американского химического общества (с 1948 г.), член Национальной академии наук США. Один из инициаторов Пагуошских конференций. Автор обращения американских учёных к президенту США о немедленном прекращении испытаний ядерного оружия (1957 г.) и петиции аналогичного содержания, направленной в ООН с подписями свыше 9 тыс. учёных различных стран (1958 г.). В 1954 г. получает Нобелевскую премию по химии, в 1962 г. – Нобелевскую премию мира, в 1970 г. – Международную Ленинскую премию «За укрепление мира между народами».

Полинг буквально накануне выступления «ершистой молодёжи» опубликовал свою статью, в которой представлял ДНК как трёхцепочечную спиральную структуру наподобие девичьей косы. В действительности ДНК является двойной спиралью. Просто тогда у Полинга был недостаточно очищенный препарат. Но и Полинг оказался отчасти прав – сейчас трёхцепочечность некоторых участков наших генов хорошо известна. Это свойство ДНК пытались одно время использовать в борьбе с раком, выключая с помощью олигонуклеотидов те или иные раковые гены (онкогены). Биологии нуклеиновых кислот долго не везло. Первую Нобелевскую премию за открытие строения нуклеотидов получил немецкий учёный Альбрехт Коссель ещё в 1910 г. А знаменитая реакция Фельгена для окрашивания ДНК была предложена накануне Первой мировой войны и усовершенствована в 1920-е годы. Тогда бы и началась новая эра биологии...

Однако биологи были уверены, что «монотонная» ДНК с её различающимися четырьмя нуклеотидами просто не могла нести генетическую информацию о миллионах самых разнообразных белков. И хотя уже применялась азбука Морзе с тремя кодирующими элементами, менталитет исследователей ещё не достиг уровня информационной эры с её двоичной системой записи («0» или «1») любой информации. Лишь к началу 1950-х годов учёные стали обращать внимание на ДНК, роль которой в передаче наследственной информации (наследственных признаков) у микроорганизмов в 1943 г. открыл американский учёный Освальд Эйвери. Исследования Эйвери продолжил Сальватор Лурия, который вместе с Максом Дельбрюком организовал лабораторию недалеко от Нью-Йорка на биостанции Колд-Спринг-Харбор.

Эдуард Сальвадор Лурия – американский микробиолог итальянского происхождения, член Национальной академии наук США (1959 г.). Медицинское образование получил в Турине (1935 г.). Работал в лаборатории Ф. Жолио-Кюри в Париже (1938–1940 гг.), затем переехал в США. Один из основоположников генетики микроорганизмов. В 1969 г. получил Нобелевскую премию вместе с М. Дельбрюком и А. Херши за работы в области молекулярной биологии [1, т. 15]. Макс Дельбрюк – американский физик, генетик, вирусолог немецкого происхождения, член Национальной академии наук США (1959 г.). В 1924–1930 гг. учился в институтах Тюбингена, Берлина, Бонна, Геттингена, где в 1930 г. получил степень доктора философии. В 1932–1937 гг. работал в институте кайзера Вильгельма. В 1937 г. эмигрировал в США и работал в Калифорнийском технологическом институте (Пасадена), с 1947 г. профессор биологии этого института [1, т. 8]. Интересно отметить, что физик Макс Дельбрюк был учеником русского генетика Н.В. Тимофеева-Ресовского в биологии и соавтором их с К. Циммером знаменитой статьи, посвящённой определению размеров гена. Лурия с Дельбрюком изучали жизненный цикл бактериофагов (вирусов бактерий), в результате чего пришли к выводу о биологической роли ДНК. Лурия послал затем своего аспиранта, американского биолога Джеймса Уотсона в Кавендишскую лабораторию в Кембридже, где в то время английские физики Морис Уилкинс и Розалинда Франклин исследовали

строение ДНК с помощью рентгена (в то время англичане лидировали в рентгеноструктурном анализе биомолекул).

В лаборатории Уилкинса работал и молодой тогда физик Фрэнсис Крик, известный в научных кругах своим скепсисом: для него не существовало никаких авторитетов, чем он и заработал репутацию скандалиста. Статью Полинга в лабораторию принёс его сын, который, кстати, помог Уотсону и Крику уяснить роль попарного комплементарного соединения азотистых оснований. Статья стала последней каплей перед озарением...

Научное сообщество не сразу признало их открытие. Сначала Нобелевскую премию за работы в области ДНК присудили в 1959 г. известным американским биохимикам Северо Очоа и Артуру Корнбергу. Очоа был первым, кто в 1955 г. синтезировал РНК, а Корнберг в 1956 г. синтезировал ДНК в пробирке (*in vitro*). В 1962 г. настал черёд Ф. Крика, Дж. Уотсона и М. Уилкинса. Розалинда Франклин к тому времени уже умерла от рака в возрасте 37 лет, иначе это был бы единственный случай в истории Нобелевских премий, когда награду вручили бы четверым, хотя это не допускается уставом. Вклад Р. Франклин в развитие рентгеноструктурного анализа ДНК был просто недооценен. В 2002 г. вышла книга Брэнды Мэдокс «Розалинда Франклин: забытая леди ДНК», в которой освещены ранее неизвестные подробности работы Розалинды и её причастности к открытию структуры ДНК.

После открытия Уотсона и Крика важнейшей проблемой стало выявление соответствия между первичными структурами ДНК и белков. Поскольку в составе белков 20 аминокислот, а нуклеотидов в ДНК всего четыре, то для записи информации о последовательности аминокислот в полинуклеотидах необходимо не менее трёх оснований. Варианты трёхбуквенных генетических кодов предложили американский физик российского происхождения Георгий Гамов и российский биолог Александр Нейфах. Гипотезы были достаточно умозрительны. В конечном счёте трёхбуквенный генетический код расшифровал Ф. Крик. Вряд ли он предполагал, что в будущем этот ход станет возможной расшифровкой генома человека. Однако два открытия позволили сдвинуть проблему с места.

В 1970 г. неизвестные широкой общественности американские учёные Говард Тёмин и Дэвид Балтимор опубликовали в журнале «Нейчер» статьи, посвящённые ферменту обратной транскриптазе. Этот фермент содержится в РНК-содержащих вирусах, в том числе и раковых. Обратная транскриптаза синтезирует ДНК на матрице РНК, поэтому является РНК-зависимой ДНК-полимеразой, т.е. осуществляет реакцию, обратную той, которую наблюдали обычно в клетке. В 1975 г. Г. Тёмин и Д. Балтимор получают Нобелевскую премию в области биологии.

Открытие обратной транскриптазы позволило выделить первые гены. Но процесс был очень трудоёмкий. Спустя 15 лет химик из Калифорнии предложил уникальную полимеразную цепную реакцию (ПЦР), сразу же ставшую знаменитой. В этой реакции фермент полимеразы ходит как «челнок» по ДНК, нарабатывая любые количества этого фрагмента [15]. В современной молекулярной биологии секвенирование (расшифровка) ДНК – это уже технологич-

ная рутина. Принцип был изобретён в 1977 г. Сэнгером и др. и произвёл революцию в методах биологических исследований. Метод секвенирования по Сэнгеру основан на том, что копирование ДНК с помощью специального фермента (ДНК-полимеразы) в пробирке можно остановить с помощью 2',3'-дидезоксинуклеозидтрифосфата – неправильного нуклеотида («буквы» ДНК), – который так устроен, что узнаётся ферментом, но зато не цепляет следующий нуклеотид [18]. ПЦР, а также новая электронная техника и компьютеры сделали реальной задачу расшифровки всего генома человека. Долгие дебаты закончились в конце сентября 1988 г, когда во главе проекта *HUGO (Human Genome Project)* – Организации по расшифровке генома человека – был поставлен учёный, открывший двойную спираль ДНК, – Джеймс Уотсон.

Журнал «Тайм» назвал Уотсона «охотником за генами». Уотсон: «Это захватывающая перспектива. Тридцать лет назад мы не могли и мечтать о том, чтобы узнать структуру генома даже мельчайшего вируса. А сегодня мы уже расшифровали геном вируса СПИДа и почти полностью прочитали геном кишечной палочки объёмом в 4,5 млн. букв генкода» [15].

Международный проект «Геном человека»

Идея этого проекта была выдвинута в США. Он стал кульминацией работы, поддержанной Министерством энергетики США, в частности, семинаров, проводившихся в 1984 и 1986 г. и последовавшими действиями Министерства энергетики [5; 6]. Затем национальные программы «Геном человека» были приняты в Великобритании, Франции, Германии, Италии, СССР. Существование многих национальных программ изучения генома человека вызвало к жизни проблемы, главными из которых были координация усилий и распространение полученных результатов. Для их решения и был создан международный проект «Геном человека» (*HUGO*). Предложение об организации международного органа, выполняющего функции координации усилий учёных разных стран в изучении генома человека, впервые было сделано на первом симпозиуме в Колд-Спринг-Харбор по картированию и секвенированию генома человека Виктором Маккьюзиком. Оно было принято, и уже в 1989 г. проект был зарегистрирован в Женеве и Делавере (США) [14].

HUGO организована по образцу академий, т.е. её члены выбираются из числа известных учёных. В её рамках созданы и действуют шесть комитетов:

- по международным школам картирования генома;
- по физическому картированию;
- по информатике;
- по картированию генома мыши;
- по этическим, юридическим и социальным аспектам;
- по интеллектуальной собственности.

Цели организации заключаются в:

- 1) координации исследований генома, в частности, организации сотрудничества между учёными во избежание конкуренции или дублирования усилий;
- 2) координации работ по изучению модельных организмов; обмену информацией и биоматериалами;

3) распространении соответствующих технологий путём организации учебных программ;

4) предоставлении информации о результатах исследований и их применение; о научных, юридических, этических, социальных и коммерческих аспектах изучения генома человека.

Столь масштабная задача, как изучение генома человека, вызвала к жизни новые формы организации научных исследований, дала мощный толчок развитию международного сотрудничества. Впервые, пожалуй, такие большие научные силы задействованы для получения базовой, справочной, информации, которая в полной мере может быть использована ещё не скоро. Вероятно, не будет преувеличением сказать, что это очень изменило психологический климат молекулярной биологии. Официально эта научная программа с участием ведущих молекулярно-биологических лабораторий США, Западной Европы, Японии и России оформилась в 1990 г. Джеймс Уотсон, бывший в то время самой популярной звездой от биологии, убедил Конгресс рискнуть и выделить деньги на расшифровку генома.

Международный проект «Геном человека» – один из самых трудоёмких и дорогостоящих в истории науки. В 1990 г. Конгресс утвердил для него бюджетное финансирование и обозначил дату реализации 2005 годом. Несмотря на то, что первоначально расходы оценивались в 3 млрд. долл., за этот период на проект было потрачено меньше, чем ожидалось, около 2,7 млрд. долл. [19]. Если в 1990 г. на него было потрачено около 60 млн. долл. в целом, то в 1998 г. одно только правительство США израсходовало 253 млн. долл. В проекте задействованы несколько тысяч учёных из более чем 20 стран. С 1989 г. в нём участвует и Россия, где по проекту работало около 100 исследовательских групп. Для изучения все хромосомы человека были поделены между странами-участницами, и, в частности, России для исследования достались 3-, 13- и 19- хромосомы [17].

Дж. Уотсон умело руководил американской частью проекта в течение первых двух лет. Он стал главой Национального центра исследований человеческого генома в Национальной организации здравоохранения США. В этот период были созданы центры по расшифровке генома, и к работе подключились лучшие и талантливейшие учёные нынешнего поколения. Однако многие относились со скепсисом к идее о том, что можно за 15 лет осуществить этот проект: тогда ещё не были созданы некоторые технологии, позволившие ускорить процесс. В 1992 г. наступил кризис: Уотсон неожиданно оставил проект после публичного спора с директором системы Национальных институтов здравоохранения Бернадией Хили о возможности патентования фрагментов генетического кода, против чего Дж. Уотсон категорически возражал [3]. Начались поиски нового директора проекта и выбор остановился на руководителе центра Мичиганского университета Фрэнсисе Коллинзе. В 1997 г. название центра, которым он стал руководить, было переименовано в Национальный институт исследований человеческого генома (*National Human Genome Research Institute*). Следующие десять лет были бешеной гонкой со множеством взлётов и падений. Иногда многообещающие методы проваливались при применении в крупном масштабе, случались трения между членами научной команды. Неко-

торые центры не выдержали заданного темпа, и их пришлось вывести из проекта. Однако бывали и моменты торжества. К 1996 г. они были готовы запустить полномасштабное секвенирование генома с применением значительно более совершенной и рентабельной технологии метод «беспорядочной стрельбы» (*shotgun technology*). Суть метода в том, что определяемую ДНК организма разбивают на множество небольших фрагментов, каждый из которых вводят в автомат, определяющий последовательность ДНК. После того, как будут определены последовательности каждого фрагмента, в действие вводят сложнейшие компьютерные программы, заново собирающие исходную последовательность. Такое интенсивное использование информационных технологий объясняет, почему многие учёные называют новую область исследований генома *биоинформационной*, а не *биомолекулярной* революцией [12].

Каковы же достижения учёных за десять с небольшим лет работы над проектом? Первым крупным успехом стало полное картирование в 1995 г. генома бактерии *Haemophilus influenzae*. Позднее были полностью описаны геномы ещё более 20 бактерий, среди которых возбудители туберкулеза, сыпного тифа, сифилиса и др. В 1996 г. картировали ДНК первой эукариотической клетки – дрожжей, а в 1998 г. впервые был картирован геном многоклеточного организма – круглого червя *Caenorhabditis elegans*. К 1998 г. установлены последовательности нуклеотидов в 30 261 гене человека, т.е. расшифрована примерно половина генетической информации человека. В определённый момент руководители международной части проекта приняли важнейшее решение, сделав обязательным условием участия немедленный доступ к полученным данным и договорившись не подавать заявок на патентование каких бы то ни было фрагментов кода. Учёным был необходим немедленный свободный и открытый доступ к информации.

Следующие три года прошли плодотворно, так что к 1999 г. они смогли ускорить процесс расшифровки. Но тут возникла новая проблема: у проекта «Геном человека» появился конкурент – частная компания. Когда проект начинался, полное секвенирование всего генома не представляло коммерческого интереса, но положение вещей менялось по мере того, как ценность получаемой информации становилась очевиднее, а себестоимость секвенирования снижалась. Завершение работ по расшифровке генома человека консорциумом учёных планировалось к 2003 г. – 50-летию открытия структуры ДНК. Однако конкуренция в этой области сказала своё слово.

В 1998 г. секвенировали около 100 тыс. пар оснований ДНК в день при стоимости 50 центов за основание, или около 36 млн. пар оснований в год. К 1999 г. усовершенствование автоматических устройств привело к пятикратному ускорению работы, поэтому к 2003 г. объём секвенирования достиг 500 млн. пар оснований в год при стоимости 50 центов за основание.

Крейг Вентер основал частную компанию «Селера дженомикс» (*Celera Genomics*), которая стала продавать генные последовательности за большие деньги. Включившись в гонку по расшифровке генома, она за один год сделала то, что у международного консорциума учёных из разных стран ушло десять лет. Это стало возможным благодаря новому методу чтения генетических последовательностей и использованию автоматизации процесса чтения.

В 1998 г. заявление только что созданной компании «Селера дженомикс» о том, что она собирается увеличить разработку метода фрагментирования ДНК на человеческий геном, в некоторых кругах было встречено скептически. Техника фрагментирования разрывает ДНК на фрагменты различных размеров, от 2,000 до 300,000 пар нуклеотидов в длину, образуя то, что называется «библиотекой» ДНК. Затем ДНК «читают» с помощью автоматического секвенатора кусками по 800 пар нуклеотидов длиной с обоих концов каждого фрагмента. С помощью сложного алгоритма сборки и суперкомпьютера, кусочки собирают воедино, после чего геном может быть реконструирован из миллионов коротких фрагментов длиной в 800 пар нуклеотидов. Успех как государственного, так и частного проектов зависел от новой, более высоко автоматизированной капиллярной секвенирующей ДНК-машины, которая называлась «Эпплайд Баосистемс-3700» (*Applied Biosystems-3700*). Она прогоняла цепочки ДНК через необычайно тонкую капиллярную трубку, а не через плоский гель, как это делали в ранних моделях секвенаторов. Ещё более критическим фактором была разработка новой, более масштабной программы сборки генома, ассемблера, который мог бы обрабатывать 30–50 миллионов последовательностей, требующихся для секвенирования всего человеческого генома. В то время такой программы не существовало. Одним из первых крупных проектов в «Селера дженомикс» стала разработка данного ассемблера, который был написан параллельно с созданием большой, высокоавтоматизированной фабрики секвенирования геномов. Разработка ассемблера велась под руководством Брайена Рамоса. Первая версия появилась в 2000 г., когда команда «Селера дженомикс» объединила силы с профессором Джеральдом Рубином для секвенирования генома фруктовой мушки *Дрозофила меланогастер* методом фрагментирования генома [6]. Собрав 130 млн. пар нуклеотидов, программа обработала, по меньшей мере, в 10 раз больше данных, чем любой ранее собранный из результатов метода фрагментирования геном. Год спустя команда «Селера дженомикс» опубликовала свою сборку 3 млрд. пар нуклеотидов человеческого генома.

Вентер, глава фирмы «Селера дженомикс» заявил, что планирует выполнить полномасштабное секвенирование генома, причём, в отличие от государственного проекта, намерен запатентовать многие гены и хранить информацию в закрытой базе данных, предоставляя доступ по подписке за значительную плату. Идея обратить информацию генома в частную собственность вызвала глубокую тревогу. Ещё сильнее беспокоила реакция Конгресса. Хотя никаких данных команда «Селера дженомикс» не представила, в Конгрессе был поднят вопрос о том, целесообразно ли дальше финансировать за счёт налогоплательщиков проект, с которым, возможно, лучше справился бы частный сектор? В своих публичных выступлениях представители фирмы упирали на более высокую эффективность собственного подхода, клеймя государственный проект как медлительный и бюрократизированный. Эти заявления были спорными, так как в проекте «Геном человека» участвовали лучшие университеты мира, талантливейшие учёные. Но пресса любит скандалы, и журналисты много писали о «состязании» в секвенировании генома, привлекая к сравнению яхту Вентера и мотоцикл Коллинза. «Авторы пасквильных статей, – пишет в своей книге Ф. Коллинз, – похоже, не поняли главного:

спор шёл вовсе не о том, кто выполнит работу быстрее или с меньшими затратами (и «Селера дженомикс», и государственный проект были хорошо подготовлены), а об идеалах. Чем должна стать последовательность генома, наше общее наследие – коммерческой услугой или общественным благом?» [5].

Двадцать центров по расшифровке генома в мире работали безостановочно. За полтора года, определяя 1000 комплементарных пар в секунду, 24 часа, 7 дней в неделю, государственный проект получил первый вариант расшифровки, покрывающий 90% всего генома человека. «Селера дженомикс» тоже сгенерировала большой объём данных, но они хранились в закрытой базе, и доступ к ним отсутствовал. В определённый момент сотрудники фирмы осознали, что могут на общих основаниях пользоваться открытой информацией, и остановили работу. В итоге «Селера дженомикс» представила расшифровку, которая более чем наполовину состояла из данных, опубликованных в рамках государственного проекта.

Вначале фирма анонсировала, что она будет добиваться патентной защиты «всего лишь 200 или 300» генов, но позднее внесла поправки, что ищет «защиту интеллектуальной собственности» на «полное описание важнейших структур», составляющих примерно 100–300 целей. Наконец, фирма подала предварительные патентные заявки на 6 500 целых или частичных генов. «Селера дженемикс» также обещала опубликовать результаты своей работы, выпуская новые данные ежеквартально (проект «Геном человека» выпускал новые данные ежедневно), однако, в отличие от проекта с государственным финансированием, фирма не давала разрешения на свободное распространение или коммерческое использование своих данных. В марте 2000 г., президент США Клинтон заявил, что последовательность генома не может быть запатентована и должна быть свободно доступна для всех исследователей. После заявления президента акции компании «Селера дженомикс» сильно упали, что потянуло вниз весь биотехнологический сектор НАСРАК, потерявший около 50 млрд. долл. рыночной капитализации за два дня. Интерес к «состязанию» становился неуместным и грозил отвлечь внимание от цели обоих проектов. В конце апреля 2000 г. состоялась встреча К. Вентера и Ф. Коллинза, которые договорились об одновременном объявлении результатов работы. Так они оказались главными действующими лицами торжества в Белом доме. Знаменательно, что именно в этот год праздновалось **100-летие** открытия Грегором Менделем фундаментальных законов наследственности.

Итак, государственная программа «Геном человека», финансируемая Министерством энергетики США и руководимая Ф. Коллинзом, выполнялась в течение десяти лет и стоила около 3 млрд. долл. Секвенирование «Селеры дженомикс», руководимое К. Вентером, заняло два года и стоило 300 млн. долларов [17].

В феврале 2001 г. два наиболее авторитетных научных журнала в мире «Мейчер» и «Сайенс» опубликовали отчёты двух научных групп, расшифровавших геном человека. В журнале «Нейчер» от 12 февраля 2001 г. приведены подробные данные о структуре генома человека, полученные международным консорциумом под руководством Фрэнсиса Коллинза, в котором работали учёные Англии, Германии, Китая, США, Франции и Японии в рамках междуна-

родной программы «Геном человека» с привлечением государственного финансирования. Эта группа выделила в ДНК особые маркеры, легко распознаваемые участки, и по ним определила нуклеотидные последовательности генома человека [7]. В течение следующих трёх лет Ф. Коллинз продолжал работу в качестве директора проекта «Геном человека». Уточнялись первоначальные данные, заполнялись пробелы, повышался уровень корректности информации, и все результаты выкладывались в общедоступные базы данных.

В журнале «Сайенс» от 16 февраля 2001 г. учёные частной фирмы «Селера дженомикс» под руководством Крэйга Вентера опубликовали результаты расшифровки генома человека, полученные с применением другой стратегии исследований, в основе которой лежит анализ последовательностей нуклеотидных оснований в коротких участках ДНК человека. Таким образом, при расшифровке генома человека были использованы два научных подхода, каждый из которых имеет свои преимущества и недостатки. Важно отметить, что получены близко совпадающие результаты, которые взаимно дополняют друг друга и свидетельствуют об их достоверности. Вопрос о точности изучения последовательностей ДНК особенно важен в отношении генома человека. В нашем геноме существует большое число повторов нуклеотидов. Кроме них в хромосомах есть теломеры, центромеры и зоны гетерохроматина, где секвенирование затруднено, и они пока исключены из исследований. Предварительный анализ опубликованных материалов по расшифровке генома человека позволяет отметить несколько особенностей. Количество генов у человека оказалось существенно меньше, чем предполагали учёные несколько лет назад, называя величины 80–100 тыс. генов. По данным, опубликованным в журнале «Нейчер», у человека около 32 тыс. генов, тогда как в геноме мухи дрозофил их 13 тыс., круглого червя нематоды – 19 100, а растения арабидопсиса – 25 тыс. генов. При сопоставлении этих величин следует иметь в виду, что расчётное число генов человека получено методами компьютерной геномики, и не у всех генов выявлены конечные продукты. Кроме того, в геноме человека действует принцип «один ген – много белков», т.е. многие гены кодируют семейство родственных, но существенно различающихся белков. Следует также иметь в виду процесс посттрансляционной модификации белков за счёт различных химических групп – ацетильных, гликозильных, метильных, фосфатных и других. Поскольку таких групп в молекуле белка много, то и разнообразие может быть практически безграничным. Другой особенностью генома человека является наличие в нём генов различных вирусов и бактерий, которые постепенно накапливались в процессе многомиллионной эволюции человека. По образному выражению академика Л.Л. Киселёва, «геном человека представляет собой молекулярное кладбище, на котором покоятся вирусные и бактериальные гены, большинство из них молчит и не функционирует» [4; 11].

В апреле 2003 г., когда отмечалось 50-летие открытия Уотсоном и Криком двойной спирали ДНК, было объявлено о достижении всех целей, ставившихся перед проектом. Ф. Коллинз испытывал огромную гордость за две с лишним тысячи учёных, совершивших этот выдающийся подвиг, – расшифровка генома и через тысячу лет будет считаться одним из главных достижений человечества.

Предварительное описание особенностей генома человека

Итак, геном человека прочитан. На 3,1 млрд. нуклеотидов кода ДНК, распределённых по 24 хромосомам, оказалось, что очень маленькая часть присутствует для кодирования белков, – всего 20–25 тыс. генов, на них приходится 1,5% генетического кода. Теоретически ожидаемое количество – 100 тыс. генов. Некоторые наблюдатели восприняли этот факт как оскорбительный для человечества. Недоумение учёных понятно: у мушки дрозофилы 13 601 ген, у круглого почвенного червя – 19 тыс., у горчицы – 25 тыс. генов. Столь малое количество генов у человека не позволяет выделить его из животного царства и считать «венцом» творения [15]. Стали высказываться мнения, что, может быть, наша сложность определяется не числом отдельных пакетов инструкций, а тем, как они используются. Может быть, компоненты, из которых мы состоим, обучились работать в многозадачном режиме. Пример: объём активного словаря у образованного носителя английского языка – около 20 тыс. слов, из которых составляются простые документы, например, памятка водителя, и «Улисс» Джеймса Джойса.

Основную часть ДНК хромосом человека занимают пустынные участки и так называемые тандемные повторы. Повторы бессмысленны и следуют друг за другом наподобие велосипедов-тандемов, отсюда и название. ДНК повторов называют паразитической или эгоистической, так как она ничего не делает в геноме, однако сохраняется и увеличивает массу хромосом. Зато там, где располагаются гены, активность ДНК и ферментов, синтезирующих её копии в виде молекул информационной РНК, повышается в 200–800 раз. Это – «горячие точки генома». В геноме человека учёные насчитали 223 гена, которые сходны с генами кишечной палочки. Зачем нам такие «древние» гены? Видимо, современные организмы унаследовали от предков какие-то фундаментальные свойства клеток и биохимические реакции, для которых необходимы соответствующие белки. Половина белков млекопитающих имеет сходство аминокислотных последовательностей с белками мухи дрозофилы. С мышью – 90%, с шимпанзе – 99%. В нашем геноме человека много последовательностей, доставшихся от ретровирусов (вирусы рака и СПИДа). А у них присутствует обратная транскриптаза. После синтеза ДНК на матрице РНК вируса с помощью обратной транскриптазы вирусный геном встраивается в ДНК хромосом клетки. Таких ретровирусных последовательностей в геноме человека довольно много. Время от времени они исключаются, в результате чего возникает рак. Недавнее открытие: для встраивания вируса необходима специфическая последовательность из 14 букв генетического кода для того, чтобы блокировать агрессивные ретровирусы и целенаправленно «внедрять» нужные гены. Такой метод используется в генотерапии.

У млекопитающих ретровирусы блокируют активность иммунных Т-лимфоцитов, ответственных за отторжение чужеродных органов и тканей. Плод развивается внутри организма матери без иммунного ответа отторжения наполовину генетически чужеродного материала (отцовского). Они активируются в геноме и накапливаются в клетках плаценты. Недавно был обна-

ружен вирус, который блокирует экспрессию ретровируса. Если вирусом-блокатором заразить беременную мышь, то мышата рождаются нормальными и в срок. Но если ввести в клетки плаценты, то происходит выкидыш плода, так как активируются Т-лимфоциты матери. Ретровирусные последовательности возникают на концах хромосом – теломерах. Теломеры состоят из одноцепочечной ДНК, которая синтезируется ферментом теломеразой на матрице РНК. Считается, что теломеры являются нашими молекулярными часами, поскольку они укорачиваются с каждым клеточным делением. Раньше считалось, что в теломерах нет генов, однако расшифровка генома показала, что генов там довольно много и они активны в детстве и в молодом возрасте, постепенно «угасая» по мере старения организма. Не так уж бездействительны tandemные повторы. В норме они имеют определённое число повторяющихся троек, пятёрок и даже семёрок букв. Но в некоторых случаях в результате мутаций число повторов начинает нарастать, что ведёт к нестабильности генома. Дело доходит даже до «поломок» концов хромосом. Фрагментация концевых участков хромосом может привести к транслокации хромосом, а также синтезу таких форм белка, которые вызывают гибель нервных клеток, как это наблюдается при наследственной хорее Гентингтона [15]. Существует и еще один механизм генетического разнообразия, который выявился только в процессе прочтения генома. Это сингулярный нуклеотидный полиморфизм, или, так называемые факторы СНП.

Полиморфизмом в генетике называют ситуацию, когда гены одного и того же признака существуют в разных вариантах. Примером полиморфизма, или, другими словами, множественных аллелей, служат группы крови, когда в одном хромосомном локусе (участке) могут находиться варианты генов А, В или О, которые определяют первую, вторую, третью и четвертую группы крови иммунологической системы АВО. Сингулярность по-латыни означает одиночество, что-то единственное. СНП – это изменение «буквы» генетического кода без «последствий для здоровья». Считается, что у человека СНП встречается с частотой 0,1%, т.е. каждый человек отличается от других одним нуклеотидом на каждую тысячу нуклеотидов в ДНК. У шимпанзе, представляющей собой более древний вид и к тому же гораздо более гетерогенный, число СНП при сравнении двух разных особей достигает 0,4% [15]. Но если различия в СНП не сказываются на здоровье особей, то чем они интересны и важны? Во-первых, изучение СНП имеет большое теоретическое значение. Именно они позволяют сравнивать возрасты популяций и определять пути их миграции. Так, например, в мужской половой хромосоме (Y) выделены 22 фактора СНП, анализ которых у 1007 европейцев позволил определить, что 80% европейских мужчин имеют сходный «СНП-паттерн», т.е. «рисунок». Это говорит о том, что тысячи поколений назад четыре пятых европейских мужчин имели общего предка. Но и практическое значение СНП велико. Возможно, не все знают, что сегодня самые распространённые лекарства эффективны не более чем для четверти населения. Минимальные генетические отличия, обусловленные СНП, определяют эффективность лекарств и их переносимость в каждом конкретном случае. Такие генетические варианты называются гаплотипами.

Например, среди астматиков довольно популярно лекарство албутерол, который взаимодействует с рецептором адреналина и подавляет приступ удушья. Однако из-за разнообразия гаплотипов людей лекарство действует не на всех, а некоторым больным оно вообще противопоказано. Это обусловлено СНП: люди с последовательностью букв в одном из генов ТЦТЦЦ (Т-тимин, Ц-цитозин) не реагируют на албутерол, если же концевой цитозин заменён на гуанин (ТЦТЦГ), то реакция есть, но частичная. Для людей же с тимином вместо концевого цитозина в этом участке – ТЦТЦТ – лекарство токсично. Благодаря расшифровке строения генома человека, медицинская генетика в ближайшее время превратится в медицинскую геномику, поскольку станет возможным установить однозначное соответствие между определенными мутациями в определённых генах и теми или иными патологиями. Уже идентифицированы сотни так называемых больных генов, чья связь с наследственными болезнями чётко доказана.

* * *

Трёхмиллиардный проект был формально запущен в 1990 г. Министерством энергетики США и Национальными институтами здравоохранения, и ожидалось, что он продлится 15 лет. Помимо США, в международный консорциум вошли генетики Китая, Франции, Германии, Японии и Великобритании.

В силу широкой международной кооперации и новых достижений в области геномики (особенно в секвенировании), а также значительных достижений в вычислительной технике, «черновик» генома был закончен в 2000 г., о чём и было объявлено. Продолжение секвенирования привело к объявлению в апреле 2003 г. о почти полном завершении работы, на два года раньше чем планировалось. В мае 2006 г. была пройдена другая веха на пути к завершению проекта, когда в журнале «Нейчер» была опубликована последовательность последней хромосомы – хромосомы 1. Вслед за расшифровкой генома человека (2000 г.) настал черёд других видов: 25 архебактерий, 270 бактерий, 46 эукариот, из них 12 протистов, 19 грибов и 13 животных – нематоды (*Caenorhabditis briggsae*), дрозофилы (*Drosophila melanogaster*), малярийного комара (*Anopheles gambiae*), шелкопряда (*Bombyx mori*), асцидии (*Ciona intestinalis*), рыб (*Tetraodon nigroviridis*, *Danio rerio*, *Takifugu rubripes*), крысы (*Rattus norvegicus*), мыши (*Mus musculus*), домашней собаки (*Canis familiaris*), шимпанзе (*Pan troglodytes*). На сегодняшний день исследовательскими группами ведётся расшифровка ещё 905 прокариотических и 569 эукариотических геномов [18].

Последовательность человеческой ДНК сохраняется в базах данных, доступных любому пользователю через Интернет. Национальный центр биотехнологической информации США (и его партнёрские организации в Европе и Японии) хранят геномные последовательности в базе данных известной как *GenBank*, вместе с последовательностями известных и гипотетических генов и белков. Другие организации, к примеру, Калифорнийский университет в Санта-Крузе и *Ensembl* поддерживают дополнительные данные и аннотации, а также мощные инструменты для визуализации и поиска в этих базах [13; 20]. Были разработаны специальные компьютерные программы для анализа дан-

ных, потому что сами данные без таких программ интерпретировать практически невозможно.

Полностью секвенированы геномы дрожжей, возбудителя малярии, дрозофилы, плоского червя и нескольких сот бактерий. Геномная информация по некоторым из них, полученная частными фирмами, остаётся закрытой в коммерческих интересах, но последовательности нуклеотидов десятков полных геномов бактерий, включая данные по микоплазмам, можно найти в открытых банках данных:

GenBank (<http://www.ncbi.nlm.nih.gov/>);

EMBL (<http://www.embl-heidelberg.de/>)

и *DDBJ* (<http://www.ddbj.nig.ac.jp/>).

Геномика исследует и сопоставляет целые геномы прокариот, геномы ядер и органелл эукариот. Развитие геномики сопряжено с прогрессом в биоинформатике, обеспечивающей корректную обработку и интерпретацию экспериментальных данных [14].

Согласно определению, которое использует Международный проект по расшифровке генома человека, геном расшифрован полностью. График истории расшифровки показывает, что большая часть человеческого генома была закончена в конце 2003 года. Однако ещё остаётся несколько регионов, которые считаются незаконченными:

- Прежде всего, центральные регионы каждой хромосомы, известные как центромеры, которые содержат большое количество повторяющихся последовательностей ДНК; их сложно секвенировать при помощи современных технологий. Центромеры имеют длину миллионы (возможно, десятки миллионов) пар нуклеотидов, и, по большому счёту, остаются несеквенированными.

- Во-вторых, концы хромосом, называемые теломерами, также состоящие из повторяющихся последовательностей, и по этой причине в большинстве из 46 хромосом их расшифровка не завершена. Точно не известно, какая часть последовательности остаётся не расшифрованной до теломер, но как и с центромерами, существующие технологические ограничения препятствуют их секвенированию.

- В-третьих, в геноме каждого индивидуума есть несколько локусов, которые содержат членов мультигенных семейств, которые также сложно расшифровать с помощью основного на сегодняшний день метода фрагментирования ДНК. В частности, эти семейства кодируют белки, важные для иммунной системы.

- Кроме перечисленных регионов ДНК, остаётся ещё несколько брешей, разбросанных по всему геному, некоторые из которых довольно крупные, но есть надежда, что все они будут закрыты в ближайшие годы.

Ф. Коллинз, руководитель государственной программы «Геном человека», сформулировал прогноз геномных исследований на ближайшие 40 лет.

2010 год. Генетическое тестирование, профилактические меры, снижающие риск заболеваний, и генная терапия до 25 наследственных заболеваний. Медсестры начинают выполнять медико-генетические процедуры. Широко доступна преимплантационная диагностика, яростно обсуждаются ограничения в

применении данного метода. В США приняты законы для предотвращения генетической дискриминации и соблюдения конфиденциальности. Не всем доступны практические приложения геномики, особенно в развивающихся странах.

2020 год. На рынке появляются лекарства от диабета, гипертонии и других заболеваний, разработанные на основе геномной информации. Терапия рака, прицельно направленная на свойства раковых клеток. Фармакогеномика становится общепринятым подходом для создания многих лекарств.

Изменение способа диагностики психических заболеваний, появление новых способов их лечения, изменение отношения общества к таким заболеваниям.

Демонстрация безопасности генотерапии на уровне зародышевых клеток при помощи технологии гомологичной рекомбинации.

2030 год. Определение последовательности нуклеотидов всего генома отдельного индивида станет обычной процедурой, стоимость которой менее 1 тыс. долларов.

Каталогизированы гены, участвующие в процессе старения.

Проводятся клинические испытания по увеличению максимальной продолжительности жизни человека.

Лабораторные эксперименты на человеческих клетках заменены экспериментами на компьютерных моделях.

Активизируются массовые движения противников передовых технологий в США и других странах.

2040 год. Все общепринятые меры здравоохранения основаны на геномике.

Определяется предрасположенность к большинству заболеваний (при/до рождения).

Доступна эффективная профилактическая медицина с учётом особенностей индивида. Болезни детектируются на ранних стадиях путем молекулярного мониторинга.

Для большинства заболеваний доступна генная терапия.

Замена лекарств продуктами генов, вырабатываемыми организмом при ответе на терапию.

Средняя продолжительность жизни достигнет 90 лет благодаря социально-экономическим мерам. Проходят серьёзные дебаты о возможности человека контролировать собственную эволюцию.

Неравенство в мире сохраняется, создавая напряженность на международном уровне [18].

Список литературы

1. Большая Советская энциклопедия. Т. 6, 8, 15. М.: Советская энциклопедия, 1975.
2. Киселёв Л.Л. Геном человека и биология XXI века // Вестник РАН, 2000. Т. 70. № 5. С. 412–424.
3. Коллинз Фрэнсис. Доказательство Бога (Аргументы учёного). М., АНФ, 2009.

4. *Adams M.D.* et al. (2000). The Genome Sequence of *Drosophila melanogaster*. *Science* 287: 2185-2195. DOI:10.1126/science.287.5461.2185. PMID 10731132.
5. *Barnhart Benjamin J.* (1989). DOE Human Genome Program ". Human Genome Quarterly 1: 1. Retrieved 2005-02-03.
6. *Cook-Deegan R.* (1989). The Alta Summit, December 1984. *Genomics* 5: 661-663. DOI:10.1016/0888-7543(89)90042-6.
7. International Human Genome Sequencing Consortium (2001). Initial Sequencing and Analysis of the Human Genome. (PDF). *Nature* 409: 860-921. DOI:10.1038/35057062.
8. *Sanger F., Niclén S., Coulson A.R.* DNA Sequencing with Chain-terminating Inhibitors. *Proc. Natl. Acad. Sci. USA.* 1977. Vol. 74. P. 5463-5467.
9. *Venter J.C.* et al (2001). The Sequence of the Human Genome. (PDF). *Science* 291: 1304-1351. DOI:10.1126/science.1058040. PMID 11181995.
10. <http://elementy.ru/trefil/21149?context=21458&discuss=21149>
11. <http://genome.ucsc.edu>
12. <http://humbio.ru/humbio/01122001/kartirovan/0000503f.htm>
13. <http://learnbiology.narod.ru/18.htm>
14. <http://lib.sportedu.ru/press/tpfk/2001N6/P60-63>
15. http://npvr.chat.ru/texts/humans_genom.htm
16. <http://www.chem.msu.su/rus/journals/chemlife/2000/genom.html>
17. <http://www.genome.gov/11006943>
18. <http://www.ensembl.org>
19. <http://www.wisegeek.com/what-is-the-cost-of-human-genome-sequencing.htm>