

ВВЕДЕНИЕ

Если век XIX по-праву вошел в историю мировой цивилизации, как Век Физики, то стремительно завершающемуся веку XX, в котором нам посчастливилось жить, по всей вероятности, уготовано место Века Биологии, а может быть и Века Генетики. Действительно, за неполных 100 лет после вторичного открытия законов Г. Менделя генетика прошла триумфальный путь от натурфилософского понимания законов наследственности и изменчивости, через экспериментальное накопление фактов формальной генетики к молекулярно-биологическому пониманию сущности гена, его структуры и функции. От теоретических построений о гене как абстрактной единице наследственности к пониманию его материальной природы как фрагмента молекулы ДНК, кодирующей аминокислотную структуру белка, до клонирования индивидуальных генов, создания подробных генетических карт человека и животных, идентификации генов, мутации которых сопряжены с тяжелыми наследственными недугами, разработки методов биотехнологии и генной инженерии, позволяющих направленно получать организмы с заданными наследственными признаками, а также проводить направленную коррекцию мутантных генов человека, то есть генотерапию наследственных заболеваний. Молекулярная генетика значительно углубила наши представления о сущности жизни, эволюции живой природы, структурно-функциональных механизмах регуляции индивидуального развития. Благодаря её успехам начато решение глобальных проблем человечества, связанных с охраной его генофонда.

Естественно, что возможность манипуляции с индивидуальными генами человека и животных еще недостаточна для понимания функции всего генома, его организации в целом, взаимодействия его частей в обеспечении всего многообразия механизмов онтогенеза, то есть развития одной клетки до целого организма. Если добавить к этому, что в геноме любого вида записана не только программа индивидуального развития, но закодирована и вся эволюция вида, то есть его филогенез, становится понятным насколько логичной и методически своевременной явилась Международная научная программа "Геном человека". Наряду с аналогичными программами для других видов (лабораторные мыши, нематоды) программа Геном человека, начатая около 10 лет назад, уже к 2 000 году позволит полностью расшифровать первичную структуру ДНК, то есть идентифицировать все гены человека, их регуляторные элементы. Захватывающая Одиссея о наследственности, которой и является эта программа, безмерно расширит наши представления о структуре и функции генома, его эволюции, откроет горизонты столь увлекательного, а, возможно, и не менее опасного направленного воздействия человека на геном растений, животных и, что особенно рискованно, на свой собственный геном. Важно осознать, что это не завтешний день фундаментальной науки, не отдаленные абстракции, но день сегодняшний. Он уже наступил и стал реальным независимо от нас, и, если не быть к нему готовым концептуально и методически, то пройдет мимо нас.

Предлагаемая вашему вниманию книга, действительно, представляет собой введение в молекулярную генетику наследственных болезней и рассчитана на достаточно широкую аудиторию медиков и биологов. Для большинства из уже состоявшихся специалистов в этих областях – это реальная возможность для самообразования, которой, увы, с годами мы так часто пренебрегаем, запутавшись в повседневных заботах. Для студентов биофаков и особенно для студентов-медиков – эта книга вполне

может рассматриваться в качестве учебного пособия по молекулярным основам медицинской генетики. Однако, и для первых и для вторых, по глубокому убеждению авторов, много лет отдавших внедрению достижений молекулярной биологии в медицинскую практику, книга может служить в качестве справочного руководства по молекулярной генетике человека. Действительно, ни одна клиническая дисциплина (за исключением, может быть, службы организации здравоохранения) не мыслима сегодня без знаний и определенных навыков по молекулярной генетике. Ни один биолог, занятый вопросами наследственности, изменчивости, онтогенеза или эволюции независимо от конкретного биологического объекта, не может игнорировать человека, как одного из пока немногих биологических видов с полностью расшифрованной структурой генома. Быстро набирающая силы молекулярная медицина, преподавание азов которой все еще явно недостаточно для будущих врачей, на самом деле представляет собой принципиально новый качественный уровень в понимании вопросов этиологии, патогенеза, а, следовательно, и лечения многих болезней, как наследственной моногенной, так и мультифакториальной природы.

По нашему мнению, не только современный врач и специалист-биолог, но и каждый образованный человек сегодня должен знать о триумфе Международного Научного Сообщества в выполнении программы Геном человека, в результате которой успешно расшифровываются все гены человека, каждый из которых, будучи выделенным из организма и проклонированным может выступать в качестве лечебного препарата для генотерапии. О том, что уже сегодня идентифицировано на генетических картах более 5 000 структурных генов и свыше 60 000 пока неизвестных смысловых и анонимных ДНК последовательностей. О том, что всего за 5 лет после первых успешных попыток введения чужеродных маркерных генов в клетки человека *in vivo*, число уже одобренных для клинических испытаний программ по генной терапии наследственных заболеваний достигло более 100! Эти итоги представляются особенно впечатляющими если учесть, что согласно данным Всемирной Организации Здравоохранения около 2,4% всех новорожденных на земном шаре страдает теми или иными наследственными нарушениями; около 40% ранней младенческой смертности и инвалидности с детства обусловлены наследственной патологией. Нельзя не упомянуть о реальных достижениях молекулярной генетики в расшифровке наследственных факторов таких бичей человечества как ишемия сердца, атеросклероз, диабет, онкологические и инфекционные заболевания. Адекватно воспринимать происходящую на наших глазах революцию в биологии и в медицине, уметь воспользоваться её заманчивыми плодами и избежать опасных для человечества соблазнов - вот что необходимо сегодня и врачам, и биологам, и представителям других смежных специальностей, и просто образованному человеку.

Именно эта цель, эта сверхзадача, поставлена перед данной монографией, восполняющей, по мнению авторов, наметившийся в отечественной научной литературе пробел в области молекулярных аспектов медицинской генетики и генетики человека. Отдельные обзоры, монографии (Шишкин, Калинин, 1993), переводная литература по молекулярной биологии и даже обстоятельные сводки, подводящие ежегодные итоги работ по программе "Геном человека" достаточно фрагментарны и касаются лишь отдельных аспектов проблем генодиагностики и генотерапии. Рассчитаны преимущественно на специалистов по молекулярной биологии. Задача данной монографии не только осветить современное положение дел в молекулярной диагностике и лечении наследственных болезней методами генной терапии, но,

главным образом, подготовить читателей, прежде всего врачей и биологов, к пониманию и восприятию этой методически и концептуально достаточно сложной области генетики.

Для достижения поставленной цели нам представлялось логичным начать изложение материала с описания структуры и современных методов анализа ДНК, с общих представлений о её клонировании, секвенировании, геномных библиотеках (Глава I). Глава II полностью посвящена структуре генома человека, новой трактовке понятия "ген", генным семействам, вариабильным структурам генома. Генетические карты, принципы их построения, функциональное и позиционное картирование, молекулярные маркеры, современные достижения в разработке физических и хромосомных карт человека и в картировании генов, ответственных за наследственные заболевания рассмотрены в Главе III. Глава IV целиком посвящена описанию молекулярных методов детекции как уже известных мутационных изменений в структурных генах, так и методов сканирования предполагаемых мутаций определенных генов. Описание прямых методов идентификации мутаций дополнены косвенными методами, основанными на молекулярном маркировании мутантных генов. Все эти методы, как прямые, так и косвенные, составляют основу молекулярной диагностики моногенных наследственных болезней, широко используются в генетике человека при построении генетических карт, исследовании проблем филогенеза, в популяционной генетике и в геномной дактилоскопии, то есть для идентификации личности. Подробному анализу внутренних (эндогенных) факторов мутагенеза, а также принципам популяционного анализа мутаций посвящена Глава V. Основные подходы, используемые при изучении экспрессии генов в модельных бесклеточных системах, на уровне отдельных клеток и целых организмов приведены в Главе VI. Принципы молекулярной диагностики наследственных болезней и, в частности, пренатальной диагностики, а также особенности выявления гетерозиготного носительства в семьях высокого риска, изложены в Главе VII. Небольшая по размеру, но важная также и для понимания принципов генной терапии Глава VIII касается искусственного создания генетических моделей наследственных заболеваний, в частности на базе трансгенных животных. Описаны используемые при этом методы направленного переноса чужеродных генов в эукариотические системы. В Главе IX изложены основы генотерапии наследственных заболеваний, рассмотрены методы доставки чужеродной ДНК в клетки человека *in vitro* и *in vivo*, преимущества и недостатки существующих векторных систем (физических, химических и биологических), их конструирование, перспективы создания "идеальных" векторных систем. Кратко рассмотрены итоги уже проведенных испытаний по генотерапии тех заболеваний, для которых Программы клинических испытаний уже одобрены или находятся на стадии эксперимента. В заключительной главе (Глава X) мы посчитали целесообразным подвести некоторые итоги и более подробно рассмотреть молекулярную диагностику трех групп наследственных заболеваний: (1) достаточно полно изученную группу лизосомных болезней накопления; (2) болезни экспансии (преимущественно нейродегенеративные заболевания), вызываемые совершенно новым ранее неизвестным типом так называемых "динамических" мутаций и (3) наиболее частые, социально значимые наследственные заболевания, по пренатальной диагностике которых молекулярными методами уже накоплен достаточно большой опыт в нашей лаборатории и в других медико-генетических центрах России.

Итак, предлагаемая монография рассчитана на достаточно широкий круг читателей, но, прежде всего, на медиков и биоло-

гов, а также специалистов смежных профессий. Нам хотелось бы думать, что книга будет особенно полезной для студентов медицинских институтов и академий, врачей курсов повышения квалификации, сотрудников медико-генетических консультаций и центров, а также для студентов-биологов, многие из которых, как показывает наш опыт, пополняют ряды специализированных диагностических лабораторий, медико-генетических центров и институтов.

ГЛАВА III

ГЕНЕТИЧЕСКИЕ КАРТЫ, ПОЗИЦИОННОЕ КЛОНИРОВАНИЕ.

Раздел 3.1 Классификация генетических карт, оценка сцепления.

Генетические карты определяют хромосомную принадлежность и взаимное расположение различных компонентов генома относительно друг друга. Возможность построения таких карт обусловлена двумя фундаментальными характеристиками генома: линейным характером локализации генов в хромосомах (это определяется линейностью молекулы ДНК) и относительной стабильностью расположения облигатных элементов генома в пределах вида. При построении генетических карт используют разные подходы. В первую очередь, к ним относятся анализ генетического сцепления на основе определения частот мейотической рекомбинации в информативных семьях и изучение особенностей наследования признаков, сцепленных с маркерными хромосомными перестройками. Во-вторых, исследование экспрессии генов или поиск специфических последовательностей ДНК в клеточных гибридах, содержащих лишь часть генома человека - одну или несколько хромосом или их фрагменты. В ряде случаев с этой целью используют механический сортинг целых хромосом и даже их относительно небольших участков. Эти приемы позволяют привязать картируемый ген к определенной хромосоме и даже к определенному фрагменту хромосомы. С помощью комплекса весьма тонких методов хромосомного анализа, прежде всего, методов гибридизации *in situ* (см. Главу I) удается картировать отдельные гены на хромосомах человека часто с точностью до одного бэнда. И, наконец, методами молекулярного анализа осуществляют физическое картирование последовательностей ДНК, локализованных в специфических участках хромосом. Затем проводят идентификацию в этих последовательностях транскрибируемых областей, то есть генов, с последующей изоляцией и клонированием соответствующих им полноразмерных молекул кДНК. Каждый из рассмотренных этапов анализа структуры генома завершается построением карт генов, различающихся по единицам измерения расстояний между отдельными элементами этих карт, масштабам, по насыщенности или степени детализации на различных участках генома. Соответственно различают карты сцепления, генетические карты, цитогенетические карты индивидуальных хромосом и физические или молекулярные карты определенных участков ДНК. Для полной молекулярной идентификации отдельных элементов генома, то есть определения их границ, структуры и нуклеотидной последовательности, необходимо совмещение всех типов карт в местах локализации этих элементов.

Первым шагом на пути построения генетических карт явля-

ется формирование групп сцепления генов, контролирующих различные наследственные признаки, и исследование их взаимного расположения в этих группах. На следующем этапе определяют соответствие между генетическими группами сцепления и цитогенетически идентифицируемыми хромосомами или их фрагментами. Цитогенетическую идентификацию хромосом проводят с использованием методов дифференциальной окраски (см. раздел 3.2). По мере появления все большего числа локализованных признаков эффективность построения генетических карт значительно возрастает, так как увеличивается число маркированных участков хромосом и, таким образом, появляется возможность комбинированного использования различных экспериментальных подходов для более подробного исследования этих участков.

Принципиальная схема картирования неизвестных генов, представленная на Рис.3.1, включает следующие этапы. 1. Выяснение группы сцепления; 2. Поиск ближайших фланкирующих маркеров; 3. Определение физической области (ДНК-последовательности), включающей искомый ген; 4. Клонирование набора фрагментов ДНК, перекрывающих исследуемую область; 5. Выделение из этого набора клонов, содержащих транскрибируемые ДНК-последовательности, предположительно соответствующие гену или его фрагменту; 6. Анализ специфических мРНК и клонирование кДНК-последовательности; 7. Секвенирование и идентификация самого гена (Wicking, Williamson, 1991). Рассмотрим подробнее эту схему.

Построение карт сцепления основано на изучении процессов расхождения и рекомбинации гомологичных хромосом в мейозе. Генетические признаки, локализованные в разных хромосомах, не сцеплены друг с другом, то есть передаются от родителей детям независимо, и частота их рекомбинации (Q) составляет 0.5. Это обусловлено случайным характером расхождения гомологичных хромосом в мейозе во время редукционного деления. Гены, локализованные в одной хромосоме, рекомбинируют за счет кроссинговера, то есть за счет обмена участками гомологичных хромосом в процессе их спаривания в мейозе (Рис.3.2). При этом порядок генов не нарушается, но в потомстве могут появиться новые комбинации родительских аллелей. Вероятность кроссинговера между двумя генами зависит от расстояния между ними. Чем ближе гены расположены друг к другу, то есть чем больше они сцеплены, тем эта вероятность меньше.

Оценку сцепления между генами проводят на основании статистического анализа сегрегации признаков в семьях с разветвленными родословными. Чаще всего при этом используют метод максимального правдоподобия (Kao, 1983), то есть подсчитывают десятичный логарифм шансов - lod (\log of the odds), где шансы (odds) выражаются как отношение вероятности наблюдаемой родословной при условии, что два гена сцеплены ($0 < Q < 0.5$), к той же вероятности при отсутствии сцепления ($Q = 0.5$). Если значение lod $> +3$, гены локализованы в одной хромосоме, причем максимально правдоподобная оценка соответствует максимальному значению lod. При значениях lod < -2 гены не сцеплены, то есть локализованы в разных хромосомах или на разных концах одной хромосомы. Статистическую обработку родословных обычно проводят с помощью компьютерных программ, наиболее известные из которых программы LIPED, CRIMAP и LINKAGE (Ott, 1985; Ott, 1991; Terwilliger, Ott, 1994). На генетических картах сцепления расстояние между генами определяется в сантиморганах (сМ). 1 сМ соответствует 1% рекомбинации. Общая длина генома че-

ловека в этих единицах составляет около 3300 сМ. Сопоставляя эту величину с размером гаплоидного набора молекул ДНК, можно заключить, что 1 сМ приблизительно эквивалентен 1 миллиону пар нуклеотидов. Такие расчеты, однако, весьма приблизительны, так как частоты рекомбинации, а значит и реальная длина одного сМ, могут сильно варьировать в различных частях генома. Существуют, так называемые, горячие точки рекомбинации, также как и районы генома, где рекомбинация подавлена (центромерные и теломерные участки хромосом, блоки конститутивного гетерохроматина и др.). Известно также, что частота рекомбинации у мужчин меньше, чем у женщин, так что общая длина мужского генома, измеренная в единицах рекомбинации, составляет лишь 3000 сМ. Таким образом, генетическое расстояние может дать лишь весьма ориентировочную информацию о физическом (реальном) расстоянии, выражаемом в парах нуклеотидов.

Раздел 3.2 Соматическая гибридизация, цитогенетический анализ, картирование анонимных последовательностей ДНК.

До начала 70-ых годов построение генетических карт человека продвигалось очень медленными темпами. Небольшой размер семей, длительный период одного поколения, ограниченное число информативных родословных и отсутствие методов эффективного цитогенетического анализа всех пар хромосом затрудняло целенаправленное картирование генов человека. Достаточно сказать, что 1-й ген человека - ген цветной слепоты был картирован на X-хромосоме в 1911г., а 1-й аутосомный ген - только в 1968г. К 1973г. на хромосомах человека было картировано всего 64 гена, а к 1994 на генетических картах человека было локализовано уже свыше 60 000 маркерных ДНК последовательностей, в том числе около 5 000 структурных генов - Рис.3.3. Столь стремительный прогресс в картировании генов человека связан с появлением новых технологий в цитогенетике, в клеточных культурах и особенно в молекулярной генетике.

Техника соматической гибридизации, то есть возможность экспериментального конструирования способных к размножению межвидовых клеточных гибридов, явилась одним из наиболее мощных инструментов для нахождения связей между группами сцепления и цитогенетически идентифицируемыми хромосомами и даже их отдельными сегментами. Гибридные клоны получают путем искусственного слияния культивируемых соматических клеток разных видов, в частности, клеток человека и различных грызунов - китайского хомячка, мыши, крысы (Шея, 1985). Культивирование таких соматических гибридов, как оказалось, сопровождается утратой хромосом человека. Так были получены панели гибридных клеточных клонов, содержащих всего одну или несколько хромосом человека и полный набор хромосом другого вида. Обнаружение человеческих белков, специфических мРНК или последовательностей ДНК в таких клонах позволяет однозначно определить хромосомную принадлежность соответствующих генов. Таким способом удалось локализовать более 250 аутосомных генов человека (Као, 1983).

Дальнейший прогресс в области генетического картирования в значительной степени ассоциируется с деятельностью многих научно-исследовательских центров и лабораторий по созданию банков клеточных культур, представляющих наиболее интересные и обширные родословные. Так, в Центре по Изучению Полиморфизма Человека (СЕРН) (Париж, Франция) была создана

уникальная коллекция перевиваемых клеточных культур от всех членов семей, многоступенчатые родословные которых насчитывают многие десятки и даже сотни индивидуумов (СЕРН-семьи) (Todd, 1992; Weissenbach et al, 1992). Перевиваемые линии клеток получали из первичных культур, трансформированных вирусом Эпштейн-Барра, после чего такие лимфобластные клетки становятся "бессмертными", то есть могут неограниченно долго поддерживаться в условиях культивирования. СЕРН-семьи представляют собой идеальные системы для генетического анализа наследственных признаков. В результате исследования этих клеточных линий определены генотипы членов СЕРН-семей одновременно по тысячам полиморфных локусов и построены соответствующие генетические карты. Кроме того, совместными усилиями многих лабораторий мира были созданы Банки Клеточных Культур, в которых поддерживаются коллекции лимфобластных линий клеток, полученных от членов наиболее информативных семей, в которых наблюдается сегрегация по различным наследственным признакам, в том числе по моногенным заболеваниям. Материал этих линий в виде клеточных клонов или образцов ДНК используется, в частности, для анализа сцепления сегрегирующих генов с известными генетическими маркерами или с вновь описанными полиморфными локусами. Таким образом, во многих случаях при картировании генов человека удается преодолеть ограничения, связанные с недостатком обширных информативных родословных.

Наконец, в начале 70-х годов появилась реальная возможность точной идентификации не только всех хромосом в кариотипе человека, но и их отдельных сегментов. Это связано с появлением методов дифференциального окрашивания препаратов метафазных хромосом, которые, по сути, произвели революцию в цитогенетике и хромосомологии. Для получения дифференциальной окраски хромосомные препараты окрашивают некоторыми флюорохромами, либо, после соответствующей протеолитической обработки или нагревания, - красителем Гимза. При этом на хромосомах выявляется характерная поперечная исчерченность, так называемые бэнды, расположение которых специфично для каждой хромосомы. На метафазных хромосомах малой степени спирализации идентифицируется около 750 таких полос, на прометафазных хромосомах 2 500 - 3 000. Следовательно, величина небольших бэндов на прометафазных хромосомах примерно соответствует 1 см на цитогенетических картах или 1 миллиону пар оснований на физических картах. Такие ориентировочные расчеты, как уже упоминалось, оказываются полезными при сопоставлении масштабов различных генетических карт.

Согласно официально утвержденной номенклатуре (ISCN, 1971) каждая хромосома человека после дифференциальной окраски может быть разделена на сегменты, нумерация которых начинается от центромерного района вверх (короткое плечо - p), либо вниз (длинное плечо - q) (Рис.3.4). Полосы в каждом сегменте также пронумерованы в аналогичном порядке. Крупные полосы зачастую подразделяются на несколько частей, соответствующих более мелким бэндам, выявляемым только при окраске малоспирализованных прометафазных хромосом. Запись положения гена на цитогенетической карте включает номер хромосомы, плечо, а также номер сегмента, бэнда и его субъединицы. Например, запись 7q21.1 означает, что ген локализован в субъединице 1, 1-го бэнда 2-го сегмента длинного плеча хромосомы 7.

Такая подробная запись особенно удобна при использовании для цитогенетического картирования метода гибридизации *in situ* (см.Главу I), позволяющего локализовать ген с точ-

ностью до одного бэнда и даже его субъединицы. Основу данного метода, как уже указывалось составляет гибридизация геномной ДНК целых хромосом на разных стадиях их спирализации с предварительно мечеными специфическими ДНК-зондами. Источником последних могут служить геномные последовательности ДНК, сцепленные с картируемым геном, либо кДНК-овые последовательности. В настоящее время из тканеспецифических библиотек генов выделено около 20 000 анонимных последовательностей кДНК, представляющих более 10 000 различных генов человека (Sikela, Auffray, 1993). Эти кДНК составляют примерно 10-15% всех кодирующих последовательностей генома. Методом гибридизации *in situ* уже идентифицировано более 2000 генов, для многих из которых пока не найдено сцепление с полиморфными маркерами (Poduslo et al., 1991). Такие кодирующие последовательности присутствуют на карте генов человека, но их пока нет на картах сцепления.

Раздел 3.3 Генетические индексные маркеры.

Как мы уже отмечали раньше, успешная локализация неизвестного наследственного признака (гена) в значительной степени определяется присутствием на карте фланкирующих маркеров, находящихся на относительно небольшом расстоянии по обе стороны от гена. В качестве генетических маркеров специфических участков хромосом могут быть использованы любые локализованные в этих участках элементы генома с высоким уровнем легко идентифицируемой популяционной изменчивости или полиморфизма. Отбор сцепленных с картируемым признаком генетических маркеров производят по результатам совместного анализа их сегрегации в информативных семьях.

Значительные успехи в геномном картировании были связаны с использованием в качестве генетических маркеров широко распространенной во многих популяциях изменчивости по изоферментному спектру различных белков. Оказалось, что многие ферменты у разных индивидуумов, в разных тканях и на разных стадиях онтогенеза могут находиться в различных изоформах. Такие варианты одного и того же белка обычно не отличаются по специфической активности, но имеют измененную электрофоретическую подвижность. Популяционный анализ изоферментов обнаружил существование полиморфизма для очень многих белковых систем, контролируемых разными генами, локализованными во многих хромосомах. Таким образом, для этих хромосом были найдены генетические маркеры, с помощью которых были идентифицированы соответствующие группы сцепления.

Совершенствование молекулярных методов анализа специфических последовательностей ДНК привело к обнаружению большого числа высокоизменчивых участков генома - полиморфных сайтов рестрикции, гипервариабельных мини- и микросателлитных последовательностей (см.Главу II) Эта изменчивость также была использована для маркировки участков хромосом с целью установления более точного взаиморасположения локусов. Вскоре после обнаружения полиморфных сайтов рестрикции были опубликованы теоретические расчеты, согласно которым от 10 до 20 таких локусов, расположенных равномерно на каждой хромосоме на расстоянии около 20 сМ друг от друга, достаточно для определения хромосомной принадлежности генов, ответственных за любой тип наследственной изменчивости (Botstein et al,1980). По этим оценкам около 200 таких индексных маркеров позволят построить карты сцепления для всех известных генов на основании анализа сегрегации соответствующих признаков в семьях с большим числом мутантов. Для опре-

деления порядка расположения генов на хромосомах и оценки генетических расстояний между ними достаточно около 400 равномерно распределенных полиморфных маркеров.

Идентификация в геноме человека большого числа полиморфных сайтов рестрикции и разработка простых методов анализа индивидуальной изменчивости по этим локусам существенно повысили возможности локализации неизвестных признаков на геномных картах. Уже к 1989г. были картированы более 2000 клонированных последовательностей ДНК, обнаруживающих полиморфизм по длине рестрикционных фрагментов (ПДФ) в различных популяциях (Kidd et al., 1989). Более 1000 из них типировано в СЕРН-коллекциях родословных. В значительной степени это анонимные последовательности, связь которых со специфическими генами не установлена. Их наименования чаще всего соответствуют названию отобранного из библиотеки генов клонна. В дальнейшем была разработана стандартная генетическая номенклатура для обозначения используемых в качестве маркеров сегментов ДНК с неизвестной функцией. Первая буква D, что значит ДНК, затем номер хромосомы, далее S для уникальных и Z для повторяющихся последовательностей и в конце номер, идентифицирующий данный зонд в определенном районе ДНК.

Применение полиморфных сайтов рестрикции в качестве генетических маркеров имеет два ограничения: сравнительно низкая информативность (частота гетерозигот не может превышать 50% - см. Главы II, Y) и неравномерное распределение ПДФ-сайтов по хромосомам. Этих недостатков практически лишены гипервариабельные STR-сайты (ди-, три- и тетра-нуклеотидные повторы). Использование микро- и минисателлитных ДНК последовательностей в качестве индексных генетических маркеров открыло новую эру в построении карт сцепления генома человека. В настоящее время работа по идентификации высокополиморфных маркеров, перекрывающих весь геном и равномерно распределенных по хромосомам практически завершена (Weissenbach et al., 1992; Reed et al., 1994). Эта система построена на базе динуклеотидных (C-A)_n повторов.

(C-A)_n*(G-T)_n представляют собой наиболее частый класс простых повторов, обнаруженных в геноме человека (за исключением An* Tn мультимеров). Такие повторы присутствуют примерно в 1% колоний из геномных библиотек, сконструированных на базе фрагментов длиной 300 - 500 п.о., которые образуются после переваривания геномной ДНК эндонуклеазой AluI. Более 90% из них оказываются полиморфными по числу копий в кластере, причем в 70% локусов присутствует более трех аллелей. В 1992г. группе французских ученых под руководством Жана Вайссенбаха удалось разработать систему идентификации с помощью ПЦР индивидуальной изменчивости в местах локализации (C-A)_n повторов и на этой основе создать геномную карту из 814 высокополиморфных индексных маркеров со средним расстоянием между ними около 5 см, получившую название Genethon-коллекции микросателлитных маркеров (Weissenbach et al., 1992). Для этого были просеквенированы более 12 000 фрагментов ДНК, выделенных из геномной AluI-библиотеки путем ее скрининга poly(dC-dA)*poly(dG-dT) ДНК-зондом. Праймеры для амплификации подбирали из последовательностей ДНК, окружающих повторы, используя для этого компьютерные программы. Для дальнейшего анализа было отобрано около 3 000 (C-A)_n-сайтов и проведена специфическая амплификация этих участков у четырех неродственных СЕРН-индивидуумов. На следующем этапе была определена хромосомная принадлежность полутысячи наиболее информативных маркеров путем их идентификации в панели из 18 соматических гибридных клонов, со-

держат разные наборы хромосом человека. Детальная карта сцепления индексных маркеров построена по результатам их генотипирования в коллекции 8 самых больших CEPH-родословных. Предложенная система маркеров перекрывает все хромосомы, а суммарное расстояние между ними соответствует примерно 90% всего генома человека.

К 1994г. число индексных STR-маркеров было увеличено до 2 066, а средний интервал между соседними локусами уменьшен до 2,9 cM (Guarun et al., 1994). В последнее время перспективы широкомасштабного картирования всего генома человека стали еще более значительны. Группой английских авторов под руководством Дж.Тодда была разработана система автоматического скринирования 254 динуклеотидных маркеров, перекрывающих весь геном человека со средним расстоянием между соседними маркерами около 13 cM. 80% этих повторов было отобрано из Genethon-коллекции маркеров, остальные - из других источников (Genome Data Base, Baltimore). Амплификацию всех 254 полиморфных сайтов проводили в 39 мультиплексных ПЦР (МПЦР), причем используемые для этих целей олигопраймеры были мечены четырьмя типами флюорохромов, так что аллели даже одинакового молекулярного веса можно было различить на электрофореграмме по цвету. В каждой из 39 МПЦР скринировали 7 - 9 STR-сайтов какой-то определенной хромосомы. Такие МПЦР-наборы были разработаны для всех 22 аутосом и для X-хромосомы. Регистрация аллелей всех STR проводилась автоматическим сканнером с использованием компьютерной программы Genotyper. Только с помощью одного автоматического сканнера удается проанализировать по этой схеме более 2.5 тысяч генотипов в день! (Reed et al., 1994). Такая система уже сегодня открывает самые широкие возможности не только для генетического картирования и создания подробных карт сцепления практически любых моногенных заболеваний, но, что особенно существенно, она делает реальной разработку стратегии картирования генов, мутации которых предрасполагают к мультифакториальным заболеваниям, таким как диабет, гипертония, инфаркт миокарда, психозы и многое другое.

Раздел 3.4 Хромосом-специфические библиотеки генов, пульсирующий гель-электрофорез.

Используя столь обширную систему молекулярных маркеров и проводя анализ сцепления на коллекциях клеточных культур или на материале информативных родословных можно довольно быстро привязать любой признак, особенно моногенный, не только к определенной хромосоме, но даже к одному бэнду, определить ближайшие фланкирующие маркеры и перейти непосредственно к позиционному клонированию с целью выделения и идентификации самого гена. В этой связи важное значение в картировании генов принадлежит молекулярно-цитогенетическим подходам, являющимся принципиально важным звеном для успешного совмещения карт сцепления и физических карт целых хромосом и их фрагментов.

Точность цитогенетического картирования определяется степенью спирализации хромосом, характером использованной метки и разрешающей способностью микроскопического оборудования. При картировании на стандартных метафазных хромосомах и использовании радиоактивно меченых зондов точность картирования ограничивается одним крупным бэндом или даже сегментом хромосомы и составляет около 5-10 миллионов п.о. При использовании биотиновой метки на прометафазных хромосомах точность картирования возрастает в среднем в 5-10 раз (до 1

миллиона п.о.), а при работе со специально приготовленными и растянутыми интерфазными хромосомами может достигать до 50 тысяч п.о. (Boehringer Mannheim Manual, 1992). Тем не менее, даже при такой разрешающей способности цитогенетическое картирование дает лишь весьма ориентировочные результаты и обычно рассматривается как 1-й этап физического картирования.

Значительно более точные результаты достигаются на 2-м этапе - этапе физического (рестрикционного) картирования. Среднее расстояние между стандартными сайтами узнавания на рестрикционных картах колеблется в пределах от 10 до 20 кб. Из-за расхождений почти на два порядка масштабов цитогенетического и молекулярного картирования прямое сопоставление этих типов физических карт практически невозможно.

Одним из способов преодоления этих трудностей является конструирование хромосом-специфических библиотек генов. Как уже упоминалось (см. Глава I, 1.5) для приготовления таких библиотек используют наборы клеточных линий соматических гибридов с отдельными хромосомами человека либо хромосомы, механически отобранные путем проточной цитометрии. Для некоторых видов молекулярного клонирования удобнее оказались библиотеки генов, построенные из субхромосомальных фрагментов. Получение таких фрагментов достигается путем целенаправленного конструирования соматических гибридов, содержащих лишь часть какой-либо хромосомы человека. Субхромосомные клоны могут быть получены и с помощью микроманипуляций, при которых механически, под контролем микроманипулятора может быть вырезан практически любой видимый фрагмент каждой хромосомы. Разработаны также молекулярные методы выделения из генома и идентификации крупных фрагментов ДНК, приближающихся по размерам к единичным хромосомным бэндам. Это стало возможным после обнаружения редкощеплящих рестриктаз, разрезающих ДНК на фрагменты длиной от сотен тысяч до миллиона пар нуклеотидов (Estivill, Williamson, 1987).

Другим важным шагом на пути клонирования и анализа больших субхромосомальных фрагментов ДНК явилась разработка методов их разделения путем гель-электрофореза в пульсирующем поле (Barlow, Lehrach, 1987; Smith, Cantor, 1986; Smith et al., 1987). В соответствии со стандартными методами электрофореза под действием однонаправленного постоянного поля в агарозном или в полиакриламидном гелях удается разделять фрагменты ДНК размером не более 3 - 5 десятков килобаз. Продвижение больших фрагментов ДНК в геле при пульсирующем изменении направления электрического поля происходит, по-видимому, за счет конформационных изменений, обусловленных скручиванием и раскручиванием молекул ДНК в момент переключения направления поля. При этом более короткие молекулы легче адаптируются к изменению условий и потому движутся в геле быстрее. Существуют различные варианты пульсирующего гель-электрофореза, главным образом, связанные с геометрическим расположением направлений полей - ортогональный, гексогональный, инверсионный. При использовании любого из этих вариантов могут быть разделены молекулы ДНК размером от 50 кб до более, чем 9 миллионов п.о. Эффективность разделения фрагментов ДНК зависит не только от их размеров, но и от условий проведения электрофореза (напряжение, температура буфера, концентрация агарозы, время одного импульса). В качестве маркеров для определения величины больших молекул ДНК используют целые хромосомы дрожжей известной молекулярной массы. В дальнейшем отбор крупных фрагментов ДНК, несущих специфические последовательности, также может быть осуществлен путем блот-гибридизации с ДНК-зондами. Разделенные и

идентифицированные фрагменты ДНК могут быть элюированы из геля и использованы для рестрикционного картирования, построения библиотек генов и для молекулярного клонирования с целью идентификации и изоляции генных последовательностей. В последнее время для изоляции крупных субхромосомальных сегментов ДНК широко используется метод клонирования в искусственных дрожжевых минихромосомах - YAC, и построения библиотек генов на основе YAC-векторов.

Раздел 3.5 Позиционное клонирование, прогулка и прыжки по хромосоме, идентификация и изоляция генов.

Мы уже упоминали, что средние размеры гена составляют около 10-30 кб, варьируя в широких пределах (см.Глава II.2.4). Единицы рекомбинации, размеры цитогенетических бэндов и субхромосомных фрагментов ДНК измеряются миллионами пар нуклеотидов, также как и размеры фрагментов ДНК, выделяемых с помощью обработки геномной ДНК редкощепящими эндонуклеазами, пульсирующего электрофореза и клонирования в дрожжевых минихромосомах. Переход от этих крупных фрагментов к последовательностям ДНК, сопоставимым с размерами гена, осуществляют с помощью молекулярного клонирования, то есть получения набора фаговых или космидных клонов, содержащих относительно небольшие последовательности, насыщающие или полностью перекрывающие крупный сегмент ДНК, предположительно содержащий идентифицируемый ген (Рис.3.5). Затем проводят упорядочивание клонов в соответствии с взаимным расположением инсертированных в них фрагментов ДНК, осуществляя одномерный молекулярный анализ этих фрагментов с целью идентификации регуляторных или кодирующих областей генов. Позднее мы подробнее остановимся на тех критериях, с помощью которых можно различить транскрибируемые и нетранскрибируемые участки генома. Для молекулярного клонирования используют различные подходы (Iannuzzi, Collins, 1990). Прежде всего, это насыщающее клонирование, то есть изоляция из хромосом-специфических библиотек нескольких сотен клонов с целью картирования различными методами инсертированных в них фрагментов ДНК и идентификации клонов с последовательностями, локализованными в заданном районе. Значительно чаще используется тактика скринирования фаговых, космидных и YAC библиотек, сконструированных из субхромосомальных сегментов ДНК, предварительно отобранных на основании сцепления с различными ДНК-маркерами. При этом методы выделения субхромосомальных сегментов ДНК могут быть самыми различными. Дальнейший поиск в библиотеках генов клонов, содержащих транскрибируемые последовательности ДНК, осуществляют достаточно трудоемкими методами, получившими название "прогулки" и "прыжков" по хромосоме.

"Прогулка" по хромосоме или скользящее зондирование (Рис.3.6) заключается в последовательном отборе клонов, содержащих частично перекрывающиеся фрагменты ДНК из определенного района генома (Rommens et al., 1989). На первом этапе проводят скрининг библиотеки с помощью маркерной ДНК, сцепленной с геном. После нахождения положительных клонов последние сами служат зондами для изоляции других клонов, содержащих перекрывающиеся последовательности ДНК. Таким образом, каждый раз отобранный фрагмент используется в качестве скринирующего ДНК-зонда для последующего поиска. В результате получают набор клонированных фрагментов ДНК, полностью перекрывающих область поиска гена. Группа подобных клонов носит название "контигов". С помощью физического кар-

тирования инсертированной ДНК в разных клонах удается точно установить степень перекрытия между соседними фрагментами и соответственно упорядочить положение клонов в "контигах". При скринировании космидных библиотек с выявлением каждого нового клона к участку ДНК, полностью перекрытому отобранными зондами, в среднем добавляется около 20 кб. Таким способом, однако, редко удается пройти более 200 - 300 кб в одном направлении из-за наличия в геноме повторяющихся и трудно клонируемых последовательностей ДНК.

Для преодоления этих ограничений и ускорения процесса поиска генных последовательностей американским исследователем Фрэнком Коллинзом, ныне президентом Программы Геном Человека, был разработан метод "прыжков" по хромосоме. Этот метод позволяет изолировать фрагменты ДНК, отстоящие в геноме друг от друга на сотни тысяч нуклеотидов (длина прыжка), не изолируя при этом все промежуточные последовательности ДНК (Collins, Weissman, 1984). Как видно на представленной схеме (Рис.3.7), прыжки начинаются со стартового зонда, то есть с последовательности, гибридизующейся со сцепленным с геном ДНК-маркером. Предварительно геномная ДНК переваривается редкощепящей рестриктазой, в результате чего образуются большие фрагменты ДНК, соответствующие по длине одному прыжку. Затем, эти фрагменты переводятся в кольцевую форму за счет искусственного присоединения к их концам небольшого маркерного гена. При этом концы рестрикционных фрагментов сближаются. Кольцевые молекулы ДНК разрезают среднещепящими рестриктазами и из пула относительно небольших фрагментов ДНК отбирают те, которые содержат маркерный ген, а, следовательно, и окружающие его концевые участки исходных крупных фрагментов. Отобранные последовательности клонируют в фаговых или космидных векторах, получая библиотеку генов концевых участков. Затем в этой библиотеке проводят скрининг клонов, содержащих стартовый зонд. Только в этих клонах комплементарные зонду последовательности соединены маркерным геном с последовательностями ДНК, отстоящими от стартового участка поиска на длину прыжка. При необходимости промежуточные сегменты ДНК также могут быть клонированы с использованием метода скользящего зондирования.

Остановимся теперь на тех критериях, по которым можно отличить сегменты ДНК, являющиеся частями генов, от любых других последовательностей (Рис.3.7.) (Lindsay, Bird, 1987; Rommens et al., 1989; Wicking, Williamson, 1991; Collins, 1992). Условно эти критерии могут быть разделены на три группы. В первой группе исследуют структурные особенности генных последовательностей. Вторая группа критериев основана на поиске функциональных участков генов. В третьем случае анализируют характер нуклеотидных последовательностей тестируемых фрагментов ДНК. Диагностику структурных участков генов осуществляют путем гибридизации с ДНК-зондами или прямым скринированием кДНК-овых библиотек. Функциональная диагностика генов включает улавливание экзонов (exon trapping), промоторных участков, поли-А сигнальных последовательностей, а также перенос генов в иные конструкции и идентификацию в них соответствующих транскриптов. И, наконец, поиск генов может быть осуществлен путем прямого секвенирования крупных фрагментов ДНК с последующим компьютерным анализом нуклеотидной последовательности и сопоставлением её с присутствующими в базах данных идентифицированными генами других видов живых существ.

Как уже отмечено ранее, кодирующие области генов, представленные в геноме уникальными последовательностями,

достаточно консервативны в процессе эволюции. Существует высокий процент гомологии в структуре ДНК между одинаковыми генами у разных видов млекопитающих. На этом факте основан, так называемый зоо-блот – скрининг клонированных последовательностей, не содержащих повторов, но дающих перекрестную гибридизацию с геномной ДНК, выделенной из разных видов животных – приматов, сельскохозяйственных животных, грызунов, птиц, рептилий. Клоны, содержащие консервативные последовательности, подвергают дальнейшему анализу на присутствие в инсертированных фрагментах ДНК CpG островков, часто маркирующих 5'-фланкирующие области генов позвоночных, особенно генов домашнего хозяйства (см.Главу II,2.4), и исследуют наличие открытых рамок считывания -ORF (open reading frames). Дальнейший поиск генов в более узком интервале может быть осуществлен с помощью компьютерного анализа соответствующей нуклеотидной последовательности ДНК. Кроме того, все клонированные ДНК из этого интервала могут быть сразу использованы для анализа РНК-транскриптов (Iannuzzi, Collins, 1990).

Важным доказательством принадлежности клонированной ДНК гену является идентификация гомологичных РНК транскриптов в тканях, где можно предполагать экспрессию этого гена. С этой целью проводят гибридизацию уже отобранных по первым двум критериям клонов ДНК с тотальной мРНК, выделенной из этих тканей, а также скринируют соответствующие кДНК-овые библиотеки. Для генов наследственных заболеваний с неизвестным первичным биохимическим дефектом библиотеки конструируют из пораженных органов и тканей. При обнаружении последовательностей кДНК, гибридизующихся с геномными зондами, их, в свою очередь, используют для зондирования библиотеки и выявления всех клонов с перекрывающимися последовательностями кДНК. К сожалению, для генов с низким уровнем экспрессии гибридизация может не дать положительных результатов.

Выделенные клоны, удовлетворяющие перечисленным критериям, с большой вероятностью содержат последовательности ДНК, являющиеся частями гена. Однако, всегда существует опасность выбора какого-то другого гена (или псевдогена), локализованного в той же области ДНК. Поэтому требуются дополнительные доказательства идентичности выбранной последовательности ДНК специфическому гену. Такие доказательства могут быть получены, например, при определении нуклеотидной последовательности кДНК и сопоставлении ее с аминокислотной последовательностью кодируемого этим геном белка. Веским доказательством в пользу правильности проведенной идентификации гена может быть обнаружение мутантных вариантов аллелей в изолированных последовательностях ДНК у больных, страдающих соответствующим наследственным заболеванием. Так, например, при идентификации гена муковисцидоза, у 70% больных в клонируемой кДНК последовательности была обнаружена однотипная мутация – делеция трех нуклеотидов – delF508. Наконец, решающим аргументом правильности идентификации нужного гена является успешно осуществленная с его помощью генокоррекция первичного биохимического дефекта, выполненная на соответствующих культурах мутантных клеток, или получение стойкого терапевтического эффекта у трансгенных животных – биологических моделей данного наследственного заболевания.

Определение размера молекул мРНК, гибридизующихся с геномными клонами, дает оценку суммарной величины гена. Эта оценка имеет важное значение для реконструирования полноразмерной кДНК. Её клонирование, по-сути, означает идентификацию гена, так как позволяет определить его границы в геномной ДНК, охарактеризовать его экзонно-интронную структуру и ре-

гуляторные элементы. Зная первичную нуклеотидную последовательность кДНК, можно с уверенностью прогнозировать аминокислотную последовательность соответствующего белка и таким образом определить первичное биохимическое звено в патогенезе соответствующего наследственного заболевания.

Описанный способ изучения молекулярных и биохимических основ наследственных заболеваний получил название обратной генетики, а сам процесс в отличие от традиционного пути от белка к гену, так называемого функционального клонирования, был назван позиционным клонированием, тем более, что термин обратной генетики уже использовался ранее для обозначения метода анализа функции гена путем направленного введения в него мутаций (Collins, 1992).

Возможность использования функционального клонирования зависит от доступности информации о белковом продукте и/или о функции соответствующего гена. Для подавляющего большинства моногенных болезней определение первичного биохимического дефекта представляет собой очень трудную задачу из-за недостаточного понимания функционирования огромного числа клеточных ферментов, сложностей их взаимодействия, низких концентраций, отсутствия эффективных методов выделения и очистки а, зачастую, даже из-за отсутствия сведений о клетках - мишенях, в которых следует искать первичный биохимический дефект. Поэтому на фоне стремительного роста данных о структуре генома человека и, прежде всего, о насыщенности генами и анонимными ДНК маркерами отдельных хромосом и их сегментов, реальные соотношения функционального и позиционного клонирования в идентификации генов, ответственных за наследственные заболевания, быстро меняются в сторону безусловного доминирования последнего.

Успех позиционного клонирования определяется возможностями картирования гена, при этом функция гена исследуется уже после его идентификации и клонирования. На рис. 3.8 представлена общая схема позиционного клонирования, заимствованная из работы Коллинза (Collins, 1992). Обычно, для нахождения положения неизвестного гена на карте сцепления используют 100 - 200 полиморфных маркеров. После обнаружения хромосомной принадлежности картируемого гена более точная локализация может быть установлена с помощью прицельного отбора дополнительных индексных маркеров из определенного цитогенетического сегмента. Картирование гена, определяющего наследственное заболевание, может быть значительно ускорено при наличии у какого-то больного цитогенетически видимой структурной перестройки в области локализации этого гена, чаще всего делеции или транслокации. Хотя такие пациенты, как правило, встречаются редко, но описание даже одного такого случая может исключить необходимость картирования гена путем последовательного анализа его сцепления с генетическими маркерами целого генома и позволит перейти непосредственно к молекулярному клонированию. Именно таким образом были идентифицированы гены хронического грануломатоза, миопатии Дюшенна, ретинобластомы, X-сцепленной глухоты, нефрофиброматоза I, аниридии и некоторых других наследственных болезней.

С другой стороны, в ряде случаев удается исключить длительный процесс молекулярного клонирования, используя метод "кандидатного гена". Разработка методов, облегчающих нахождение транскрибируемых областей генома, улавливание экзонов и регуляторных участков генов, секвенирование и картирование методами гибридизации *in situ* большого количества анонимных кДНК последовательностей, изолированных из тканеспецифи-

ческих библиотек, все это в комплексе приводит к значительному увеличению степени насыщенности различных сегментов хромосом известными генными последовательностями, среди которых и осуществляют поиск гена-кандидата. Большая роль в этих исследованиях принадлежит также мутантным генетическим линиям животных, моделирующим различные наследственные заболевания человека (см Глава VIII). Значительное сходство нуклеотидных последовательностей кодирующих участков гомологичных генов млекопитающих и человека, наличие большого числа консервативных групп сцепления с наборами идентичных генов позволяют успешно вести параллельные исследования геномов человека и других животных, значительно ускоряющие эффективность поиска и молекулярного анализа индивидуальных генов человека (Dietrich et al., 1994; Copeland et al, 1993).

Молекулярная идентификация генов открывает широкие возможности для анализа тканеспецифической регуляции их экспрессии в процессе развития организма на всех уровнях от транскрипции до трансляции. Следующим этапом молекулярного анализа является генотипирование мутаций и исследование тех нарушений в структуре, локализации или в ферментативной активности соответствующих белков, которые возникают в результате изменений нуклеотидных последовательностей ДНК. Эти проблемы более подробно освещены в следующих разделах книги. Отметим только, что в настоящее время подобные исследования стали возможны для многих сотен наследственных заболеваний человека, для которых идентифицированы геномные последовательности ДНК, соответствующие генам, и проклонированы полноразмерные кДНК-последовательности.

Раздел 3.6 Каталог генов и генных болезней МакКьюсика. Международная программа "Геном человека".

Огромный вклад в систематизацию и обобщение информации о генетических картах хромосом человека, локализации и функциях отдельных генов, и о структуре генома в целом, вносят исследования, проводимые на протяжении последних 30 лет в Университете Джона Хопкинса в Балтиморе под руководством профессора Виктора МакКьюсика. Результатом этих исследований является систематическое, с 2-х-годовым интервалом между последними пятью публикациями, издание энциклопедий, содержащих сводные данные о всех картированных генах человека и связанных с ними наследственных болезнях под названием: "Менделеево наследование у человека: каталог человеческих генов и генетических болезней" ("Mendelian inheritance in man. Catalogs of autosomal dominant, autosomal recessive, and X-linked phenotypes"). Эти издания содержат современные хромосомные карты генов человека и для каждого локуса обобщенные в виде отдельных статей сведения о характере наследования, функциях и размерах генов; методах их картирования и идентификации; кодируемых продуктах; мутантных аллелях, полиморфизмах и внутригенных повторах; фенотипических проявлениях мутаций, их связи с наследственными заболеваниями, а также о природе основного дефекта, включая патогенез и патофизиологию заболевания. Все статьи снабжены исчерпывающими литературными ссылками. Сводные таблицы по картированным локусам с различным типом наследования и по генам наследственных заболеваний составлены либо в соответствии с их хромосомной локализацией, либо в алфавитном порядке по названиям генов или по наименованиям соответствующих генных болезней. Отдельно представлены данные по клонированным генам, для которых известен первичный молекулярный дефект. При этом

количество различных идентифицированных мутантных вариантов для разных генов колеблется от одного до нескольких сотен. Издания содержат также список доступных мутантных клеточных линий.

Каждому локализованному менделирующему локусу в этой энциклопедии присвоен шестизначный номер (MIM), первая цифра которого определяет характер наследования: 1 - для аутосомно-доминантных генов, 2 - для аутосомно-рецессивных, 3 и 4 - для генов, локализованных в X- и в Y-хромосомах, соответственно, 5 - для митохондриальных генов. Четыре цифры, следующие после точки непосредственно за шестизначным номером, предназначены для кодирования различных мутантных вариантов данного локуса. Издания выпускаются как в печатной форме, так и в компьютерном варианте (OMIM) на дискетах или на компакт-дисках. В последнем случае они снабжены программами, позволяющими осуществлять поиск по любой позиции и проводить постоянное обновление энциклопедии текущей информацией. Программы OMIM совместимы с другими базами генетических данных, в первую очередь, с GDB (Genome Data Base), содержащей полную информацию (включая последовательности ДНК) обо всех картированных генах, ДНК-маркерах и ДНК-зондах человека, а также и с GenBank - полной базой данных всех известных нуклеотидных последовательностей ДНК.

В последнем 11-ом издании энциклопедии МакКьюсика содержатся сведения о 6678 картированных менделирующих локусах человека (McKusick, 1994). Из них 4458 генов с аутосомно-доминантным характером наследования, 1730 - с аутосомно-рецессивным, 412 генов локализовано в X-хромосоме, 19 - в Y-хромосоме и 59 - в митохондриальной ДНК. Для более чем 2800 картированных генов определена их функция. С моногенными заболеваниями связано 770 картированных локусов, а общее число нозологических форм, для которых гены картированы, включает 933 заболевания. При этом более 420 генов наследственных болезней уже клонированы и для каждого из этих генов описано от одного до нескольких сотен мутантных вариантов аллелей, характеризующихся различным фенотипическим проявлением.

Различные хромосомы и их участки картированы с разной степенью детализации. На самой крупной по размерам хромосоме 1 картировано вдвое меньше генов, чем на X-хромосоме (200 и 400 соответственно). Плотность уже картированных генов в разных хромосомах очень неравномерна. Так, хромосома 19 содержит 178 генов, тогда как хромосома 13 только 40, при этом первая больше второй. Хромосомы 17 и 18 примерно равны по величине, но на первой уже картировано 180 генов, а на второй - только 26. На хромосоме 2 картировано примерно такое же количество генов (около 175), как и на втрое меньше её по размерам хромосоме 17. Существенные различия в числе картированных генов отмечаются и внутри различных участков хромосом. К примеру, 19 из 43 генов хромосомы 21 локализованы в сегменте 21q22.3, составляющем лишь 20% длинного плеча. Область 9q34 занимает 10% хромосомы 9, но содержит 27% генов - 38 из 141 (Antonarakis, 1994). Число подобных примеров неравномерного распределения картированных генов по хромосомам может быть значительно увеличено.

Более 10 лет тому назад был полностью просеквенирован митохондриальный ген (Anderson et al., 1981), состоящий из 16 569 нуклеотидов и содержащий 37 генов, 22 из которых это гены транспортных РНК, 2 гена рибосомальной РНК и 13 белковых генов, кодирующих субъединицы комплексов окислительного фосфорилирования (OXPHOS). Следует отметить, что 56 субъеди-

ниц этого комплекса кодируется ядерными генами (McKusick: 1994). Митохондриальная ДНК очень плотно насыщена кодирующими участками, так как митохондриальные гены не содержат интронов и имеют очень ограниченные размеры некодирующих фланкирующих ДНК. В настоящее время описано достаточно много болезней, связанных с мутациями в митохондриальном геноме, и все они развиваются вследствие нарушений в системе окислительного фосфорилирования.

Мы уже упоминали о том, что в настоящее время проклонировано около 20 000 анонимных последовательностей кДНК, выделенных из тканеспецифических библиотек генов и представляющих около 10-15% всех генов человека. Хотя этих последовательностей пока нет на картах генов, секвенирование, сопоставление с компьютерными базами данных и гибридизация *in situ* позволят уже в самое ближайшее время провести их идентификацию и локализацию (McKusick, Amberger, 1993).

Следует отметить, что каждый картированный ген и полиморфный локус сами по себе автоматически становятся точками отсчета в геноме, то есть молекулярными маркерами. Наряду с этим, продолжается интенсивное насыщение генома новыми молекулярными маркерами типа STS (sequence tagged sites) и микросателлитными повторами типа STR (short tandem repeats) (см. Главу II). К сентябрю 1994г Genome Database (GDB) включила 6691 STR-сайтов и 3 752 из них (56%) имели уровень гетерозиготности более 60%. Карты сцепления для индексных маркеров сконструированы, в основном, по результатам генотипирования сорока CEPH референтных семей (см.Глава II,2.3). Среднее расстояние между соседними маркерами варьирует от 2 сМ для хромосомы 21 до 5 сМ для самых крупных хромосом с очень небольшим числом участков в геноме с расстоянием между маркерами большим, чем 10 сМ. GDB содержит 672 гена, локализованных на картах сцепления индексных маркеров, из общего числа 3485 клонированных генов (Guaraya et al., 1994). Созданные в последние годы достаточно подробные геномные карты сцепления молекулярных маркеров в масштабах 13, 0; 5, 0 и даже 2, 9 сантиморганид; автоматизация процесса генотипирования маркерных микросателлитных (STR) аллелей; большое число уже картированных структурных генов, анонимных ДНК-последовательностей значительно упрощают и, главное, ускоряют процесс генетического картирования. Если в 1992г. в распоряжении исследователей было только 814 динуклеотидных полиморфных сайтов (Weissenbach et al.,1992), то уже к маю 1994 г. их число возросло до 3 300 (Guyaraya et al.,1994) , а к концу года - до 5 000- 6 000 (Shmitt, Goodfellow, 1994). Столь же быстрыми темпами нарастает число молекулярных маркеров и в геноме лабораторных мышей (Service, 1994). По всей видимости, человек и лабораторная мышь будут первыми млекопитающими с полностью расшифрованными геномами.

Картирование генов человека и выяснение первичной нуклеотидной последовательности человеческого генома составляют основные, взаимосвязанные задачи Международной программы "Геном Человека". Официально эта научная программа с участием ведущих молекулярно-генетических лабораторий США, Западной Европы, России и Японии оформилась в 1990г. Однако, задолго до приобретения официального статуса, в этих странах проводились важные молекулярные исследования по изучению генома человека и картированию его генов. История отечественной программы началась в 1987г. Её инициатором и безусловным лидером в течение многих лет был академик А.А.Баев. По его настоянию в 1989г. она стала одной из ведущих Государственных научно-технических программ СССР. Основные разделы этой

программы как в России, так и во всем мире включают три главных направления научных исследований: 1. Картирование и секвенирование генома; 2. Структурно-функциональное изучение генома; 3. Медицинскую генетику и генотерапию (Баев, 1990; 1994).

Предполагалось, что основной раздел программы, касающийся секвенирования всего генома, то есть выяснения первичной последовательности всей молекулы ДНК одной клетки человека длиной около 1,5 метров, состоящей из 3.5×10^9 нуклеотидов, будет завершен уже к 2 005 году. Однако, серьезные технические усовершенствования этого трудоемкого процесса, его автоматизация и резкое снижение себестоимости (от 1\$ США за один шаг в 1990г. до 0,2\$ в 1995г.) позволяют надеяться, что эта гигантская молекула, несущая информацию о всей программе индивидуального развития человека и его эволюции будет полностью расшифрована уже к 2 000 году! (Marshall, 1995).

Естественно, что в итоге этой работы будут идентифицированы и все гены человека, то есть будет точно определено их число, взаиморасположение на генетической карте и структурно-функциональные особенности. Предполагается, что осуществление этого проекта, помимо колоссальных теоретических обобщений для фундаментальных наук, окажет огромное влияние на понимание патогенеза, предупреждение и лечение наследственных болезней, значительно ускорит исследование молекулярных механизмов, лежащих в основе развития очень многих моногенных нарушений, будет способствовать более эффективному поиску генетических основ мультифакториальных заболеваний и наследственной предрасположенности к таким широко распространенным болезням человека как атеросклероз, ишемия сердца, психиатрические и онкологические заболевания.

ГЛАВА X.

МОЛЕКУЛЯРНО-ГЕНЕТИЧЕСКАЯ ХАРАКТЕРИСТИКА НЕКОТОРЫХ МОНОГЕННЫХ ЗАБОЛЕВАНИЙ.

Раздел 10.1. Хромосомная локализация и принципы классификации генов наследственных болезней.

Раздел 8.1 Хромосомная локализация и принципы класси-

К настоящему времени на хромосомах человека картировано около 800 генов, мутации которых приводят к различным наследственным заболеваниям. Количество моногенных заболеваний, для которых известна локализация контролирующего гена, еще больше и приближается к 950 за счет существования аллельных серий, то есть групп болезней, клинически сильно отличающихся друг от друга, но обусловленных мутациями в одном и том же гене (см.Глава IV). Для всех этих заболеваний принципиально возможна пренатальная диагностика с использованием косвенных методов молекулярного анализа (см.Главу VII).

Более половины картированных генов клонировано и охарактеризовано методами молекулярного анализа. Для каждого из этих генов описаны мутантные варианты среди соответствующих групп больных, причем количество идентифицированных аллелей в разных генах может колебаться от одного до нескольких сотен (см.ниже). Молекулярное генотипирование мутации позволяет проводить прямую пренатальную диагностику соответствующего наследственного заболевания в семьях высокого риска (см.Главу VII).

Число генов наследственных болезней, локализованных на

каждой хромосоме приведено на Рис. 10.1. В среднем, на каждой из них к 1995г. идентифицировано около 30 таких структурных генов. Обращает на себя внимание неравномерный характер распределения этих генов. Так, хромосомы 1 и 2 имеют примерно одинаковые размеры (хромосома 2 даже несколько крупнее), однако, число уже картированных генов, связанных с наследственными заболеваниями, на хромосоме 1 в 3 раза меньше, чем на хромосоме 2. Наибольшее число таких генов (больше 100) картировано на X-хромосоме. Это, по-видимому, можно объяснить гемизиготным проявлением мутаций генов X-хромосомы в компаунде гоносом XY у мужчин. Вместе с тем, анализ приведенных данных (Рис. 10.1) свидетельствует и о феномене различной насыщенности разных хромосом структурными генами. Наибольшая плотность структурных генов свойственна хромосомам 1, 3, 7, 9, 17, 22, X. Значительно меньшая - хромосомам 2, 13, 18, 21, Y (Antonarakis, 1994). Неслучайно, дисбаланс некоторых из хромосом 2-й группы часто совместим с постнатальным развитием (синдром Дауна - трисомия 21; синдром Эдвардса - трисомия 18; синдром Патау - трисомия 13). По-видимому, это связано со сравнительно низкой плотностью структурных генов в этих хромосомах, а также с отсутствием в них генов, контролирующих ранние стадии развития. Напротив, сравнительно слабая насыщенность известными генами хромосом 2 и 15 в сочетании с редкостью их дисбаланса даже в абортном материале, может рассматриваться в пользу наличия в этих хромосомах "ранних генов", контролирующих начальные стадии онтогенеза человека: гаметогенез, ранний эмбриогенез. Мутации таких генов отмечаются селекцией уже на этих ранних стадиях, а потому не обнаруживаются постнатально. Стремительный рост данных о генетической информации, заключенной в каждой хромосоме, распределении в ней структурных и регуляторных генов, их взаимодействии с надмолекулярными структурами хромосом (гетерохроматином), межхромосомных взаимодействиях и феномене геномного импринтинга открывает широкие возможности на новом методическом и концептуальном уровне подойти к проблеме хромосомного (геномного) контроля ранних стадий развития человека - основной проблемы цитогенетики развития млекопитающих (Баранов, 1984; 1990; 1992; Dyban, Baranov, 1987).

Другое положение, которое следует напомнить в вводной части этой главы касается специфичности мутационных повреждений каждого структурного гена. Как указывалось ранее (см.Глава V), несмотря на наличие общих закономерностей в мутационных процессах, спектр мутаций для каждого гена, равно как и сами структурные гены - уникальны. Причины этой уникальности кроются в особенностях первичной структуры ДНК каждого гена, в частности, обогащенности CG нуклеотидами, его размерах, наличии прямых и обращенных повторов, присутствию внутри гена ДНК последовательностей, гомологичных внегеномным участкам, что может приводить к нарушениям процессов рекомбинации в мейозе и т.д. Для каждого идентифицированного гена, мутации которого приводят к наследственным заболеваниям, разработаны эффективные методы молекулярной диагностики, как правило, направленные на генотипирование наиболее частых мутаций этого гена. Реже для этих же целей используется непрямой метод диагностики с помощью молекулярных маркеров (см.Глава VII).

Цитогенетические карты представляют собой один из способов однозначной и объективной систематизации генов. Для практических целей медико-генетического консультирования и дифференциальной диагностики моногенных заболеваний подобная классификация не всегда удобна, так как при составлении карт

генов никак не учитывается информация об особенностях кодируемых генами продуктов или о фенотипическом проявлении мутантных аллелей. В медицинских целях чрезвычайно важно иметь представление о группах генов, кодирующих функционально и структурно родственные белки, или контролирующие заболевания со сходной клинической картиной. Однако, далеко не всегда классификация по клиническим параметрам может быть проведена однозначно по ряду причин. Во-первых, большое число моногенных наследственных заболеваний носит синдромальный характер и, зачастую, не удается выделить группу ведущих клинических симптомов. Во-вторых, многие болезни отличаются высоким уровнем фенотипической гетерогенности, связанной либо со специфичностью мутационных повреждений, либо с различиями в окружающих условиях и/или в генетическом фоне (см. Главу IV). Кроме того, многие болезни, вызванные мутациями в разных генах, могут протекать сходным образом и, основываясь только на клинических симптомах, трудно провести дифференциальную диагностику подобных заболеваний. Поэтому наиболее объективная классификация моногенных наследственных болезней с известными первичными биохимическими дефектами проводится на основе классификации соответствующих генопродуктов с учетом их участия в определенных метаболических циклах.

В данной заключительной главе мы попытаемся проиллюстрировать на ряде примеров теоретические положения, изложенные в предыдущих главах. В качестве примеров будут приведены краткие молекулярно-генетические характеристики некоторых классов хорошо изученных и достаточно распространенных моногенных наследственных болезней. Большинство из этих заболеваний в той или иной мере изучаются в соответствующих научных центрах России, а их диагностика в медико-генетических центрах страны проводится не только по клиническим параметрам, но и с обязательным учетом результатов молекулярного и/или биохимического обследования.

Раздел 10.2. Метаболические дефекты лизосомных ферментов. Болезни накопления.

В качестве примера наиболее полно и всесторонне изученных заболеваний мы выбрали группу болезней, обусловленных наследственными дефектами лизосомальных гидролаз. В Табл. 10.1 представлены данные о наследовании и встречаемости лизосомных болезней, хромосомной локализации и структуре соответствующих генов, кодируемых ими продуктами и идентифицированных мутантных аллелях. Даны также ссылки на основные работы по картированию соответствующих генов, их клонированию и идентификации мажорных (то есть наиболее частых) мутаций. Таблица составлена по материалам Каталога наследственных болезней В. МакКьюсика 1994 г. (McKusik, 1994) и дополнена необходимыми литературными данными.

Таблица 10.1 Молекулярно-генетические основы лизосомных болезней (N) - примечания, представленные в конце таблицы).

| -----Т-----Т-----Т-----Т-----Т----- |
|--|
| -----Т-----Т-----Т-----Т-----Т----- |
| Синдромы 1), номер по Встречаемость, Типы и количество му- Литература |
| |
| МакКьюсику; хромосом- белок, размеры таций 5), мажорные мута (локализация и |
| структура |
| ная локализация; ген в аминокисло- ции - в скобках указаны генов, клонирование |
| кДНК, |

2); размеры 3); экзоны; тах 4) | частоты аллелей у б-ных | идентификация мутаций). |

-----+

N-ацетил-альфа-D-га- | Очень редко 6) | Миссенс - 2: | Wang et al., 1990
| лактозаминидазы деф.; | | E325K - Шиндлера болезнь | Desnick, 1991
| Шиндлера; Канзаки б-нь | Ацетилгалактоз | R329W - Канзаки болезнь |
| 104170; 22q11; | аминидаза, аль- | |
| NAGA.2; кДНК-2.2 кб | фа-N-; 411 | |
|

-----+

Ангиокератома Фабри; | 1 : 40 000 | Миссенс -31; делеции (от | Bishop et al., 1988
| дистопический липидоз | | 1 н. до неск.экз.) -11; | Kornreich et
al., 1990 |
301500; Xq22; | Галактозидаза | сплайс.-5 (из них 3 с | Davies et al.,
1993 |
GLA.50; 12 кб, 7 экз. | альфа; 429 | дел.экз); инс., дупл.-3 | Eng et al., 1993
|

-----+

Аспартилглюкозамину- | Более 100 случ | Миссенс -5; делеции -4; | Ikonen et al., 1991
| рия, | в Финляндии | инсерции -2; | Ikonen et al., 1992
| 208400; 4q23-q27; | Аспартилглюкоз | C163S - мажорная мута- |
| AGA.11; | аминидаза; 346 | ция в Финляндии (98%) | |
|

-----+

Вольмана б-нь; гиперхо | Более 70 случ. | Делеция 72 н., возникшая | Anderson et
al., 1991 |
лестерин-гипертригли- | | в результате сплайсинга | Anderson et
al., 1994 |
церидемия, | Лизосомальная | вой мутации -обнаружена | Maslen et al., 1993
| 278000; 10q24-q25; | кислая липаза | в 2 случаях; миссенс -2; | Klima et al., 1993
| LIPA.4; 36 кб; 10 экз. | -А | инсерция 1 н. - 1. | |
|

-----+

Галактосиалидоз, | Преим. в Японии | Делеция экз.7 (сплайс.) | Galjard et al., 1988
| | Протективный | -мажорная у взрослых в | Wiegant et al., 1991
| 256540; 20q13.1; | белок бета га- | Японии; миссенс -6; | Takano et al., 1991
| PPGB.7; мРНК - 2 кб | лактозидазы 452 | F412V - в 2 случаях | Zhou et al., 1991
|

-----+

Ганглиозидоз GMI; му- | Неизвестно | Миссенс -10; дуплик.-2; | Oshima et al., 1988
|

кополисахаридоз 1YB, | |I51T и R201C-мажорные в|Noshimoto et
al.,1991 |
230500; 3p21.33; |Галактозидаза |в Японии, R482H и W273L|Suzuki et al.,1993
|
GLB1.12; кДНК - 2кб |бета-1; 677|мажорные в Европе |Mosna et al.,1992
|
-----+-----+-----+-----+-----+-----+-----+-----+-----+
-----+
Ганглиозидоз GM2-I, |1 : 300 000; у|Миссенс-34;дел.-8;инс.-|Proia et al.,1987
|
варианты В,В1 и псев-|евреев 1:3 000|2;сплайс.-8; Мажорные:у|Myerowitz et
al.,1988 |
до АВ;Тяя-Сакса б-знь|Гексозаминида-|евреев-инс.4 н.-70%,спл|Arpaia et al.,1988
|
272800; 15q23-q24; |за А, альфа |айс.-20%;G269S-взр.-20%|Navon et al.,1989
|
HEXA.52; 35 кб, 14экз| 529|не евреи - R247W - 32% |Triggs-Raine et
al.,1992|
-----+-----+-----+-----+-----+-----+-----+-----+-----+
-----+
Ганглиозидоз GM2, тип|1 : 300 000 |Миссенс -5; делеции -2;|Proia: 1988
|
II,Зандхоффа болезнь,| |инсерции - 2; Мажорные:|Neote et al.,1990
|
268800; 5q13; |Гексозаминида-|16-кб делеция 1-5 экз.-|Wakamatusi et
al.,1992 |
HEXB.9; 40 кб, 14экз.|за В,бета; 556|27%;делеция 50кб; P417K|McInnes et al.,1992
|
-----+-----+-----+-----+-----+-----+-----+-----+-----+
-----+
Ганглиозидоз-GM2, АВ |Очень редко |Миссенс -3: C107R; |Heng et al.,1993
|
вариант, | |R169P; C138R (1 пациент|Schroder et
al.,1993 |
272750; q31.3-q33.1;|GM2-активатор-|гомозиготен) |
|
GM2A.3; |ный белок; 193| |
|
-----+-----+-----+-----+-----+-----+-----+-----+-----+
-----+
Гоше болезнь; глико- |1:600 у евреев|Миссенс -30;инс.-1;деле|Sorge et al.,1985
|
сфинголипидоз, |изол. в Швеции|ции-3;сплайс.-1; Мажор-|Sorge et al.,1987
|
230800; 1q21; |Глюкоцеребро- |ные (98%):N370S, L444P,|Beutler et al.,1992
|
GBA.36; |зидаза; 644|R463C,84insG;IVS2+1 G-A|Horowitz et
al.,1994 |
-----+-----+-----+-----+-----+-----+-----+-----+-----+
-----+
Лейкодистрофия глобо-|1 : 50 000 в |Нонсенс мутация:E369TER|Zlotogora et
al.,1990 |
ид-клеточная, Краббе,|Швеции | |Sakai et al.,1994
|
245200; 14q21-q31; |Галактозилцера| |
|
GALC.1; кДНК -3.78 кб|мидаза 669| |
|
-----+-----+-----+-----+-----+-----+-----+-----+-----+
-----+

| | | | |
|---|-------------------------------------|---|--|
| Лейкодистрофия метахроматическая, 250100; 22q13.31-qter ARSA.12; 8 экз А | 1 : 100 000 | Миссенс -7; делеции -2; сплайс. -2; регулят. -1 Арилсульфатаза Мажорные:Р426L и сплайс 507 2 -70%, регулят. - 1-2% | Stein et al.,1989 Polten et al.,1991 |
| -----+ | | | |
| Лейкодистрофия мета- хроматическая, SAP1 деф.; Гоше б-нь, 176801; 10q21-q22; PSAP.6; 20 кб, 13экз. | Редко | Миссенс -4: T23I,T216I, C241S,F385C; инсерция 33 н. -1; Просапозин регуляторная мутация в 511 инициирующем кодоне -1 | Rorman et al.,1992 Wenger et al.,1989 |
| -----+ | | | |
| Лизосомальной кислой ал.,1988 фосфатазы деф., 171650; 11p12-p11; ACP2.; кДНК-2.1 кб | | Кислая фосфата за 2, лизосом- ная 423 | Pohlmann et |
| -----+ | | | |
| Липидоз сфингомиелино ал.,1989 вый;Ниманна-Пика бо- лезнь, тип А/В, ал.,1992 257200; 11p15.4-p15.1 SMPD1.11; ал.,1992 | Редко | Миссенс - 8; делеции -3 Мажорные:тип А евреи: R496L,L302P,дел.1н.Р330 Сфингомиелина- в комплексе 65%; тип В за 629 Сев.Африка R608del->80% | Quintern et Levran et al.,1991 Schuchman et Suchi et al.,1992 Takahashi et |
| -----+ | | | |
| Ниманна-Пика болезнь, тип С, 257220; 18p; NPC.; | Очень редко | | Carstea et al.,1993 Kurimasa et al,1993 |
| -----+ | | | |
| Ниманна-Пика болезнь, тип D, 257250; | Очень редко | | Winsor et al.,1978 |

| | | | |
|--|---|---|---|
| Маннозидоз, альфа В, лизосомальный, 248500; 19p13.2-q12; MANB.; | 50-100 случаев Лизосомальная альфа-D-манно- зидаза В | | Kaneda et al., 1987 |
| Маннозидоз, бета, 248510; chr.4?; MANB1.; | Очень редко Лизосомальная бета маннози- даза | | Lundin, 1987 |
| MASA синдром (сложная спастическая парапле- гия), 303350; al., 1992 MASA.; | Редко Маннозосвязава- ющий лектин248 | | Schrandner-Stumpel et al., 1990 Rosenthal et al., 1992 |
| Мукополисахаридоз I; Гурлера синдром; Шейе, 252800; 4p16.3; al., 1993 IDUA.9; 19 кб, 14 экз. | 1:100 000, 1:600 000-Шейе Альфа-L-идуро- нидаза 653 | Нонсенс -4; миссенс -3; сплайс. -1; дел. 1н. -1 Мажорные: W402X (31%), Q70X (15%), P533R (3%) | Scott et al., 1991 Scott et al., 1992 Moskowitz et al., 1993 Scott et al., 1993 |
| Мукополисахаридоз II; Хантера синдром, 309900; Xq28; IDS.29; 24 кб, 9 экз. | 1:70 000 в Из- раиле Идуронат-2- сульфатаза 550 | 20% -крупные делеции, из них 4,5 % -всего гена; делеции 1-3н-7; миссенс- 13; нонсенс -4; сплайс.-5 | Wilson et al., 1990 Bunge et al., 1993 Flomen et al., 1993 Hopwood et al., 1993 |
| Мукополисахаридоз IIIA, Санфилиппо синд- ром А, 252900; | 1 : 24 000 в Нидерландах (все типы А-D) | | Kresse et al., 1971 |

| | | | |
|---|------------------------------|--|--|
| -----+ | | | |
| Мукополисахаридоз IIIIB, Санфилиппо синдром В, 252920; Chr.17? | Наиболее часто на юге Европы | | Pande et al., 1992 |
| -----+ | | | |
| Мукополисахаридоз III С, Санфилиппо синдром С, 252930; Chr.14 или 21?; | Редко | | Zaremba et al., 1992 |
| -----+ | | | |
| Мукополисахаридоз III С, Санфилиппо синдром D, 252940; 12q14; GNS.; | Редко | N-ацетил-глюкозоамин-6-сульфатаза 552 | Robertson et al., 1988a Robertson et al., 1988b |
| -----+ | | | |
| Мукополисахаридоз IYA Моркио синдром, 253000; 16q24.3; GALNS.4; | 1 : 300 000 | Миссенс - 3: N204K, A138V, R386C; Галактозамин-6-делеция 2 н. - 1 | Tomatsu et al., 1991 Tomatsu et al., 1992 Masuno et al., 1993 |
| -----+ | | | |
| Мукополисахаридоз YII Марото-Лами синдром, 253200; 5q11-q13; ARSB.5; 1992 | Редко | Миссенс - 4: C137V, C117R, L236P, C405Y; Арилсульфатаза B 533; делеция 1н. - 1 | Litjens et al., 1989 Schuchman et al., 1990 Jin et al., 1992 Litjens et al., 1992 |
| -----+ | | | |
| Мукополисахаридоз YII Слая синдром, 253220, 7q21.11; | Очень редко | Миссенс - 5: A619V, R382C, R216W, R354W, B611W | Guise et al., 1985 Oshima et al., 1987 Miller et al., 1990 |

| | | | | |
|-----------------------|----------------|-----------------------------|-------------------------|---------------------|
| GUSB.5; 21 кб, 12экз. | нидаза | 651 | | Tomatsu et al.,1991 |
| -----+ | | | | |
| Сиалидоз типы I и II; | 50-100 случаев | | | Oohira et al.,1986 |
| липомукополисахаридоз | | | | Klein et al.,1986 |
| 256550; 6p21.3; | Нейраминида- | | | Sasagasako et |
| al.,1993 | NEU.; | за-1 | | |
| -----+ | | | | |
| Фукозидоз, | 30-60 случаев | Нонсенс - 5:Q351X-мажор | | Fowler et al.,1986 |
| | | Фукозидаза аль- | ная (20%), E375X, Q77X, | Kretz et al.,1992 |
| 230000; 1p34; | фа-L-1,ткане- | W382X, Y211X;делеции -4 | | Roychoudhuri et |
| al.,1988 | | | | |
| FUCA1.10; 23 кб, 8экз | вая | 461 (экз2-1,1н-3);сплайс.-1 | | Seo et al.,1993 |
| -----+ | | | | |
| ----- | | | | |

- 1) - через ";" указаны различные наименования болезней либо аллельные варианты
- 2) - после наименования гена через "." указано количество идентифицированных мутантных аллелей к июлю 1994г.
- 3) - размеры генов указаны в килобазах (кб), иногда вместо размеров гена указаны размеры кДНК или мРНК
- 4) - размеры белка указаны в правом нижнем углу
- 5) - при количестве мутаций, меньшем 5, все они указываются после знака ":", при большем количестве перечисляются только типы мутаций
- 96) - частоты заимствованы из сводной таблицы Scriver et al., 189.
Сокращения - взр.- взрослые; дел.- делеция; дупл.- дупликация; инс. - инсерция; н.- нуклеотиды; сплайс. - замена нуклеотидов в донорном или акцепторном сайтах сплайсинга или мутации в интронных областях, создающие новый сайт сплайсинга; экз.- экзон

Из таблицы 10.1 следует, что для 29 из 31 лизосомных болезней определена хромосомная локализация гена, причем в 26 случаях - абсолютно однозначно. 23 лизосомных гена клонированы. Для 20 лизосомных заболеваний описаны различные мутантные аллели, что подтверждает правильность идентификация соответствующих генов (см.Главу III). Для 12 заболеваний найдены мажорные мутации в различных популяциях. Для 8 заболеваний общее количество идентифицированных мутаций пока не превысило шести и, возможно, мажорные мутации для них еще будут идентифицированы.

Спектр мутаций в разных лизосомных генах очень разнообразен. Так, при болезни Фабри наряду с явным преобладанием миссенс мутаций обнаружено 14 внутригенных перестроек размерами от 0.4 до 8 кб, многие из которых имеют точки разрыва в экзоне 2 - области локализации большого числа Alu-повторов. Сам ген GLA содержит 12 различных Alu-элементов, составляющих около 30% его длины. В местах разрывов часто обнаруживаются короткие прямые и обращенные 2-6 нуклеотидные повторы.

Одним из возможных механизмов возникновения структурных перестроек в данном гене может быть незаконная Alu-Alu рекомбинация или, что более вероятно, рекомбинация между короткими повторами. Участие Alu-элементов предполагается также при возникновении 16-кб делеции промоторной области и первых 5-и экзонов гена HEXB – мажорной мутации при болезни Зандхоффа. Нарушение процесса рекомбинации является, по-видимому, причиной возникновения очень большого числа крупных и мелких делеций в IDS-гене при синдроме Хантера. Высокая концентрация CpG динуклеотидов рассматривается как возможный эндогенный механизм возникновения мажорной мутации среди евреев-ашкенази мутации P330FS в гене SPDM1 при болезни Ниманна-Пика типа В, так как эта делеция возникает в районе, где из 10 остатков 9 составляют цистеины.

Делеции целых экзонов или инсерции интронных областей возникают сравнительно часто в результате точковых мутаций в донорных или акцепторных сайтах сплайсинга. Примерами являются мажорная в Японии сплайсинговая мутация, сопровождающаяся делецией 7-го экзона гена- PPGB, приводящая к галактосиалидозу и сплайсинговая мутация IVS2+1, обуславливающая вырезание экзона 2 гена GBA при болезни Гоше. Появление в результате точковой мутации в интронной области нового сайта сплайсинга также может сопровождаться структурными перестройками. Такова, в частности, природа 33-нуклеотидной инсерции в гене PSAP при метахроматической лейкоцистозии, обусловленной дефицитом SAP1; 24-кб инсерции в гене HEXB при болезни Зандхоффа и 5-нуклеотидной инсерции в гене IDUA при синдроме Шейе. Важно отметить, что подобные мутации совместимы с образованием небольшого числа функционально активных мРНК, следствием чего является относительно более мягкое течение соответствующих форм заболеваний.

В некоторых случаях возникновению мутаций может способствовать наличие псевдогена. Молекулярный анализ псевдогена, тесно сцепленного с геном GBA, показал, что, по крайней мере, 4 мутации, обнаруженные у пациентов с болезнью Гоше, присутствуют в норме в псевдогене. Это 2 мажорные мутации – L444P и IVS2+1 и еще 2 миссенс мутации в 10-м экзоне (A456P и V460V). Подобное сходство несомненно указывает на фундаментальную роль псевдогена в образовании мутаций в GBA-гене. В то же время само по себе присутствие псевдогена не является мутагенным фактором, особенно если сам ген и его псевдоген локализованы в разных хромосомах, как, например, в случае генов GM2A и FUCA1, псевдогены которых находятся в хромосоме 3 и в области 2q31-q32, соответственно.

Для двух лизосомных болезней – фукозидоза и синдрома Гурлера, мажорными являются нонсенс мутации. Более того, при фукозидозе все известные к настоящему времени мутации приводят к полному отсутствию продукта FUCA1-гена. Так, наряду с мажорной мутацией Q351X, представленной в 20% хромосом у больных фукозидозом, описаны еще 4 нонсенс мутации и 4 делеции со сдвигом рамки считывания. При синдроме Гурлера две мажорные нонсенс мутации – W402X и Q70X, составляют около 50% всех известных мутантных аллелей гена IDUA. Кроме того, при этом заболевании зарегистрированы еще 4 минорные по частоте нонсенс мутации и 1 делеция со сдвигом рамки считывания. 3 миссенс мутации и интронная мутация, создающая дополнительный сайт сплайсинга в гене IDUA, не приводят к полному блоку синтеза идуронидазы и реализуются в виде синдрома Шейе. Уместно заметить, что оба заболевания – синдром Гурлера и синдром Шейе, являются классическим примером фенотипического разнообразия, обусловленного существованием аллель-

ных серий (см.Главу IV). Такой спектр крайне тяжелых мутаций нельзя объяснить только повышенной частотой их возникновения. Более вероятным представляется предположение о том, что кодируемые FUCA1- и IDUA-генами белки обнаруживают определенную устойчивость к небольшим повреждениям и сохраняют функциональную активность при определенных аминокислотных заменах, то есть миссенс аллели в этих генах проявляют себя как нейтральные мутации и не приводят к развитию заболеваний.

Хорошо известно, что распространение мутаций в популяциях определяется не только, и не столько повышенной частотой их возникновения, но многими другими популяционно-генетическими факторами и, в первую очередь, связано с эффектом основателя (см.Главу V). Типичным следствием эффекта основателя, как известно, является наличие различных мажорных по частоте мутаций одного и того же гена у пациентов разных изолятов и этнических групп. Подобная картина наблюдается, в частности, при ганглиозидозе GMI. Так, в Японии мажорными при этом заболевании являются миссенс мутации I51T и R201C, тогда как в Европе преобладают мутации R482H и W273L. Эффектом основателя можно объяснить высокую частоту аспартилглюкозаминурии в Финляндии, так как в 98% случаев у пациентов финского происхождения заболевание обусловлено присутствием одной и той же миссенс мутации C163S, резко уменьшающей активность аспартилглюкозаминидазы. Интересно отметить, что эта мутация у больных находится в сильном неравновесном сцеплении с другой миссенс мутацией в AGA-гене - R161Q, являющейся, в свою очередь, редкой формой полиморфизма. Невозможно, однако, исключить возможность комбинированного влияния этих двух мутаций на фенотип.

Яркие примеры этнических различий по частоте и спектру мажорных мутаций выявляются при анализе таких лизосомных болезней накопления как болезнь Тея-Сакса, Ниманна-Пика и болезнь Гоше. Прежде всего, эти заболевания особенно распространены среди евреев-ашкенази, среди которых они встречаются в десятки раз чаще, чем в других популяциях европейского или азиатского происхождения. Наличие специфических мажорных мутаций для всех трех заболеваний у 70 - 95% всех пациентов еврейского происхождения скорее всего нельзя объяснить только эффектом основателя. Генетический дрейф, селективное преимущество гетерозигот, характер миграции, социальные и религиозные особенности, обуславливающие ассортативность образования супружеских пар - вот те факторы, которые, по всей видимости, лежат в основе этих различий. В этой связи интересно отметить, что среди пациентов других национальностей мажорные мутации гомологичных генов, как правило, иные, чем у евреев-ашкенази. Так, болезнь Ниманна-Пика типа В часто встречается среди жителей стран, расположенных в западной части Северной Африки. Однако, в 80% случаев заболевание связано с делецией R608 в SMPD1-гене, которая не является мажорной среди евреев-ашкенази.

На примере лизосомных болезней могут быть хорошо прослежены корреляции между типами мутаций и клиническими особенностями заболеваний. Выше уже упоминалось об аллельных вариантах гена IDUA, приводящих либо к синдрому Гурлера, либо к синдрому Шейе. Разные миссенс мутации в гене NAGA приводят к болезни Шиндлера или к болезни Канзаки (Табл. 10.1.). Важное значение для анализа молекулярных основ патогенеза заболеваний имеют специфические мутации с поздней фенотипической манифестацией (так называемые взрослые формы). Такие мутации обнаружены в соответствующих генах при болез-

нях Тея-Сакса и галактосиалидозе. Очень интересен случай различного фенотипического проявления на разном расовом генетическом фоне одной и той же мутации - мажорной 16-кб делеции, обнаруживаемой у 27% пациентов с детской формой болезни Зандхоффа (McInnes et al., 1992; Sidransky et al., 1994). В частности, в одной франко-канадской семье эта мутация в компаунде с миссенс мутацией P417L, описанной впервые в Японии у пациента с подростковой формой заболевания, проявлялась как взрослая форма с очень мягким течением заболевания.

В ряде случаев удалось проанализировать молекулярную природу совместного влияния двух аллелей одного гена на фенотип. К примеру, при некоторых формах метахроматической лейкодиетрофии трудность молекулярной диагностики заболевания связана с существованием, так называемого, псевдодефицитного аллеля ARSA-гена. Этот полиморфный аллель встречается в популяциях с достаточно высокой частотой, так что гомозиготы по нему составляют 1 - 2% всего населения. Оказалось, что псевдодефицитный аллель представляет из себя сочетание двух мутаций в цис-положении. Одна из них - 3'-концевая регуляторная мутация в первом сайте после стоп кодона, изменяет сигнал полиаденилирования. Другая - миссенс мутация в 6-м экзоне, приводит к потере сайта N-гликозилирования. Попутно отметим, что для гена ARSA (также как и для IDUA-гена) обнаружен альтернативный сплайсинг, в результате которого в фибробластах и печени образуются 2 различных типа мРНК, размером 2.1 кб и 3.9 кб, соответственно. У гомозигот по псевдодефицитному аллелю в фибробластах отсутствует 2.1 кб мРНК, при этом клинических проявлений заболеваний не наблюдается. Однако, при наличии S96F мутации в ARSA-гене на фоне псевдодефицитного аллеля развивается тяжелая форма лейкодиетрофии.

В заключении раздела кратко рассмотрим состояние проблемы генотерапии лизосомных заболеваний. В литературе отсутствуют сообщения об успешных клинических испытаниях программ генотерапевтического лечения этих заболеваний, однако, по крайней мере, для некоторых лизосомных болезней такие программы уже разработаны и утверждены (см. Главу IX, Табл. 9.2). Имеются сведения о положительных результатах таких исследований на культурах мутантных клеток и на модельных животных. Так, в опытах *in vitro* был осуществлен успешный ретровирусный перенос нормальной кДНК гена GBA в культурах мутантных фибробластов (Sorge et al., 1987) и в культурах клеток крови пациентов с болезнью Гоше (Fink et al., 1990), в результате чего была достигнута коррекция глюкоцереброзидазной активности. Такая же коррекция метаболических дефектов при болезни Ниманна-Пика и при синдроме Хантера была достигнута путем введения в соответствующие мутантные линии клеток нормальных кДНК генов SMPD1 и IDS соответственно. При этом активность идуронат-2-сульфатазы после ретровирусной трансдукции *in vitro* оказалась существенно выше нормальной и рекомбинантный фермент активно участвовал в метаболизме глюкозамногликанов. Генотерапия первичного биохимического дефекта при мукополисахаридозе VII (синдром Слая) была получена как *in vitro*, путем ретровирусного переноса нормального гена GUSB в мутантные фибробласты человека, так и *in vivo* на собаках и мышах. При этом у больных собак нормальный белок (бета-глюкуронидаза) не только экспрессировался, но появлялся в лизосомах и восстанавливал процессинг специфических глюкозаминогликанов (Wolf et al., 1990). Введение этого же гена (GUSB) в мутантные стволовые клетки мышей приводило к длительной экспрессии бета-глюкуронидазы, снижению лизосомального накопления в печени и мозге и частичной коррекции

болезни у трансгенных животных (Wolf et al., 1992). В другом эксперименте GUBS-кДНК вводили в культивируемые мутантные фибробласты кожи мышей и затем трансдуцированные клетки имплантировали подкожно мутантным мышам. У всех животных наблюдали экспрессию введенного гена и полное исчезновение лизосомальных отложений в печени и в мозге (Sly, 1993). Полученные результаты подтверждают принципиальную возможность лечения, по крайней мере, некоторых лизосомных болезней с помощью методов генной терапии.

Раздел 10.3. Болезни экспансии, вызванные "динамическими" мутациями.

Обнаруженный в 1991г. новый тип так называемых динамических мутаций и связанные с ними наследственные заболевания частично рассматривались нами в Главе IV. Однако их уникальность, необычный механизм экспрессии, особенности наследования, быстрый рост нозологий, обусловленных подобными нарушениями в последовательности ДНК, и, как оказалось, достаточно широкая распространенность (см.Табл.9.2) делают целесообразным их более подробный анализ.

Как упоминалось, этот тип мутаций пока найден только у человека и не зарегистрирован ни у одного вида млекопитающих или других хорошо изученных организмов (дрозофила, нематоды и пр.). Суть мутаций заключается в нарастании числа триплетных повторов, расположенных в регуляторной или в кодирующей, а значит и в транслируемой части генов. Впервые такой тип мутации был обнаружен при молекулярном анализе синдрома фрагильной (ломкой) X-хромосомы, наследственная передача которой не подчинялась обычным менделевским законам. В дальнейшем аналогичные динамические мутации были описаны и при 7 других наследственных заболеваниях, контролируемых генами, расположенными на разных хромосомах - Таблица 10.2. Вместе с тем, все нижеперечисленные нозологии имеют ряд общих признаков, позволяющих объединить их в одну самостоятельную группу. Прежде всего, для триплетных повторов, экспансия которых блокирует функцию гена, характерен выраженный популяционный полиморфизм, причем число аллелей может варьировать от единиц до нескольких десятков. Другой их особенностью является доминантный тип наследования, характерный как для X-сцепленных генов, так и для генов, находящихся на аутосомах. Особенностью практически всех болезней "экспансии" является также эффект антиципации (ожидания), смысл которого заключается в нарастании тяжести симптомов заболевания в последующих поколениях, что, как оказалось, является результатом накопления ("экспансии") исходного числа триплетов после того как их количество возрастает больше нормального. Характерными для этих нозологий являются и особенности их передачи потомству: для некоторых заболеваний типична передача по материнской (Fra-X, миотоническая дистрофия), а для других - преимущественно по отцовской линии (хорея Гентингтона). Практически для всех "динамических" мутаций характерно поражение головного мозга и особенно подкорковых структур, причем тяжесть заболевания и его начало четко коррелируют с числом повторов. Молекулярный анализ этих генов позволяет предполагать определенное сходство механизма экспансии триплетов, которая, по всей вероятности, происходит в митозе, затрагивает чаще аллели с начально большим числом повторов, при этом нередко сигналом экспансии является утрата негомологичного триплета, в норме разделяющего цепочку монотонных повторов. Вместе с тем, патогенетические механизмы проявле-

ния мутаций экспансии принципиально различны. В случае различных вариантов FRAX мутаций наблюдается блок экспрессии соответствующих генов вследствие стабильного метилирования области CpG островка в промоторной части генов. При миотонической дистрофии нарушение экспрессии, по-видимому связано с ошибками взаимодействия транскрибируемой нити ДНК с нуклеосомами. В случае остальных сугубо нейродегенеративных заболеваний (хорея Гентингтона, спинально-бульбарная мышечная атрофия и др.) экспрессия гена не нарушена, однако, образующийся белковый продукт с необычно длинной полиглутаминовой цепочкой каким-то образом нарушает процессы нормального метаболизма нервных клеток подкорковых отделов мозга.

Таким образом, причиной повреждающего действия одних "динамических" мутаций является блок генной экспрессии, то есть потеря функции (loss-of-function mutation), тогда как другие мутации того же типа, связанные с нейродегенеративными заболеваниями, ведут к появлению белковых продуктов с аномальными функциями (мутации типа -gain-of-function). Интересно отметить, что помимо динамических мутаций для каждого названного гена обнаружены единичные точковые мутации, число которых крайне невелико. Для каждой болезни "экспансии" разработан свой вариант диагностики, основанный на ПЦР. Амплификация области триплетных повторов и дальнейший электрофоретический анализ синтезированных продуктов позволяет определить число повторов, то есть провести генотипирование аллелей. Вместе с тем, при числе повторов более 200, амплификация с помощью ПЦР обычно не достигается. В этих случаях размеры участка повторов определяются методом блот-гибридизации с соответствующими ДНК зондами. Например, используются зонды StB12.3, Oх1.9 или Oх 0.55 в случае синдрома FRAXA; зонд сDNA25 в случае миотонической дистрофии.

Подробнее с этой интересной группой заболеваний можно ознакомиться в ряде обзоров (Willems, 1994; Баранов и др., 1993; Иллариошкин и др., 1995).

Таблица 10.2. Болезни экспансии, вызванные динамическими мутациями.

| Болезнь, номер по МакКьюсику (MIM) | Ген, локализация | Триплет | Норма | Премутация | Мутация | Литература |
|------------------------------------|------------------|---------|-------|------------|---------|-------------------------|
| Синдром ломкой X-хромосомы; | FMR1, FRAXA | (CGG)n | 5-50 | 50-90 | >90 | Rousseau et al., 1991 |
| мосомы; | 309550; Xq27.3 | | | | | Hirst et al., 1991 |
| Синдром ломкой X-хромосомы тип 2; | FMR2, FRAXE | (CGG)n | 6-25 | 25-200 | >200 | Knight et al., 1994 |
| мосомы тип 2; | 309548; Xq27.3 | | | | | |
| Миотоническая дистрофия; | DM, MP-1 | (CTG)n | 5-10 | 19-30 | >30 | Shelbourne et al., 1992 |

| | | | | | | | | |
|---|--------|-------------|--|----------|--|-------|--|------------------------|
| фия; | 160900 | 19q13.3 | | | | | | Wieringa,1994 |
| -----+ | | | | | | | | |
| Хорея Гентингтона; Dis. | | HD, IT-15 | | (CAG)n | | 6-37 | | 37-121 Huntington's |
| | | 143100 | | 4p16.3 | | | | |
| Collab.Res.Group,1993 | | | | | | | | |
| -----+ | | | | | | | | |
| Спинально-мозжечковая атаксия тип 1; al.,1993 | | SCA1 | | (CAG)n | | 6-39 | | 41-81 Orr et al.,1993 |
| | | 164400 | | 6p21.3 | | | | Chung et |
| | | | | | | | | |
| -----+ | | | | | | | | |
| Денто-рубральная-палли- al.,1994 | | DRPLA, B-37 | | (CAG)n | | 7-34 | | 54-75 Koide et |
| до-люисовая дегенерация al.,1994 | | 12pter-p12 | | | | | | Nagafuchi et |
| | | 125370 | | | | | | |
| -----+ | | | | | | | | |
| Спинально-бульбарная al.,1991 | | AR | | (CAG)n | | 12-33 | | 40-62 La Spada et |
| мышечная атрофия; al.,1991 | | 313200 | | Xq11-q12 | | | | |
| -----+ | | | | | | | | |
| Спинально-мозжечковая al.,1994 | | MJD | | (CAG)n | | 13-36 | | 68-79 Kawaguchi et |
| дегенерация Мачадо- Джозефа | | 14q32.1 | | | | | | |
| -----+ | | | | | | | | |

Раздел 10.4. Моногенные наследственные болезни, диагностируемые молекулярными методами в России.

Сводка, представленная в таблице 10.3, составлена на основании анализа работ основных отечественных лабораторий и публикаций, связанных с проблемой молекулярной диагностики наследственных болезней. Сводка не является исчерпывающей и включает преимущественно те заболевания для которых возможна или уже проводится диагностика на внутриутробных стадиях развития (Баранов, 1994).

Таблица 10.3. Моногенные наследственные болезни, диагностируемые молекулярными методами и доступные пренатальной диагностике в России.

| -----Т----- | | -----Т----- | |
|-------------|------------------------------|--------------------|--|
| N пп | Болезни | Медицинские центры | |
| +-----+ | | | |
| 1 | Муковисцидоз | ИАГ;ИЭМ РЦМГ; ТИМГ | |
| 2 | Миодистрофия Дюшенна/Беккера | ИАГ;РЦМГ; ТИМГ | |
| 3 | Гемофилия А | ИАГ; ГНЦ; | |
| 4 | Гемофилия В | ИАГ; ГНЦ | |

| | | |
|----|--|---------------------|
| 5 | Фенилкетонурия | ИАГ; ГНЦ; ПМА; РЦМГ |
| 6 | Синдром ломкой X-хромосомы | ИАГ; |
| 7 | Миотоническая дистрофия | ИАГ |
| 8 | Болезнь Виллебранда | ИАГ; ГНЦ |
| 9 | Хорея Гентингтона | ИАГ; РЦМГ; НИИН |
| 10 | Болезнь Леш-Нихана | ИАГ |
| 11 | Спинально-бульбарная мышечная атрофия | ИАГ; НИИН |
| 12 | Гепато-лентиккулярная дегенерация | РЦМГ; ИАГ |
| 13 | Болезнь Хантера | ИАГ |
| 14 | Адрено-генитальный синдром | ЦОЗМиР; РЦМГ |
| 15 | Атаксия Фридрейха | ГНЦ |
| 16 | в-Талассемия | ГНЦ; ПМА |
| 17 | Болезнь Верднига-Гоффмана | РЦМГ |
| 18 | Дефицит альфа-1-антитрипсина | ИЭМ |
| 19 | Семейная гиперхолестеринемия | ИЭМ; ПМА |
| 20 | Предрасположенность к инсулин-зависимому диабету | ПМА |
| 21 | Дефицит ацил-СоА дегидрогеназы | ПМА |

ИАГ - Институт Акушерства и Гинекологии РАМН, Санкт Петербург
 ИЭМ - Институт Экспериментальной Медицины РАМН, Санкт Петербург
 ПМА - Педиатрическая Медицинская Академия, Санкт Петербург
 РЦМГ- Российский Центр Медицинской Генетики РАМН, Москва
 ГНЦ - Гематологический Научный Центр МЗ РФ, Москва
 ТИМГ- Томский Институт Медицинской Генетики, Томск
 НИИН- Научно-исследовательский Институт Неврологии РАМН, Москва
 ЦОЗМиР- Центр Охраны Здоровья Матери и Ребенка, Москва

Молекулярные характеристики некоторых из приведенных в таблице заболеваний уже были рассмотрены в разделах 10.1 и 10.2. Отдельные аспекты, касающиеся идентификации соответствующих генов, их картирования, клонирования и сканирования мутаций уже упоминались в качестве примеров при изложении соответствующих вводных глав монографии (см. Главы II, III, IV, V). В этой части нам представляется целесообразным суммировать эти разрозненные сведения и охарактеризовать те наиболее частые, социально значимые моногенные заболевания, в отношении которых у нас накоплен достаточно большой собственный опыт молекулярных исследований. Список этих нозологий, их частоты и основные молекулярные характеристики приведены в Таблице 10.4.

Таблица 10.4. Основные молекулярно-генетические характеристики моногенных болезней, диагностируемых в Лаборатории пренатальной диагностики ИАГ РАМН, Санкт-Петербург.

(примечания в конце таблицы те же, что в Табл.10.1).

-----Т-----Т-----Т-----

Синдромы 1), номер по |Встречаемость, |Типы и количество му- |Литература
 |
 МакКьюсику; хромосом- |белок, |таций 4), мажорные мута | (локализация и
 структура |
 ная локализация; ген |размеры в ами- |ции -в скобках указаны |генов, клонирование
 кДНК, |
 2), размеры 3), экзоны |ноукислотах |частоты аллелей у б-ных |идентификация
 мутаций). |

-----+-----+-----+-----+-----
 -----+
 Муковисцидоз; врожден|1:2500- Европа|Точковые-преобладающие;|Rommens et al.,
 1989 |
 ное отсутствие vas de|1:3800 -Россия|небольшие дел.и дупл.; |Kerem et al., 1989
 |
 ferens, |CF-трансмемб- |Мажорные:delF508-30-90%|Riordan et al.,
 1989 |
 219700; 7q31.2 |ранний регуля-|W1272X-2-33%,3732delA -|Горбунова и др.,
 1991 |
 CFTR.500; 260кб,27экз|тор 1480|4%,394delTT,G542X,R117N|Баранов и др., 1995
 |

-----+-----+-----+-----+-----
 -----+
 Миопатия Дюшенна; Бек|1:3500 мальчи-|Делеции протяженные - |Kunkel et al.,1985
 |
 кера;кардиомиопатия |ков |60%; дубликации - 6-7%;|Kunkel et al.,1986
 |
 делеционная, | |делеции нескольких н. -|Koenig et al.,1988
 |
 310200; Xp21.2 |Дистрофин |7; нонсенс - 9; сплайс.|Roberts et
 al.,1992a;b |
 DMD.21; 2000кб,73экз | 3685|-3; миссенс -1; инс.-1 |Ahn et al.,1993
 |

-----+-----+-----+-----+-----
 -----+
 Гемофилия А, |1:6500 мальч. |Делеции экз. - 31; мис-|Wood et al, 1984
 |
 фактора YIII деф.; |Фактор YIII |сенс - 21; нонсенс -8 |Toole et al., 1984
 |
 306700; Xq28 |свертываемости|Мажорные: инверсия 26 -|Higuchi et al.,1991
 |
 F8C.66; 186 кб, 26экз|крови 2351|25 экзонов - 45% семей |Naylor et al., 1993
 |

-----+-----+-----+-----+-----
 -----+
 Гемофилия В, Кристма-|1:20000 мальч.|Миссенс и нонсенс более|Kurachi et al.,
 1982 |
 са, фактора IX деф.; |Фактор IX |60%; сплайс.-10%; регу-|Jagadeeswaran et
 al,1984|
 306900, Xq27.1-q27.2 |свертываемости|лят.- 3.5%; делеции -до|Green et al., 1991
 |
 F9.400; 34 кб, 8 экз |крови 461|40% при тяжелых формах |Giannelli et al.,
 1993 |

-----+-----+-----+-----+-----
 -----+
 фон Виллебранда бо- |1: 5 - 20 000 |Тип I и II-миссенс; Ма-|Lynch et al., 1985
 |
 лезнь, |Фактор Y111R |жорные:R543W,R545C,V553|Bonthorn et al.,
 1986 |
 193400; 12pter-p12 |свертываемости|M,R578Q.Тип III-делеции|Cooney et al., 1991
 |
 F8VWF.22; 178кб,52экз|крови |1 н.в 28экз; нонсенс -4|Randi et al., 1991
 |

-----+-----+-----+-----+-----
 -----+
 Фенилкетонурия; гипер|1 : 10 -15 000|Миссенс -62%; нонсенс -|Guttler, Woo, 1986
 |
 фенилаланинемия, мяг-|Фенилаланин- |13%;сплайс.-13%;делеций|DiLella et al.,
 1986 |

| | | | |
|---|--|-------------------------|--------------------|
| кая, 261600; 12q24.1 гидроксилаза 9%; Мажорные: IVS12+1, Cotton, 1990 | | | |
| | | | |
| РАН.70; 90 кб, 13 экз | 452 R408W,R261Q,R158Q,IVS10 | Барановская и др. | 1995 |
| -----+ | -----+ | -----+ | -----+ |
| -----+ | | | |
| Леш-Нихана синдром; | Миссенс -53%;небольшие | Jolly et al., 1983 | |
| | | | |
| HPRT-родств. подагра, Гипоксантин- | структ.перестройки -40% | Edwards et al., | 1990 |
| 308000; Xq26-q27.2 | фосфорибозил сплайс. -5%; нонсенс-2% | Rossiter et al., | 1991 |
| HPRT.100; 44 кб,9 экз | трансфераза217 | Мажорные: R170TER (15%) | Sculley et al., |
| 1992 | | | |
| -----+ | -----+ | -----+ | -----+ |
| -----+ | | | |
| Гепато-лентикулярная | Миссенс -15; делеции/ | Bull et al., 1993 | |
| | | | |
| дегенерация Вильсона- | инсерция -14;Мажорные - | Petruchin et al., | 1993 |
| Коновалова, | Медь-транспор- | H714Q - 31% в Америке, | Tanzi et al., 1993 |
| | | | |
| 277900; 13q14.3-q21.1 | тирующая АТФа 22% в России; 1 н. дел. | Thomas et al., 1995 | |
| | | | |
| АТР7В.34; | за Р типа 1434 | H1070G1-28%;G11267L 10% | |
| | | | |
| -----+ | -----+ | -----+ | -----+ |
| -----+ | | | |

Многим из приведенных в Табл.9.4 заболеваний посвящены обстоятельные обзоры, руководства и монографии, включающие наряду с клиническими данными специальные разделы, посвященные молекулярной природе патологического процесса, характеристике гена, его мутаций, их фенотипических проявлений, а так же последним достижениям и перспективам лечения, включая генную терапию. Некоторые из этих публикаций уже были упомянуты в предыдущих разделах книги, другие будут процитированы в соответствующих подразделах этой главы. Следует так же упомянуть, что все приведенные нозологии были первыми моногенными болезнями, которые были исследованы молекулярными методами в нашей стране и для которых, в итоге, были разработаны и оптимизированы с учетом этнических и национальных особенностей аллельного полиморфизма, частот и паттерна мутаций оптимальные схемы молекулярной обследования семей высокого риска с целью пренатальной диагностики и выявления гетерозиготного носительства. Практически все эти исследования проводились в рамках основных научных программ ГКНТ "Геном Человека" и "Приоритетные направления генетики". Обзор отечественных работ по молекулярной диагностике наследственных болезней представлен в ряде научных сводок (Баранов, 1992; 1994; Varanov, 1993; Баранов и др., 1994; Евграфов, 1993).

10.4.1 Муковисцидоз.

Муковисцидоз (кистозный фиброз поджелудочной железы) – самое распространенное моногенное наследственное заболевание у представителей белой расы. Это первая "генная" болезнь, молекулярные основы которой определены методами позиционного клонирования без использования каких-либо данных о структурных перестройках в области локализации предполагаемого гена.

Напомним, что обнаружение пациентов с транслокациями или крупными делециями, затрагивающими исследуемый ген, значительно облегчает его картирование и идентификацию, как это было при миопатии Дюшенна, ретинобластоме или хроническом грануломатозе. В области локализации гена муковисцидоза хромосомные перестройки, практически, отсутствуют. Подробно об истории открытия гена, его структуре, клонировании, анализе функций, идентификации мутаций можно ознакомиться в следующих отечественных работах (Гембицкая, Баранов, 1991; Баранов, Гинтер, 1994).

Белковый продукт гена -трансмембранный регуляторный белок муковисцидоза -ТРБМ (cystic fibrosis transmembrane regulator -CFTR), как оказалось, является белком хлорных каналов, регулирующих обмен ионов хлора через апикальные мембраны всех эпителиальных клеток организма. В зависимости от типа мутаций, их локализации функция гена ТРБМ при муковисцидозе может быть полностью или частично нарушенной.

К началу 1995г. в гене ТРБМ идентифицировано более 500 мутаций, несколько делеций и дупликаций. Наиболее распространенной у жителей Западной Европы и Северной Америки является мутация delF508, приводящая к отсутствию фенилаланина в 508 положении белка ЕРМБ. В этих странах частота delF508 находится в пределах 70-85%. В Европе частота мутации обнаруживает определенный градиент распространения с севера на юг и с запада на восток, достигая 85% в Дании она уменьшается до 50% в Италии и до 20-30% в Турции. В Европейской части России она составляет около 50% всех мутантных (CF) хромосом. У евреев-ашкенази доминирующей по частоте является мутация W1282X (33%).

Биохимический анализ клеток пациентов, больных муковисцидозом, позволил выяснить фенотипические особенности экспрессии мутантных аллелей гена ТРБМ, установить определенный градиент патологических нарушений на мембранном, клеточном и организменном уровнях в зависимости от типа мутации, её локализации и структуры аномального белкового продукта (Баранов, Гинтер, 1994). Так, было установлено, что некоторые CFTR- мутации, в том числе и delF508, не препятствуя трансляции, нарушают процессинг белка так, что он не достигает апикальной мембраны и не создает хлорный поток. Этим объясняется тяжелая клиника муковисцидоза при подобных нарушениях (Sheppard et al., 1993). Другие мутации (R117H, R334W и R347P), выявленные при более мягких формах заболевания, не затрагивают процессинга белка, хлорный канал формируется, но функционирует менее эффективно, чем в норме. Сопоставление процессинга, локализации и функционирования белка в 3-х линиях L-клеток, трансдуцированных нормальным CFTR-геном и двумя его аллельными вариантами, показало, что дефект одной из мажорных мутаций - G551D, связан с частичной инактивацией функции хлорного канала, в то время как delF508, как оказалось, обладает комбинированным эффектом - нарушает локализацию и стабильность белка и снижает эффективность его функционирования в качестве хлорного канала (Yang et al., 1993). Интересно, что мутация R117H в гомозиготном состоянии, чаще в компаунде с другими аллелями, в том числе с delF508, обнаруживается у пациентов с врожденным отсутствием семявыводящих канальцев (vas deferens). При этом клиника муковисцидоза у таких пациентов, как правило, отсутствует или очень стерта. Это еще один из примеров фенотипического разнообразия, обусловленного аллельными сериями.

В настоящее время в России доступны идентификации по наличию известных мутаций около 65% больных муковисцидозом

(Иващенко, 1992; Потапова, 1994). Диагностическая ценность основных мажорных мутаций в порядке убывания их частоты следующая: delF508 (50%), 3732delA (4,3%), W1282X (2,4%), 394delTT (2,1%); G542X (2,0%) (Иващенко, 1992). Согласно нашим данным (Баранов, 1994) молекулярная диагностика муковисцидоза прямыми методами возможна примерно в 55-60% семей. В случае непрямой диагностики используются полиморфные сайты локусов ДНК, расположенные в непосредственной близости (IRP, D7S8, D7S23, MET) или внутри самого гена CFTR - минисателлитные последовательности в интронах 6, 8, 17b гена CFTR (Zielenski et al., 1991; Morral et al., 1991; Chehab et al., 1991; Агбангла и др., 1992; Potapova et al., 1994). При отсутствии мажорных мутаций и информативности по полиморфным сайтам молекулярные исследования могут быть дополнены биохимическим анализом амниотической жидкости на содержание ферментов микроворсинок кишечника плода (Гобунова и др., 1989; Баранов и др., 1991).

Методами направленного мутагенеза (gene targeting см. Главу VIII) в различных лабораториях США и Великобритании осуществлено успешное конструирование трансгенных моделей муковисцидоза на мышах (Clarke, et al., 1992; Colledge et al., 1992; Dorin et al., 1992; Snouwaert et al., 1992). Выявлены признаки сходства и некоторые межвидовые различия проявления мутации гена CFTR у мышей и человека. Показано, что различные мутации этого гена по-разному реализуются в фенотипе CFTR дефицитных мышей. Так, в одной из трансгенных линий отмечено преимущественное поражение легких, у других - поджелудочной железы и кишечника. В одной мутантной линии наблюдается гибель большого числа зародышей от причин, сходных с мекониальным илеусом. Таким образом, эти линии представляют собой идеальные модели для исследования молекулярных основ патогенеза муковисцидоза, а также для разработки и испытания терапевтических мероприятий, включая генотерапию. Как уже отмечалось в предыдущей главе, программы генотерапии муковисцидоза реализуются по крайней мере в 7 центрах США и двух центрах Западной Европы (Великобритания и Франция). Уже успешно осуществлены не только апробации генноинженерных конструкций на мутантных культурах клеток и модельных животных, но начаты и успешно проводятся клинические испытания генотерапевтического лечения муковисцидоза на 70 пациентах. Исследования по генотерапии муковисцидоза, начатые в нашей стране пока проводятся на уровне клеточных культур.

10.4.2 Миодистрофия Дюшенна.

Миодистрофия Дюшенна - сцепленная с полом мышечная дистрофия; выделяют две клинические формы: тяжелую - миодистрофию Дюшенна (МД) и гораздо более мягкую - миодистрофию Беккера (МБ). Ген миодистрофии Дюшенна (DMD) - один из самых крупных известных генов человека, кодирует белок дистрофин, входящий в состав сарколеммы мышечного волокна. При МД дистрофин либо полностью отсутствует, либо деградирует вскоре после синтеза. При форме Беккера, как правило, дистрофин присутствует, хотя и в измененном, чаще всего в укороченном состоянии.

В соответствии с современными представлениями (Ahn, Kunkel, 1993) огромный DMD-ген находится под контролем сложной системы регуляции транскрипции и сплайсинга. По крайней мере, 5 независимых промоторов осуществляют альтернативную специфическую транскрипцию первых экзонов в разных тканях и

на разных стадиях эмбрионального развития. 3 промотора экспрессируют полноразмерную молекулу дистрофина, в то время как 2 осуществляют экспрессию последних доменов взаимоисключающим способом. Высоко консервативные последовательности 6-и экзонов, кодирующих С-конец белка, альтернативно сплайсируются, образуя несколько структурно различающихся форм дистрофина, осуществляющих различные функции. Так, идентифицирована 6.5-кб мРНК, транскрибируемая с DMD-гена и являющаяся, по-видимому, его основным продуктом в не-мышечных тканях, включая мозг (Bar et al., 1990). Соответствующий белок значительно отличается от дистрофина и его уровень в некоторых не-мышечных тканях сопоставим с количеством дистрофина в мышцах. Описан также 4.8-кб транскрипт того же локуса, экспрессирующийся во многих типах тканей, но особенно в Шванновских клетках проводящей нервной системы, где сам дистрофин отсутствует (Blake et al., 1992). Этот белок получил название апо-дистрофин-1. Клонирован и секвенирован еще один 2.2-кб транскрипт DMD-гена, кодирующий аподистрофин-3, экспрессирующийся в позднем эмбриогенезе (Tinsley et al., 1993).

В 60% случаев в гене DMD у больных мальчиков обнаруживаются протяженные делеции, захватывающие от одного до нескольких соседних экзонов и сосредоточенные, обычно, в двух "горячих" районах - в области 5'-конца гена (экзоны 6-19) и в 3'-конце (экзоны 40-53), при этом 30% делеций локализованы в проксимальной части гена и 70% - в дистальной. Проксимальные делеции возникают, по-видимому, в раннем эмбриогенезе и имеют больше шансов стать "семейными" мутациями. Дистальные делеции возникают позднее и чаще встречаются, как изолированные случаи. Считается, что в спорадических случаях повторный риск рождения больного ребенка при проксимальной делеции составляет 30%, тогда как при дистальной - только 4% (Passos-Bueno et al., 1992).

Отмечены популяционные особенности паттерна делеций в разных европейских популяциях, а также в популяциях России и стран СНГ (Baranov et al., 1993). У некоторых пациентов делетирован не только весь ген, но достаточно протяженные соседние области. Очень часто концы делеций локализованы в центральной части дистрофинового гена. Так в интроне 44, протяженностью 160-180 кб, расположено около 30% точек разрыва всех делеций. Показано, что проксимальный конец одной из делеций в интроне 43 расположен внутри транспозон-подобного элемента, принадлежащего TNE-1 семейству ретротранспозонов (Pizzuti et al., 1992). Описан еще один пациент, у которого делеция оканчивается в TNE-1 элементе. Эти элементы, размером 2.3 кб, фланкированные 350-нуклеотидными LTR, повторены около 10 000 раз в геноме человека. Гипотеза нестабильности DMD-гена, вызванная присутствием транспозон-подобного элемента, привлекается также для объяснения нескольких случаев обнаружения различий в молекулярном дефекте у пациентов, принадлежащих одной и той же родословной, то есть имеющих мутацию общего происхождения (Miciak et al., 1992). В 2-х случаях дупликаций из 8, для которых был проведен сиквенс концевых участков, показано присутствие Alu-элементов. В остальных 6 случаях рекомбинация осуществлялась между не-гомологичными последовательностями (Hu, Worton, 1992).

Прямой корреляции между тяжестью течения заболевания и протяженностью делеции не отмечается, но различия между формами Дюшенна и Беккера, в общем случае, связаны с наличием или отсутствием сдвига рамки считывания. Ген дистрофина может быть вовлечен также в другие структурные перестройки -

дупликации, транслокации. Так, примерно 5% мутаций гена дистрофина составляют дупликации и около 35%-точечные мутации, преимущественно, микроделеции (от 1 до нескольких нуклеотидов), а также нонсенс мутации. Относительно высокий процент нонсенс мутаций в гене дистрофина связывают с присутствием большого количества глутаминовых триплетов, мутации в котором часто приводят к образованию стоп-кодона. Крайне редки миссенс мутации. Считается, что в 30% семей с МД и МБ мутации спонтанного происхождения и возникают преимущественно в оогенезе, в остальных семьях это наследственные формы.

Разработаны очень эффективные методы диагностики делеций в DMD-гене, основанные на мультиплексной ПЦР. Одновременное тестирование 18 экзонов + промоторной части гена позволяет выявить до 98% всех крупных делеций гена (Baranov et al., 1993). Для обнаружения гетерозиготного носительства делеций используется метод RT-PCR, то есть изоляция эктопической DMD-мРНК из клеток крови, обратная транскрипция, амплификация кДНК, рестрикционный анализ и секвенирование (Roberts et al., 1991). Для этой же цели применяется метод иммунодетекции мутантного белка в белковом лизате мышц и на гистологических срезах биоптатов мышц (Arahata et al., 1989). При отсутствии идентифицируемой делеции применяют косвенный метод ДНК-диагностики с использованием внутригенных полиморфных сайтов: pERT87-8/TaqI; pERT87-15/BamHI; 124/PstI; 16intron/TaqI, и аллельных вариантов динуклеотидных СА повторов в интронах 49 и 50 - STR-49; STR-50 (Мальшева и др., 1991; 1992; Евграфов, Макаров, 1987; Евграфов и др., 1990).

Серьезную проблему для диагностики гетерозиготного носительства и, следовательно, для медико-генетического консультирования представляют случаи, так называемого, гонадного мозаицизма - присутствие в соматических клетках гонады женщины-носителницы мутаций гена дистрофина, что, в свою очередь, приводит к появлению нескольких генераций (клонов) половых клеток (ооцитов) с мутантным и нормальным генами дистрофина. Предполагается, что такие мутации могут возникать еще на уровне первичных половых клеток, то есть на ранних стадиях внутриутробного развития будущей матери. По ориентировочным оценкам примерно 6-7% всех спорадических случаев являются следствием гонадного мозаицизма у матери. Оценить величину аберрантного клона ооцитов практически не представляется возможным. В связи с этим прогноз в отношении здоровья следующего ребенка в семьях, в которых у матери больного МД не удается определить гетерозиготное носительство, весьма затруднен. Эмпирически определено, что при наличии спорадического случая рождения ребенка с МД и при отсутствии прямых молекулярных доказательств гетерозиготного носительства мутации гена дистрофина у матери риск повторного рождения больного ребенка может достигать 14% (Essen et al., 1992).

У многих пациентов с миопатией Дюшенна при иммуногистохимическом окрашивании мышц обнаруживаются редкие дистрофин-положительные волокна. Однако, при использовании антител с антигенными детерминантами, кодируемыми делетированным участком гена, окрашивания не происходит и это позволяет отвергнуть гипотезу соматического мозаицизма. Наиболее вероятный механизм такого явления - возникновение второй соматической делеции в мышечных клетках, устраняющей сдвиг рамки считывания, вызванный основной делецией. В результате мутантный ген может транскрибироваться с образованием стабиль-

ного, хотя и аномального дистрофина (Klein et al., 1992).

Имеется модельная линия мышей с миодистрофией Дюшенна - mdx. Эта модель была получена в результате отбора мутации, спонтанно возникшей в C57BL/10 линии (Bulfield et al., 1984). В мышцах и в мозгу у мышей этой линии обнаруживается резко сниженное количество Dmd-мРНК, однако дистрофин полностью отсутствует. Несмотря на это, никаких видимых клинических аномалий у mdx-мышей не наблюдаются. Идентифицирована нонсенс мутация в mdx-гене, в результате которой у мышей транслируется лишь 27% дистрофинового полипептида. Обнаружена также сплайсинговая мутация, приводящая к вырезанию экзона 7 DMD-гена у собак (Sharp et al., 1992).

В серии экспериментов на mdx-мышцах доказана принципиальная возможность генокоррекции МД. Появление дистрофина человека в саркомере мышечных волокон mdx-мышей наблюдали после введения ретровирусных или аденовирусных генноинженерных конструкций, содержащих полноразмерную кДНК гена дистрофина (14 кб) или его делетированную, но функционально активную форму, так называемый мини-ген (6.3 кб) (Wells et al., 1992; Cox et al., 1993; Acsadi et al., 1995). После внутривенного введения фрагментов Dmd-гена в составе рекомбинантного аденовируса наблюдали длительное присутствие экзогенной ДНК в скелетных и сердечных мышцах животных (Srataford-Perricaudet et al., 1992). Несмотря на эти очевидные успехи, проблема генноинженерной коррекции МД еще далека от своего решения. До сих пор ведутся оживленные дебаты исследователей о перспективности генной терапии МД по сравнению с клеточной терапией (пересадка здоровых эмбриональных миобластов). Пока не утверждена ни одна программа клинических испытаний генотерапевтического подхода МД (см. Главу IX). Основная сложность проблемы генокоррекции - необходимость обеспечения системы эффективной доставки гена дистрофина в миофибриллы не только скелетных мышц, но, что особенно важно - в мышцы сердца и диафрагмы. В 1995г. исследования по генотерапии МД начаты в рамках программы Геном человека и в нашей стране.

10.4.3 Гемофилия А.

Гемофилия А - сцепленное с полом заболевание, вызванное наследственным дефектом фактора VIII, важнейшего звена в системе свертывания крови (см. Табл. 10.4). Комплекс фактора VIII с молекулярным весом более 1 миллиона состоит из 2-х компонентов. Главный компонент - VIIIС, кодируется геном F8C, локализованным в X-хромосоме. С VIIIС нековалентно связан фактор Виллебранда - VIIIР, кодирующийся аутосомным геном. Фактор Виллебранда стабилизирует фактор VIII и регулирует его активность.

Ген F8C - одним из очень крупных генов человека; содержит 26 экзонов (размером от 69 до 3106 нуклеотидов). Общая длина интронов составляет 177 кб; около 20% этой ДНК приходится на интрон 22 (32.4 кб). мРНК гена F8C размером 9 009 нуклеотидов включает 5'нетранслируемую последовательность (150 п.о.), 3'нетранслируемую последовательность (1 806 п.о.) и кодирующий фрагмент (7053 п.о.). Внутри гена F8 в интроне 22 локализовано еще два других структурных гена неизвестной природы - F8A и F8B, что было обнаружено методами молекулярный анализ. Ген F8A, целиком локализованный в интроне 22 гена F8C, не содержит интронов и транскрибируется в направлении, противоположном фактору VIII (3'-5'). Первый экзон гена F8B также расположен в интроне 22, а следующие

его области распределены до экзона 23-26; транскрибируется он в том же направлении, что и ген F8C (5'-3'). Оба гена экспрессируются во всех тканях (Levinson et al., 1990; Freije, Schlessinger, 1992; Lakich et al., 1993). Интрон 22 оказался необычным и в том отношении, что содержит CpG-островок на расстоянии около 10 кб от экзона 22 - предположительное место локализации бинаправленного промотора для генов F8A и F8B. Оказалось, что на расстоянии примерно 500 кб в 5'-направлении от гена F8 находятся еще 2 транскрибируемые копии гена F8A (Lakich et al., 1993).

Во время процессинга первичного белкового продукта гена F8C от исходного пептида из 2 351 аминокислотных остатка отщепляется последовательность в 335 аминокислотных остатка. В плазме крови фактор VIII существует в виде металлозависимого гетеродимера, состоящего из C-концевой легкой цепи (80 кД) и N-концевой тяжелой цепи (200 кД). Половина всех больных с гемофилией A не имеют фактора VIII, 5% - имеют нормальное количество нефункционирующего белка и в остальных случаях активность белка сохранена, но его количество резко снижено (McGinnis et al., 1993).

Изолированные случаи гемофилии A составляют 30%; 70% - семейные варианты. Показано, что мутации в гене F8 возникают в сперматогенезе в 3 - 5 раз чаще, чем в оогенезе (Rosendaal et al., 1990; Brocker-Vriends et al., 1991). Это означает, что в 80-86% спорадических случаев матери являются носителями мутации, возникшей в зародышевых клетках их отца. Кроме того, около 14% матерей, не являющихся носителями мутации, могут быть соматическими или гонадными мозаиками, так что вероятность повторного рождения больного ребенка у них также повышена.

Около 10% всех идентифицированных мутаций в гене F8 являются делециями одного или нескольких смежных экзонов. Примерно 5% всех мутаций составляют короткие делеции и дубликации гена, остальные мутации - точковые замены (Antonarakis et al., 1995). Почти половина миссенс мутаций идентифицирована в домене A2. Показано, что 35% всех известных мутаций локализовано в CpG динуклеотидах, причем свыше 90% из них представляют собой C-T или G-A транзиции (Cooper, Youssoufian, 1988). Подобные мутации в кодирующих районах встречаются в 42 раза чаще, чем это можно было бы ожидать на основании случайного характера мутагенеза. Для подавляющего большинства мутаций гена F8C характерно практически полное отсутствие "горячих" точек: каждая семья высокого риска по гемофилии A имеет свою собственную мутацию. Исключение составляет группа обнаруженных сравнительно недавно протяженных инверсий интрона 22, захватывающих экзоны 1-22 и полностью блокирующих функцию гена. Такие инверсии, как оказалось, присутствуют в 45% семей с тяжелой формой гемофилии A (Lakich et al., 1993). Причиной инверсий в этой области гена является гомологичная рекомбинация между идентичными последовательностями гена F8A, расположенного в интроне 22 F8C-гена, и другими копиями этого же гена, находящимися на расстоянии 500 кб от 5'конца гена F8 (см. выше).

Помимо инверсий и точечных мутаций в гене FVIII зарегистрированы несколько случаев инсерционного мутагенеза, связанных с перемещением в геноме транспозонподобных элементов типа LINE (см. Главу II). У двух пациентов неродственного происхождения был идентифицирован инсертированный в экзоне 14 F8C-гена длинный элемент LINE-1 (L1) (Kazazian et al., 1988). В обеих семьях это были мутации de novo. L1 последовательности представляют собой специфическое для ге-

нома человека семейство длинных, размером от 2 до 4 кб, повторяющихся элементов, распределенных по всем хромосомам и состоящее, примерно, из 100 000 копий. Было показано что оба L1 элемента, инсертированные в F8C-ген, родственны ретротранспозону, локализованному на хромосоме 22 (Dombroski et al., 1991). В третьей семье инсертированный в интроне 10 F8C-гена L1 элемент не был связан с болезнью. Все 3 L1 элемента имели открытые рамки считывания, а соответствующие реконструируемые аминокислотные последовательности были высоко идентичны друг другу с уровнем гомологии, превышающим 98%. Таким образом, были получены еще одни косвенные подтверждения существования ряда функциональных L1 элементов, кодирующих 1 или несколько белков, необходимых для их ретротранспозиции.

Прямая диагностика протяженных инверсий в гене F8 осуществляется путем блот-гибридизации с ДНК зондом p482.6 с последующей рестрикцией эндонуклеазами Bcl1, Dra1, Nco1 (Lakich et al., 1993). В остальных случаях, в силу отсутствия мажорных мутаций в гене F8C, чаще всего используют косвенные методы молекулярной диагностики. С помощью ПЦР анализируют полиморфные динуклеотидные CA-повторы экзона 13, HindIII полиморфизм в интроне 19; HbaI полиморфизм в интроне 22 и внегенный полиморфизм локуса DXS52 (St14/TaqI) (Асеев и др., 1989; Aseev et al., 1994; Сурин и др., 1990).

Учитывая наличие функционально активной формы белка фактора VIII в плазме крови генноинженерные подходы в терапии этого заболевания направлены на получение в чистом виде полноценного белкового продукта (заместительная терапия), либо на введение в организм большого соответствующей кДНК, обеспечивающей синтез FVIII и его поступление в кровь. Осуществленное 10 лет назад выделение и клонирование кДНК этого гена сделало реальным оба эти подхода. Имеются сообщения о получении трансгенных животных (коз), в геном которых введен ген фактора VIII. Они могут быть использованы как продуценты полноценного белкового продукта. Генная терапия этого заболевания находится на стадии экспериментальных разработок (см.Главу IX). Успешно осуществлена трансдукция фибробластов человека *in vitro* с помощью ретровирусного вектора. Основная проблема в данном направлении заключается в выборе эффективного промотора и подборе клеток, в которых экспрессия гена могла быть достаточно длительной. В настоящее время найдены невирусные промоторы, обеспечивающие эффективную и длительную экспрессию гена фактора VIII *in vivo*. В качестве возможных клеток-мишеней используют мышечные клетки, фибробласты, гепатоциты и клетки эндотелия сосудов. В 1994г. методом направленного мутагенеза (см.Главу VIII) получены трансгенные модели гемофилии А на мышах. Есть все основания считать, что клинические испытания генотерапии этого заболевания начнутся уже в ближайшем будущем.

10.4.4 Гемофилия В.

Гемофилия В - сцепленное с полом заболевание, вызванное наследственным дефектом фактора IX - важного компонента средней фазы внутреннего каскада свертывания крови. Белок (фактор IX) - гликопротеин, состоит из 415 аминокислотных остатков, объединенных в 8 доменов, синтезируется в виде молекулы-предшественника клетками печени. В плазме крови фактор IX находится в виде гетеродимера, состоящего из 2-х полипептидных цепей - легкой (L) и тяжелой (H), ковалентно связанных между собой одним дисульфидным мостиком. Фактор IX

циркулирует в виде неактивного зимогена до тех пор, пока не произойдет протеолитическое высвобождение его активирующего пептида, что позволяет ему принять конформацию активной сериновой протеазы. Его роль в свертывании крови связана с активацией фактора X посредством взаимодействий с ионами кальция, фосфолипидами мембраны и фактором VIII.

Ген фактора IX транскрибируется в гепатоцитах с образованием мРНК размером 1 383 п.о. Для гена F9 характерна высокая частота возникновения мутаций - $4.1 \cdot 10^{-6}$ за поколение. Также как и при гемофилии А мутации значительно чаще возникают в сперматогенезе, чем в оогенезе (Montandon et al., 1992). Считается, что вероятность получения мутации от отца в 11 раз выше, чем от матери. Это означает, что в изолированном случае вероятность гетерозиготного носительства мутации у матери составляет более 80%. Обнаружена четкая корреляция между возрастом отца и вероятностью получения от него новой мутации в гене F9. Так, средний возраст отца в момент рождения дочери - носительницы новой мутации, составляет около 42 лет (King et al., 1992).

К 1994 г идентифицировано около 400 мутаций в гене гемофилии В. Подавляющее большинство из них замены нуклеотидов, приводящие к заменам аминокислот или к образованию стоп-кодонов. Характерно, практически, полное отсутствие выраженных мажорных мутаций и доминирующих областей повышенной частоты мутирования. Только одна мутация - I397T, встретилась в 7 семьях. Около 42% точечных мутаций возникает в CpG динуклеотидах (Bottema et al., 1993). Показано, что частота G-A или C-T транзиций в CpG сайтах в 24 раза выше, чем в других местах гена (Koeberl et al., 1990). Кроме того, в CpG динуклеотидах гена F9 в 7.7 раз чаще возникают трансверсии (A-T, A-C, G-T или G-C). Это объясняется тем, что содержание (G+C) в кодирующих областях F9-гена составляет 40% (Bottema et al., 1991).

В 40% случаев при тяжелых, ингибиторных формах гемофилии В у пациентов обнаруживаются делеции различной протяженности. Около 10% точечных мутаций локализовано в донорных или акцепторных сайтах сплайсинга или создают новые сайты сплайсинга внутри интронов. В одной семье разрушение гена произошло в результате инсерции Alu-элемента в экзон 5 (Vidaud et al., 1993). Описано 13 точечных мутаций в промоторной области гена F9. Именно с такими мутациями связана Лейденовская (Leiden) форма заболевания, при которой к возрасту половозрелости наступает улучшение многих клинических показателей и, в частности, исчезает кровотокающий диатез. Объясняется это тем, что мутации в промоторной области могут приводить к переключению конститутивной экспрессии гена на стероид-гормон-зависимую, нарушая связывание гепатоцитарного ядерного фактора 4 (HNF-4), принадлежащего к суперсемейству транскрипционных факторов для рецепторов стероидных гормонов.

Гемофилия В была использована как модель для выработки стратегии генетического консультирования при моногенных заболеваниях, обладающих выраженной мутационной гетерогенностью (Giannelli et al., 1992). Основой такой стратегии является составление национальных баз данных молекулярных дефектов и специфических методов их диагностики. В частности, основываясь на подобной информации, авторы провели характеристику мутаций в группе из 170 неродственных индивидуумов с гемофилией В шведского и английского происхождения и только в одном случае им не удалось идентифицировать мутацию.

Молекулярная диагностика гемофилии В проводится как непрямыми так и прямыми методами. Непрямая диагностика осно-

вана на анализе методом ПЦР внутригенных полиморфных сайтов: Taq1 (в положении 11 109-11 113); инсерционного полиморфизма в интроне А (рестриктазы Hinf1 и Dde1) ; Taq1 в интроне F в положении 72. Метод ПДФ анализа информативен только у 60-70% всех семей с гемофилией В (Aseev et al., 1994; Сурин и др., 1994). Прямая диагностика гемофилии В включает амплификацию геномных фрагментов гена фактора IX с последующей детекцией ошибок комплементации методом mismatch detection (см.Главу IV) и прямое секвенирование продуктов амплификации (Montadont et al.1990).

Сравнительно небольшие размеры гена, присутствие белкового генопродукта в сыворотке крови и наличие естественных биологических моделей способствовали быстрому прогрессу исследований по генотерапия гемофилии В, которая в настоящее время уже включена в программы клинических испытаний. Успешная трансдукция и коррекция генетического дефекта получена в опытах *in vitro* и *in vivo* на самых различных модельных системах (Culver, 1994; Gerrard et al., 1993). Так, при введении полноразмерной кДНК в составе ретровирусного вектора в первичные культуры кератиноцитов человека наблюдали экспрессию F9 и секрецию биологически активного фактора IX. После трансплантации этих трансдуцированных клеток *nu/nu* мышам человеческий фактор IX в небольшом количестве появлялся в кровотоке и сохранялся там в течение недели (Gerrard et al.,1993). На собаках, страдающих гемофилией В, осуществлена трансдукция гепатоцитов *in vivo* путем прямой инфузии рекомбинантного ретровирусного вектора в портальную вену. При этом наблюдали устойчивую экспрессию фактора IX в течение более 5 месяцев и улучшение биохимических показателей свертываемости крови (Kau et al.,1993). Имеется сообщение об успешной коррекции гемофилии В в Китае в 1992г. Двум больным мальчикам в кожу спины трансплантировали культуру аутологичных фибробластов, предварительно трансдуцированных *ex vivo* рекомбинантной кДНК гена FVIII. Несмотря на определенный скептицизм в оценке этого достижения со стороны специалистов, нет сомнения в том, что успешная генотерапия гемофилии В – событие самого ближайшего будущего.

10.4.5 Болезнь Виллебранда.

Болезнь Виллебранда- аутосомно-доминантное (при некоторых формах рецессивное) заболевание, обусловленное наследственным дефицитом белка VIII_R, родственного фактору VIII_C свертывания крови (см.Гемофилия А) и известного, как фактор фон Виллебранда. Этот большой гликопротеин синтезируется клетками эндотелия, в которых специфическая VIII_R-мРНК составляет 0.3%, и поступает в кровь в виде двух мультимеров с молекулярными весами от 850 кД до 20 миллионов дальтон. Фактор VIII_R осуществляет взаимодействие между стенкой сосудов и тромбоцитами, регулируя их адгезию в местах повреждения эндотелия. Фактор VIII_R участвует также в регуляции синтеза и секреции фактора VIII_C и стабилизирует комплекс фактора VIII.

Различают 7 типов болезни Виллебранда – I, IIA-III и III (Zimmerman, Ruggeri, 1987). При типе I снижена концентрация всех мультимеров в плазме, но их качество не нарушено. Генетически эта форма заболевания подразделяется на рецессивные и доминантные варианты. Типы IIC и III – рецессивны. Тип II характеризуется качественными аномалиями фактора VIII_R, выражающимися в уменьшении способности формировать большие мультимеры, (типы IIA и IIC) или в увеличении ско-

рости их выведения из плазмы (тип IIB).

Ген F8VWF достаточно протяженный и состоит из 52 экзонов, размерами от 40 до 1379 п.о. (Mancuso et al., 1989). Величина интронов варьирует в огромных пределах (от 100 до 20 000 пар нуклеотидов). Сигнальный пептид и пропептид кодируются первыми 17 экзонами, в то время как зрелая субъединица VIIIIR- фактора и 3'нетранслируемая область - остальными 35 экзонами. Внутри гена идентифицированы повторяющиеся последовательности, включая 14 Alu-элементов и полиморфный TСТА повтор размером около 670 п.о. в интроне 40. Районы гены, кодирующие гомологичные домены, имеют сходную структуру. На хромосоме 22q11-q13 обнаружен псевдоген длиной 21-29 кб, соответствующий экзонам 23-34 F8VWF-гена (Mancuso et al., 1991). Идентифицированные в нем сплайсинговые и нонсенс мутации препятствуют образованию функционального транскрипта.

Наибольшее число мутаций идентифицировано при типе II болезни Виллебранда. Подавляющее большинство из них - замены аминокислот, чаще всего происходящие в результате транзиций в CpG динуклеотидах (Cooney et al., 1991; Randi et al., 1991; Donner et al., 1992). Мутации при болезни типа IIA кластеризованы в A2 домене, где предположительно локализован сайт протеолитического отщепления, в то время, как при типе IIB - в домене, обеспечивающим взаимодействие с тромбоцитарным гликопротеиновым комплексом (Ib-IX рецептором). Большая группа мутаций при форме заболевания IIB локализована в сегменте из 11 аминокислот внутри единственного дисульфидного изгиба (loop), соединяющего цистеины в 509 и 695 положениях. При форме заболевания Нормандского типа, мимикриющей гемофилию А, фактор Виллебранда структурно и функционально нормален, за исключением того, что нарушено его взаимодействие с фактором VIII. У таких пациентов действительно идентифицируются миссенс мутации, расположенные в области гена, кодирующей сайты связывания фактора VIIIIR с фактором VIIIC (Mazurier, 1992).

Тип III представляет собой наиболее тяжелую форму заболевания, при которой фактор VIIIIR, как правило, отсутствует. Получены доказательства, что такие пациенты являются гомозиготами или компаундами по нонсенс мутациям, обнаруживаемым в одной дозе у больных типа I (Zhang et al., 1992). При этом же типе заболевания выявлен кластер мутаций со сдвигом рамки считывания, возникающих в результате делеции одного из 6 цитозинов в положении 2679-2684 экзона 18. Именно такая мутация была обнаружена в семье, зарегистрированной впервые фон Виллебрандом в 1926 году. У некоторых членов этой родословной как было установлено недавно, она находилась в компаунде с мутацией P1266L, возникшей в результате рекомбинации между геном F8VWF и псевдогеном (см. выше) (Zhang et al., 1993).

Выбор адекватного метода молекулярной диагностики болезни Виллебранда в значительной мере предопределяется правильностью предшествующей клинической и лабораторной диагностики и результатами медико-генетического консультирования, позволяющей достаточно четко определить характер наследования заболевания в семье высокого риска и установить его форму. К сожалению, достичь этого далеко не всегда возможно, а отсутствие характерных мажорных мутаций значительно снижает эффективность прямой молекулярной диагностики. Вместе с тем, по крайней мере, в некоторых популяциях (Финляндия, Швеция) обнаружены "горячие" точки мутаций, которыми являются экзоны 18 и 42, при типе II болезни Виллебранда (Holmberg et al., 1993). В популяциях России такие "горячие" точки пока не обнаружены, хотя исследования в этом направлении ведутся

(Асеев, Шауи Абдельрхани, 1995). Значительно более перспективной на современном этапе представляется непрямая диагностика. В промоторной части гена, в интронах 15, 17, 23, 40, 41, 49, а также в экзонах 26, 35, 39 идентифицированы многочисленные полиморфные сайты рестрикции с достаточно высоким уровнем полиморфизма. Особенно перспективным для диагностики является полиморфизм интрона 40, представляющий собой две области варьирующих по числу тетрамерных повторов ТСТА на расстоянии 212 п.о. Амплификация этой части интрона 40 с помощью ПЦР, рестрикция *AluI* с последующим электрофоретическим разделением позволяет идентифицировать до 98 аллельных вариантов этого полиморфизма (Merster et al., 1991). Столь выраженный полиморфизм позволяет с высокой эффективностью маркировать мутантную хромосому (ген) и проследить её передачу в потомстве.

Сведения о генотипах болезни Виллебранда в доступной литературе не обнаружены.

10.4.6 Фенилкетонурия.

Фенилкетонурия (ФКУ) – одно из наиболее частых аутосомно-рецессивных заболеваний, обусловленных наследственным дефектом фенилаланингидроксилазы, приводящим при отсутствии своевременной терапии к тяжелой умственной отсталости. В Европе один больной ребенок встречается в среднем среди 10 – 17 000 новорожденных. В Ирландии и Шотландии частота ФКУ достигает 1 на 4500 новорожденных (DiLella et al., 1986). Распространена ФКУ также в Польше и в Белоруссии. В России частота заболевания колеблется в пределах 1 : 8 – 10 000. Очень важна ранняя диагностика ФКУ, так как при своевременном назначении пациенту диеты, не содержащей фенилаланин, умственная ограниченность, как правило, не развивается или имеет очень стертые формы. Разработаны биохимические скринирующие тесты диагностики ФКУ у новорожденных.

Гидроксилирование фенилаланина является достаточно сложным процессом, в котором участвуют, по крайней мере, 3 фермента. Фенилаланингидроксилаза (ПАН), гомополимерный фермент, состоящий из субъединиц с молекулярным весом 52 кД, продуцируется клетками печени и регулирует превращение L-фенилаланина в L-тирозин. Его дефицит приводит к накоплению фенилаланина в сыворотке крови. Гиперфенилаланинемия может возникать также при дефиците дигидроптеридинредуктазы и при дефектах синтеза биоптерина. Однако, эти заболевания, хотя и сопровождаются снижением активности ПАН, значительно отличаются от классической ФКУ и не корректируются диетой, лишенной фенилаланина.

ПАН-ген транскрибируется в гепатоцитах с образованием мРНК размером 2.4 кб. Наиболее распространенный тип мутаций – однонуклеотидные замены (миссенс, нонсенс, мутации в сайтах сплайсинга), причем часто эти мутации являются результатом транзиций в 22-х обнаруженных в ПАН-гене CpG динуклеотидах. Крупных структурных перестроек не найдено, хотя имеется небольшой процент точечных делеций. Отмечается неравномерный характер внутргеновой локализации мутаций (Scriver et al., 1989). Так, наибольшее число миссенс мутаций встречается в центральной части гена: в экзоне 7, кодирующем участок связывания белка с кофактором, где расположено 5 CpG дуплетов, а также в экзонах 9 и 12. Преимущественный район локализации делеций – экзоны 1, 2 и 3.

Внутри ПАН-гена локализовано более 10 полиморфных сайтов рестрикции, причем распределения гаплотипов по этим мар-

керам среди представителей разных рас и этнических групп значительно различаются. Обнаружено сильное неравновесие по сцеплению между определенными мутациями в PАН-гене и гаплотипами по внутригенным сайтам рестрикции. Так, каждая из 5-и наиболее частых в европейских популяциях мутаций ассоциирована только с одним из более, чем 70 гаплотипов по 8 рестрикционным полиморфизмам (EisenSmith et al., 1992). Мажорная в западно-европейских популяциях сплайсинговая мутация в донорном сайте 12-го интрона сцеплена с гаплотипом 3 (DiLella et al., 1986). В то же время другая мутация в экзоне 12 - R408W, наиболее распространенная на востоке Европы, в частности в Белоруссии и России, и не найденная в Японии и Китае, связана с гаплотипом 2 (DiLella et al., 1987). Мажорная в Европе мутация R158Q в 40% сцеплена с гаплотипом 4, наиболее частым среди жителей Японии и Китая. Распространенная в Турции сплайсинговая мутация в интроне 10, приводящая к 9-и-нуклеотидной инсерции, ассоциирована с "южными" гаплотипами 6, 10 и 36.

Сопоставление частот различных гаплотипов по полиморфным сайтам рестрикции и мутаций в PАН-гене в разных популяциях, национальностях и этнических группах позволяет сделать вывод, что большинство из них, или даже все, произошли уже после дивергенции рас. Распространение мажорных мутаций гена PАН в различных популяциях и этнических группах связано с эффектом основателя. По некоторым оценкам эти мутации возникли однократно от нескольких сотен до нескольких тысяч лет тому назад. Однако, в ряде случаев распределение мутаций не может быть объяснено в генетических терминах, сопоставимых с демографической историей. Несомненно доказанными являются примеры независимого и рекуррентного возникновения в разных популяциях таких мутаций, как R261Q или R158Q. Высокие популяционные частоты специфических мутаций в PАН-гене связаны, по-видимому, не только с эффектом основателя и/или с существованием эндогенных механизмов повышенного мутагенеза, но и с преимуществом гетерозигот. Высказано предположение, что носительство PАН - мутаций повышает устойчивость организма к токсическому эффекту охратоксина А, продуцируемого некоторыми видами грибковой плесени (*Aspergillus*, *Penicillium*), развивающимися при хранении зерна и других продуктов (Woolf, 1986). Предполагается, что беременные женщины, гетерозиготные пл PАН -мутациям имеют меньшую вероятность абортa, индуцированного действием этих микотоксинов. Возможно, высокая частота ФКУ в Ирландии и Шотландии частично может быть объяснена мягким и влажным климатом этих стран, способствующем росту таких грибов.

В медицинской практике используется как прямая, так и косвенная диагностика мутаций в PАН-гене. Разработан очень быстрый и эффективный метод ПЦР/StyI-диагностики самой частой в России (более 70%) мутации R408W (Ivaschenko, Varanov, 1993; Иващенко и др., 1993). Диагностика других мажорных мутаций в PАН-гене осуществляется методами ПЦР+АСО, аллель-специфической амплификации (ARMS), методом однонитчатого конформационного полиморфизма (SSCP) (см. Главу IV). При первичном обследовании семьи чрезвычайно удобно использовать три полиморфные нейтральные мутации в кодонах 232, 245 и 385, сцепленные в Кавказских популяциях с определенными ПДРФ-гаплотипами, а значит и со специфическими мутантными аллелями. Каждая из этих мутаций создает новый сайт рестрикции и поэтому их аллельное состояние может быть легко протипировано с помощью амплификации и рестрикции (Kalaydjieva et al., 1991). При анализе семьи, в которой отсутствуют легко

идентифицируемые прямыми методами мутации, молекулярная диагностика может быть проведена с помощью внутригенных полиморфных сайтов рестрикции. Удобен, в частности, Msp1-полиморфизм в 8-м экзоне, анализ которого может быть осуществлен методом ПЦР/рестрикции (Wedmeyer et al., 1993). В последнее время появились данные о наличии высокополиморфных сайтов внутри интронов гена PАН, которые оказались особенно удобными для молекулярного маркирования мутантных аллелей (Goltzov et al. 1994).

Генотипирование ФКУ успешно осуществлена в опытах *in vitro* и в настоящее время находится на стадии экспериментальной разработки (Табл.9.2. Глава IX).

10.4.7 Синдром Леш-Нихана.

Синдром Леш-Нихана - рецессивное сцепленное с полом заболевание, обусловленное наследственной недостаточностью гипоксантин-гуанин фосфорибозилтрансферазы (HPRT) и сопровождающееся тяжелыми поражениями центральной нервной системы. Фермент HPRT участвует в регуляции метаболизма пуринов, контролируя превращение гуанина и инозина в соответствующие рибонуклеотиды. Ген HPRT экспрессируется во всех типах клеток с образованием мРНК размером 654 п.о.. Культивируемые линии клеток, дефектные по HPRT, устойчивы к 8-азагуанину и 6-тиогуанину, и таким образом, могут быть отобраны на соответствующих селективных средах. Гетерозиготные носители мутаций по HPRT-гену могут быть легко выявлены по наличию 2-х типов клеток - устойчивых и чувствительных к 8-азагуанину, в первичной культуре фибробластов или в клетках волосающих луковиц. В большинстве мутантных клеточных линий количество мРНК нормально, а белок отсутствует. У части пациентов хотя и транскрибируется достаточно много мРНК, но в этих молекулах обнаруживаются структурные и функциональные аномалии. В небольшом проценте случаев у больных не удается выявить ни белка, ни мРНК.

В 15% хромосом у больных с синдромом Леш Нихана ген HPRT вовлечен в крупные структурные перестройки, которые могут быть выявлены методами Саузерн или Нозерн блот-гибридизации. Синдром Леш Нихана одно из первых моногенных наследственных заболеваний, для которых была проведена молекулярная идентификация точечных мутантных аллелей. Именно на этой модели впервые был разработан и опробован метод анализа мутаций, основанный на расщеплении РНК-ДНК гибридов рибонуклеазой А в местах негомологичноно спаривания (метод расщепления рибонуклеазой А - см.Главу VI, Gibbs, Caskey, 1987). Комбинация методов блот-гибридизации и расщепления рибонуклеазой А позволяет выявить до 50% мутаций. В настоящее время в гене HPRT найдено более 100 спорадических мутаций, половина которых - однонуклеотидные замены типа миссенс, нонсенс и в сайтах сплайсинга. Около 40% мутантных хромосом имеют структурные аномалии, в том числе крупные делеции, нехватки отдельных экзонов и микроделеции одного или нескольких нуклеотидов. В HPRT-гене, практически, отсутствуют мутации, доминирующие по частоте в каких-либо популяциях. Исключение составляет нонсенс мутация R170TER, которая составляет около 15% всех нуклеотидных замен (Gibbs et al., 1989). Также как и при гемофилиях мутации гена HPRT чаще возникают в сперматогенезе, чем в оогенезе. Вероятность мутирования возрастает с возрастом отца. Идентифицировано 3 HPRT-псевдогена в хромосомах 3, 5 и 11 (Stout, Caskey, 1984).

Описаны редкие случаи синдрома Леш Нихана у гетерози-

готных девочек. При этом, как правило, болезнь развивается вследствие неслучайной инактивации X-хромосомы, не содержащей мутации (Ogasawara et al., 1989). Однако, у 3-х женщин - облигатных носительниц мутаций в HPRT-гене, селективный тест не выявил присутствия мутантных клеток в культивируемых фибробластах и волосных луковицах. В связи с этим высказано предположение, что определенные мутации гена HPRT находятся в неравновесном сцеплении с неидентифицированной летальной мутацией в X-хромосоме, что и приводит к селекции клона клеток только с одной (мутантной или немутантной по гену HPRT) X-хромосомой (Marcus et al., 1992).

Молекулярная диагностика болезни Леш-Нихана возможна прямыми и непрямыми методами. Прямой вариант основан на проведении обратной транскрипции мРНК, ее амплификации, SSCP-анализе одноцепочечных ДНК фрагментов с их последующим секвенированием (см.Глава VI). Косвенная диагностика предусматривает маркирование мутантной хромосомы при помощи полиморфных сайтов (в частности, локуса DXS52 - зонд St14/TaqI).

Как мы уже отмечали (Главы VII,VIII), первая трансгенная животная модель наследственного заболевания человека, сконструированная путем направленного переноса мутаций в культивируемые эмбриональные стволовые клетки, была получена для синдрома Леш-Нихана (Hooper et al., 1987; Kuehn et al., 1987). На этой модели впервые была проведена генокоррекция наследственного дефекта *in vivo*. Эти успехи в значительной степени связаны с существованием селективных сред, позволяющих вести автоматический отбор мутантных клеток. Вообще, синдром Леш-Нихана представляет собой идеальную систему не только для изучения пуринового метаболизма, но и для решения многих теоретических вопросов биологии и медицины (Seegmiller, 1989; Maraus et al., 1993; Boyel et al., 1993). Сложность генокоррекции заболевания, однако, заключается в необходимости обеспечения эффективной доставки гена HPRT (или его кДНК) непосредственно в мутантные нервные клетки. Эта проблема еще не решена. Поэтому реальные клинические программы генотерапии этого заболевания на сегодняшний день отсутствуют (см.Главу IX).

10.4.8 Болезнь Вильсона-Коновалова.

Болезнь Вильсона-Коновалова (БВК) - гепатолентикулярная дегенерация - аутосомно-рецессивное заболевание, обусловленное наследственным дефектом одной из медь-транспортирующих АТФаз. У больных резко снижена концентрация основного медь-содержащего белка плазмы крови - церулоплазмينا и в меньшей степени - цитохромоксидазы, еще одного белка, участвующего в метаболизме меди. Выделяют, по крайней мере, 3 формы БВК (Cox et al., 1972). При редкой атипичной форме, предположительно Германского происхождения, у гетерозигот содержание церулоплазмينا снижено, по крайней мере, в два раза. При двух других, типичных формах - славянской и ювенильной, содержание церулоплазмينا у гетерозигот находится в пределах нормы. Славянский тип БВК характеризуется сравнительно поздним началом и преимущественно неврологической симптоматикой. Ювенильная форма чаще встречается в Западной Европе и ведущими в этиологии заболевания являются печеночные нарушения. Среди евреев-ашкенази встречается БВК с поздним началом и почти нормальным содержанием церулоплазмينا в сыворотке крови больных.

Ген БВК, идентифицированный в 1993г. независимо сразу в

2х лабораториях США, представляет собой медь-транспортирующую АТФазу Р типа с 6-ю металл-связывающими районами. Ген имеет 60% гомологию по нуклеотидному составу с ранее идентифицированным геном АТФ-азы (АТР7А), мутантном при болезни Менкеса (Bull et al., 1993; Petruchin et al., 1993; Tanzi et al., 1993). По аналогии с геном болезни Менкеса, также обусловленной нарушением транспорта меди, ген БВК назван АТР7В. Два пациента с БВК оказались гомозиготными по 7-нуклеотидной делеции в кодирующей области гена АТР7В, что доказывало его идентичность гену БВК (Petruchin et al., 1993). Ген экспрессируется в клетках печени, мозга, почках, лимфоузлах. Типичным для экспрессии АТР7В оказался альтернативный сплайсинг двух и более экзонов центральной части гена (6, 7, 8, 12 и 13).

Кодируемый АТР7В-геном белок содержит несколько мембранных доменов, АТФ-консенсусную последовательность, сайт фосфорилирования и, по крайней мере, 2 медь-связывающих сайта. В мозге, печени, почках и лимфоузлах обнаружены изоформы белка, соответствующие продуктам альтернативного сплайсинга гена АТР7В. Их назначение и функции пока неизвестны. В гене АТР7В идентифицированы полиморфные микросателлитные маркеры, а также около 10 полиморфных сайтов рестрикции. В настоящее время в гене АТР7В идентифицированы более 30 мутаций, в том числе 14 мелких делеций/инсерций, 2 - нонсенс мутации, 15 - миссенс мутаций, 3 - сплайсинговые мутации. Диагностическую ценность для европейцев представляют мутации His1070Gln и Gly1267Lys, зарегистрированные в 28% и 10% всех мутантных хромосом, соответственно (Thomas et al., 1995).

В заключении данного раздела представляется целесообразным кратко рассмотреть другие достаточно частые моногенные заболевания, для которых показана и проводится молекулярная диагностика, в том числе и пренатальная, в других медико-генетических центрах России и, прежде всего, в Лаборатории молекулярной диагностики Института клинической генетики РАМН (Москва).

10.4.9 Адено-генитальный синдром.

Адено-генитальный синдром - (врожденный дефицит 21-гидроксилазы) - достаточно распространенное аутосомно-рецессивное заболевание. Частота "классических" форм 1:10 000 новорожденных, "неклассической" - около 1% в популяции. В зависимости от характера нарушения функции гена и, соответственно клинических проявлений "классическая форма" подразделяется на два варианта: 1. летальная сольтеряющая форма; 2. нелетальная - вирулизирующая форма, связанная с избытком андрогенов (Morel, Miller, 1991).

В локусе 6p21.3, внутри сложного супергенетического комплекса HLA идентифицированы два тандемно расположенных 21-гидроксилазных гена - функционально активный СУР21В и псевдоген - СУР21А, неактивный вследствие делеции в 3-м экзоне, инсерции со сдвигом рамки считывания в 7-м экзоне и нонсенс мутаций - в 8-м экзоне. Ген и псевдоген разделены смысловой последовательностью гена С4В, кодирующей 4-й фактор комплемента. Оба гена состоят из 10 экзонов, имеют длину 3,4 кб и отличаются только по 87 нуклеотидам. Высокая степень гомологии и тандемное расположение указывают на общность эволюционного происхождения этих генов. Любопытно отметить, что такие же тандемно расположенные гены 21-гидроксилазы (называемые также P450c21) обнаружены и у других млекопитаю-

ших, причем у мышей, в отличие от человека, активен только ген CYP21A, но не CYP21B, тогда как у крупного рогатого скота функционально активны оба гена.

Белок- 21-гидроксилаза (P450c21- микросомальный цитохром 450) обеспечивает превращение 17-гидроксипрогестерона в 11-дезоксикортизол и прогестерона - в дезоксикортикостерон. В первом случае возникает дефицит глюкокортикоидов и, прежде всего, кортизола, что в свою очередь стимулирует синтез АКТГ, и ведет к гиперплазии коры надпочечников (вирилирующая форма). Нарушение превращения прогестерона в дезоксипрогестерон ведет к дефициту альдостерона, что в свою очередь нарушает способность почек удерживать ионы натрия и приводит к быстрой потере соли плазмой крови (соль теряющая форма).

Как и в случае гемофилии А, наличие рядом с кодирующим геном гомологичной ДНК последовательности зачастую ведет к нарушениям спаривания в мейозе и, как следствие этого, к конверсии генов (перемещения фрагмента активного гена на псевдоген), либо к делеции части смыслового гена. В обоих случаях функция активного гена нарушается. На долю делеций приходится около 40% мутаций, на долю конверсий - 20% и примерно 25% составляют точечные мутации. Согласно отечественным данным в случае наиболее тяжелой сольтеряющей формы АГС, на долю конверсий приходится более 20% мутантных хромосом, на долю делеций - около 10% (Evgrafov et al., 1995).

Непрямая диагностика АГС возможна с помощью типирования тесно сцепленных с геном CYP21B аллелей HLA A и HLA B генов, а также аллелей гена HLA DQA1. Прямая ДНК диагностика АГС основана на амплификации с помощью ПЦР отдельных фрагментов генов CYP21B и CYP21A, их рестрикции эндонуклеазами NaeIII или RsaI и анализе полученных фрагментов после электрофореза (Evgrafov et al., 1995).

10.4.10 Спинальная мышечная атрофия.

Спинальная мышечная атрофия (СМА) - аутосомно-рецессивное заболевание, характеризуется поражением моторных нейронов передних рогов спинного мозга, в результате чего развиваются симметричные параличи конечностей и мышц туловища. Это - второе после муковисцидоза наиболее частое летальное моногенное заболевание (частота 1: 6 000 новорожденных).

СМА подразделяется на три клинические формы. Тип I. Острая форма (болезнь Верднига-Гоффмана), проявляется в первые 6 месяцев жизни и приводит к смерти уже в первые два года; Тип II. Средняя (промежуточная) форма, пациенты не могут стоять, но обычно живут более 4-х лет; Тип III. Ювенильная форма (болезнь Кугельберга-Веландера) - прогрессирующая мышечная слабость после 2-х лет. Все три формы представляют собой аллельные варианты мутаций одного гена SMN (survival motor neurons), картированного в локусе D5S125 (5q13) и идентифицированного методом позиционного клонирования (см.Главу III) в 1995г (Lefebvre et al. 1995). В этой пока единственной работе показано, что ген SMN размером всего 20 000 п.о.состоит из 8 экзонов. мРНК этого гена содержит 1 700 п.о. и кодирует ранее неизвестный белок из 294 аминокислотных остатков с молекулярным весом 32 КилоДальтона.

Ген дублирован. Его копия (возможно вариант псевдогена) располагается несколько ближе к центромере и отличается от гена SMN наличием 5-и точечных мутаций, позволяющих отличить оба гена путем амплификации экзонов 7 и 8 и их исследованием методом SSCP анализа (см.Главу IV). Ген назван svCD541, по аналогии с первоначальным вариантом названия для

теломерной копии, т.е. часть гена SMN, tBCD541. Ген cBCD541 экспрессируется, но в отличие от гена SMN его сДНК подвергается альтернативному сплайсингу с утратой экзона 7. Отсутствие гена SMN (tBCD541) у 93% больных (213 из 229), его разорванная (interrupted) структура у 13 обследованных пациентов (5.6%) и наличие серьезных мутаций у оставшихся 3-х больных дали основание именно данную теломерную копию гена считать ответственной за заболевание. Существенно отметить, что центромерная копия гена обнаружена у 95.05% больных, тогда как отсутствует она только у 4,4% 4 пациентов 0.

В непосредственной близости от теломерного конца гена SMN идентифицирован еще один ген - ген белка-ингибитора запрограммированной гибели нейронов (neuronal apoptosis inhibitory protein -NAIP). При тяжелых клинических формах СМА (Тип I), обусловленных делециями, по-видимому, нередко происходит утрата гена NAIP.

Согласно гипотезе авторов СМА возникает при гомозиготном состоянии мутаций (обычно-делеций) в гене SMN, при этом различия между формами СМА определяются двумя основными факторами: 1. числом копий гена cBCD541 (две - в случае Типа I и четыре (возникающих вследствие конверсии между SMN и cBCD541) - в случае Типа III), 2. наличием или отсутствием гена NAIP. Среди всех обследованных СМА-больных не обнаружены случаи одновременной делеции обоих гомологичных генов - SMN (tBCD541) и cBCD541, что указывает, по мнению авторов, на то, что такая aberrация должна проявляться как доминантная леталь еще в эмбриогенезе.

Некоторые положения этой, безусловно, основополагающей работы французских авторов, по-видимому, еще требуют уточнения, однако, уже сейчас она сделала возможной прямую молекулярную диагностику СМА у 98,6% больных. С этой целью проводится амплификация экзона 7, который отсутствует у подавляющего большинства больных. Нормальный экзон 7 (ген SMN) дифференцируют от мутантного варианта (ген cBCD541) с помощью SSCP анализа. При необходимости возможна косвенная диагностика - ПЦР анализ динуклеотидных (CA) повторов ДНК локусов D5S125; D5S112; D5S127; ПДРФ-анализ с фланкирующими ДНК-зондами MU, 105-153RA; 153-6741 GT.

10.4.11 Атаксия Фридрейха.

Атаксия Фридрейха (АФ) - сравнительно редкое (1 : 22 -25 000) аутосомно-рецессивное заболевание, характеризующееся прогрессивной дегенерацией нервных клеток мозжечка. Ген АФ не идентифицирован, но достаточно точно картирован на хромосомных (9q13-q21) и физических картах ДНК-маркеров. Наиболее тесное сцепление гена АФ показано для локуса D9S5 (зонд 26P). Сконструированы космидные библиотеки и составлены подробные физические карты области 4-го геномной ДНК хромосомы 9, включающей локус D9S7 и, предположительно, ген АФ. Определено положение гена АФ по отношению к другим фланкирующим молекулярным маркерам (Fujita et al., 1991; Wilkes et al., 1991). В настоящее время известно, по крайней мере, 5 таких ДНК маркеров: GS4, MCT-112, GS2 -дистальные и микросателлитные маркеры FD1 (на расстоянии 80 кб) и MLS1 (на расстоянии 150 кб) - проксимальные. Изучены особенности аллельного полиморфизма этих систем для различных популяций Западной Европы. Для всех 5 молекулярных маркеров выяснены гаплотипы, сцепленные с заболеванием. Гаплотипы обоих микросателлитных маркеров оказались в абсолютном генетическом

неравновесии с АФ, что доказывает их весьма близкое расположение на генетической карте по отношению к мутантному гену АФ (Pianese et al., 1994).

Диагностика АФ пока возможна только непрямими методами. ПДРФ анализ с помощью ДНК-зондов на дистальные полиморфные сайты, либо ПЦР анализ полиморфизма проксимальных по отношению к гену АФ микросателлитных маркеров MLS1 или FD1.

Нами рассмотрены лишь некоторые моногенные наследственные болезни, условно разделенные на три подгруппы, исходя, главным образом, из того насколько они изучены с молекулярно-генетических позиций, их актуальности для пренатальной диагностики и в какой мере они важны для медико-генетической службы нашей страны. Более того, исторически сложилось так, что именно такие заболевания как муковисцидоз, миодистрофия Дюшенна, гемофилия А, фенилкетонурия, то есть 4 0социально наиболее значимые, раньше других генных болезней стали предметом детального молекулярного анализа в нашей лаборатории и в других медико-генетических центрах и научно-практических подразделениях России (см. Баранов, 1991, 1994; Baranov, 1993; Евграфов, Макаров, 1987).

Естественно, что рассмотренными нозологиями отнюдь не исчерпывается список тех болезней, которые являются объектами молекулярных исследований в нашей стране. Например, из обзора выпали такие моногенные 4 0болезни как гиперхолестеринемия, гемоглобинопатии, дефицит альфа-1 антитрипсина, митохондриальные болезни. 4 0Для многих из них разработаны и широко применяются эффективные методы молекулярной диагностики, ведутся исследования по генотерапии. 4 0Мы не касались также работ проводимых, 4 0главным образом, 4 0в возглавляемой профессором Е.И.Шварцем лаборатории молекулярной диагностики ПИЯФ РАН и посвященных молекулярному анализу мультифакториальных заболеваний, 4 0таких как диабет, гипертония, ишемия сердца. Результаты этих 4 0исследований 4 0будут, по-видимому, предметом следующих обзоров и монографий.

ГЛАВА I

СТРУКТУРА И МЕТОДЫ АНАЛИЗА ДНК.

Раздел 1.1 Общие представления, центральная догма, генетический код.

Универсальная генетическая субстанция или "энциклопедия жизни", ДНК, содержит информацию, необходимую для синтеза белков и нуклеиновых кислот, присутствующих во всех типах клеток как про- так и эукариот. Дезоксирибонуклеиновые кислоты (ДНК) - это нитевидные молекулы, состоящие из четырех расположенных в варьирующем порядке нуклеотидов: пуринов - аденина и гуанина, и пиримидинов - цитозина и тимина, соединенных в полинуклеотидную цепь с остовом из чередующихся остатков сахара - дезоксирибозы, и фосфата. Последовательность нуклеотидов ДНК или пар оснований составляет информационную емкость молекулы, определяя порядок синтеза и аминокислотную последовательность белков в соответствии с универсальным для всех живых существ трехбуквенным - триплетным, генетическим кодом (Табл.1.1). Дезоксирибонуклеиновые кислоты представляют собой единственный тип молекул, способных к самовоспроизводству или репликации, что и обеспечивает преемственность

генетической информации в ряду поколений. Записывается последовательность ДНК слева направо (5' - 3') первыми главными буквами соответствующих нуклеотидов, являющихся одновременно единицами измерения молекулы. Размеры ДНК могут меняться в гигантских пределах от нескольких нуклеотидов до миллиардов пар оснований (п.о.). В качестве единиц измерения размеров ДНК используются также килобазы (kb) и мегабазы (mb) - последовательности, соответствующие тысячи и миллиону пар оснований, соответственно.

ДНК могут существовать как в виде однонитевых, так и в виде двухнитевых молекул. Двухнитевые или двухцепочечные молекулы образуются за счет химического комплементарного спаривания между аденином и тиминном (А - Т) и между гуанином и цитозином (Г - Ц). Эти водородные связи между парами нуклеотидов достаточно непрочные, так что цепи ДНК могут легко диссоциировать - разделяться, и ассоциировать - соединяться, при изменении температуры или солевых концентраций. При каждом цикле ассоциации - диссоциации или, как еще говорят, отжиге - плавлении, будет точно воспроизводиться двухнитевая структура - дуплекс, устойчивость которого определяется соответствием нуклеотидных пар. Наиболее устойчивы структуры, представленные полностью комплементарными нитями ДНК. Процесс образования дуплексов носит название гибридизации. Способность к комплементарному спариванию оснований - одно из самых замечательных свойств ДНК, определяющих возможность ее саморепликации и точного выбора специфических участков активации молекулы в процессе считывания генетической информации. Это свойство широко используется в молекулярной биологии для поиска и идентификации нужных последовательностей в огромных молекулах ДНК при использовании в качестве зондов ее сравнительно небольших меченых фрагментов.

У человека большая часть ДНК - 3.2 миллиарда пар оснований, находится в ядрах клеток в виде 46 плотно упакованных, суперскрученных за счет взаимодействий с ядерными белками структур, называемых хромосомами. Сравнительно небольшая часть ДНК - около 5%, пристывает в митохондриях - органеллах цитоплазмы, обеспечивающих процессы дыхания и энергетического обмена клеток эукариот. В большинстве соматических клеток ДНК представлена в двух копиях - по одной в каждой хромосоме. Таким образом, в клетках присутствуют 23 пары хромосом, 22 из которых гомологичны друг другу - аутосомы, и одна пара (X и Y) - половые хромосомы. Наличие Y хромосомы определяет мужской пол особи. При записи нормального кариотипа индивидуума указывается общее число хромосом и тип половых хромосом. Таким образом, нормальный кариотип мужчины - 46,XY, а женщины - 46,XX. В процессе гаметогенеза происходит случайное расхождение гомологичных хромосом в мейозе и в каждой зрелой половой клетке - гамете, остается только 23 хромосомы, то есть гаплоидный набор хромосом. При этом в каждой гамете сохраняется лишь одна половая хромосома - гоносома. В яйцеклетках это X хромосома, тогда как сперматозоиды с равной вероятностью несут как X, так и Y хромосому, то есть пол будущей особи детерминирован геномом сперматозоида. При оплодотворении диплоидный набор хромосом восстанавливается. В соответствии с современными представлениями геном человека состоит из 25 хромосом, 22 из которых аутосомы, 2 половые хромосомы и одна митохондриальная. В каждой клетке присутствует порядка 1000 митохондрий, а в каждом митохондрионе содержится около 10 кольцевых митохондриальных хромосом, сходных с хромосомами бактерий. Таким образом, в клетках присутствует около 1000 копий митохондриальных хро-

мосом.

В хромосомах эукариот ДНК находится в двухнитевой форме, что обеспечивает возможность ее точной репликации при каждом цикле деления клетки. Одна нить кодирующая или смысловая, комплементарная ей нить - антисмысловая. Декодирование информации, заключенной в молекуле ДНК, или процесс транскрипции, осуществляется за счет избирательного синтеза молекул РНК, комплементарных определенным участкам ДНК, так называемых первичных РНК транскриптов. Транскрибируемые участки ДНК носят название генов. Рибонуклеиновые кислоты (РНК) по своей структуре очень сходны с молекулами ДНК. Они также состоят из четырех нуклеотидов, только одно из пиримидиновых оснований - тимин, заменено на урацил и в сахарозном остове вместо дезоксирибозы представлена рибоза. Молекулы РНК существуют только в однонитевой форме, но могут образовывать дуплексы с молекулами ДНК. После синтеза молекулы РНК претерпевают достаточно сложную модификацию - процессинг. При этом происходят изменения в концевых участках молекул и вырезаются области, гомологичные интронам - некодирующим частям гена. Этот процесс называется сплайсингом. В результате из первичных РНК транскриптов образуются молекулы информационной или матричной РНК (мРНК), представляющие собой непрерывную последовательность нуклеотидов, гомологичную только экзонам - смысловым участкам гена. Молекулы мРНК в виде рибонуклеопротеиновых гранул выходят из ядра в цитоплазму и соединяются с рибосомами, где происходит процесс трансляции - синтез полипептидной цепи. Трансляция мРНК происходит в точном соответствии с генетическим кодом, согласно которому последовательность из трех нуклеотидов РНК - кодон, соответствует определенной аминокислоте или сигналу терминации синтеза полипептидной цепи (Табл.1.1). Реализация генетического кода осуществляется с участием 20-ти типов транспортных РНК (тРНК), единственных нуклеиновых кислот, содержащих в своем составе наряду с нуклеотидами одну из аминокислот. тРНК имеют кленообразную форму, в хвостовой части молекулы расположена определенная аминокислота, в точном соответствии с последовательности из трех нуклеотидов в области, называемой антикодоном. Прохождение мРНК по рибосоме является сигналом приближения к рибонуклеопротеидному комплексу той тРНК, у которой последовательность нуклеотидов в антикодоне комплементарна кодирующему триплету мРНК. Таким образом транспортируется соответствующая аминокислота и осуществляется последовательный синтез полипептидной цепи. Митохондрии имеют свою автономную систему белкового синтеза: рибосомальные РНК, мРНК и транспортные РНК.

Генетический код универсален для всех живых существ - это одно из его главных свойств. Небольшие отличия в структуре кода найдены только для митохондриальной ДНК. Так в митохондриальном генетическом коде стоп кодонами являются триплеты АГА и АГЦ, кодирующие аргинин в ядерной ДНК (Табл.1.1). Универсальность генетического кода служит наиболее веским аргументом в пользу гипотезы об едином источнике возникновения жизни на земле и о филогенетическом родстве всех видов живых существ. Кроме того, именно это свойство обеспечивает возможность прочтения в любых модельных клеточных системах искусственно введенной генетической информации, сконструированной из фрагментов ДНК разного видового происхождения. Таким образом, вся геновая инженерия основана на универсальности генетического кода. Другим свойством генетического кода является его вырожденность, заключающаяся в том, что все аминокислоты кроме двух кодируются несколькими

вариантами триплетов. Действительно, из 64 возможных комбинаций нуклеотидных триплетов РНК три соответствуют терминирующим кодонам - ochre, amber и opal, остальные варианты (61) кодируют 20 аминокислот, причем триплеты, кодирующие одну и ту же аминокислоту, как правило, различаются по третьему нуклеотиду в кодоне. Таким образом, зная нуклеотидную последовательность кодирующего участка ДНК, можно однозначно прогнозировать аминокислотную последовательность соответствующего полипептидного фрагмента, тогда как одна и та же аминокислотная последовательность может кодироваться различным образом. При этом, число возможных вариантов кодирующих ДНК резко возрастает с увеличением длины полипептида.

На следующем этапе полипептидные цепи транспортируются к специфическим органеллам клетки и модифицируются с образованием зрелого функционально активного белка. В некоторых случаях информация с молекул РНК может обратно транскрибироваться в молекулы ДНК. В частности, при обратной транскрипции мРНК образуются молекулы комплементарной ДНК - кДНК, в которой в зависимости от полноты процесса представлены частично или полностью все смысловые кодирующие последовательности гена. Рассмотренная схема реализации однонаправленного потока информации ДНК-РНК-Белок составляет основу центральной молекулярно-биологической догмы - рис.1.1.

Более детально с процессами репликации, транскрипции, процессинга и трансляции можно ознакомиться в многочисленных руководствах по молекулярной биологии, цитологии и генетике (Стент, Кэлиндер, 1981; Зенгер, 1987; Льюин, 1987).

1.2 Выделение ДНК, ее синтез и рестрикция.

ДНК может быть изолирована из любого типа тканей и клеток, содержащих ядра. Этапы выделения ДНК включают быстрый лизис клеток, удаление с помощью центрифугирования фрагментов клеточных органелл и мембран, ферментативное разрушение белков и их экстрагирование из раствора с помощью фенола и хлороформа, концентрирование молекул ДНК путем преципитации в этаноле. Из 1 грамма сырой ткани или из 10^9 клеток обычно получают 2 миллиграмма ДНК. У человека ДНК, чаще всего, выделяют из лейкоцитов крови, для чего собирают от 5 до 20 мл венозной крови в стерильную пробирку с раствором, препятствующим коагуляции (например, с глюкециром или гепарином). Затем отделяют лейкоциты и разрушают клеточные и ядерные мембраны добавлением буферных растворов, содержащих денатурирующие агенты. Наилучшие результаты при выделении ДНК дает применение протеиназы-К с последующей фенол - хлороформной экстракцией разрушенных белков. ДНК осаждают в этаноле и растворяют в буферном растворе. Оценку качества экстрагированной ДНК проводят на основании измерения оптической плотности раствора ДНК в области белкового и нуклеинового спектров поглощения. В чистых образцах ДНК соотношение $A(260)/A(280) > 1.8$. В противном случае процедуру очистки необходимо повторять, так как для успешного использования и хранения ДНК белки должны быть полностью удалены. Более подробно с методами выделения и очистки ДНК из различных тканей можно ознакомиться в работах и руководствах, приведенных в конце книги (Маниатис и др., 1984; Дейвис, 1990; Горбунова и др., 1991).

В процессе сложного и многообразного функционирования различные участки хромосом и ДНК претерпевают разнообразные регулируемые и, в основе своей, обратимые изменения. Эти модификации осуществляются с помощью специальных белков - фер-

ментов. Описание ферментативного аппарата репликации, транскрипции, репарации – системы защиты и восстановления поврежденных участков ДНК, рекомбинации, то есть обмена участками гомологичных хромосом и ДНК, далеко выходит за рамки нашего изложения. Мы кратко ознакомимся только с двумя классами ферментов ДНК – полимеразы и рестриктазы, особенно важными для понимания основ современной молекулярной диагностики.

Ферменты, осуществляющие синтез ДНК, называются ДНК-полимеразами. И в бактериальных клетках, и в клетках эукариот содержатся три различные формы ДНК-полимераз, все они обладают синтезирующей активностью и способны удлинять цепи ДНК в направлении 5' – 3', последовательно наращивая по одному нуклеотиду к 3'-ОН концу, причем точность синтеза определяется специфичностью спаривания оснований. Таким образом, для работы ДНК-полимеразы необходима однонитевая матричная ДНК с двухнуклеотидным участком на 3'-конце молекулы. Кроме того, в среде должны присутствовать четыре типа трифосфатов (dATP, dCTP, dGTP и dTTP) – молекул, состоящих из основания –А,С,Г или Т, сахара – дезоксирибозы (d) и трех фосфатных остатков (P). В клетках эукариот репликацию осуществляет ДНК-полимераза альфа, а в клетках *E. coli* – ДНК-полимераза β . ДНК-полимеразы обладают различными активностями, в том числе и экзонуклеазной в направлении 3' – 5', что позволяет им исправлять – репарировать, дефекты, допущенные при подборе комплементарных оснований. ДНК-полимераза β *E. coli* способна инициировать репликацию в месте разрыва ДНК и замещать гомологичный участок в двойной цепи ДНК. Это свойство используется для введения в ДНК меченых нуклеотидов методом ник-трансляции.

Открытие бактериальных ферментов, обладающих эндонуклеазной активностью – рестриктонных эндонуклеаз или рестриктаз, значительно продвинуло исследование структуры ДНК и возможности генноинженерного манипулирования с молекулами ДНК. *In vivo* эти ферменты участвуют в системе распознавания и защиты "своих" и уничтожении чужеродных ДНК. Рестриктазы узнают специфические последовательности из 4 – 6, реже 8 – 12 нуклеотидов в двухцепочечной молекуле ДНК и разрезают ее на фрагменты в местах локализации этих последовательностей, называемых сайтами рестрикции. Количество образующихся рестриктонных фрагментов ДНК определяется частотой встречаемости сайтов рестрикции, а их размер – характером распределения этих сайтов по длине исходной молекулы ДНК. Чем чаще расположены сайты рестрикции, тем короче фрагменты ДНК после рестрикции. В настоящее время известно более 500 различных типов рестриктаз бактериального происхождения, причем каждый из этих ферментов узнает свою специфическую последовательность. Рестриктазы выделяют путем биохимической очистки из различных видов бактерий и обозначают тремя буквами, соответствующими первым трем буквам латинского названия вида бактерий, и римской цифрой, соответствующей хронологии открытия этого фермента у данного вида бактерий. В зависимости от частоты встречаемости сайтов рестрикции в молекуле ДНК различают три класса рестриктаз часто-, средне- и редкощепящие. Естественно, что рестриктазы, узнающие длинные специфические последовательности (8-12 п.о.), как правило, являются редкощепящими (например *NorI*), а узнающие короткие (4-5 п.о.) – частощепящими (*TaqI*, *EcoRI*).

Сайты рестрикции могут быть использованы в качестве генетических маркеров ДНК. Действительно, образующиеся в результате рестрикции фрагменты ДНК могут быть упорядочены по

длине путем электрофореза в агарозном или полиакриломидном геле, и тем самым может быть определена их молекулярная масса, а, значит, и физическое расстояние между сайтами. Напомним, что обычным методом выявления ДНК в геле, также как и РНК, является ее специфическое окрашивание, чаще всего этидиумом бромидом, и просмотр геля в проходящем ультрафиолете. При этих условиях места локализации ДНК имеют красную окраску. При использовании для рестрикции нескольких эндонуклеаз с последующим электрофоретическим анализом перекрывающихся аддитивных по длине фрагментов ДНК можно добиться полного упорядочивания сайтов узнавания для каждого из ферментов относительно друг друга и каких-то иных маркеров, присутствующих в исследуемой молекуле ДНК. Процесс этот называется физическим картированием и является обязательным элементом анализа плазмидных, вирусных, бактериальных ДНК и относительно небольших фрагментов ДНК эукариот. На рис.1.2. представлен простейший пример такого картирования в том случае, когда в исследуемой молекуле ДНК присутствует по одному сайту рестрикции для двух эндонуклеаз. После обработки исходной ДНК отдельно каждой из рестриктаз образуется два фрагмента, соответствующих по длине расстоянию от концов молекулы ДНК до сайтов рестрикции. При совместной обработке обеими эндонуклеазами на электрофореграмме появляется новый фрагмент, размер которого соответствует расстоянию между сайтами рестрикции. Очевидно, что эти данные еще не позволяют однозначно определить положение сайтов рестрикции по отношению к концам молекулы ДНК. Однако, достаточно знать расположение хотя бы одного маркера для того, чтобы произвести точное физическое картирование исходной молекулы ДНК независимо от количества локализованных в ней сайтов рестрикции.

При обработке тотальной геномной эукариотической ДНК, в частности ДНК человека, часто- или среднешепящими эндонуклеазами образуется так много фрагментов различной длины (в среднем, порядка 1 миллиона), что их не удастся разделить с помощью электрофореза, то есть не удастся визуально идентифицировать отдельные фрагменты ДНК на электрофореграмме. После электрофореза рестрицированной геномной ДНК получается равномерное окрашивание по всей длине геля - так называемый шмер. Идентификация нужных фрагментов ДНК в таком геле возможна только путем гибридизации с мечеными ДНК-зондами. Это достигается при помощи метода блот-гибридизации по Саузерну.

1.3. Блот-гибридизация по Саузерну, гибридизация *in situ*.

Одним из наиболее эффективных методов идентификации определенных молекул ДНК среди электрофоретически разделенных фрагментов является ставший уже классическим метод блот-гибридизации по Саузерну, по фамилии автора Edward Southern, предложившего данный метод в 1975г. Последовательные этапы данного метода представлены на Рис.1.3. Суть метода заключается в том, что геномная ДНК подвергается рестрикции одной или несколькими рестриктазами, после чего образующиеся фрагменты разделяются по молекулярному весу в агарозном или акриламидном гелях. Затем ДНК подвергается денатурации *in situ* и переносится с геля на плотный носитель (обычно нитроцеллюлозный фильтр или нейлоновую мембрану). Сам перенос (блоттинг) осуществляется за счет действия капиллярных сил, электрического поля или вакуума. Фиксированную на фильтре ДНК гибридизуют с радиоактивно меченым ДНК или РНК зондом. Методом автордиографии определяют положение искомого фрагмента геномной ДНК на электрофореграмме. Блот-гибридиза-

ция - высоко чувствительный метод идентификации специфических последовательностей ДНК. При достаточно длительной экспозиции (в течение несколько дней) и при высокой удельной радиоактивности ДНК-зонда (более 10^9 расп/ мин/микроГ) этот метод позволяет выявлять менее, чем 0,1 пикоГ ДНК. Так при использовании зонда размерами в несколько сот оснований уникальная последовательность в 1 000 п.о. может быть выявлена в 10 микроГ геномной рестрицированной ДНК в виде отдельной полосы на радиоавтографе после его экспозиции в течение 12 часов. Метод позволяет работать и с очень короткими олигонуклеотидными зондами (20 п.о.), однако требует особенно хорошего мечения и длительной экспозиции фильтра. Необходимость работы с чистыми препаратами ДНК, применение высокомеченых радиоактивных зондов, длительность и трудоемкость всей процедуры делают её весьма дорогостоящей. Тем не менее, в ряде случаев и сегодня метод не потерял своего значения в том числе и для диагностики генных болезней. В последнее время для этих целей нередко используют различные варианты нерадиоактивного мечения или окраску ДНК азотно-кислым сербром.

Гибридизация с меченым ДНК-зондом препаратов ДНК или РНК, нанесенных капельно на твердый матрикс без предварительной рестрикции и электрофореза носит название дот- или слот-гибридизации в зависимости от конфигурации пятна ДНК на фильтре, округлой или продолговатой, соответственно. На рис.1.3 также изображены последовательные этапы этих методов. Попутно отметим, что метод гибридизации ДНК-зондов с электрофоретически разделенными молекулами РНК носит название Нозерн блот, тогда как Вестерн блот или иммуноблот - это связывание электрофоретически разделенных белков, фиксированных на фильтрах, с мечеными антителами. Название этих методов - дань уважения молекулярных генетиков профессору Саузерну, внесшему неоценимый вклад в разработку экспериментальных подходов, используемых для анализа ДНК.

В ряде случаев для проведения гибридизации с ДНК зондами не требуется предварительного выделения и очистки ДНК. Процедуру гибридизации можно проводить не только на геле, на фильтрах или в растворе, но и на гистологических или хромосомных препаратах. Этот метод носит название гибридизации *in situ*. Вариант метода, при котором в качестве зондов используются препараты ДНК или РНК, меченые флюорохромами, называется FISH (fluorescein *in situ* hybridization). Меченый ДНК-зонд наносят на препараты дифференциально окрашенных и подготовленных для гибридизации (денатурированных) метафазных хромосом. Предварительная обработка хромосом направлена на облегчение доступа зонда к геномной ДНК. Важное значение имеет также подбор условий, максимально способствующих процедуре гибридизации. После отмывки несвязавшихся молекул ДНК и нанесения светочувствительной эмульсии (при использовании радиоактивной метки), либо проведения соответствующей обработки (при использовании биотин- или флюоресцеин-меченых ДНК зондов) места локализации последовательностей ДНК, комплементарных использованному ДНК-зонду, можно непосредственно наблюдать в микроскоп в виде характерных точек над соответствующими участками определенных хромосом (Рис.1.4).

Гибридизация *in situ*, является одним из наиболее эффективных методов картирования комплементарных ДНК-зонду последовательностей ДНК на хромосомах. Эта методика особенно эффективна при исследовании распределения по геному повторяющихся последовательностей ДНК, клонированных последовательностей ДНК анонимного происхождения, при определении не

только хромосомной принадлежности, но и внутри-хромосомной локализации уникальных генов в тех случаях, когда имеются соответствующие ДНК-зонды. При этом разрешающая способность метода может достигать нескольких хромосомных бэндов. Согласно последним данным, в экспериментах на специально приготовленных и растянутых интерфазных хромосомах человека разрешающая способность метода FISH может достигать 50 kb, что составляет всего около 1/20 величины среднего хромосомного бэнда. Проблемы взаиморасположения клонированных фрагментов ДНК даже в пределах одного хромосомного локуса также с успехом решаются методом FISH.

Гибридизация *in situ* между молекулами РНК и кДНК-овыми зондами, проводимая на гистологических препаратах, является одним из наиболее эффективных методов анализа тканеспецифического распределения и внутриклеточной локализации мРНК (Манк, 1990). Подробно с этим и другими современными методом молекулярного и цитогенетического анализа, а также с их многочисленными модификациями и вариантами можно ознакомиться в серии работ, руководств и обзоров (Маниатис и др., 1984; Дейвис, 1990; Sambrook et al., 1989).

1.4 ДНК-зонды, клонирование, векторные системы.

ДНК-зондом может служить любая однонитевая ДНК ограниченного размера, используемая для поиска комплементарных последовательностей в молекуле большего размера или среди пула разнообразных молекул ДНК. В ряде случаев в качестве зондов используют искусственным образом синтезированные олигонуклеотидные последовательности ДНК, размер которых обычно не превышает 30 нуклеотидов. Зондом также могут служить выделенные из генома последовательности ДНК. Однако значительно чаще такие последовательности предварительно клонируют, чтобы иметь возможность получать их в любое время и в неограниченном количестве. Клонирование предполагает встраивание (инсерцию) чужеродной экзогенной ДНК в векторную молекулу ДНК, обеспечивающую проникновение этой конструкции в бактериальные клетки хозяина (Рис 1.5). Химерные молекулы ДНК, составленные из фрагментов разного происхождения, носят название рекомбинантных ДНК. В качестве клонирующих векторов используют модифицированные плазмиды, фаги, космиды, ретровирусы, а также некоторые другие генетические конструкции. Размеры клонированных ДНК-зондов составляют от сотен до нескольких тысяч нуклеотидов, что определяется, главным образом, способностью вектора удерживать чужеродный фрагмент ДНК. Особенно широко применяют в качестве векторов плазмидную ДНК.

Плазмиды – это небольшие кольцевые двухцепочечные молекулы ДНК, которые могут присутствовать в различном числе копий в бактериальных клетках. Открытие плазмид связано с изучением генетической природы антибиотикоустойчивости. Оказалось, что именно плазмиды могут нести гены, сообщающие клеткам устойчивость к различным антибиотикам, и потеря чувствительности инфекционных бактерий к их действию как раз и происходит за счет отбора тех штаммов, в которых имеются плазмиды с соответствующей генетической информацией. Заметим, что присутствие плазмиды в бактериальной клетке вовсе не обязательно для обеспечения ее жизнедеятельности, так как при отсутствии антибиотиков в среде обитания бактерий штаммы, не содержащие плазмид, вполне жизнеспособны. Плазмиды имеют автономную систему контроля репликации, обеспечивающую поддержание их количества в клетке на определенном уровне –

от одного до нескольких сотен плазмидных геномов на клетку. Обычно для клонирования выбирают плазмиды с ослабленным контролем репликации, что позволяет им накапливаться в клетке в большом числе копий. Конструирование плазмидных клонирующих векторов заключается во внесении изменений в систему контроля репликации и в добавлении или вырезании генов антибиотикоустойчивости или удобных для клонирования иных генетических элементов: специфических сайтов рестрикции, инициации и регуляции транскрипции и т.п. Чаще всего для клонирования используют плазмиды pBR322, ColE1 или их производные.

Кольцевую молекулу плазмидной ДНК можно легко перевести в линейную форму путем единичного разрыва в месте локализации уникального сайта рестрикции. Присоединение (встраивание, инсерция) фрагмента чужеродной ДНК к концам линейной молекулы осуществляется с помощью специфических ферментов-лигаз, после чего гибридная плаزمида вновь принимает кольцевую форму. Разработаны достаточно простые и эффективные методы трансформации бактерий, то есть искусственного введения плазмид в бактериальные клетки. При этом, присутствующие в плаزمиде гены антибиотикоустойчивости используют в качестве маркеров трансформированных бактерий для их отбора на соответствующих селективных средах. При размножении трансформированных бактерий происходит увеличение числа копий инсертированного фрагмента ДНК. Таким образом, этот чужеродный для бактерий генетический материал может быть получен, практически, в любых количествах. Выделенная из бактерий плазмидная ДНК или изолированный из плазмиды инсертированный фрагмент могут быть в дальнейшем использованы в качестве ДНК-зондов.

Для некоторых целей в качестве клонирующих векторов оказалось удобнее использовать фаги - бактериальные вирусы. Фаговая ДНК существует только в линейной форме, поэтому при ее рестрикции образуются два фрагмента, которые сшиваются с чужеродной ДНК с образованием химерного фага. Чисто технически эта операция проще, чем инсерция в плазмиду. Однако, размеры встраиваемой ДНК ограничены пакующей способностью головки фага. Поэтому при конструировании вектора вырезают последовательности фаговой ДНК, не имеющие критического значения для жизнеобеспечения фага. Такой бактериофаг может существовать только в том случае, если в него встроена чужеродная ДНК, по размерам сопоставимая с вырезанной фаговой ДНК. Наиболее удачные конструкции векторов были получены на основе фага лямбда - лямбда gt10, лямбда gt11, лямбда Zap.

Многие проблемы молекулярной генетики успешно решаются с использованием экспрессионных векторов, содержащих в своем составе регуляторные последовательности, обеспечивающие синтез чужеродных белков в клетках хозяина. Так в случае лямбда gt11 фаги могут быть выращены в, так называемых, репликативных условиях, обеспечивающих экспрессию инсертированной ДНК. Так как обычно ДНК встраивают в район локализации маркерного гена, позволяющего вести селекцию химерных фагов, то экспрессироваться будет слитый белок, в котором часть полипептидной цепи будет соответствовать маркерному белку, а часть цепи будет транслироваться в соответствии с информацией, заключенной во встроеном фрагменте ДНК. Этот белок может быть идентифицирован путем детекции фрагмента маркерного белка либо с помощью антител к специфическим участкам, кодируемым чужеродной ДНК.

В последнее время большое распространение получило клонирование в космидах - конструкциях, объединяющих в себе преимущества плазмид и фагов. Космиды получены на основе

плазмид, но в них введены генетические элементы фага лямбда, отвечающие за упаковку ДНК в фаговой частице. Такие векторы могут существовать не только в виде плазмид, но и в виде фаговых частиц *in vitro*. Космиды обладают большей клонирующей способностью по сравнению с плазмидными и фаговыми векторами и могут нести до 40-45 тысяч пар оснований инсертированной ДНК. Все вышеперечисленные векторы используются для клонирования в прокариотических системах.

Векторы, пригодные для направленного переноса в эукариотические клетки, конструируют на основе прокариотических или дрожжевых плазмид - единственных плазмид, найденных в клетках эукариот, а также используют различные эукариотические вирусы, чаще всего ретровирусы, аденовирусы или аденоассоциированные вирусы. При использовании плазмид в качестве клонирующих векторов в них вводят вирусные последовательности, ответственные за начало репликации. Введение векторов в эукариотические клетки часто осуществляют путем ко-трансформации, то-есть одновременно вводят плазмиду и сегмент чужеродной ДНК. Векторные последовательности, введенные в клетки эукариот, могут сохраняться там в течение нескольких дней в виде суперскрученных кольцевых молекул - эписом. В редких случаях возможна интеграция экзогенной ДНК в хромосомную ДНК. В этих случаях введенные последовательности устойчиво сохраняются в геноме клеток хозяина и наследуются по менделевскому типу (см. Глава VIII).

Для клонирования субхромосомальных фрагментов ДНК, содержащих целые гены, разработана система дрожжевых минихромосом. Искусственные дрожжевые хромосомы (YAC - artificial yeast chromosomes) конструируют на основе плазмидных векторов, содержащих в своем составе известные центромерные и теломерные последовательности хромосом дрожжей, необходимые для поддержания и репликации векторов в клетках хозяина. Такие системы способны удерживать фрагменты чужеродной ДНК размером в несколько сотен тысяч и даже миллионов пар оснований.

Остановимся коротко на методах введения векторов в клетки хозяина. Но прежде всего, определим основные термины. Как уже упоминалось, введение плазмидной ДНК в бактериальные клетки называется трансформацией. Если перенос генов осуществляется с помощью фага, то говорят о трансдукции. Процесс введения экзогенной ДНК в эукариотические клетки называется трансфекцией. Все эти методы основаны на подборе условий, облегчающих прохождение плазмидной или фаговой ДНК через клеточные и ядерные мембраны. Для повышения проницаемости мембран используют два разных подхода. В первом случае проводят обработку векторной ДНК и клеток хозяина буферными растворами, повышающими проницаемость клеточных и ядерных мембран (метод кальций-фосфатной преципитации, DEAE-декстран-опосредованная трансфекция). Во втором случае используют краткосрочное физическое воздействие на клетки для создания в мембранах микропор, проходимых для макромолекул ДНК (метод электропорации - воздействие высоковольтным электрическим полем, "бомбардировка" частицами золота и т.п.). Более подробно проблемы векторов и методы генетической трансформации (трансдукции) рассмотрены в Главе IX. Вопросам молекулярного клонирования также посвящена обширная литература (Гловер, 1988; 1989; Шишкин, Калинин, 1992; Маниатис и др., 1984; Дейвис, 1990; Sambrook et al., 1989).

1.5 Геномные и к-ДНК-овые библиотеки генов, их скрининг.

Рассмотрим более подробно методы выделения и идентификации фрагментов ДНК, необходимых для анализа или для использования в качестве ДНК-зондов. Основным источником этих фрагментов являются искусственным образом сконструированные библиотеки генов, в которых осуществляют поиск или скрининг нужных последовательностей ДНК разными методами в зависимости от специфических особенностей этих последовательностей. Библиотека генов это полный набор клонированных перекрывающихся фрагментов ДНК, полученных в результате рестрикции или механического разрезания тотальной ДНК, выделенной из какого-либо специфического источника. В зависимости от происхождения ДНК различают геномные и кДНК-овые библиотеки генов. Для конструирования геномных библиотек используют ДНК, выделенную из тканей, культур клеток, из отдельных хромосом или из их фрагментов. При создании кДНК-овых библиотек выделяют тотальную мРНК из тканей или культивируемых клеток, в которых заведомо экспрессируются интересные исследователя гены. На следующем этапе методом обратной транскрипции (РНК-ДНК) синтезируют кДНК. Затем её разрезают и упаковывают в выбранный для клонирования вектор. Схема конструирования геномных и кДНК-овых библиотек представлена на рис.1.6. Как видно на схеме в геномных библиотеках присутствуют не только кодирующие последовательности генов, но также несмысловые внутригенные последовательности - интроны и межгенные участки ДНК, причем удельный вес некодирующих фрагментов ДНК значительно выше. кДНК-овые библиотеки состоят только из кодирующих - экзонных, областей генов. На более удобный размер инсертируемой ДНК сопоставим со средним размером гена млекопитающих и составляет 15 - 25 тысяч пар оснований (kb). Оптимальный по размеру набор перекрывающихся последовательностей геномной ДНК человека получается после ее переваривания частощепящими рестриктазами *Sau3a* или *Mbo1*. Информационная емкость каждой библиотеки, то есть количество клонов с различными инсертированными фрагментами ДНК, определяется размерами исходного генома и необходимостью присутствия каждой его последовательности хотя бы в одном клоне. Поэтому достаточно представительные геномные библиотеки млекопитающих обычно содержат не менее $8 \cdot 10^5$ - 10^6 различных клонов.

Чаще библиотеки конструируют на основе фаговых или космидных клонирующих векторов, так как в таком виде легче хранить большие количества химерных ДНК. Для создания библиотек генов человека особенно удобны векторы, полученные на основе фага лямбда, такие как EMBL3 или EMBL4. Пакующая способность этих векторов от 9 до 23 kb, они содержат много удобных клонирующих сайтов, так что для инсерции ДНК могут быть использованы разные рестриктазы. Кроме того, эти векторы не содержат последовательностей плазмид, наиболее часто используемых для клонирования : pBR322 и ColE1. Это позволяет проводить отбор нужных клонов с помощью фаговой ДНК, не вырезая предварительно инсертированный в нее фрагмент. Для создания библиотек клонов, содержащих большие районы ДНК, используется технология искусственных дрожжевых хромосом -YAC. Последние представляют собой крупные (до 1 млн п.о.) фрагменты геномной ДНК человека, сшитые с центромерными районами хромосом дрожжей. После идентификации в таких библиотеках нужных клонов с инсертированными фрагментами чужеродных ДНК последние могут быть субклонированы в фаговых или космидных библиотеках.

Скрининг библиотек проводят путем гибридизации на фильтрах с олигонуклеотидными, кДНК-овыми или любыми иными

ДНК-зондами, а также с помощью антител, если библиотека сконструирована на основе экспрессионного вектора (рис.1.7). Для этого химерные фаги, составляющие библиотеку, высевают на плотно растущий в чашках Петри газон бактерий таким образом, чтобы образовались отдельные литические бляшки в результате инфицирования клеток одним рекомбинантным фагом. Все культуры дублируют путем отпечатка - реплики, на другие чашки Петри. Затем на исходные культуры накладывают фильтры и переносят на них растущие и лизированные колонии, проводят их разрушение, фиксацию белков и ДНК на фильтре и блот гибридизацию с меченым ДНК-зондом или иммуноблот с мечеными антителами (для экспрессионных библиотек). После отмывки фильтров от несвязавшихся меченых зондов и радиоавтографии на рентгеновской пленке проявятся темные пятна в местах локализации колоний, содержащих в инсертированном фрагменте ДНК последовательности, комплементарные зонду, или специфические антигены. Отбор положительных колоний фагов на дублированных культурах производят именно в тех местах, где произошло потемнение пленки. Чтобы избежать возможного загрязнения, отобранные колонии размножают и вновь подвергают скринингованию. Обычно, инсертированную ДНК изолируют из бактериофага и субклонируют в плазмидном векторе, позволяющем наращивать большие количества этой ДНК.

1.6 Секвенирование последовательностей ДНК.

Следующими этапами анализа отобранного и клонированного ДНК фрагмента являются его физическое картирование и определение нуклеотидной последовательности, то есть секвенирование. Методология секвенирования достаточно проста и заключается в том, чтобы получить серию комплементарных молекул ДНК, различающихся по длине на одно основание. На практике, однако, определение нуклеотидной последовательности протяженных молекул ДНК представляет собой весьма трудоемкую задачу. Существует два основных метода секвенирования: химическое - метод Максама-Гильберта и дидезоксисеквенирование - метод Сэнджера. В первом случае используют химическое расщепление ДНК по одному основанию, во втором - синтезируют нужную цепь ДНК *in vitro*, специфически останавливая синтез на заданном основании.

Чаще при секвенировании используют метод Сэнджера, так как он более надежный и простой в исполнении. Принцип данного метода показан на рис.1.8. На первом этапе ДНК денатурируют, чтобы получить одностранные молекулы. Затем добавляют секвенирующий праймер - искусственно синтезированную олигонуклеотидную последовательность, комплементарную определенному участку исходной молекулы ДНК. Создают условия для гибридизации праймера, то есть для образования двухцепочечного участка, и инициируют синтез ДНК, добавляя в реакционную смесь ДНК-полимеразу и трифосфаты - dATP, dCTP, dGTP и dTTP, один из которых является радиоактивным. Синтез ведут в четырех параллельных пробирках, в каждую из которых добавляют один из специфических дидезоксинуклеотидов или терминаторов - ddATP, ddCTP, ddGTP или ddTTP. При встраивании ddNTP на место соответствующего нуклеотида синтез ДНК прекращается. Таким образом, в каждой из пробирок получают набор различающихся по длине радиоактивно меченых фрагментов ДНК с одним и тем же специфическим для данной пробирки дидезокситерминатором на конце молекулы. После одновременного электрофоретического разделения этих фрагментов на четырех соседних дорожках и радиоавтографии, размер синтезированных фрагментов

может быть определен, а значит и определена локализация дидезоксинуклеотидов и порядок соответствующих им нуклеотидов в исходной молекуле ДНК (Рис.1.9). На каждом секвенирующем геле может быть определена первичная последовательность всего около 500 пар оснований. В своем первоначальном варианте этот метод является достаточно трудоемким и дорогостоящим. Достаточно заметить, что цена одного звена (шага) в цепи ДНК в Международной Программе "Геном человека" еще несколько лет назад оценивалась в 1\$.

В качестве модификаций метода Сэнджера используют предварительное клонирование ДНК в векторах, сконструированных на основе фага M13, для получения протяженных однонитчатых участков ДНК, которые могут быть непосредственно секвенированы без денатурации и праймирования. Очень эффективной оказалась система векторов mp8 и mp9, позволяющих вести секвенс одновременно в обоих направлениях и прочитывать участки большей протяженности. Интенсивно разрабатывается методология автоматического ДНК секвенирования. Особенно перспективным для массового секвенирования в автоматическом режиме оказалось применение меченых различными флюорохромами дидезоксинуклеотидов. В этом варианте каждому из нуклеотидов соответствует свой цвет полосы в геле, который легко учитывается автоматически. Этот метод нашел особенно широкое применение в реализации программы Геном человека. По последним данным, разработка автоматических секвенаторов позволила снизить стоимость одного звена до 0,5\$ и резко повысить эффективность этого процесса. Так, на 30 автоматических секвенаторах фирмы Applied Biosystems за одну неделю работы в 1994г. можно было просеквенировать до 1 млн пар оснований.

В последние годы активно разрабатываются принципиально новые, более эффективные и экономичные методы секвенирования. Особенно перспективным на сегодняшний день представляется метод секвенирования путем гибридизации исследуемой последовательности ДНК с набором олигонуклеотидов, так называемой олигонуклеотидной матрицей (Мирзабеков, 1990; Барский и др., 1994; Southern, et al, 1992). Наиболее удобными для этой цели являются наборы матриц - чипы, в ячейках которых пришиты (иммобилизованы) октануклеотиды. Суть данного подхода в том, что пришивается набор всех возможных вариантов перестановок из 4-х стандартных нуклеотидов (А, Г, Ц, Т) определенной длины, при этом в случае октануклеотидов количество возможных вариантов нуклеотидов составляет 65536. Секвенируемый фрагмент ДНК метят радиоактивным фосфором и гибридизуют с октануклеотидной матрицей. Фрагмент ДНК гибридизуется только с теми октануклеотидами, последовательности которых комплементарны его участкам. Таким образом, определяется набор всех возможных октануклеотидов, присутствующих в исследуемом фрагменте ДНК. После этого, при помощи специальной компьютерной обработки упорядочивается расположение этих октамеров в изучаемом фрагменте ДНК. Этот перспективный метод позволяет значительно ускорить время секвенирования за счет автоматизации процесса.

1.7 Полимеразная цепная реакция.

До недавнего времени гибридизация с ДНК-зондами и клонирование являлись единственными способами поиска и выделения специфических геномных или к-ДНК-овых последовательностей ДНК с целью их дальнейшего исследования. Не говоря уже о большой трудоемкости этих методов, они имеют ряд принципиальных недостатков и ограничений. Во-первых, исследуемые

фрагменты значительно превосходят по длине ДНК-зонды и слишком велики для прямого молекулярного анализа. Концы этих последовательностей не могут быть выбраны произвольно, так как определяются наличием соответствующих рестрикционных сайтов в исследуемой молекуле ДНК. Во-вторых, для проведения успешной рестрикции и гибридизации необходимо большое количество хорошо очищенной высокомолекулярной геномной ДНК (не менее 10 микроГ на одну реакцию). Такое количество ДНК обычно получают из 3–5 мл крови. Часто это количество оказывается слишком большим, если учесть, что речь идет о детях и о тяжело больных людях. Кроме того, необходимость немедленного использования собранной крови для выделения ДНК или хранения ее до использования при -20 C затрудняет исследование нетранспортабельных пациентов или родственников, проживающих на отдаленном расстоянии от исследовательских центров. В-третьих, для геномной гибридизации, как правило, необходимы радиоактивные ДНК-зонды с высокой удельной активностью, не менее 10^{18} – 10^{19} имп./мин/мкг., причем они должны быть использованы в течение очень короткого периода после их приготовления. Кроме того, работа с радиоактивным материалом требует соблюдения необходимой техники безопасности и наличия специально оборудованного изотопного блока, что возможно лишь для некоторых диагностических центров. В-четвертых, необходимость длительной экспозиции автографов значительно удлиняет время получения результатов, что также ограничивает использование методов блот-гибридизации для пренатальной диагностики плода, специфика которой во многом определяется сроком беременности.

Предложенный в 1983г. американским исследователем Карри Муллисом, удостоенным за это изобретение Нобелевской премии в 1993г., альтернативный метод анализа геномной ДНК – метод полимеразной цепной реакции (ПЦР), явился эпохальным открытием молекулярной биологии нашего века. Метод ПЦР или специфической амплификации ДНК позволяет избирательно синтезировать *in vitro* относительно небольшие участки ДНК, длиной от нескольких десятков до нескольких сотен пар нуклеотидов, реже до 1000 – 2000 п.о., используя в качестве матрицы любые образцы ДНК, содержащие амплифицируемую последовательность. Необходимым условием для проведения ПЦР является знание нуклеотидной последовательности амплифицируемой области ДНК, так как специфический выбор этого участка осуществляют путем гибридизации матричной ДНК с двумя искусственно синтезированными праймерами – олигонуклеотидными последовательностями ДНК, длиной, обычно, от 15 до 30 п.о., комплементарными 3'-концам амплифицируемого участка на смысловой и антисмысловой нитях ДНК, соответственно. Таким образом, расстояние между праймерами определяет длину синтезируемых молекул. В качестве матрицы для синтеза может быть использован любой тип ДНК – геномная ДНК отдельных индивидуумов различных видов про- и эукариот; ДНК, выделенная из культур клеток, бактериальных клонов, библиотек генов или из других источников. Для проведения специфической амплификации не требуется больших количеств матричной ДНК и, в принципе, достаточно даже одной молекулы (Li et al., 1988). Успех в разработке метода ПЦР в значительной мере связан с использованием в качестве фермента, обеспечивающего синтез ДНК, термофильной ДНК полимеразы, выделенной из бактерий, живущих в горячих источниках, и поэтому устойчивой к действию высоких температур (Kogan et al., 1987).

Принципиальная схема полимеразной цепной реакции показана на Рис.1.10. На первом этапе исследуемая двухнитевая

матричная ДНК переводится в однонитевую форму путем ее нагревания в течение нескольких минут до температуры, превышающей +95 - +98 С. Дальнейшая схема проведения ПЦР достаточно проста и заключается в чередовании циклов (1) гибридизации или отжига ДНК с праймерами, (2) синтеза последовательности, комплементарной матричной ДНК, и (3) денатурации образовавшихся двухнитевых структур. При гибридизации, достигаемой за счет понижения температуры реакционной смеси до +30 - +50 С, происходит образование двухнитевого участка ДНК в строго специфичных областях, комплементарных праймерам. При температуре, оптимальной для работы ДНК-полимеразы - +60 - +70 С, начинается синтез ДНК в направлении 5' - 3' с двухнитевого участка, образованного праймерами. Затем при нагревании раствора примерно до +80 - +90 С синтез прекращается и двухнитевой участок между матричными и вновь синтезированными молекулами ДНК расплавляется (денатурирует). В первом цикле олигопраймеры гибридизуются с исходной матричной ДНК, а затем в последующих циклах и с вновь синтезированными молекулами ДНК по мере их накопления в растворе. В последнем случае синтез ДНК заканчивается не при изменении температурного режима, а по достижении ДНК-полимеразой границы амплифицированного участка, что и определяет, в конечном счете, размер вновь синтезированного участка ДНК с точностью до одного нуклеотида.

Таким образом, при каждом цикле происходит двухкратное увеличение числа синтезированных копий выбранного для амплификации участка ДНК и содержание продуктов амплификации в растворе нарастает в геометрической прогрессии. Каждая из процедур цикла определяется двумя параметрами - температурой реакции и ее длительностью, меняющейся в диапазоне от десятков секунд до 1 - 3 минут, так что полный цикл длится от одной до несколько минут. За 25 - 30 циклов число синтезированных копий ДНК достигает нескольких миллионов. ПЦР обычно проводят в автоматическом режиме, используя для этого специальные приборы - термоциклеры или амплификаторы ДНК. Такой прибор позволяет задавать нужное количество циклов и выбирать оптимальные временные и температурные параметры для каждой процедуры цикла из большой и часто непрерывной шкалы возможных вариантов. В техническом исполнении ПЦР очень проста. Все компоненты реакции - матричную ДНК, олигопраймеры, смесь трифосфатов и термофильную ДНК полимеразу добавляют в специфический буфер (часто одномоментно), добавляют небольшой объем вазелинового масла для предотвращения раствора от высыхания, затем помещают пробирку с реакционной смесью в автоматический термоциклер и выбирают необходимую программу циклической смены температуры и длительности каждого шага реакции. Общий объем реакционной смеси, обычно, составляет от 10 до 50 - 100 микролитров. Выбор оптимального режима работы определяется длиной и специфичностью амплифицируемого участка. Следует подчеркнуть, что, реально, для каждой полимеразной реакции в зависимости от типа праймеров, длины амплифицированного фрагмента, особенностей его первичной нуклеотидной структуры, а также от типа используемой термофильной ДНК-полимеразы оптимальные режимы температуры и состав амплификационной смеси могут существенно варьировать и зачастую подбираются эмпирически.

С помощью ПЦР можно непосредственно исследовать места локализации предполагаемых мутаций или полиморфных сайтов, а также изучать наличие любых других специфических особенностей ДНК. Подбор системы олигопраймеров производят на основании анализа нуклеотидной последовательности амплифици-

руемого участка. Разработаны различные варианты автоматического поиска последовательностей ДНК, использование которых в качестве праймеров оптимизирует процедуру амплификации специфического участка геномной ДНК. Для того, чтобы гибридизация проходила строго специфично, праймеры не должны содержать повторяющихся последовательностей ДНК. Мы уже отмечали, что для проведения ПЦР достаточно минимального количества ДНК, вплоть до одной молекулы. Кроме того, различные органические компоненты клеток, такие как белки, липиды, углеводы, фрагменты клеточных органелл или мембран заметно не препятствуют процессу амплификации. Поэтому в качестве источника матричной ДНК может быть использован любой, даже деструктурированный биологический материал, сохранивший в своем составе достаточно полный набор фрагментов исходных молекул ДНК. Для специфической амплификации наряду с очищенной ДНК используют высушенные на фильтровальной бумаге пятна крови, небольшие кусочки тканей, например, такие как ворсины хориона, смывы из полости рта, луковицы корней волос, культуры клеток и сливы среды с клеточных культур, содержащие не прикрепившиеся и не давшие роста клетки, соскребы с цитогенетических препаратов и, что особенно удивительно, полученные при патологоанатомическом анализе и длительно хранившиеся фиксированные ткани (Higuchi, R. et al., 1988). Более подробные сведения о постановке ПЦР можно получить в ряде специальных руководств и инструкций.

Существуют различные модификации ПЦР, которые используются в зависимости от конкретных целей проведения реакции или от характера последующего молекулярного анализа амплификатов. Так, для трудноамплифицируемых участков ДНК (содержащих различные повторяющиеся последовательности или необычные структурные элементы), а также в тех случаях, когда матричная ДНК присутствует в следовых количествах, ПЦР проводят в два раунда, используя в качестве матричной ДНК на втором этапе амплификации, или как еще говорят при доамплификации, продукты ПЦР, синтезированные в первом раунде. Часто в этих случаях для повышения специфичности праймирования используют систему, так называемых вмонтированных (nested) праймеров, то есть при доамплификации в качестве праймеров выбирают последовательности, локализованные внутри амплифицируемого в первом раунде участка ДНК.

В ряде случаев удобно проводить мультиплексную ПЦР, то есть одновременную амплификацию нескольких участков матричной ДНК. Можно получать меченые продукты ПЦР, добавляя в реакционную смесь меченые трифосфаты. Особого внимания заслуживает возможность проведения ПЦР с молекулами кДНК. На основе этой реакции разработаны методы анализа экспрессии генов и получения больших количеств кДНК. Уместно заметить, что реакцию амплификации можно проводить не только в растворах, но и непосредственно на хромосомных препаратах, при этом в случае использования меченых нуклеотидов продукты амплификации будут гибридизоваться и выявлять комплементарные им участки ДНК на хромосомах (метод PRINS - polymerase reaction in situ). До настоящего времени доступными амплификации были участки ДНК, не превышающие по длине 5 kb. В последнее время, благодаря внесению ряда кардинальных усовершенствований (особый подбор праймеров, использование сразу двух различных ДНК полимераз, трицинового буфера, специального температурного режима полимеразных циклов), возможно проведение амплификации фрагментов ДНК, достигающих 35 kb. В частности, таким образом удалось амплифицировать плазмиду со вставкой 8.5 kb, равной целому геному вируса HIV 1 (Cheng et

al., 1994). И это еще не предел. В дальнейшем мы неоднократно будем возвращаться к описанию различных модификаций ПЦР, так как этот метод по праву стал один из основных в молекулярной диагностике наследственных болезней (Erlich, 1993; Rolfs et al., 1993; RT-PCR, Methods and Applications, 1991).

Возможность очень точного и специфичного выбора участка ДНК для амплификации, небольшие размеры синтезируемых молекул, их огромное количество чрезвычайно облегчают молекулярный анализ продуктов ПЦР. Как правило, амплифицированную ДНК можно непосредственно наблюдать в проходящем ультрафиолете в виде красной полосы после электрофоретического концентрирования и обычного окрашивания геля этидиумом бромидом. Кроме того, возможна идентификация этих молекул путем блот-гибридизации со специфическими олигонуклеотидными зондами. При наличии сайтов рестрикции в амплифицированных участках ДНК после их обработки эндонуклеазами и электрофореза количество и положение окрашенных полос на геле изменяется в соответствии с положением этих сайтов (Рис. 1.11). Путем электрофореза могут быть выявлены небольшие отклонения в размерах синтезированных молекул, которые возникают за счет делеций или инсерций нескольких нуклеотидов в исходной матричной молекуле ДНК; конформационные изменения в одонитевых молекулах ДНК, возникающие при замене оснований; а также структурные изменения в дуплексах между нормальными и мутантными вариантами амплифицированных фрагментов ДНК. И, наконец, возможно определение полной нуклеотидной последовательности синтезированных молекул с идентификацией всех мутаций в исследуемом районе матричной ДНК (см. Главу IV). Разработаны варианты ПЦР, при которых синтезируются преимущественно одонитевые фрагменты ДНК, что в последующем значительно облегчает их секвенирование. В дальнейшем мы более подробно рассмотрим методы генотипирования мутаций с помощью ПЦР, позволяющие производить полный молекулярный анализ определенных участков ДНК у отдельных индивидуумов. При этом отпадает необходимость визуализации малых количеств ДНК путем их гибридизации с мечеными геномными или кДНК-зондами. Это не означает, конечно, что методы блот-гибридизации и клонирования потеряли свою актуальность. Без их использования невозможна молекулярная идентификация генов, предшествующая разработке методов молекулярной диагностики с использованием ПЦР.

ГЛАВА VI.

ЭКСПЕРИМЕНТАЛЬНЫЙ АНАЛИЗ ЭКСПРЕССИИ ГЕНОВ.

Раздел 6.1 Дифференциальная активность генов, выбор адекватных биологических моделей.

Молекулярная идентификация гена, нарушение работы которого приводит к развитию наследственного заболевания, создает предпосылки для дальнейшего более подробного анализа генетических и биохимических основ патогенеза и разработки на этой базе наиболее эффективных методов лечения. Начальные этапы решения поставленной задачи включают в себя исследование механизмов тканеспецифической экспрессии и регуляции активности генов в нормальных клетках, оценку клинического

выражения различных типов нарушений гена, выявление первичного биохимического дефекта, а также сопоставление молекулярных основ работы генов в нормальных и мутантных клетках.

Естественно, что в различных тканях организма экспрессируется не все, а лишь определенные группы генов. Исключение составляют лишь так называемые гены домашнего хозяйства (*house-keeping genes*), генопродукты которых обеспечивают жизнедеятельность всех типов клеток (см. Главу II). По весьма ориентировочным оценкам в тканях млекопитающих и человека работают в среднем около 2-3% всех генов, в клетках печени - основной биохимической лаборатории организма - около 5%, тогда как в клетках мозга - примерно 9-10% (Корочкин, 1977). Это означает, что в различных соматических клетках эукариот транскрибируется от 5 до 20 тысяч генов (Льюин, 1987). Значительная часть контролируемых ими белков необходима для обеспечения жизнедеятельности самих клеток. В процессе онтогенеза и клеточной дифференцировки в разных тканях организма происходит избирательная активация многих других специфических генов, что, в конечном итоге, обуславливает значительные межклеточные различия в наборе белков и в скорости их синтеза.

Контроль генной активности осуществляется за счет дифференциальной транскрипции и процессинга РНК в клеточных ядрах, различной стабильности мРНК в цитоплазме, избирательной трансляции мРНК. Дифференциальная экспрессия генов, конечным результатом которой является синтез функционально активного белка, предполагает не только адекватную регуляцию генной активности, но и полноценность всех последующих этапов, включая сам белковый продукт, его устойчивость, способность к посттрансляционным модификациям, правильную локализацию и корректное взаимодействие с другими компонентами клетки. Решающее значение для успешного анализа всего этого сложного комплекса имеет выбор адекватных биологических моделей, поиск и целенаправленное конструирование которых представляет вполне самостоятельную научную задачу.

Наиболее доступными модельными системами для анализа экспрессии генов *in vitro* являются культуры клеток. Для клонирования, генноинженерного манипулирования, направленного введения сайт специфических мутаций, получения большого количества клонированных последовательностей ДНК, специфических молекул мРНК, а также белкового продукта гена обычно используют генетически хорошо изученные прокариотические системы (Хеймс, Хиггинс, 1987). Для исследования процессов трансляции, посттрансляционных модификаций белка, его внутриклеточной локализации и функционирования чаще используют культуры клеток эукариот и, в частности, специфические культуры клеток человека. Особая роль в изучении начальных этапов развития патологического процесса, обусловленного присутствием генных мутаций, а также в разработке терапевтических методов, включая генноинженерную коррекцию метаболического дефекта, принадлежит культурам мутантных клеток. Это могут быть первичные или перевиваемые культуры клеток, полученные из специфических тканей больного человека, либо выделенные из тканей линейных животных, служащих генетической моделью наследственного заболевания.

Идентификация гомологичных генов у экспериментальных животных во многих случаях значительно облегчает и ускоряют исследование функциональной активности нормальных и мутантных генов человека. Большая роль в изучении молекулярных механизмов развития патологических процессов *in vivo* принадлежит генетическим линиям животных. Это могут быть линии, по-

лученные в результате отбора спонтанно возникших или индуцированных мутаций, а также искусственно сконструированные модели на базе трансгенных животных, в геном которых введен чужеродный ген или фрагмент ДНК. Рассмотрим основные экспериментальные подходы, используемые для анализа экспрессии генов.

Раздел 6.2 Анализ регуляторных элементов гена, изоляция и исследование мРНК, искусственные транскрипционные системы.

Регуляция экспрессии генов в цепочке ДНК - РНК - белок может осуществляться на различных молекулярных уровнях. В соответствии с этим исследования дифференциальной активности генов в разных типах клеток и тканей включают оценку работы контролирующих элементов генов, анализ молекул РНК на всех этапах от появления первичного транскрипта до зрелой мРНК и изучение соответствующего белкового продукта, включая его процессинг (созревание), внутриклеточную локализацию, тканеспецифическое распределение.

Исследования регуляторных цис-действующих элементов генома, таких как промоторы, инхансеры, участки ДНК, подавляющие транскрипцию, являются составной частью анализа молекулярной структуры любого гена. Идентификацию таких элементов проводят с использованием разнообразных современных методов молекулярной генетики. В частности, последовательности ДНК в 5'-фланкирующей области гена, ответственные за тканеспецифическую индукцию генной активности, могут быть локализованы путем исследования транскрипции в различных линиях клеток при введении в них генов с искусственными делециями этих участков ДНК. Для оценки активности идентифицированных регуляторных последовательностей их сливают с чужеродными хорошо изученными неиндуцибельными клонированными генами, так называемыми "репортерами". Такие генетические конструкции в составе векторных последовательностей вводят в культивируемые клетки эукариот и наблюдают за изменением уровня экспрессии. В качестве "репортера" часто используют ген хлорамфеникол-ацетил-трансферазы (CAT-ген), который в естественных условиях экспрессируется только в клетках прокариот. Сам фермент (CAT) обладает высокой активностью, что позволяет не только легко обнаруживать ее минимальные количества в клетке, но и с высокой точностью проводить количественную оценку. Для повышения чувствительности анализ экспрессии химерных генов часто проводят в культуре фибробластов почек африканской зеленой мартовки (COS-клетки). Эти клетки модифицированы таким образом, что в них после трансфекции происходит амплификация копий сконструированных определенным образом эписом (внехромосомных генетических конструкций) (см. Главу X), что ведет к значительному усилению сигналов экспрессии введенных генов (трансгенов). Перенос генов (трансгеноз) может быть осуществлен и на уровне целого организма, в частности, зиготы. Полученные в результате подобных манипуляций трансгенные животные могут быть также использованы в качестве модельной системы для анализа механизмов тканеспецифической активации генов *in vivo*.

Матричная РНК является наиболее удобным объектом для изучения регуляции транскрипции генов и посттранскрипционных модификаций РНК. Тотальная клеточная РНК состоит на 90 - 95% из рибосомальных и транспортных РНК, тогда как доля транслируемых или poly(A)+ РНК не превышает 5% (Льюин, 1987). При этом, концентрация РНК-транскриптов индивидуальных генов

среди всех молекул мРНК, в среднем, колеблется в пределах от 0.01% до 0.001% (Гайцхоки, 1978). Поэтому для обнаружения индивидуальных типов мРНК должны использоваться высокочувствительные методы. Обычным методом идентификации мРНК на тканевом и клеточном уровнях является гибридизация *in situ* РНК- или ДНК-зондов с молекулами мРНК на гистологических срезах (Хаффнер, Уиллисон, 1990). В качестве ДНК-зондов используют клонированные последовательности кДНК и синтетические олигонуклеотиды. После инкубации меченых зондов на цитологических препаратах с последующей тщательной отмывкой несвязавшихся молекул положение комплементарных РНК-последовательностей в клетках определяют радиоавтографическими, либо в случае биотинового мечения - иммуногистохимическими методами. Оптимальные условия гибридизации дают возможность не только выявлять присутствие специфических мРНК, но и определить их внутриклеточную локализацию (Манк, 1990; Хаффнер, Уиллисон, 1990; Boehringer, Manual, 1994).

Анализ индивидуальных РНК включает изоляцию из тканей пула неповрежденных биологически активных мРНК и идентификацию среди них специфических молекул путем использования различных вариантов ДНК-РНК гибридизации. Для генов с высоким уровнем транскрипции могут быть пригодны дот или слот блоты (см. Главу I). Когда источником РНК служат клетки, которые не могут быть получены в большом количестве, используют цитоплазматический дот-блот. При этом целые клетки лизируют, и фиксируют непосредственно на тех мембранах, на которых проводят гибридизацию. Значительно большей чувствительностью обладает, так называемый Northern blot (нозерн-блот) - гибридизация с ДНК-зондами на фильтрах предварительно сконцентрированных и фракционированных путем электрофореза молекул РНК (Sambrook et al., 1987). Электрофорез проводят в агарозе с добавлением формальдегида, денатурирующего РНК. В этих условиях скорость продвижения молекул РНК через гель находится в логарифмической зависимости от длины последовательности, что позволяет точно определить размер РНК транскрипта. Основная масса РНК на геле представлена в виде двух доминирующих бэндов, соответствующих двум типам рибосомальной РНК - 28S и 18S. Все молекулы мРНК сконцентрированы в плохо различимой, слабо окрашенной области геля, в которой отдельные типы мРНК могут быть выявлены только путем гибридизации с соответствующими ДНК-зондами. Нозерн-блот имеет то преимущество, что при электрофорезе могут быть разделены молекулы РНК, дающие перекрестную гибридизацию с ДНК-зондом. Кроме того, характер электрофоретического разделения позволяет визуально оценить качество изолированной РНК. При очень низких концентрациях специфических мРНК или в тех случаях, когда ДНК-зонды дают перекрестную гибридизацию с другими компонентами (не мРНК), проводят обогащение изолированной тотальной РНК транслируемыми мРНК путем отбора на колонках фракций, содержащих поли-А "хвосты". Для этого выделенную РНК пропускают через короткую колонку с пришитыми поли-Т олигонуклеотидными последовательностями и высокой концентрацией солей в буферном растворе, так чтобы молекулы мРНК, содержащие поли-А "хвосты", задерживались на колонке. При снижении концентрации солей в буфере происходит расплавление А-Т дуплексов и высвобождение молекул мРНК. Таким способом доля этих молекул в определенных солевых фракциях может быть увеличена на два порядка. Конечно, поли-А селекция применима в тех случаях, когда имеется достаточно большое количество тотальной РНК. Другим методом для исследования структуры РНК транскриптов является S1-анализ. В этом случае ДНК-РНК гиб-

ридизацию ведут в растворе, куда и добавляют S1 нуклеазу для переваривания одонитевых несвязавшихся молекул как ДНК, так и РНК, после чего проводят электрофоретическую очистку дуплексов, которые затем элюируют из геля для последующего анализа (Sambrook et al., 1989). Этот метод очень удобен для анализа стартовых сайтов и 3'-концов генов, для определения направления транскрипции и картирования интронов.

Уровень мРНК в клетке определяется несколькими кинетическими параметрами – скоростью первичного синтеза, эффективностью процессинга РНК-транскриптов и периодом полураспада зрелых молекул мРНК. Последний параметр определяют по динамике исчезновения мРНК после добавления к клеткам актиномицина D, специфическим образом супрессирующего транскрипцию.

Исследование механизмов транскрипции и процессинга первичных РНК-транскриптов проводят *in vitro* с использованием искусственным образом сконструированных транскрипционных систем (Manley et al., 1986; Dignam et al., 1983; Gutierrez-Hartmann et al., 1987; Хэймс, Хиггинс, 1987). Для этого могут быть выбраны два различных подхода. В первом случае изолируют ядра и в качестве транскрипционной матрицы используют неповрежденный хроматин. Синтез РНК проводят с добавлением всех необходимых реагентов и, в частности, трифосфатов, в один из которых (обычно в урацил) вводят радиоактивную метку. При этом вновь синтезированные молекулы РНК оказываются мечеными. Выбор специфических молекул РНК проводят путем ДНК-РНК гибридизации, однако, в отличие от ранее описанных методов анализа мРНК, используют немеченые кДНК-зонды, предварительно нанесенные на фильтры. Большим достоинством этой транскрипционной системы является ее максимальная приближенность к естественным процессам. При втором подходе транскрипция ведется с клонированных фрагментов ДНК, а ядерные экстракты служат источником ферментов и регуляторных белков.

Раздел 6.3 Анализ трансляции, ДНК-экспрессионные системы.

Традиционные методы анализа регуляции трансляции и посттрансляционных модификаций белков основаны на использовании модельных систем, представляющих собой цитоплазматические свободные от мРНК безядерные экстракты клеток, содержащие рибосомальный аппарат, транспортные РНК, набор аминокислот и ферментов, необходимых для трансляции и процессинга белков (Хэймс, Хиггинс, 1987; Клеменс, 1987). После добавления к такой системе специфической мРНК происходит синтез соответствующей полипептидной цепи *in vitro*. При введении меченых аминокислот в систему вновь синтезированные белки после электрофоретической очистки могут быть идентифицированы путем радиоавтографии либо иммунологическими методами, при наличии соответствующих антител (Клеменс, 1987). Однако, для значительного числа моногенных наследственных заболеваний первичный биохимический дефект неизвестен, а следовательно, не идентифицированы и мРНК транскрипты. Биохимическое изучение многих белков затруднено из-за их минорного содержания и отсутствия эффективных методов выделения и очистки. Последнее обстоятельство в значительной мере относится к нерасворимым белкам, ассоциированным с мембранными структурами клеток.

ДНК-экспрессионные системы, то есть клеточные культуры, синтезирующие чужеродные белки, являются очень мощным

средством анализа структуры, функции и синтеза белков (Sambrook et al., 1989). Такие системы конструируют на основе экспрессионных векторов, содержащих в своем составе сильные промоторы и регуляторные последовательности, обеспечивающие высокий, но в то же время регулируемый уровень экспрессии. Кодирующие последовательности чужеродных генов инсертируют (вставляют) с помощью соответствующих генно-инженерных приемов в область действия этих промоторов. Конечно, такие системы должны содержать и трансляционные сигналы, в частности, сайты связывания рибосом, обеспечивающие работу рибосомального аппарата клеток хозяина. В некоторых случаях экспрессионные векторы вводят в мутантные по протеазным генам клеточные культуры, с тем чтобы предотвратить деградацию чужеродных белков в клетках.

Существует три типа экспрессионных систем - бактериальные, сконструированные обычно на основе *E. coli*, дрожжевые и экспрессионные культуры клеток млекопитающих. Каждая из этих систем имеет свои преимущества и недостатки. Бактериальные системы наиболее удобны для клонирования, обладают высоким уровнем экспрессии (до 1-2 грамм белка на литр культуры) и их используют, обычно, для производства большого количества чистого белка, необходимого для получения антител или для фармацевтических целей. Удобны также эти системы для введения изменений в различные районы полипептидной цепи путем сайт-направленного мутагенеза в нуклеотидной последовательности чужеродной ДНК. Получение и исследование таких "мутантных" белков очень важно для оценки функциональной значимости различных участков белка.

Уровень экспрессии чужеродных белков в дрожжевых клетках вдвое, а в клетках млекопитающих в десятки раз ниже, чем в бактериальных. Однако, в бактериальных клетках отсутствуют ферментативные системы, обеспечивающие процессинг эукариотических белков. Поэтому эукариотические системы удобнее использовать для изучения посттрансляционных модификаций белка - гликозилирования, то есть присоединения к полипептидной цепи углеводных остатков; скручивания белка с образованием третичной структуры, часто, за счет возникновения дисульфидных связей; и N-концевых модификаций, стабилизирующих структуру белка. В ДНК-экспрессионных системах может быть синтезировано достаточно много белка, чтобы получить его в кристаллической форме и исследовать пространственную структуру и функциональное назначение отдельных доменов (Хэймс, Хиггинс, 1987).

Использование экспрессионных библиотек для изоляции кодирующих последовательностей гена рассматривалось ранее (см. Глава II). После секвенирования кДНК можно, исходя из генетического кода, прогнозировать аминокислотный состав белка и произвести компьютерный поиск в банке данных гомологичных последовательностей в составе белков с уже известной структурой и функциями. Выявление родственных белков, близких по своему полипептидному составу, значительно ускоряет и облегчает дальнейший молекулярный анализ функционирования исследуемого белка в клетке. Аминокислотная последовательность белка позволяет прогнозировать его третичную структуру, идентифицировать домены, оценивать функциональную значимость целого белка и отдельных его компонентов. Не менее важным практическим следствием этих данных является также возможность получения антител к строго специфичным участкам белка. Для этого могут быть использованы два подхода - биохимический и молекулярно-генетический. В первом случае для иммунизации используют искусственно синтезированные полипептиды,

которые пришивают к белковой молекуле-носителю (гаптену). Размеры таких полипептидов, обычно, не превышают 30 аминокислот – они не могут быть очень большими из-за высокой стоимости и трудоемкости синтеза длинных молекул. При втором подходе экзонные участки гена инсертируют в экспрессионный вектор в область, кодирующую селективируемый белок. В результате экспрессии такой конструкции получают слитый белок, в котором наряду с аминокислотной последовательностью селективируемого маркера содержится определенный фрагмент исследуемого белка. Эту химерную молекулу и используют для иммунизации животных и получения моновалентных или моноклональных антител. При наличии антител могут быть применены различные иммунологические подходы для анализа тканеспецифического и внутриклеточного распределения белка, исследования его модификаций, а также для получения нативного белка в препаративных количествах.

Следующим шагом на пути анализа молекулярных механизмов регуляции экспрессии гена является идентификация тех нарушений в структуре, локализации и активности молекул мРНК и белка, которые возникают вследствие генетических мутаций. Мы уже упоминали об огромном значении культур мутантных клеток для подобных исследований. Однако, многие патологические процессы, протекающие в организме больного, не могут быть исследованы *in vitro*. С другой стороны, возможности получения необходимого количества клеток и тканей пациента и испытания *in vivo* различных схем лечения значительно ограничены. Поэтому для многих наследственных болезней эффективность изучения основ патогенеза существенным образом зависит от наличия адекватных биологических моделей. Способы конструирования таких моделей подробно изложены в Главе VIII.

ГЛАВА II.

ГЕНОМ ЧЕЛОВЕКА, СТРУКТУРА ГЕНОВ.

Раздел 2.1. Определение генома и его основных элементов.

Термин геном используется для обозначения полной генетической системы клетки, определяющей характер онтогенетического развития организма и наследственную передачу в ряду поколений всех его структурных и функциональных признаков. Понятие генома может быть применено к таксономической группе, виду, отдельной особи, клетке, микроорганизму или вирусу. Так, можно говорить о структуре генома эукариот и прокариот, сравнивать геномы разных видов, изучать особенности строения генома у конкретных индивидуумов или следить за изменениями, происходящими в геноме специфических клеток в процессе их онтогенетической дифференцировки. Часто геном определяется как генетическая информация, заключенная в молекулах ДНК одной клетки. Однако, такие факты, как отсутствие связи между количеством ДНК в расчете на гаплоидный геном и таксономическим статусом видов, а также многочисленные примеры существования огромных различий в содержании ДНК между близкородственными видами (так называемый "С-парадокс") свидетельствуют о том, что далеко не все участки ДНК связаны с информационными функциями. Понятия генома и ДНК в значительной степени тождественны, так как основные принципы организации и функционирования генома целиком определяются свойствами ДНК. Присущие этим молекулам

потенциальные возможности практически неограниченного структурного разнообразия определяют все многообразие мира живых существ, как на уровне межвидовых, так и индивидуальных различий в пределах одного вида (Баев и др., 1990; Ратнер, 1985).

Процесс эволюции и дифференцировки отдельных видов, как правило, сопровождался накоплением изменений в структуре генома. Это касается, прежде всего, таких параметров, как локализация и характер упаковки ДНК в клетках; количество ДНК, приходящееся на гаплоидный геном; типы, соотношение и функции кодирующих и не кодирующих нуклеотидных последовательностей; регуляция экспрессии генов; межпопуляционная вариабельность и филогенетический консерватизм первичной структуры генома. В пределах одного вида основные параметры генома достаточно постоянны, а внутривидовое разнообразие обеспечивается за счет мутационной изменчивости, то есть выпадения, вставки или замены нуклеотидов на сравнительно небольших участках ДНК. Чаще всего такие изменения касаются неэкспрессируемых элементов генома (интронов, псевдогенов, межгенных спэйсерных участков ДНК и т.д.).

Геномы эукариот, по-существу, можно рассматривать как мультигеномные симбиотические конструкции, состоящие из облигатных и факультативных элементов (Golubovsky, 1995). Основу облигатных элементов составляют структурные локусы, количество и расположение которых в геноме достаточно постоянно. Присутствие в хромосомах некоторых видов повторяющихся ДНК, амплифицированных участков, ретровирусных последовательностей, псевдогенов, также как наличие в клетке эписом, ретротранскриптов, ампликонов, дополнительных В-хромосом и различных цитосимбионтов (вирусов, бактерий, простейших) является не строго обязательным, их количество и положение может значительно варьировать, то есть эти элементы являются факультативными. В то же время участие факультативных элементов в наследственной передаче признаков, в формировании мутационной изменчивости и в эволюционных преобразованиях видов несомненно доказано. Кроме того, существует непрерывный переход от одних состояний к другим за счет инсерции экстрахромосомных ДНК в хромосомы и выстраивания транспозоноподобных мобильных элементов из хромосом. Следовательно, несмотря на значительные отличия факультативных последовательностей от облигатных по характеру основных информационных процессов (репликации, транскрипции, трансляции и сегрегации), они также должны рассматриваться, как важнейшие элементы генома.

Остановимся теперь более детально на основных принципах организации генома человека. В каждой диплоидной клетке с 46 хромосомами содержится около 6 пикограмм ДНК, а общая длина гаплоидного набора из 23 хромосом составляет $3.5 \cdot 10^9$ пар нуклеотидов (Као, 1985). Этого количества ДНК достаточно для кодирования нескольких миллионов генов. Однако, по многим независимым оценкам истинное число структурных генов находится в пределах от 50 000 до 100 000. В разделе 2.4 изложены современные подходы, используемые для подсчета общего количества генов, из которых следует, что наиболее вероятная оценка их числа составляет около 80 000. Сопоставляя это значение со средними размерами гена и соотношением между величиной их экзонных и интронных областей, можно заключить, что кодирующие последовательности ДНК занимают не более 10-15% всего генома (McKusick, Ruddle, 1977). Таким образом, основная часть молекул ДНК не несет информации об аминокислотной последовательности белков, составляющих основу любого живого организма, и не кодирует структуру рибосомаль-

ных, транспортных, ядерных и других типов РНК. Функции этой "избыточной" (junk) ДНК не ясны, хотя ее структура изучена достаточно подробно. Предполагается, что эта ДНК может участвовать в регуляции экспрессии генов и в процессинге РНК, выполнять структурные функции, повышать точность гомологичного спаривания и рекомбинации, способствовать успешной репликации ДНК и, возможно, является носителем принципиально иного генетического кода с неизвестной функцией.

Наиболее общая характеристика генома может быть получена с помощью анализа кинетики реассоциации молекул ДНК. Динамика плавления геномной ДНК обнаруживает присутствие по крайней мере трех различающихся по химической сложности фракций (Льжин, 1987; Газарян, Тарантул, 1983). Быстро ренатурирующая фракция ДНК состоит из относительно коротких высокоповторяющихся последовательностей; в промежуточную фракцию входит множество умеренно повторяющихся ДНК - более протяженных, но представленных меньшим числом копий; медленно ренатурирующая фракция объединяет в себе уникальные последовательности ДНК, встречающиеся в геноме не более 1-2 раз.

С помощью молекулярного анализа проведена идентификация основных классов повторяющихся последовательностей ДНК, составляющих более 35% всего генома человека и включающих сателлитную ДНК, инвертированные повторы, умеренные и низкочисленные повторы, а также мини- и микросателлитные последовательности ДНК. Классификация этих типов повторов достаточно условна и основана, главным образом, на двух характеристиках: длине повторяющихся коровых единиц, которая может варьировать от 1-2 до более, чем 2000 п.о., и числе их копий, также меняющихся в очень широких пределах - от десятка до миллиона на гаплоидный геном. Не менее важными характеристиками различных классов повторяющихся ДНК являются нуклеотидная последовательность "коровых" единиц повтора, специфичность их организации, хромосомная локализация, внутри- и межвидовая стабильность, а также возможные функции этих типов ДНК.

Раздел 2.2. Повторяющиеся последовательности ДНК.

Сателлитная ДНК это класс высокоповторяющихся последовательностей, составляющих около 10% всего генома человека (Као, 1985). При центрифугировании геномной ДНК в градиенте плотности CsCl эти последовательности образуют четыре отдельных сателлитных пика с различными средними значениями плавучей плотности. Методом гибридизации *in situ* показано присутствие сателлитной ДНК преимущественно в центромерных, теломерных и гетерохроматиновых районах большинства хромосом, при этом характер гибридизации сходен для всех четырех групп и не зависит от принадлежности ДНК-зондов к семействам повторов, образующих различные сателлитные пики. В каждой из этих групп, однако, присутствует небольшое количество последовательностей, имеющих специфическую хромосомную локализацию. Так например, около 40% длинного плеча Y хромосомы составляет семейство последовательностей, тандемно повторяющихся более 3000 раз и не найденных в других хромосомах.

Выделяют три основных типа сателлитной ДНК: (1) короткие - от 2 до 20 п.о., стабильные тандемные повторы с кратностью в несколько десятков тысяч раз, которые иногда перемежаются с неповторяющимися последовательностями; (2) кластеры более протяженных повторов, слегка различающихся по нуклеотидной последовательности; (3) сложные, достигающие нескольких со-

тен пар нуклеотидов, повторяющиеся последовательности различной степени гомологии (Газарян, Тарантул, 1983). К последнему типу относятся альфа-сателлитные или альфоидные ДНК, среди которых найдено много хромосом-специфических последовательностей. Размеры повторяющихся "коровых" единиц альфоидной ДНК составляют около 170-200 п.о. В геноме человека и других приматов эти мономеры организованы в кластеры по 20 и более "коровых" единиц. После расщепления рестриктазой BamHI в альфоидной ДНК выявляется серия фрагментов, длиной около 2 000 п.о., в составе которых обнаруживаются альфоидные последовательности, специфичные для гетерохроматиновых районов разных хромосом человека (1, 3, 4, 5, 9, 6, 7, 11, 17, 19, X). В некоторых случаях эти повторы гомологичны двум разным хромосомам (9 и 15, 13 и 21, 18 и X) (Willard, Waye, 1987). Хромосом-специфические последовательности сателлитной альфоидной ДНК нашли широкое применение в молекулярной цитогенетике в качестве ДНК-зондов, удобных для маркирования индивидуальных хромосом в метафазных и интерфазных клетках человека (Юров, 1987). Предполагается, что сателлитные ДНК играют важную роль в поддержании структур хромосом и, возможно, в их спаривании в процессе мейоза (Charlesworth et al., 1994).

Особое место среди сателлитных ДНК занимают микро- и минисателлитные последовательности, представляющие собой многочисленную группу рассеянных по всему геному относительно коротких tandemных повторов. Микросателлиты - это класс динуклеотидных повторов. Размер повторяющихся единиц в минисателлитных последовательностях может меняться от 3 - 4 до 10 - 15 нуклеотидов. Отличительной особенностью микро- и минисателлитов является наличие среди них большого количества участков, варьируемых по числу копий в кластере.

Инвертированные или обращенные повторы составляют до 5% генома. Они состоят из двух тождественных копий длиной около 300 п.о., ориентированных в противоположных направлениях на одной нити ДНК и лежащих на расстоянии от нуля до десятка тысяч пар нуклеотидов друг от друга (в среднем - 1,6 кб). Около 1/3 обращенных повторов не разделены промежуточными последовательностями и носят название палиндромов. Среднее расстояние между двумя различными парами инвертированных повторов около 12 кб, и их распределение по геному носит случайный характер. Комплементарные пары легко ассоциируют при отжиге, образуя шпильчатые структуры с дуплексной ножкой и однонитевой петлей, длина которой соответствует расстоянию между парой обращенных повторов. Вследствие этого оказываются приближенными достаточно удаленные друг от друга участки ДНК, что важно для работы ряда ферментов, обеспечивающих процессы репликации и транскрипции. Важно отметить и то, что однонитевой участок ДНК, образующий петлю, становится доступным для действия нуклеазы S1, специфически разрушающей однонитевую ДНК.

Группа умеренно повторяющихся последовательностей очень гетерогенна по длине и числу копий и составляет около 20% генома человека. Как правило, они распределены дисперсно по всем хромосомам, причем относительно короткие последовательности ДНК до 500 п.о., так называемые короткие диспергированные повторы - Sine, повторяются более 100 000 раз, в среднем, через каждые 2.2 кб. Число копий более длинных диспергированных последовательностей - Line, не превышает 10 000. Умеренные повторы найдены во всех структурных компонентах генома за исключением кодирующих областей генов. Два главных семейства умеренных повторов, Alu и Kpn1, занимают, по крайней мере, 10% генома и практически столько же занято

несколькими сотнями других семейств повторов этого класса.

Основной единицей Alu семейства является короткая последовательность из 300 п.о., повторенная в геноме человека несколько сот тысяч раз, в среднем, через каждые 5 кб или через каждые 2 - 3 диспергированных повтора, принадлежащих другим семействам (Као, 1985; Льюин, 1987). При этом, кластеры Alu-повторов, как правило, лежат внутри R-дисков метафазных хромосом - Гимза отрицательных (G- дисков). Распределение Alu-повторов по геному весьма неравномерно как между хромосомами, так и по их длине. Так, в хромосомах 14, 16, 21 Alu-последовательности концентрируются в области центромеры, а в хромосомах 4, 19, 20, X и Y выраженные кластеры Alu повторов не найдены. Члены Alu семейства не полностью идентичны друг другу. Однако, все они содержат сайт рестрикции для фермента AluI и имеют димерную структуру, то есть состоят из двух прямых повторов длиной около 130 п.о. с богатой аденином вставкой из 31 нуклеотида во втором мономере. Каждая Alu последовательность фланкирована прямыми повторами длиной от 7 до 20 п.о., различными для разных членов семейства и имеющими большую степень гомологии с транспозоноподобными элементами про- и эукариот. Некоторые члены Alu семейства могут транскрибироваться с помощью фермента РНК-полимеразы 111. Предполагается, что при определенных условиях образующиеся при этом молекулы РНК могут обратно транскрибироваться, что в свою очередь, может привести к появлению в клетках Alu-содержащих кДНК, обладающих свойствами ретропазонов, то есть способных инсертироваться в геномную ДНК. В литературе описаны случаи инсерционного мутагенеза Alu повторов, приводящие к гемофилии В и нейрофиброматозу типа 1. Однако, частота таких событий, по-видимому, невелика. Предполагается, что короткие умеренные повторы, подобные Alu семейству, участвуют в регуляции транскрипции, в процессинге РНК и в инициации репликации ДНК. Кроме того, обнаружена высокая степень гомологии Alu последовательностей с одним из видов низкомолекулярной РНК (7 S РНК), участвующей в секреции белков.

К числу Line повторов относится Kpn1 семейство, которое состоит из более длинных и значительно более гетерогенных последовательностей, рассеянных по всему геному. В ряде случаев члены Kpn1 семейства группируются в кластеры, образуя более длинные структуры, повторяющиеся несколько тысяч раз. Для некоторых членов этого семейства также доказана возможность инсерции кДНК-овых копий Kpn1- РНК-транскриптов в геномную ДНК и возникновение мутаций. Такое явление было обнаружено в одном случае гемофилии А. Некоторые Kpn1- последовательности не только транскрибируются, но и способны транслироваться (Charlesworth et al., 1994).

Раздел 2.3 Мультигенные семейства, псевдогены, онкогены.

Многие гены человека повторены в геноме от нескольких единиц до нескольких сотен раз и образуют мультигенные семейства (Газарян, Тарантул, 1983; Босток, Самнер, 1981; Као, 1985; Льюин, 1987). Эти гены обычно сгруппированы в кластеры в определенных районах одной, либо нескольких хромосом. Во многих мультигенных семействах наряду с функционально активными генами содержатся псевдогены - мутационно измененные последовательности, не способные транскрибироваться или продуцирующие функционально неактивный генный продукт. Примерами мультигенных семейств могут служить гены рибосомальных,

транспортных и ядерных РНК, гены альфа- и бета-глобинов, тубулинов, миоглобина, актина, интерферона и многих других. В ряде случаев, возможна избирательная амплификация некоторых семейств генов в процессе их экспрессии, как, например, генов рибосомальных РНК. При этом число способных транскрибироваться копий генов увеличивается за счет их избирательной амплификации в сотни и даже тысячи раз, что сопровождается лавинообразным нарастанием доли соответствующего генопродукта в клетках. Особое место среди мультигенных семейств занимают супергены - очень большие кластеры из сотен функционально и структурно родственных генов, расположенных в сегментах отдельных хромосом. Классическим примером супергена может служить HLA комплекс, контролирующий главные антигены гистосовместимости. Он занимает район более 6000 кб на коротком плече хромосомы 6p21 и состоит из серии тесно сцепленных генов, ответственных за синтез множества белков, включающих клеточные поверхностные антигены, молекулы иммунного ответа и некоторые компоненты комплемента. К супергенным семействам относятся три комплекса расположенных на разных хромосомах мультигенов, контролирующих синтез тяжелых и легких цепей иммуноглобулинов. Интересно, что в процессе дифференцировки В лимфоцитов, продуцирующих иммуноглобулины, происходит структурная перестройка этих семейств. При этом отдельные последовательности ДНК элиминируются, тогда как другие сливаются, так что структура генов иммуноглобулинов в зрелых В лимфоцитах значительно отличается от исходной, то есть от той, которая наблюдается в зародышевых клетках.

Одной из важных структурных особенностей генома человека является наличие так называемых псевдогенов, уникальных последовательностей, очень сходных по своей структуре с определенными нормальными генами, но в силу присутствия в кодирующих последовательностях целого ряда мутаций не способных транскрибироваться или правильно транслироваться с образованием структурно и функционально активного продукта. Псевдогены обнаружены для многих генов. Их количество варьирует от одной до нескольких десятков копий на геном и в этом случае они, как правило расположены тандемно. Иногда псевдогены тесно сцеплены с нормальными генами, во многих случаях псевдогены и гены локализованы в разных хромосомах. Для некоторых моногенных заболеваний идентифицированы мутантные аллели, сходные с мутациями в соответствующих псевдогенах. В этих случаях обсуждается возможная роль псевдогенов в спонтанном мутационном процессе.

В геноме человека присутствуют также нуклеотидные последовательности, гомологичные генам некоторых вирусов. Впервые эти последовательности были идентифицированы в геноме вирусов, индуцирующих развитие опухолей у животных и человека, и потому они были названы онкогенами. Гомологичные последовательностям в геноме человека носят название протоонкогенов. В настоящее время уже идентифицировано более 100 протоонкогенов. Белковые продукты протоонкогенов, по-видимому, играют важную роль в нормальной пролиферации клеток особенно на ранних стадиях эмбрионального развития, контролируя клеточный цикл и выбор геномной программы развития клетки. При возникновении специфических мутаций в протоонкогенах, а также при нарушениях регуляции их работы, выражающихся в гиперпродукции или в экспрессии в нетипичном месте или в несвойственный момент жизнедеятельности клетки, они начинают вести себя как онкогены, стимулируя неконтролируемое размножение и пролиферацию определенных клеточных клонов, что и может, в конечном счете, привести к формированию опухоли.

Раздел 2.4 Современное определение понятия "ген", транскрипция, регуляторные элементы генов.

Около 10-15% генома человека представлено уникальными транскрибируемыми последовательностями, составляющими основу структурных генов (Льюин, 1987). В настоящее время в понятие "ген" включается не только его транскрибируемая область - экзоны + интроны, но также фланкирующие последовательности - лидерная, предшествующая началу гена, и хвостовая нетранслируемая область, расположенная на 3' конце гена (Рис.2.1). В отличие от генов прокариот гены человека редко представлены одной непрерывной последовательностью и в подавляющем большинстве имеют прерывистую структуру. Относительно короткие кодирующие участки - экзоны, чередуются с длинными интронами, которые транскрибируются и входят в состав первичного РНК-продукта, но затем при процессинге первичного РНК-транскрипта они вырезаются и не участвуют в трансляции. Процесс вырезания интронов из первичных транскриптов получил название сплайсинга. Таким образом, в зрелой мРНК интронные области отсутствуют, а экзоны составляют непрерывную кодирующую последовательность. Размеры зрелых мРНК нередко в десятки раз меньше первичных РНК-транскриптов и, соответственно, размеров самого гена.

Согласно классическим представлениям ген - это локус, на хромосоме, мутации в котором реализуются на уровне фенотипа. В молекулярной биологии ген трактуется как ассоциированный с регуляторными последовательностями фрагмент ДНК, соответствующий определенной единице транскрипции. Следовательно, представления о гене формальных генетиков далеко не полностью тождественны его физической единице и соотношения между этими двумя понятиями достаточно запутанные. Отметим некоторые причины этих противоречий. Известно, что мутации одного гена могут приводить к совершенно разным и даже в ряде случаев к комплементарным фенотипам. Результаты прямого секвенирования генома свидетельствуют о присутствии в нем значительного большего числа генов, чем можно ожидать от результатов мутационного анализа. Одна и та же последовательность ДНК в геноме может кодировать несколько различных белков, что достигается за счет так называемого альтернативного сплайсинга (образование разных мРНК из одного первичного РНК-транскрипта). В крупных интронах ряда генов обнаружены смысловые последовательности других генов ("ген в гене"), считываемые в противоположном направлении. Транскрипционные единицы генома могут перекрываться за счет наличия разных промоторов. Наконец, благодаря соматической рекомбинации структура транскрибируемых последовательностей некоторых генов может быть различной в разных клонах клеток одного организма (Т-клеточные рецепторы). Ситуация с определением понятия "ген" еще больше осложняется, если в это понятие включать многочисленные регуляторные последовательности. Возникает вопрос: "Как далеко от гена могут располагаться эти последовательности, чтобы их можно было включать в структуру гена?". Для многих целей оказывается удобным введенное в последнее время понятие "считаемый ген"- counting gene. Последний рассматривается как отдельная транскрибируемая единица ДНК или её часть, которая может транслироваться в одну или несколько взаимосвязанных аминокислотных последовательностей. Поэтому последовательность, дающая две транскрипционные единицы за счет альтернативного сплайсинга и, как следствие, два разных белка учитывается как один ген.

Однако, если степень гомологии двух генопродуктов, имеющих общий транскрибируемый участок, невелика, то эти последовательности расцениваются как два разных гена (Fields et al., 1994).

По своему функциональному назначению гены могут быть разделены на две группы. Группа I представлена генами, кодирующими собственно белки; группа II - генами, контролирующими синтез рибосомальных, транспортных и ядерных РНК. По характеру экспрессии гены также могут быть подразделены на две группы - гены "домашнего хозяйства" (housekeeping genes), продукты которых необходимы для обеспечения жизнедеятельности любого типа клеток, и тканеспецифические гены, обеспечивающие специализированные функции клеток, то есть гены, функционально активные только в определенных типах клеток (тканей) и только на определенных стадиях онтогенеза, так называемые гены терминальной дифференцировки.

Считается, что средние размеры гена человека составляют, примерно, от 10 до 30 кб. Однако, эта величина может колебаться от нескольких десятков до миллионов пар нуклеотидов. Согласно последним данным, самый маленький из известных генов - MCC-7, имеет размеры всего 21 п.о. (Gonzales-Pastor et al., 1994), а самый большой - ген дистрофина - 2.2 мегабаз. Гены отделены друг от друга протяженными промежутками - спейсерами, содержащими в своем составе большое количество повторяющихся последовательностей ДНК и нетранскрибируемые уникальные последовательности.

Рассмотрим более подробно современные данные о числе генов в геноме человека. Полученные разными методами оценки этой величины приведены в Табл.2.1. Исходя из размера генома (около 3 000 миллионов п.о.), и среднего размера одного гена порядка 10 - 30 тыс.п.о. (Gilbert, 1992), общее число генов должно быть порядка 100 000. При этом, как уже указывалось, распределение генов на хромосомах крайне неравномерно. Установлено, что более 90% генов находится в Гимза -отрицательных районах метафазных хромосом, в так называемых R-бэндах (Antequera, Bird, 1993). Проведенные недавно прямые исследования методом секвенирования показывают, что в R-бэндах их число достигает 43-50 на один бэнд, а в Гимза-положительных районах хромосом (G бэндах), - только 1-2 на 75-80 килобаз, то есть всего в геноме можно ожидать около 70 000 генов. Оценка числа генов по доле транскрибируемой части генома (всего 12%) дает совсем маленькую величину - 20 000 генов (Wagner et al., 1993). Методом исследования кинетики реассоциации РНК в культуре клеток число генов оценивается между 20 - 40 000 (Lewin, 1990). В то же время в клетках мозга число различных мРНК по некоторым данным достигает 97 000 (Wagner et al., 1993). Наиболее точные подсчеты числа генов человека проведены в последнее время и основаны на оценке числа CpG островков (Табл.2.1). Известно, что в промоторной области всех генов "домашнего хозяйства" и примерно у 40 % генов терминальной дифференцировки, обеспечивающих специализированные функции дифференцированных клеток, находятся области коротких динуклеотидных CpG повторов. Разработаны молекулярные методы точной регистрации этих участков, которые показали, что их число в геноме около 45 000, откуда число структурных генов оценивается в 67-70 000 (Antequera, Bird, 1993). Практически такое же число генов (64 000) определено недавно методом учета маркерных экспрессирующихся последовательностей (expressed sequence tags) (Fields et al., 1994). Таким образом, суммируя вышеизложенное, можно сделать вывод, что в геноме человека содержится, в среднем, около 70-80 000

отдельных транскрибируемых ДНК последовательностей, то есть генов.

Транскрипция гена начинается с 5' конца первого экзона, где расположен сайт инициации транскрипции. Определенной гомологии между стартовыми сайтами разных генов не наблюдается, но чаще всего они начинаются с нуклеотида А, окруженного пириимидиновыми основаниями. На границах между экзонами и интронами имеются консервативные канонические последовательности, играющие существенную роль в обеспечении точности вырезания интронов во время сплайсинга РНК. Все интронные последовательности начинаются с динуклеотида GT и заканчиваются динуклеотидом AG, называемыми, соответственно, донорными и акцепторными сайтами сплайсинга. На 3' конце многих структурных генов идентифицирована поли(А)-сигнальная последовательность (ААТААА), участвующая в процессе модификации первичного РНК транскрипта и ответственная за альтернативный сплайсинг мРНК, обеспечивающий синтез разных зрелых мРНК с одного и того же первичного РНК транскрипта.

Транскрипция генов осуществляется с помощью фермента РНК-полимеразы. Около 50-70% клеточного синтеза РНК обеспечивается РНК-полимеразой I, локализованной в ядрышках и ответственной за синтез генов рибосомальной РНК. РНК-полимераза II обеспечивает транскрипцию генов, кодирующих собственно структурные белки. Этот фермент локализован в ядре (но не в ядрышках). На его долю приходится 20 - 40% синтеза РНК. РНК-полимераза III контролирует синтез ядерных и транспортных РНК (Льюин, 1987). На 1-м этапе РНК-полимераза связывается с двухнитевым участком ДНК и, расплетая его, делает доступным спаривание смысловой нити ДНК с рибонуклеотидами (Рис. 2.2). После того как первый нуклеотид РНК инкорпорируется в сайт инициации транскрипции, полимеразы начинают продвигаться по нити ДНК в направлении 5'-3', расплетая двойные нити ДНК впереди себя и заплетая их позади. Этот процесс продолжается до достижения терминирующего сигнала, представляющего собой один или несколько терминирующих кодонов. Затем молекулы РНК и фермента высвобождаются и двойная спираль (дуплекс) ДНК полностью восстанавливается.

Для правильного начала синтеза РНК необходимо точное взаимодействие РНК-полимеразы с молекулой ДНК. Этот процесс контролируется промотором - специальной регуляторной последовательностью ДНК размерами около 75 п.о., локализованной, как правило, в 5'-фланкирующей области гена. Иногда под контролем одного промотора считывается несколько генов с образованием единого первичного РНК-транскрипта. Промоторные области различных генов довольно разнообразны по своему нуклеотидному составу. Однако, почти для всех промоторов характерно наличие консервативной последовательности из 7 оснований на расстоянии 19-27 нуклеотидов слева от сайта инициации транскрипции. Это, так называемый, ТАТА-бокс (блок Хогнесса), обеспечивающий корректное расположение РНК полимеразы по отношению к стартовому сайту. На расстоянии 70 - 80 п.о. в направлении 5'-конца от начала транскрипции часто расположена другая консервативная последовательность из 9 п.о. - СААТ-бокс, контролирующей начальное связывание РНК-полимеразы. Мутации в ТАТА- или в СААТ-боксах могут существенно влиять на скорость синтеза РНК. В 5'-фланкирующей области гена на расстоянии до тысячи пар оснований от начала его кодирующей части располагаются другие регуляторные последовательности, так называемые инхансеры (усилители), способные резко увеличивать продукцию гена за счет увеличения скорости транскрипции. Эти контролируемые элементы могут

работать независимо от их ориентации по отношению к сайту инициации. Для некоторых генов найдены участки ДНК, подавляющие транскрипцию, а также так называемые аттенюаторы (ослабители) – последовательности, лежащие между сайтом инициации транскрипции и собственно геном. Они могут блокировать продвижение РНК-полимеразы. Благодаря такому сложному механизму контроля, достигается очень тонкая и эффективная регуляция экспрессии генов практически на всех этапах транскрипции, трансляции и образования функционально зрелого белка. Эти механизмы более детально рассмотрены в других разделах.

Раздел 2.5 Изменчивость генома, полиморфные сайты рестрикции, ПДРФ-анализ.

Кодирующие и регуляторные области структурных генов на более консервативны в процессе эволюции, так как мутации в них подвержены давлению жесткого естественного отбора. Действительно, небольшие изменения в этих последовательностях, даже замена одного основания, делеция или инсерция нескольких нуклеотидов, могут привести к прекращению синтеза белка или к потере его функции, что, как правило, драматическим образом сказывается на жизнеспособности особей, несущих подобные мутации. Однако, около 90% генома человека состоит из некодирующих последовательностей, подобных сателлитным ДНК, умеренным повторам, интронам и спейсерным промежуткам между генами. Эти участки значительно более изменчивы и содержат множество, так называемых нейтральных мутаций или полиморфизмов, не имеющих фенотипического выражения и не оказывающих заметного влияния на жизнеспособность или репродуктивные свойства особей и, таким образом, не подверженных прямому давлению естественного отбора. Полиморфные локусы являются удобными генетическими маркерами. На основе анализа родословных можно проследить их наследование в ряду поколений, проанализировать сцепление друг с другом, с известными генами и с анонимными последовательностями ДНК, то есть использовать в качестве обычных менделевских признаков в классическом генетическом анализе. Информативность полиморфных локусов определяется уровнем их генетической изменчивости в различных популяциях.

Экспериментально легко выявляются два варианта геномного полиморфизма. Количественные изменения в области локализации мини- и микросателлитных последовательностей ДНК и качественные замены отдельных нуклеотидов, приводящие к появлению полиморфных сайтов рестрикции. В первом случае изменчивость по числу повторенных "коровых" единиц создает серию аллелей, характер и частота которых уникальны для каждого вариативного локуса. Полиморфизм в сайтах рестрикции связан с присутствием точковых нейтральных мутаций, локализованных, как правило, в уникальных последовательностях некодирующих участков ДНК. Подобные мутации в силу вырожденности генетического кода (см.Глава 1) могут возникать и в кодирующих последовательностях генов. Спонтанные мутации, возникающие в сайтах узнавания для определенных рестриктаз, делают их резистентными к действию этих ферментов. Аналогичным образом, при таких заменах могут создаваться новые сайты рестрикции. Показано, что полиморфные локусы встречаются во всех хромосомах с частотой приблизительно один полиморфный сайт на 300-500 п.о. Этот тип изменчивости ДНК был выявлен и использован для молекулярной маркировки специфических участков генома исторически раньше по сравнению с вариативными сателлитными повторами (Botstein et al., 1980).

Мутационная изменчивость в сайтах рестрикции может быть легко обнаружена по изменению длины рестрикционных фрагментов ДНК, гибридизующихся со специфическими ДНК-зондами. Анализ полиморфизма длины рестрикционных фрагментов, так называемый ПДФ-анализ (Restriction Fragment Length Polymorphism - RFLP analysis), включает следующие этапы: выделение геномной ДНК, ее рестрикцию специфической эндонуклеазой, электрофоретическое разделение образующихся фрагментов ДНК и идентификацию фрагментов ДНК, содержащих полиморфный сайт рестрикции, путем блот-гибридизации по Саузерну (см. Главу 1). При отсутствии рестрикции в полиморфном сайте на электрофореграммах или радиоавтографах (в зависимости от типа меченая ДНК-зонда) будет выявляться один крупный фрагмент, соответствующий по длине последовательности ДНК между двумя соседними константными сайтами рестрикции для той же эндонуклеазы. При наличии рестрикции в полиморфном локусе на электрофореграмме будет присутствовать меньший по размерам фрагмент, равный расстоянию между полиморфным сайтом рестрикции и одним из ближайших константных сайтов рестрикции. С каким именно из двух фрагментов, прилегающих к полиморфному локусу, будет происходить гибридизация зависит от локализации используемого для анализа ДНК-зонда. В частном случае возможна гибридизация одновременно с двумя соседними рестрикционными фрагментами, если выбранный ДНК-зонд комплементарен последовательности, содержащей полиморфный сайт рестрикции. Однако, такие зонды очень редко используются на практике, так как длина рестрикционных фрагментов обычно в десятки раз больше длины ДНК-зондов и далеко не всегда удается выделить и проклонировать фрагмент ДНК, содержащий полиморфный сайт рестрикции. Поэтому в дальнейшем для простоты изложения мы будем рассматривать только более общую ситуацию и считать, что при отсутствии рестрикции в полиморфном сайте ДНК-зонд гибридизуется с одним длинным фрагментом, а при наличии рестрикции гибридизующийся фрагмент имеет меньшую длину. Таким образом, при анализе ДНК особей, в обеих хромосомах которых присутствует сайт рестрикции в полиморфной области, на электрофореграмме будет выявлен только один бэнд в нижней области геля, соответствующий более короткому фрагменту ДНК. У особей, гомозиготных по мутации, изменяющей полиморфный сайт рестрикции, будет наблюдаться один бэнд в верхней части геля, соответствующий фрагменту большей длины, тогда как у гетерозигот проявятся оба эти бэнды (Рис. 2.3а).

ПДФ-анализ может быть значительно упрощен в том случае, если возможна специфическая амплификация участка ДНК, содержащего полиморфный сайт рестрикции. Тестирование состояния этого локуса возможно путем проведения ПЦР и рестрикции амплифицированного фрагмента. При отсутствии сайта узнавания в исследуемой области ДНК размеры амплифицированного фрагмента не изменятся после его обработки соответствующей эндонуклеазой, тогда как при полном соответствии полиморфной области сайту рестрикции образуются два фрагмента меньшей длины (Рис. 2.3б). У гетерозигот будут присутствовать 3 фрагмента, один из которых по длине будет соответствовать размеру амплификата до рестрикции, плюс 2 маленьких фрагмента с той же суммарной длиной. Таким образом, как и в случае использования для анализа блот-гибридизации по Саузерну, трем возможным вариантам генотипа будут соответствовать три различных варианта электрофореграмм.

Необходимо подчеркнуть, что первичная идентификация полиморфных сайтов рестрикции, сцепленных с определенными генами, возможна только при наличии соответствующих ДНК-зон-

дов. Дальнейшая тактика заключается в поиске рестрикционной нуклеазы, выявляющей полиморфизм. С этой целью используют широкий набор эндонуклеаз для рестрикции геномной ДНК, выделенной из группы неродственных индивидуумов, представляющих собой репрезентативную выборку популяции. Затем проводят ПДРФ-анализ отдельно для каждой из рестриктаз с набором имеющихся ДНК-зондов. После обнаружения полиморфизма в одной из популяций аналогичные исследования проводят в других популяциях.

Следует учитывать, что полиморфные сайты рестрикции всегда представляют собой двухаллельную систему, поэтому даже при самой высокой вариабильности такого сайта, число гетерозиготных по данному локусу особей в популяции не будет превышать 50%. Между тем, только при наличии полиморфного сайта в гетерозиготном состоянии можно отличить мутантный аллель от нормального, то есть осуществить его молекулярную маркировку. Следовательно, информационная емкость такого полиморфизма относительно невелика (см. Главы V и VII).

Раздел 2.6 Вариабильные микро- и минисателлитные ДНК.

Гипервариабильные сателлитные повторы являются гораздо более информативными маркерами по сравнению с полиморфными сайтами рестрикции, так как представляют собой мультиаллельные системы с уровнем гетерозиготности, достигающим 70 - 90%. Кроме того, оказалось, что количество высокоизменчивых микро- и минисателлитных последовательностей в геноме человека, по-видимому, превышает несколько десятков тысяч, они достаточно плотно и равномерно расположены в каждой из хромосом. Так более 90% из 5000 идентифицированных (С-А)-повторов являются полиморфными, причем в большинстве из них уровень гетерозиготности значительно превышает 50% (Weissenbach et al., 1992). Среди гипервариабильных мини-сателлитных последовательностей различают варьирующие по числу тандемные повторы - VNTR, три-, тетра- и пентануклеотидные повторы, а также некоторые другие классы повторов. Общее число высокополиморфных минисателлитных последовательностей в геноме превышает 1500 (Armour et al., 1990; Charlesworth et al., 1994). VNTR (variable number tandem repeats - варьирующие по числу тандемные повторы) - характеризуются наличием 10 - 15 - ти нуклеотидных "коровых" последовательностей, сходных с контролируемыми элементами рекомбинации E.coli (Jeffreys et al., 1985). С помощью ДНК-зондов, сконструированных на основе тандемно повторяющихся "коровых" последовательностей, можно анализировать индивидуальную изменчивость в единичных или множественных высокополиморфных локусах, несущих однотипную коровую последовательность. Примером может служить VNTR, обнаруженная в интроне миоглобинового гена, включающая четыре тандемных повтора из 33 п.о., фланкированных прямыми повторами из 9 п.о. Полученные из этого района ДНК-зонды с успехом используются для идентификации личности методом ДНК-фингерпринта, так как вероятность совпадения аллелей у двух неродственных индивидуумов по всем гипервариабильным локусам, гибридизующихся с этим ДНК-зондом, значительно меньше 10^{-7} (Jeffreys et al., 1991). Высокополиморфные VNTR найдены также в гене инсулина, в области глобинового псевдогена и в X-хромосоме, содержащей последовательности, гомологичные ДНК вируса гепатита В (Nakamura et al., 1985). В последнее время многочисленные VNTR-последовательности обнаружены в хромосомах 1, 2q, 14, 16, 17, 19 (O'Brien, 1992). Поиск новых гипервариабильных локусов основан на системати-

ческом скрининге клонированных последовательностей ДНК с помощью искусственно синтезированных "коровых" зондов. Особенно удобны для такого скрининга геномные библиотеки, сконструированные на основе предварительно фракционированных по величине MboI или Sau3A1 фрагментов ДНК, так как большие фрагменты обогащены длинными и более вариабильными минисателлитными последовательностями (Armour, et al., 1990). Хорошие результаты были получены также при скринировании космидных библиотек с помощью G-богатых синтетических ДНК-зондов и использование хемолюминесцентного метода в сочетании с ПЦР (Decorte, Cassiman, 1990). В некоторых случаях с этой целью применяют также последовательности ДНК, выделенные из фага M13, имеющего в своем составе сходные структуры (Armour et al., 1990).

Одним из вариантов минисателлитных последовательностей являются тримерные и тетрамерные короткие tandemные повторы -STR (short tandem repeats), которые стабильно наследуются и также обладают высоким уровнем полиморфизма по числу коровых единиц (Edwards et al., 1991). В X хромосоме такие повторы обнаружены через каждые 300 - 500 кб, тогда как в других частях генома они встречаются на расстоянии 10 кб друг от друга. По некоторым оценкам их частота может достигать 1/20 кб (Charlesworth et al., 1994). Возможно, истинное число три- и тетрамерных повторов в геноме человека еще больше, так как они обнаружены во многих генах. Подобно другим повторам STR обычно встречаются в некодирующих частях генов. В последнее время, однако, обнаружена группа структурных генов, несущих тримерные повторы в регуляторных и даже в транслируемых частях генов. Изменения числа этих внутригенных повторов в сторону их увеличения может приводить к нарушению функции этих генов вплоть до полного блока экспрессии и быть причиной ряда тяжелых наследственных заболеваний - болезни экспансии (см. Главу X). В качестве зондов для выявления коротких tandemных повторов используют синтетические олигонуклеотиды, построенные из простых повторов трех или четырех нуклеотидов.

В последнее время для обнаружения различных классов микросателлитных последовательностей и STR-повторов применяют эффективные методы, основанные на использовании ПЦР (Edwards et al., 1991). Для этих локусов характерно большое число аллелей, которые различаются по числу коровых единиц. Продукты амплификации этих сайтов могут отличаться друг от друга числом ди- три- или тетра-нуклеотидных повторов. Для удобства идентификации различных аллелей полимеразную цепную реакцию проводят в присутствии меченых нуклеотидов, а для электрофоретического разделения продуктов амплификации используют специальные секвенирующие гели. Многие STR-локусы настолько полиморфны, что нашли применение в судебной медицине для идентификации личности. Для повышения достоверности результатов с этой целью используют мультиплексные варианты ПЦР, позволяющие устанавливать генотип индивидуума одновременно по нескольким STR-сайтам (Weber, May, 1994; Асеев и др., 1995). Возможность точного генотипирования личности с использованием полимеразной цепной реакции, то есть на минимальном количестве ДНК, является важным методическим преимуществом STR перед ранее рассмотренными VNTR, для анализа которых чаще используют блот-гибридизацию по Саузерну.

Высокая частота коротких tandemных повторов, уникальность комбинаций числа тетра-, три- и димеров в разных сайтах в сочетании с их большой вариабильностью и сравнительной легкостью идентификации аллелей позволяет широко использо-

вать эти повторы для генетического и физического картирования в качестве наиболее удобных индексных маркеров геномных ДНК-последовательностей (см. Главу III). Такие маркерные сайты получили название STS (sequence tagged sites). В настоящее время в геноме человека уже идентифицировано около 10 000 STS, подавляющее большинство которых представляет собой тандемные повторы 2 - 4 нуклеотидов. Благодаря выраженной индивидуальной специфичности и достаточно стабильному менделевскому типу наследования STS-сайты нашли широкое применение и в молекулярной диагностике генных болезней, прежде всего в качестве молекулярных маркеров для идентификации мутантных хромосом в семьях высокого риска (см. Главу VII). Наличие большого числа гипервариабильных микро- и минисателлитных последовательностей ДНК является характерной особенностью генома человека. Аналогичные последовательности, обнаруженные в геноме приматов, значительно более однородны, что доказывает возможность существенного увеличения вариабильности этих участков ДНК за сравнительно короткий эволюционный промежуток (Юров, 1988; Gray et al., 1991).

Сведения о мутабельности высокополиморфных последовательностей в геноме человека весьма противоречивы. Показано, однако, что в наиболее вариабильных минисателлитных локусах частота мутаций может достигать 5% на гамету (Jeffreys et al., 1988). Предполагается, что одной из главных функций гипервариабильных микро- и минисателлитных последовательностей ДНК может быть контроль гомологичной рекомбинации в мейозе. На культурах клеток показано стимулирующее влияние минисателлитных последовательностей ДНК на гомологичную рекомбинацию. Так, инсерция синтезированной последовательности, составленной на основе гипервариабильных минисателлитов в геномную ДНК приводит к более, чем 10-кратному увеличению числа реципрокных обменов, причем степень этого влияния обратно пропорциональна расстоянию между STR и сайтом рекомбинации (Wahls et al., 1990). Вместе с тем, многие авторы обращают внимание на достаточно высокую стабильность минисателлитных аллелей, что позволяет их широко использовать как для генетического маркирования, так и для популяционных исследований и идентификации личности методом ДНК-фингерпринта (Decorte, Cassiman 1993; Edwards et al., 1991; Иванов, 1989).

Для многих мутаций, локализованных в некодирующих частях генома, характерны высокие уровни популяционного полиморфизма. Необходимо, однако, подчеркнуть, что эта изменчивость не затрагивает общей структуры генома, определяющей различия между видами. Более того, сопоставление первичных нуклеотидных последовательностей сравнительно протяженных секвенированных участков ДНК (области Т-рецепторных генов длиной около 100 кб) обнаружило сохранение высокой степени гомологии не только в кодирующих, но и, что особенно удивительно, в некодирующих частях этих последовательностей. Если учесть, что эволюционно человек и мышь разделены почти 80 миллионами лет эволюции, эти данные рассматриваются как свидетельство функциональной значимости некодирующих частей этих генов. По-видимому, далеко не всякие мутации в некодирующих районах ДНК являются нейтральными и в определенных случаях они могут отрицательно влиять на жизнеспособность. К сожалению, в настоящее время ничего или почти ничего неизвестно о функциях некодирующих ДНК-последовательностей. Высказывалось даже предположение, что их единственной функцией является репликация. Отсюда возникло представление об "эгоистической" или "паразитической" ДНК. Конечно, полностью

исключить наличие подобных паразитических последовательностей ДНК в любом геноме нельзя. Тем ни менее, представляется маловероятным, что значительная часть генома человека, также как и других видов, относится к эгоистической ДНК. По-видимому, наши знания о роли некодирующей или, как еще говорят, "избыточной" ДНК все еще явно недостаточны. Стабильность структурной организации генома в пределах вида свидетельствует скорее о важной эволюционной роли некодирующих ДНК-последовательностей и об их участии в процессах онтогенеза. Можно предполагать, что ответ на этот интригующий вопрос в какой-то мере будет получен при расшифровке и сравнении полной первичной нуклеотидной последовательности геномов у животных разных видов и, прежде всего, у человека и мыши, где прогресс в секвенировании геномной ДНК особенно значителен (см. Главу III). Уместно заметить, что проведенный недавно компьютерный анализ генома человека позволяет предполагать наличие в его некодирующей части особого, пока еще непонятого генетического кода, смысл и значение которого остаются загадочными (?).

Раздел 2.7 Мобильность генома, облигатные и факультативные элементы генома.

До сих пор мы рассматривали основные структурные элементы генома человека, положение которых в соответствии с представлениями классической генетики достаточно постоянно. Начиная с 50-х годов стали накапливаться данные о существовании большого числа мобильных генетических элементов, присутствие которых в геноме не является обязательным, а их топография и количество может варьировать в различных клетках, тканях и у разных индивидуумов (McClintock, 1984; Berg, Howe, 1989). У прокариот такие элементы получили название транспозонов. Их структура и функции достаточно хорошо изучены. Отличительной особенностью мобильных элементов является способность существовать как в интегрированном с хромосомой виде, так и в виде отдельных макромолекул - эписом, плазмид, вирусных частиц. Почти 50 различных семейств мобильных элементов описано у дрозофилы. Вместе эти последовательности составляют около 12% гаплоидного набора (Golubovsky, 1995). В геноме млекопитающих содержится до 50 000 диспергированных копий ретропозона LINE размером около 6500 пар оснований. Семейство Alu-повторов, содержащее от 300 до 500 тысяч копий, также относится к числу мобильных элементов генома (Charlesworth et al., 1994). Явление лизогенеза, то есть присутствие вирусных последовательностей в составе ДНК человека и наличие фрагментов генов человека в вирусных геномах, служит одним из примеров мобильности ДНК и возможности "горизонтальной" передачи наследственно закрепленных признаков между видами. Мобильные ДНК, как правило, относятся к факультативным элементам. Как уже отмечалось, не существует четких границ между облигатными и факультативными элементами генома, так как возможен взаимный переход от одного состояния к другому. Структурные локусы или сегменты хромосом могут трансформироваться в факультативные элементы за счет амплификации, интеграции в мобильные элементы или путем образования цитоплазматических ретротранскриптов. Обратный переход от факультативных элементов к облигатным осуществляется посредством инсерций, транспозон-индуцированных перестроек и обратной транскрипции.

Факультативные элементы существуют в геноме как популяции информативных макромолекул. Изменения, возникающие в них

под воздействием внешних факторов, носят совершенно иной характер по сравнению с классическими мутациями в структурных локусах. Для описания изменений в факультативных элементах предложен термин "вариации" (Голубовский, 1985). Этот термин впервые использован Жакобом и Воллманом для описания поведения эписом (Jacob, Wollman, 1961). Вариации могут приводить к изменениям на генотипическом уровне, то есть к мутациям, вследствие простого перемещения факультативных элементов или сдвига в соотношении между факультативными и облигатными элементами. В этих случаях мутации встречаются одновременно у многих индивидуумов. Подобные изменения упорядочены, могут происходить сразу во многих локусах и отличаются высокой сайт-специфичностью. Локализация структурных перестроек, возникающих в результате вариаций, предопределена первоначальной топографией факультативных элементов на хромосомах. И наконец, сами вариации могут быть индуцированы обычными "не-мутагенными" факторами, такими как температура или межлинейные кроссы (Golubovsky, 1995). Факультативные элементы могут рассматриваться как оперативная память генома, так как во многих случаях спонтанное возникновение мутаций в облигатных элементах опосредовано их активацией. Считается, в частности, что инсерционный мутагенез является причиной спонтанного возникновения 70% видимых мутаций в природных популяциях дрозофилы. Однако, у человека пока зарегистрированы лишь единичные случаи возникновения мутаций вследствие перемещения мобильных элементов генома (Vidaud et al., 1993).

Раздел 2.8 Изохоры, метилирование, гиперчувствительные сайты.

Перечисленные выше компоненты генома не случайным образом связаны с последовательностями нуклеотидов. И в этом смысле можно говорить о существовании в геноме человека структур более высокого иерархического порядка. Примером служат изохоры - длинные, в среднем, свыше 300 кб сегменты ДНК, гомогенные по композиции оснований или по GC-уровням. 62% генома состоит из GC-бедных изохор и в них локализовано около 34% генов, 31% генома представлен GC-богатыми изохорами, содержащими 38% генов, и в 3% изохор, обогащенных GC-последовательностями (так называемых H3 изохор), находится 28% генов (Mouchiroud et al., 1991; Saccone et al., 1993). Таким образом, существуют относительно небольшие участки ДНК, в которых плотность генов в 10-20 раз выше, чем в остальных последовательностях.

Другой общей чертой генома человека является то, что *in vivo* значительная доля цитозинового остатков в молекуле ДНК метилирована, то-есть находится в форме 5-метилдезоксицитидина. Экспериментальное изучение характера метилирования основано на сопоставлении рестрикционных фрагментов, образующихся после обработки ДНК эндонуклеазами, для которых сайты узнавания одинаковы и содержат в своем составе цитозин, но действуют эти ферменты по-разному, в зависимости от того, находится ли это основание в метилированном состоянии или нет. В частности, рестриктазы - Msp1 и Hpa11, узнают последовательность CCGG, но в отличие от Msp1, Hpa11 не расщепляет ДНК в тех сайтах, где внутренний CpG динуклеотид метилирован. Некоторые сегменты генома, особенно это относится к повторяющимся последовательностям, полностью метилированы в местах 5'-CCGG-3' и частично метилированы в 5'-GCGC-3' - сайтах рестрикции для Hha1. В других сегментах наблюдается

характерный рисунок частичного метилирования в 5'-CCGG-3' последовательностях (Behn-Krappa et al., 1991). Различные индивидуумы, независимо от их этнического происхождения, практически не различаются по характеру метилирования ДНК в одних и тех же типах тканей, тогда как в процессе онтогенетической дифференцировки происходят значительные изменения рисунков метилирования. В перевиваемых культурах клеток опухолевого происхождения число метилированных сайтов резко уменьшено.

Высказано предположение о наличии прямой связи между метилированием ДНК и состоянием генетической активности в клетках. Существует класс белков, которые специфическим образом связываются с метилированными участками ДНК, делая их недоступными для действия ряда ферментов, в том числе, возможно, и для полимераз. Получено много прямых экспериментальных доказательств роли метилирования ДНК в инактивации эукариотических промоторов, а, значит, и в регуляции активности генов. Напротив, гипометилирование промоторной области генов, в особенности CpG островков, как правило, свидетельствует о функциональной активности генов. Показано, что необычные структуры в молекуле ДНК, также как экзогенная ДНК, инкорпорированная в процессе генетической трансформации, нередко подвергаются метилированию. Известно, что метилирование играет важную роль в инактивации X хромосомы у самок, в регуляции экспрессии генов в процессе развития, а также непосредственно вовлечено в феномен хромосомного (геномного) импринтинга, связанного с различиями пенетрантности некоторых аллелей в зависимости от их происхождения, то есть прохождения через материнский или отцовский гаметогенез (Баранов, 1991).

В GC-богатых изомерах локализовано большое количество CpG островков - последовательностей от 500 до 2000 п.о., характеризующихся очень высоким содержанием гуанина и цитозина ($G+C > 60\%$), представленных в виде кластеров неметилированных CpG дуплетов и, так называемых, G/C боксов - локусов, родственных сайту узнавания для одного из транскрипционных факторов Sp1 - (G)4C(G)4C (Lindsay, Bird, 1987; Bird, 1986; Aissani, Bernardi, 1991). CpG острова содержат много сайтов узнавания для чувствительной к метилированию эндонуклеазы HpaII, а также сайты для редкощепящих рестриктаз, узнающих неметилированные CpG дуплеты. В частности, более 80% Nor1-сайтов связано с CpG-богатыми островками. Как правило, CpG островки локализованы в 5'-фланкирующих последовательностях, 5'-экзонах и 5'-интронах всех изученных хаузкипинг-генов и 40% тканеспецифических генов. CpG островки являются характерной особенностью транскрибируемых участков генома. Их идентификация в клонированных последовательностях геномных библиотек существенно облегчает поиск конкретных структурных генов (см. раздел 2.4). Наибольшая плотность CpG островков наблюдается в теломерных участках хромосом 1, 9, 15, 16, 17, 19, 20, 22 (Antonarakis, 1994). Точные молекулярные методы регистрации CpG островков показали, что их число в геноме человека приближается к 45000 (Antequera, Bird, 1993).

Можно также отметить существование в геноме человека сайтов, гиперчувствительных к действию ДНК-азы 1 и структурно отличающихся от основной массы хроматина. Присутствие таких сайтов показано для многих генов млекопитающих и, по-видимому, это необходимое, но не достаточное условие их экспрессии. Локализация гиперчувствительных сайтов может меняться в процессе развития и под действием гормонов. В неко-

торых случаях эти участки маркируют положение транскрипционных регуляторных элементов генома, действующих как в положительном, так и в отрицательном направлениях. В других случаях это области функционально активных генов, находящихся в деспирализованном состоянии и имеющих однонитевую структуру. Именно такие однонитевые участки ДНК особенно высокочувствительны к ДНК-азе 1. На этом их свойстве основан метод ник-трансляции *in situ*, позволяющий непосредственно на хромосомных препаратах визуализировать функционально активные районы хромосом. С этой целью хромосомные препараты обрабатывают ДНК-азой 1, после чего непосредственно на них с помощью ДНК-полимеразы проводят синтез ДНК в присутствии меченых нуклеотидов. При этом метка включается преимущественно только в те участки хромосом, где находятся функционально активные гены (Verma, Babu, 1989).

ГЛАВА II.

ГЕНОМ ЧЕЛОВЕКА, СТРУКТУРА ГЕНОВ.

Раздел 2.1. Определение генома и его основных элементов.

Термин геном используется для обозначения полной генетической системы клетки, определяющей характер онтогенетического развития организма и наследственную передачу в ряду поколений всех его структурных и функциональных признаков. Понятие генома может быть применено к таксономической группе, виду, отдельной особи, клетке, микроорганизму или вирусу. Так, можно говорить о структуре генома эукариот и прокариот, сравнивать геномы разных видов, изучать особенности строения генома у конкретных индивидуумов или следить за изменениями, происходящими в геноме специфических клеток в процессе их онтогенетической дифференцировки. Часто геном определяется как генетическая информация, заключенная в молекулах ДНК одной клетки. Однако, такие факты, как отсутствие связи между количеством ДНК в расчете на гаплоидный геном и таксономическим статусом видов, а также многочисленные примеры существования огромных различий в содержании ДНК между близкородственными видами (так называемый "С-парадокс") свидетельствуют о том, что далеко не все участки ДНК связаны с информационными функциями. Понятия генома и ДНК в значительной степени тождественны, так как основные принципы организации и функционирования генома целиком определяются свойствами ДНК. Присущие этим молекулам потенциальные возможности практически неограниченного структурного разнообразия определяют все многообразие мира живых существ, как на уровне межвидовых, так и индивидуальных различий в пределах одного вида (Баев и др., 1990; Ратнер, 1985).

Процесс эволюции и дифференцировки отдельных видов, как правило, сопровождался накоплением изменений в структуре генома. Это касается, прежде всего, таких параметров, как локализация и характер упаковки ДНК в клетках; количество ДНК, приходящееся на гаплоидный геном; типы, соотношение и функции кодирующих и не кодирующих нуклеотидных последовательностей; регуляция экспрессии генов; межпопуляционная вариабельность и филогенетический консерватизм первичной структуры генома. В пределах одного вида основные параметры генома достаточно постоянны, а внутривидовое разнообразие обеспечи-

вается за счет мутационной изменчивости, то есть выпадения, вставки или замены нуклеотидов на сравнительно небольших участках ДНК. Чаще всего такие изменения касаются не-экспрессируемых элементов генома (интронов, псевдогенов, межгенных спейсерных участков ДНК и т.д.).

Геномы эукариот, по-существу, можно рассматривать как мультигеномные симбиотческие конструкции, состоящие из облигатных и факультативных элементов (Golubovsky, 1995). Основу облигатных элементов составляют структурные локусы, количество и расположение которых в геноме достаточно постоянно. Присутствие в хромосомах некоторых видов повторяющихся ДНК, амплифицированных участков, ретровирусных последовательностей, псевдогенов, также как наличие в клетке эписом, ретротранскриптов, ампликонов, дополнительных В-хромосом и различных цитосимбионтов (вирусов, бактерий, простейших) является не строго обязательным, их количество и положение может значительно варьировать, то есть эти элементы являются факультативными. В то же время участие факультативных элементов в наследственной передаче признаков, в формировании мутационной изменчивости и в эволюционных преобразованиях видов несомненно доказано. Кроме того, существует непрерывный переход от одних состояний к другим за счет инсерции экстрахромосомных ДНК в хромосомы и выстраивания транспозоноподобных мобильных элементов из хромосом. Следовательно, несмотря на значительные отличия факультативных последовательностей от облигатных по характеру основных информационных процессов (репликации, транскрипции, трансляции и сегрегации), они также должны рассматриваться, как важнейшие элементы генома.

Остановимся теперь более детально на основных принципах организации генома человека. В каждой диплоидной клетке с 46 хромосомами содержится около 6 пикограмм ДНК, а общая длина гаплоидного набора из 23 хромосом составляет $3.5 \cdot 10^9$ пар нуклеотидов (Као, 1985). Этого количества ДНК достаточно для кодирования нескольких миллионов генов. Однако, по многим независимым оценкам истинное число структурных генов находится в пределах от 50 000 до 100 000. В разделе 2.4 изложены современные подходы, используемые для подсчета общего количества генов, из которых следует, что наиболее вероятная оценка их числа составляет около 80 000. Сопоставляя это значение со средними размерами гена и соотношением между величиной их экзонных и интронных областей, можно заключить, что кодирующие последовательности ДНК занимают не более 10-15% всего генома (McKusick, Ruddle, 1977). Таким образом, основная часть молекул ДНК не несет информации об аминокислотной последовательности белков, составляющих основу любого живого организма, и не кодирует структуру рибосомальных, транспортных, ядерных и других типов РНК. Функции этой "избыточной" (junk) ДНК не ясны, хотя ее структура изучена достаточно подробно. Предполагается, что эта ДНК может участвовать в регуляции экспрессии генов и в процессинге РНК, выполнять структурные функции, повышать точность гомологичного спаривания и рекомбинации, способствовать успешной репликации ДНК и, возможно, является носителем принципиально иного генетического кода с неизвестной функцией.

Наиболее общая характеристика генома может быть получена с помощью анализа кинетики реассоциации молекул ДНК. Динамика плавления геномной ДНК обнаруживает присутствие по крайней мере трех различающихся по химической сложности фракций (Льюин, 1987; Газарян, Тарантул, 1983). Быстро ренатурирующая фракция ДНК состоит из относительно коротких высокопов-

торящихся последовательностей; в промежуточную фракцию входит множество умеренно повторяющихся ДНК – более протяженных, но представленных меньшим числом копий; медленно ренатурирующая фракция объединяет в себе уникальные последовательности ДНК, встречающиеся в геноме не более 1-2 раз.

С помощью молекулярного анализа проведена идентификация основных классов повторяющихся последовательностей ДНК, составляющих более 35% всего генома человека и включающих сателлитную ДНК, инвертированные повторы, умеренные и низкокопийные повторы, а также мини- и микросателлитные последовательности ДНК. Классификация этих типов повторов достаточно условна и основана, главным образом, на двух характеристиках: длине повторяющихся коровых единиц, которая может варьировать от 1-2 до более, чем 2000 п.о., и числе их копий, также меняющихся в очень широких пределах – от десятка до миллиона на гаплоидный геном. Не менее важными характеристиками различных классов повторяющихся ДНК являются нуклеотидная последовательность "коровых" единиц повтора, специфичность их организации, хромосомная локализация, внутри- и межвидовая стабильность, а также возможные функции этих типов ДНК.

Раздел 2.2. Повторяющиеся последовательности ДНК.

Сателлитная ДНК это класс высокоповторяющихся последовательностей, составляющих около 10% всего генома человека (Као, 1985). При центрифугировании геномной ДНК в градиенте плотности CsCl эти последовательности образуют четыре отдельных сателлитных пика с различными средними значениями плавучей плотности. Методом гибридизации *in situ* показано присутствие сателлитной ДНК преимущественно в центромерных, теломерных и гетерохроматиновых районах большинства хромосом, при этом характер гибридизации сходен для всех четырех групп и не зависит от принадлежности ДНК-зондов к семействам повторов, образующих различные сателлитные пики. В каждой из этих групп, однако, присутствует небольшое количество последовательностей, имеющих специфическую хромосомную локализацию. Так например, около 40% длинного плеча Y хромосомы составляет семейство последовательностей, тандемно повторяющихся более 3000 раз и не найденных в других хромосомах.

Выделяют три основных типа сателлитной ДНК: (1) короткие – от 2 до 20 п.о., стабильные тандемные повторы с кратностью в несколько десятков тысяч раз, которые иногда перемежаются с неповторяющимися последовательностями; (2) кластеры более протяженных повторов, слегка различающихся по нуклеотидной последовательности; (3) сложные, достигающие нескольких сотен пар нуклеотидов, повторяющиеся последовательности различной степени гомологии (Газарян, Тарантул, 1983). К последнему типу относятся альфа-сателлитные или альфоидные ДНК, среди которых найдено много хромосом-специфических последовательностей. Размеры повторяющихся "коровых" единиц альфоидной ДНК составляют около 170-200 п.о. В геноме человека и других приматов эти мономеры организованы в кластеры по 20 и более "коровых" единиц. После расщепления рестриктазой BamHI в альфоидной ДНК выявляется серия фрагментов, длиной около 2000 п.о., в составе которых обнаруживаются альфоидные последовательности, специфичные для гетерохроматиновых районов разных хромосом человека (1, 3, 4, 5, 9, 6, 7, 11, 17, 19, X). В некоторых случаях эти повторы гомологичны двум разным хромосомам (9 и 15, 13 и 21, 18 и X) (Willard, Wayne, 1987).

Хромосом-специфические последовательности сателлитной альфа-идной ДНК нашли широкое применение в молекулярной цитогенетике в качестве ДНК-зондов, удобных для маркирования индивидуальных хромосом в метафазных и интерфазных клетках человека (Юров, 1987). Предполагается, что сателлитные ДНК играют важную роль в поддержании структур хромосом и, возможно, в их спаривании в процессе мейоза (Charlesworth et al., 1994).

Особое место среди сателлитных ДНК занимают микро- и минисателлитные последовательности, представляющие собой многочисленную группу рассеянных по всему геному относительно коротких tandemных повторов. Микросателлиты - это класс динуклеотидных повторов. Размер повторяющихся единиц в минисателлитных последовательностях может меняться от 3 - 4 до 10 - 15 нуклеотидов. Отличительной особенностью микро- и минисателлитов является наличие среди них большого количества участков, вариабильных по числу копий в кластере.

Инвертированные или обращенные повторы составляют до 5% генома. Они состоят из двух тождественных копий длиной около 300 п.о., ориентированных в противоположных направлениях на одной нити ДНК и лежащих на расстоянии от нуля до десятка тысяч пар нуклеотидов друг от друга (в среднем - 1,6 кб). Около 1/3 обращенных повторов не разделены промежуточными последовательностями и носят название палиндромов. Среднее расстояние между двумя различными парами инвертированных повторов около 12 кб, и их распределение по геному носит случайный характер. Комплементарные пары легко ассоциируют при отжиге, образуя шпильчатые структуры с дуплексной ножкой и однонитевой петлей, длина которой соответствует расстоянию между парой обращенных повторов. Вследствие этого оказываются приближенными достаточно удаленные друг от друга участки ДНК, что важно для работы ряда ферментов, обеспечивающих процессы репликации и транскрипции. Важно отметить и то, что однонитевой участок ДНК, образующий петлю, становится доступным для действия нуклеазы S1, специфически разрушающей однонитевую ДНК.

Группа умеренно повторяющихся последовательностей очень гетерогенна по длине и числу копий и составляет около 20% генома человека. Как правило, они распределены дисперсно по всем хромосомам, причем относительно короткие последовательности ДНК до 500 п.о., так называемые короткие диспергированные повторы - Sine, повторяются более 100 000 раз, в среднем, через каждые 2.2 кб. Число копий более длинных диспергированных последовательностей - Line, не превышает 10 000. Умеренные повторы найдены во всех структурных компонентах генома за исключением кодирующих областей генов. Два главных семейства умеренных повторов, Alu и Kpn1, занимают, по крайней мере, 10% генома и практически столько же занято несколькими сотнями других семейств повторов этого класса.

Основной единицей Alu семейства является короткая последовательность из 300 п.о., повторенная в геноме человека несколько сот тысяч раз, в среднем, через каждые 5 кб или через каждые 2 - 3 диспергированных повтора, принадлежащих другим семействам (Као, 1985; Льюин, 1987). При этом, кластеры Alu-повторов, как правило, лежат внутри R-дисков метафазных хромосом - Гимза отрицательных (G- дисков). Распределение Alu-повторов по геному весьма неравномерно как между хромосомами, так и по их длине. Так, в хромосомах 14, 16, 21 Alu-последовательности концентрируются в области центромеры, а в хромосомах 4, 19, 20, X и Y выраженные кластеры Alu повторов не найдены. Члены Alu семейства не полностью идентичны друг другу. Однако, все они содержат

сайт рестрикции для фермента AluI и имеют димерную структуру, то есть состоят из двух прямых повторов длиной около 130 п.о. с богатой аденином вставкой из 31 нуклеотида во втором мономере. Каждая Alu последовательность фланкирована прямыми повторами длиной от 7 до 20 п.о., различными для разных членов семейства и имеющими большую степень гомологии с транспозоноподобными элементами про- и эукариот. Некоторые члены Alu семейства могут транскрибироваться с помощью фермента РНК-полимеразы 111. Предполагается, что при определенных условиях образующиеся при этом молекулы РНК могут обратно транскрибироваться, что в свою очередь, может привести к появлению в клетках Alu-содержащих кДНК, обладающих свойствами ретропазонов, то есть способных инсертироваться в геномную ДНК. В литературе описаны случаи инсерционного мутагенеза Alu повторов, приводящие к гемофилии В и нейрофиброматозу типа 1. Однако, частота таких событий, по-видимому, невелика. Предполагается, что короткие умеренные повторы, подобные Alu семейству, участвуют в регуляции транскрипции, в процессинге РНК и в инициации репликации ДНК. Кроме того, обнаружена высокая степень гомологии Alu последовательностей с одним из видов низкомолекулярной РНК (7 S РНК), участвующей в секреции белков.

К числу Line повторов относится Kpn1 семейство, которое состоит из более длинных и значительно более гетерогенных последовательностей, рассеянных по всему геному. В ряде случаев члены Kpn1 семейства группируются в кластеры, образуя более длинные структуры, повторяющиеся несколько тысяч раз. Для некоторых членов этого семейства также доказана возможность инсерции кДНК-овых копий Kpn1- РНК-транскриптов в геномную ДНК и возникновение мутаций. Такое явление было обнаружено в одном случае гемофилии А. Некоторые Kpn1- последовательности не только транскрибируются, но и способны транслироваться (Charlesworth et al., 1994).

Раздел 2.3 Мультигенные семейства, псевдогены, онкогены.

Многие гены человека повторены в геноме от нескольких единиц до нескольких сотен раз и образуют мультигенные семейства (Газарян, Тарантул, 1983; Восток, Самнер, 1981; Као, 1985; Льюин, 1987). Эти гены обычно сгруппированы в кластеры в определенных районах одной, либо нескольких хромосом. Во многих мультигенных семействах наряду с функционально активными генами содержатся псевдогены - мутационно измененные последовательности, не способные транскрибироваться или продуцирующие функционально неактивный генный продукт. Примерами мультигенных семейств могут служить гены рибосомальных, транспортных и ядерных РНК, гены альфа- и бета-глобинов, тубулинов, миоглобина, актина, интерферона и многих других. В ряде случаев, возможна избирательная амплификация некоторых семейств генов в процессе их экспрессии, как, например, генов рибосомальных РНК. При этом число способных транскрибироваться копий генов увеличивается за счет их избирательной амплификации в сотни и даже тысячи раз, что сопровождается лавинообразным нарастанием доли соответствующего генопродукта в клетках. Особое место среди мультигенных семейств занимают супергены - очень большие кластеры из сотен функционально и структурно родственных генов, расположенных в сегментах отдельных хромосом. Классическим примером супергена может служить HLA комплекс, контролирующий главные антигены гистосовместимости. Он занимает район более 6000 кб на ко-

ротком плече хромосомы bр21 и состоит из серии тесно сцепленных генов, ответственных за синтез множества белков, включающих клеточные поверхностные антигены, молекулы иммунного ответа и некоторые компоненты комплемента. К супергенным семействам относятся три комплекса расположенных на разных хромосомах мультигенов, контролирующих синтез тяжелых и легких цепей иммуноглобулинов. Интересно, что в процессе дифференцировки В лимфоцитов, продуцирующих иммуноглобулины, происходит структурная перестройка этих семейств. При этом отдельные последовательности ДНК элиминируются, тогда как другие сливаются, так что структура генов иммуноглобулинов в зрелых В лимфоцитах значительно отличается от исходной, то есть от той, которая наблюдается в зародышевых клетках.

Одной из важных структурных особенностей генома человека является наличие так называемых псевдогенов, уникальных последовательностей, очень сходных по своей структуре с определенными нормальными генами, но в силу присутствия в кодирующих последовательностях целого ряда мутаций не способных транскрибироваться или правильно транслироваться с образованием структурно и функционально активного продукта. Псевдогены обнаружены для многих генов. Их количество варьирует от одной до нескольких десятков копий на геном и в этом случае они, как правило, расположены тандемно. Иногда псевдогены тесно сцеплены с нормальными генами, во многих случаях псевдогены и гены локализованы в разных хромосомах. Для некоторых моногенных заболеваний идентифицированы мутантные аллели, сходные с мутациями в соответствующих псевдогенах. В этих случаях обсуждается возможная роль псевдогенов в спонтанном мутационном процессе.

В геноме человека присутствуют также нуклеотидные последовательности, гомологичные генам некоторых вирусов. Впервые эти последовательности были идентифицированы в геноме вирусов, индуцирующих развитие опухолей у животных и человека, и потому они были названы онкогенами. Гомологичные последовательности в геноме человека носят название протоонкогенов. В настоящее время уже идентифицировано более 100 протоонкогенов. Белковые продукты протоонкогенов, по-видимому, играют важную роль в нормальной пролиферации клеток особенно на ранних стадиях эмбрионального развития, контролируя клеточный цикл и выбор геномной программы развития клетки. При возникновении специфических мутаций в протоонкогенах, а также при нарушениях регуляции их работы, выражающихся в гиперпродукции или в экспрессии в нетипичном месте или в несвойственный момент жизнедеятельности клетки, они начинают вести себя как онкогены, стимулируя неконтролируемое размножение и пролиферацию определенных клеточных клонов, что и может, в конечном счете, привести к формированию опухоли.

Раздел 2.4 Современное определение понятия "ген", транскрипция, регуляторные элементы генов.

Около 10-15% генома человека представлено уникальными транскрибируемыми последовательностями, составляющими основу структурных генов (Льюин, 1987). В настоящее время в понятие "ген" включается не только его транскрибируемая область - экзоны + интроны, но также фланкирующие последовательности - лидерная, предшествующая началу гена, и хвостовая нетранслируемая область, расположенная на 3' конце гена (Рис.2.1). В отличие от генов прокариот гены человека редко представлены одной непрерывной последовательностью и в подавляющем большинстве имеют прерывистую структуру. Относительно короткие

кодирующие участки - экзоны, чередуются с длинными интронами, которые транскрибируются и входят в состав первичного РНК-продукта, но затем при процессинге первичного РНК-транскрипта они вырезаются и не участвуют в трансляции. Процесс вырезания интронов из первичных транскриптов получил название сплайсинга. Таким образом, в зрелой мРНК интронные области отсутствуют, а экзоны составляют непрерывную кодирующую последовательность. Размеры зрелых мРНК нередко в десятки раз меньше первичных РНК-транскрипов и, соответственно, размеров самого гена.

Согласно классическим представлениям ген - это локус, на хромосоме, мутации в котором реализуются на уровне фенотипа. В молекулярной биологии ген трактуется как ассоциированный с регуляторными последовательностями фрагмент ДНК, соответствующий определенной единице транскрипции. Следовательно, представления о гене формальных генетиков далеко не полностью тождественны его физической единице и соотношения между этими двумя понятиями достаточно запутанные. Отметим некоторые причины этих противоречий. Известно, что мутации одного гена могут приводить к совершенно разным и даже в ряде случаев к комплементарным фенотипам. Результаты прямого секвенирования генома свидетельствуют о присутствии в нем значительного большего числа генов, чем можно ожидать от результатов мутационного анализа. Одна и та же последовательность ДНК в геноме может кодировать несколько различных белков, что достигается за счет так называемого альтернативного сплайсинга (образование разных мРНК из одного первичного РНК-транскрипта). В крупных интронах ряда генов обнаружены смысловые последовательности других генов ("ген в гене"), считываемые в противоположном направлении. Транскрипционные единицы генома могут перекрываться за счет наличия разных промоторов. Наконец, благодаря соматической рекомбинации структура транскрибируемых последовательностей некоторых генов может быть различной в разных клонах клеток одного организма (Т-клеточные рецепторы). Ситуация с определением понятия "ген" еще больше осложняется, если в это понятие включать многочисленные регуляторные последовательности. Возникает вопрос: "Как далеко от гена могут располагаться эти последовательности, чтобы их можно было включать в структуру гена?". Для многих целей оказывается удобным введенное в последнее время понятие "считаемый ген" - counting gene. Последний рассматривается как отдельная транскрибируемая единица ДНК или её часть, которая может транслироваться в одну или несколько взаимосвязанных аминокислотных последовательностей. Поэтому последовательность, дающая две транскрипционные единицы за счет альтернативного сплайсинга и, как следствие, два разных белка учитывается как один ген. Однако, если степень гомологии двух генопродуктов, имеющих общий транскрибируемый участок, невелика, то эти последовательности расцениваются как два разных гена (Fields et al., 1994).

По своему функциональному назначению гены могут быть разделены на две группы. Группа I представлена генами, кодирующими собственно белки; группа II - генами, контролирующими синтез рибосомальных, транспортных и ядерных РНК. По характеру экспрессии гены также могут быть подразделены на две группы - гены "домашнего хозяйства" (housekeeping genes), продукты которых необходимы для обеспечения жизнедеятельности любого типа клеток, и тканеспецифические гены, обеспечивающие специализированные функции клеток, то есть гены, функционально активные только в определенных типах клеток

(тканей) и только на определенных стадиях онтогенеза, так называемые гены терминальной дифференцировки.

Считается, что средние размеры гена человека составляют, примерно, от 10 до 30 кб. Однако, эта величина может колебаться от нескольких десятков до миллионов пар нуклеотидов. Согласно последним данным, самый маленький из известных генов - МСС-7, имеет размеры всего 21 п.о. (Gonzales-Pastor et al., 1994), а самый большой - ген дистрофина - 2.2 мегабаз. Гены отделены друг от друга протяженными промежутками - спейсерами, содержащими в своем составе большое количество повторяющихся последовательностей ДНК и нетранскрибируемые уникальные последовательности.

Рассмотрим более подробно современные данные о числе генов в геноме человека. Полученные разными методами оценки этой величины приведены в Табл.2.1. Исходя из размера генома (около 3 000 миллионов п.о.), и среднего размера одного гена порядка 10 - 30 тыс.п.о. (Gilbert, 1992), общее число генов должно быть порядка 100 000. При этом, как уже указывалось, распределение генов на хромосомах крайне неравномерно. Установлено, что более 90% генов находится в Гимза -отрицательных районах метафазных хромосом, в так называемых R-бэндах (Antequera, Bird, 1993). Проведенные недавно прямые исследования методом секвенирования показывают, что в R-бэндах их число достигает 43-50 на один бэнд, а в Гимза-положительных районах хромосом (G бэндах), - только 1-2 на 75-80 килобаз, то есть всего в геноме можно ожидать около 70 000 генов. Оценка числа генов по доле транскрибируемой части генома (всего 12%) дает совсем маленькую величину - 20 000 генов (Wagner et al., 1993). Методом исследования кинетики реассоциации РНК в культуре клеток число генов оценивается между 20 - 40 000 (Lewin, 1990). В то же время в клетках мозга число различных мРНК по некоторым данным достигает 97 000 (Wagner et al., 1993). Наиболее точные подсчеты числа генов человека проведены в последнее время и основаны на оценке числа CpG островков (Табл.2.1). Известно, что в промоторной области всех генов "домашнего хозяйства" и примерно у 40 % генов терминальной дифференцировки, обеспечивающих специализированные функции дифференцированных клеток, находятся области коротких динуклеотидных CpG повторов. Разработаны молекулярные методы точной регистрации этих участков, которые показали, что их число в геноме около 45 000, откуда число структурных генов оценивается в 67-70 000 (Antequera, Bird, 1993). Практически такое же число генов (64 000) определено недавно методом учета маркерных экспрессирующихся последовательностей (expressed sequence tags) (Fields et al., 1994). Таким образом, суммируя вышеизложенное, можно сделать вывод, что в геноме человека содержится, в среднем, около 70-80 000 отдельных транскрибируемых ДНК последовательностей, то есть генов.

Транскрипция гена начинается с 5' конца первого экзона, где расположен сайт инициации транскрипции. Определенной гомологии между стартовыми сайтами разных генов не наблюдается, но чаще всего они начинаются с нуклеотида А, окруженного пиримидиновыми основаниями. На границах между экзонами и интронами имеются консервативные канонические последовательности, играющие существенную роль в обеспечении точности вырезания интронов во время сплайсинга РНК. Все интронные последовательности начинаются с динуклеотида GT и заканчиваются динуклеотидом AG, называемыми, соответственно, донорными и акцепторными сайтами сплайсинга. На 3' конце многих структурных генов идентифицирована поли(А)-сигнальная после-

довательность (AATAAA), участвующая в процессе модификации первичного РНК транскрипта и ответственная за альтернативный сплайсинг мРНК, обеспечивающий синтез разных зрелых мРНК с одного и того же первичного РНК транскрипта.

Транскрипция генов осуществляется с помощью фермента РНК-полимеразы. Около 50-70% клеточного синтеза РНК обеспечивается РНК-полимеразой I, локализованной в ядрышках и ответственной за синтез генов рибосомальной РНК. РНК-полимераза II обеспечивает транскрипцию генов, кодирующих собственно структурные белки. Этот фермент локализован в ядре (но не в ядрышках). На его долю приходится 20 - 40% синтеза РНК. РНК-полимераза III контролирует синтез ядерных и транспортных РНК (Льжин, 1987). На 1-м этапе РНК-полимераза связывается с двухнитевым участком ДНК и, расплетая его, делает доступным спаривание смысловой нити ДНК с рибонуклеотидами (Рис. 2.2). После того как первый нуклеотид РНК инкорпорируется в сайт инициации транскрипции, полимеразы начинают продвигаться по нити ДНК в направлении 5'-3', расплетая двойные нити ДНК впереди себя и заплетая их позади. Этот процесс продолжается до достижения терминирующего сигнала, представляющего собой один или несколько терминирующих кодонов. Затем молекулы РНК и фермента высвобождаются и двойная спираль (дуплекс) ДНК полностью восстанавливается.

Для правильного начала синтеза РНК необходимо точное взаимодействие РНК-полимеразы с молекулой ДНК. Этот процесс контролируется промотором - специальной регуляторной последовательностью ДНК размерами около 75 п.о., локализованной, как правило, в 5'-фланкирующей области гена. Иногда под контролем одного промотора считывается несколько генов с образованием единого первичного РНК-транскрипта. Промоторные области различных генов довольно разнообразны по своему нуклеотидному составу. Однако, почти для всех промоторов характерно наличие консервативной последовательности из 7 оснований на расстоянии 19-27 нуклеотидов слева от сайта инициации транскрипции. Это, так называемый, ТАТА-бокс (блок Хогнесса), обеспечивающий корректное расположение РНК полимеразы по отношению к стартовому сайту. На расстоянии 70 - 80 п.о. в направлении 5'-конца от начала транскрипции часто расположена другая консервативная последовательность из 9 п.о. - СААТ-бокс, контролирующей начальное связывание РНК-полимеразы. Мутации в ТАТА- или в СААТ-боксах могут существенно влиять на скорость синтеза РНК. В 5'-фланкирующей области гена на расстоянии до тысячи пар оснований от начала его кодирующей части располагаются другие регуляторные последовательности, так называемые инхансеры (усилители), способные резко увеличивать продукцию гена за счет увеличения скорости транскрипции. Эти контролирующие элементы могут работать независимо от их ориентации по отношению к сайту инициации. Для некоторых генов найдены участки ДНК, подавляющие транскрипцию, а также так называемые аттенюаторы (ослабители) - последовательности, лежащие между сайтом инициации транскрипции и собственно геном. Они могут блокировать продвижение РНК-полимеразы. Благодаря такому сложному механизму контроля, достигается очень тонкая и эффективная регуляция экспрессии генов практически на всех этапах транскрипции, трансляции и образования функционально зрелого белка. Эти механизмы более детально рассмотрены в других разделах.

Раздел 2.5 Изменчивость генома, полиморфные сайты рестрикции, ПДРФ-анализ.

Кодирующие и регуляторные области структурных генов наиболее консервативны в процессе эволюции, так как мутации в них подвержены давлению жесткого естественного отбора. Действительно, небольшие изменения в этих последовательностях, даже замена одного основания, делеция или инсерция нескольких нуклеотидов, могут привести к прекращению синтеза белка или к потере его функции, что, как правило, драматическим образом сказывается на жизнеспособности особей, несущих подобные мутации. Однако, около 90% генома человека состоит из некодирующих последовательностей, подобных сателлитным ДНК, умеренным повторам, интронам и спейсерным промежуткам между генами. Эти участки значительно более изменчивы и содержат множество, так называемых нейтральных мутаций или полиморфизмов, не имеющих фенотипического выражения и не оказывающих заметного влияния на жизнеспособность или репродуктивные свойства особей и, таким образом, не подверженных прямому давлению естественного отбора. Полиморфные локусы являются удобными генетическими маркерами. На основе анализа родословных можно проследить их наследование в ряду поколений, проанализировать сцепление друг с другом, с известными генами и с анонимными последовательностями ДНК, то есть использовать в качестве обычных менделевских признаков в классическом генетическом анализе. Информативность полиморфных локусов определяется уровнем их генетической изменчивости в различных популяциях.

Экспериментально легко выявляются два варианта геномного полиморфизма. Количественные изменения в области локализации мини- и микросателлитных последовательностей ДНК и качественные замены отдельных нуклеотидов, приводящие к появлению полиморфных сайтов рестрикции. В первом случае изменчивость по числу повторенных "коровых" единиц создает серию аллелей, характер и частота которых уникальны для каждого вариабильного локуса. Полиморфизм в сайтах рестрикции связан с присутствием точковых нейтральных мутаций, локализованных, как правило, в уникальных последовательностях некодирующих участков ДНК. Подобные мутации в силу вырожденности генетического кода (см.Глава 1) могут возникать и в кодирующих последовательностях генов. Спонтанные мутации, возникающие в сайтах узнавания для определенных рестриктаз, делают их резистентными к действию этих ферментов. Аналогичным образом, при таких заменах могут создаваться новые сайты рестрикции. Показано, что полиморфные локусы встречаются во всех хромосомах с частотой приблизительно один полиморфный сайт на 300-500 п.о. Этот тип изменчивости ДНК был выявлен и использован для молекулярной маркировки специфических участков генома исторически раньше по сравнению с вариабильными сателлитными повторами (Botstein et al., 1980).

Мутационная изменчивость в сайтах рестрикции может быть легко обнаружена по изменению длины рестрикционных фрагментов ДНК, гибридизующихся со специфическими ДНК-зондами. Анализ полиморфизма длины рестрикционных фрагментов, так называемый ПДФ-анализ (Restriction Fragment Length Polymorphism -RFLP analysis), включает следующие этапы: выделение геномной ДНК, ее рестрикцию специфической эндонуклеазой, электрофоретическое разделение образующихся фрагментов ДНК и идентификацию фрагментов ДНК, содержащих полиморфный сайт рестрикции, путем блот-гибридации по Саузерну (см.Главу 1). При отсутствии рестрикции в полиморфном сайте на электрофореграммах или радиоавтографах (в зависимости от типа мечения ДНК-зонда) будет выявляться один крупный фрагмент, соответствующий по длине последовательности ДНК между двумя

соседними константными сайтами рестрикции для той же эндонуклеазы. При наличии рестрикции в полиморфном локусе на электрофореграмме будет присутствовать меньший по размерам фрагмент, равный расстоянию между полиморфным сайтом рестрикции и одним из ближайших константных сайтов рестрикции. С каким именно из двух фрагментов, прилегающих к полиморфному локусу, будет происходить гибридизация зависит от локализации используемого для анализа ДНК-зонда. В частном случае возможна гибридизация одновременно с двумя соседними рестриционными фрагментами, если выбранный ДНК-зонд комплементарен последовательности, содержащей полиморфный сайт рестрикции. Однако, такие зонды очень редко используются на практике, так как длина рестриционных фрагментов обычно в десятки раз больше длины ДНК-зондов и далеко не всегда удастся выделить и проклонировать фрагмент ДНК, содержащий полиморфный сайт рестрикции. Поэтому в дальнейшем для простоты изложения мы будем рассматривать только более общую ситуацию и считать, что при отсутствии рестрикции в полиморфном сайте ДНК-зонд гибридизуется с одним длинным фрагментом, а при наличии рестрикции гибридизующийся фрагмент имеет меньшую длину. Таким образом, при анализе ДНК особей, в обеих хромосомах которых присутствует сайт рестрикции в полиморфной области, на электрофореграмме будет выявлен только один бэнд в нижней области геля, соответствующий более короткому фрагменту ДНК. У особей, гомозиготных по мутации, изменяющей полиморфный сайт рестрикции, будет наблюдаться один бэнд в верхней части геля, соответствующий фрагменту большей длины, тогда как у гетерозигот проявятся оба эти бэнда (Рис. 2.3а).

ПДРФ-анализ может быть значительно упрощен в том случае, если возможна специфическая амплификация участка ДНК, содержащего полиморфный сайт рестрикции. Тестирование состояния этого локуса возможно путем проведения ПЦР и рестрикции амплифицированного фрагмента. При отсутствии сайта узнавания в исследуемой области ДНК размеры амплифицированного фрагмента не изменятся после его обработки соответствующей эндонуклеазой, тогда как при полном соответствии полиморфной области сайту рестрикции образуются два фрагмента меньшей длины (Рис.2.3б). У гетерозигот будут присутствовать 3 фрагмента, один из которых по длине будет соответствовать размеру амплификата до рестрикции, плюс 2 маленьких фрагмента с той же суммарной длиной. Таким образом, как и в случае использования для анализа блот гибридизации по Саузерну, трем возможным вариантам генотипа будут соответствовать три различных варианта электрофореграмм.

Необходимо подчеркнуть, что первичная идентификация полиморфных сайтов рестрикции, сцепленных с определенными генами, возможна только при наличии соответствующих ДНК-зондов. Дальнейшая тактика заключается в поиске рестриционной нуклеазы, выявляющей полиморфизм. С этой целью используют широкий набор эндонуклеаз для рестрикции геномной ДНК, выделенной из группы неродственных индивидуумов, представляющих собой репрезентативную выборку популяции. Затем проводят ПДРФ-анализ отдельно для каждой из рестриктаз с набором имеющихся ДНК-зондов. После обнаружения полиморфизма в одной из популяций аналогичные исследования проводят в других популяциях.

Следует учитывать, что полиморфные сайты рестрикции всегда представляют собой двухаллельную систему, поэтому даже при самой высокой вариабельности такого сайта, число гетерозиготных по данному локусу особей в популяции не будет превышать 50%. Между тем, только при наличии полиморфного

сайта в гетерозиготном состоянии можно отличить мутантный аллель от нормального, то есть осуществить его молекулярную маркировку. Следовательно, информационная емкость такого полиморфизма относительно невелика (см. Главы V и VII).

Раздел 2.6 Вариабильные микро- и минисателлитные ДНК.

Гипервариабильные сателлитные повторы являются гораздо более информативными маркерами по сравнению с полиморфными сайтами рестрикции, так как представляют собой мультиаллельные системы с уровнем гетерозиготности, достигающим 70 - 90%. Кроме того, оказалось, что количество высокоизменчивых микро минисателлитных последовательностей в геноме человека, по-видимому, превышает несколько десятков тысяч, они достаточно плотно и равномерно расположены в каждой из хромосом. Так более 90% из 5000 идентифицированных (С-А)-повторов являются полиморфными, причем в большинстве из них уровень гетерозиготности значительно превышает 50% (Weissenbach et al., 1992). Среди гипервариабильных мини-сателлитных последовательностей различают варьирующие по числу тандемные повторы - VNTR, три-, тетра- и пентануклеотидные повторы, а также некоторые другие классы повторов. Общее число высокополиморфных минисателлитных последовательностей в геноме превышает 1500 (Armour et al., 1990; Charlesworth et al., 1994). VNTR (variable number tandem repeats - варьирующие по числу тандемные повторы) - характеризуются наличием 10 - 15 - ти нуклеотидных "коровых" последовательностей, сходных с контролирующими элементами рекомбинации *E.coli* (Jeffreys et al., 1985). С помощью ДНК-зондов, сконструированных на основе тандемно повторяющихся "коровых" последовательностей, можно анализировать индивидуальную изменчивость в единичных или множественных высокополиморфных локусах, несущих однотипную коровую последовательность. Примером может служить VNTR, обнаруженная в интроне миоглобинового гена, включающая четыре тандемных повтора из 33 п.о., фланкированных прямыми повторами из 9 п.о. Полученные из этого района ДНК-зонды с успехом используются для идентификации личности методом ДНК-фингерпринта, так как вероятность совпадения аллелей у двух неродственных индивидуумов по всем гипервариабильным локусам, гибридизующихся с этим ДНК-зондом, значительно меньше 10^{-7} (Jeffreys et al., 1991). Высокополиморфные VNTR найдены также в гене инсулина, в области глобинового псевдогена и в X-хромосоме, содержащей последовательности, гомологичные ДНК вируса гепатита В (Nakamura et al., 1985). В последнее время многочисленные VNTR- последовательности обнаружены в хромосомах 1, 2q, 14, 16, 17, 19 (O'Brien, 1992). Поиск новых гипервариабильных локусов основан на систематическом скрининге клонированных последовательностей ДНК с помощью искусственно синтезированных "коровых" зондов. Особенно удобны для такого скрининга геномные библиотеки, сконструированные на основе предварительно фракционированных по величине MboI или Sau3A1 фрагментов ДНК, так как большие фрагменты обогащены длинными и более вариабильными минисателлитными последовательностями (Armour, et al., 1990). Хорошие результаты были получены также при скринировании космидных библиотек с помощью G-богатых синтетических ДНК-зондов и использование хемолюминесцентного метода в сочетании с ПЦР (Decorte, Cassiman, 1990). В некоторых случаях с этой целью применяют также последовательности ДНК, выделенные из фага M13, имеющего в своем составе сходные структуры (Armour et al., 1990).

Одним из вариантов минисателлитных последовательностей являются тримерные и тетрамерные короткие тандемные повторы -STR (short tandem repeats), которые стабильно наследуются и также обладают высоким уровнем полиморфизма по числу коровых единиц (Edwards et al., 1991). В X хромосоме такие повторы обнаружены через каждые 300 - 500 кб, тогда как в других частях генома они встречаются на расстоянии 10 кб друг от друга. По некоторым оценкам их частота может достигать 1/20 кб (Charlesworth et al., 1994). Возможно, истинное число три- и тетрамерных повторов в геноме человека еще больше, так как они обнаружены во многих генах. Подобно другим повторам STR обычно встречаются в некодирующих частях генов. В последнее время, однако, обнаружена группа структурных генов, несущих тримерные повторы в регуляторных и даже в транслируемых частях генов. Изменения числа этих внутригенных повторов в сторону их увеличения может приводить к нарушению функции этих генов вплоть до полного блока экспрессии и быть причиной ряда тяжелых наследственных заболеваний - болезни экспансии (см. Главу X). В качестве зондов для выявления коротких тандемных повторов используют синтетические олигонуклеотиды, построенные из простых повторов трех или четырех нуклеотидов.

В последнее время для обнаружения различных классов микросателлитных последовательностей и STR-повторов применяют эффективные методы, основанные на использовании ПЦР (Edwards et al., 1991). Для этих локусов характерно большое число аллелей, которые различаются по числу коровых единиц. Продукты амплификации этих сайтов могут отличаться друг от друга числом ди- три- или тетра-нуклеотидных повторов. Для удобства идентификации различных аллелей полимеразную цепную реакцию проводят в присутствии меченых нуклеотидов, а для электрофоретического разделения продуктов амплификации используют специальные секвенирующие гели. Многие STR-локусы настолько полиморфны, что нашли применение в судебной медицине для идентификации личности. Для повышения достоверности результатов с этой целью используют мультиплексные варианты ПЦР, позволяющие устанавливать генотип индивидуума одновременно по нескольким STR-сайтам (Weber, May, 1994; Асеев и др., 1995). Возможность точного генотипирования личности с использованием полимеразной цепной реакции, то есть на минимальном количестве ДНК, является важным методическим преимуществом STR перед ранее рассмотренными VNTR, для анализа которых чаще используют блот-гибридизацию по Саузерну.

Высокая частота коротких тандемных повторов, уникальность комбинаций числа тетра-, три- и димеров в разных сайтах в сочетании с их большой вариабильностью и сравнительной легкостью идентификации аллелей позволяет широко использовать эти повторы для генетического и физического картирования в качестве наиболее удобных индексных маркеров геномных ДНК-последовательностей (см. Главу III). Такие маркерные сайты получили название STS (sequence tagged sites). В настоящее время в геноме человека уже идентифицировано около 10 000 STS, подавляющее большинство которых представляет собой тандемные повторы 2 - 4 нуклеотидов. Благодаря выраженной индивидуальной специфичности и достаточно стабильному менделевскому типу наследования STS-сайты нашли широкое применение и в молекулярной диагностике генных болезней, прежде всего в качестве молекулярных маркеров для идентификации мутантных хромосом в семьях высокого риска (см. Главу VII). Наличие большого числа гипервариабильных микро- и минисателлитных последовательностей ДНК является характерной особен-

ностью генома человека. Аналогичные последовательности, обнаруженные в геноме приматов, значительно более однородны, что доказывает возможность существенного увеличения вариабельности этих участков ДНК за сравнительно короткий эволюционный промежуток (Юров, 1988; Gray et al., 1991).

Сведения о мутабельности высокополиморфных последовательностей в геноме человека весьма противоречивы. Показано, однако, что в наиболее вариабельных минисателлитных локусах частота мутаций может достигать 5% на гамету (Jeffreys et al., 1988). Предполагается, что одной из главных функций гипервариабельных микро- и минисателлитных последовательностей ДНК может быть контроль гомологичной рекомбинации в мейозе. На культурах клеток показано стимулирующее влияние минисателлитных последовательностей ДНК на гомологичную рекомбинацию. Так, инсерция синтезированной последовательности, составленной на основе гипервариабельных минисателлитов в геномную ДНК приводит к более, чем 10-кратному увеличению числа реципрокных обменов, причем степень этого влияния обратно пропорциональна расстоянию между STR и сайтом рекомбинации (Wahls et al., 1990). Вместе с тем, многие авторы обращают внимание на достаточно высокую стабильность минисателлитных аллелей, что позволяет их широко использовать как для генетического маркирования, так и для популяционных исследований и идентификации личности методом ДНК-фингерпринта (Decorte, Cassiman 1993; Edwards et al., 1991; Иванов, 1989).

Для многих мутаций, локализованных в некодирующих частях генома, характерны высокие уровни популяционного полиморфизма. Необходимо, однако, подчеркнуть, что эта изменчивость не затрагивает общей структуры генома, определяющей различия между видами. Более того, сопоставление первичных нуклеотидных последовательностей сравнительно протяженных секвенированных участков ДНК (области Т-рецепторных генов длиной около 100 кб) обнаружило сохранение высокой степени гомологии не только в кодирующих, но и, что особенно удивительно, в некодирующих частях этих последовательностей. Если учесть, что эволюционно человек и мышь разделены почти 80 миллионами лет эволюции, эти данные рассматриваются как свидетельство функциональной значимости некодирующих частей этих генов. По-видимому, далеко не всякие мутации в некодирующих районах ДНК являются нейтральными и в определенных случаях они могут отрицательно влиять на жизнеспособность. К сожалению, в настоящее время ничего или почти ничего неизвестно о функциях некодирующих ДНК-последовательностей. Высказывалось даже предположение, что их единственной функцией является репликация. Отсюда возникло представление об "эгоистической" или "паразитической" ДНК. Конечно, полностью исключить наличие подобных паразитических последовательностей ДНК в любом геноме нельзя. Тем ни менее, представляется маловероятным, что значительная часть генома человека, также как и других видов, относится к эгоистической ДНК. По-видимому, наши знания о роли некодирующей или, как еще говорят, "избыточной" ДНК все еще явно недостаточны. Стабильность структурной организации генома в пределах вида свидетельствует скорее о важной эволюционной роли некодирующих ДНК-последовательностей и об их участии в процессах онтогенеза. Можно предполагать, что ответ на этот интригующий вопрос в какой-то мере будет получен при расшифровке и сравнении полной первичной нуклеотидной последовательности геномов у животных разных видов и, прежде всего, у человека и мыши, где прогресс в секвенировании геномной ДНК особенно

значителен (см.Главу III). Уместно заметить, что проведенный недавно компьютерный анализ генома человека позволяет предполагать наличие в его некодирующей части особого, пока еще непонятого генетического кода, смысл и значение которого остаются загадочными (?).

Раздел 2.7 Мобильность генома, облигатные и факультативные элементы генома.

До сих пор мы рассматривали основные структурные элементы генома человека, положение которых в соответствии с представлениями классической генетики достаточно постоянно. Начиная с 50-х годов стали накапливаться данные о существовании большого числа мобильных генетических элементов, присутствие которых в геноме не является обязательным, а их топография и количество может варьировать в различных клетках, тканях и у разных индивидуумов (McClintock, 1984; Berg, Howe, 1989). У прокариот такие элементы получили название транспозонов. Их структура и функции достаточно хорошо изучены. Отличительной особенностью мобильных элементов является способность существовать как в интегрированном с хромосомой виде, так и в виде отдельных макромолекул - эписом, плазмид, вирусных частиц. Почти 50 различных семейств мобильных элементов описано у дрозофилы. Вместе эти последовательности составляют около 12% гаплоидного набора (Golubovsky, 1995). В геноме млекопитающих содержится до 50 000 диспергированных копий ретропозона LINE размером около 6500 пар оснований. Семейство Alu-повторов, содержащее от 300 до 500 тысяч копий, также относится к числу мобильных элементов генома (Charlesworth et al., 1994). Явление лизогении, то есть присутствие вирусных последовательностей в составе ДНК человека и наличие фрагментов генов человека в вирусных геномах, служит одним из примеров мобильности ДНК и возможности "горизонтальной" передачи наследственно закрепленных признаков между видами. Мобильные ДНК, как правило, относятся к факультативным элементам. Как уже отмечалось, не существует четких границ между облигатными и факультативными элементами генома, так как возможен взаимный переход от одного состояния к другому. Структурные локусы или сегменты хромосом могут трансформироваться в факультативные элементы за счет амплификации, интеграции в мобильные элементы или путем образования цитоплазматических ретротранскриптов. Обратный переход от факультативных элементов к облигатным осуществляется посредством инсерций, транспозон-индуцированных перестроек и обратной транскрипции.

Факультативные элементы существуют в геноме как популяции информативных макромолекул. Изменения, возникающие в них под воздействием внешних факторов, носят совершенно иной характер по сравнению с классическими мутациями в структурных локусах. Для описания изменений в факультативных элементах предложен термин "вариации" (Голубовский, 1985). Этот термин впервые использован Жакобом и Воллманом для описания поведения эписом (Jacob, Wollman, 1961). Вариации могут приводить к изменениям на генотипическом уровне, то есть к мутациям, вследствие простого перемещения факультативных элементов или сдвига в соотношении между факультативными и облигатными элементами. В этих случаях мутации встречаются одновременно у многих индивидуумов. Подобные изменения упорядочены, могут происходить сразу во многих локусах и отличаются высокой сайт-специфичностью. Локализация структурных перестроек, возникающих в результате вариаций, предопределе-

на первоначальной топографией факультативных элементов на хромосомах. И наконец, сами вариации могут быть индуцированы обычными "не-мутагенными" факторами, такими как температура или межлинейные кроссы (Golubovsky, 1995). Факультативные элементы могут рассматриваться как оперативная память генома, так как во многих случаях спонтанное возникновение мутаций в облигатных элементах опосредовано их активацией. Считается, в частности, что инсерционный мутагенез является причиной спонтанного возникновения 70% видимых мутаций в природных популяциях дрозофилы. Однако, у человека пока зарегистрированы лишь единичные случаи возникновения мутаций вследствие перемещения мобильных элементов генома (Vidaud et al., 1993).

Раздел 2.8 Изохоры, метилирование, гиперчувствительные сайты.

Перечисленные выше компоненты генома не случайным образом связаны с последовательностями нуклеотидов. И в этом смысле можно говорить о существовании в геноме человека структур более высокого иерархического порядка. Примером служат изохоры - длинные, в среднем, свыше 300 кб сегменты ДНК, гомогенные по композиции оснований или по GC-уровням. 62% генома состоит из GC-бедных изохор и в них локализовано около 34% генов, 31% генома представлен GC-богатыми изохорами, содержащими 38% генов, и в 3% изохор, обогащенных GC-последовательностями (так называемых H3 изохор), находится 28% генов (Mouchiroud et al., 1991; Saccone et al., 1993). Таким образом, существуют относительно небольшие участки ДНК, в которых плотность генов в 10-20 раз выше, чем в остальных последовательностях.

Другой общей чертой генома человека является то, что *in vivo* значительная доля цитозиновых остатков в молекуле ДНК метилирована, то-есть находится в форме 5-метилдезоксцитидина. Экспериментальное изучение характера метилирования основано на сопоставлении рестрикционных фрагментов, образующихся после обработки ДНК эндонуклеазами, для которых сайты узнавания одинаковы и содержат в своем составе цитозин, но действуют эти ферменты по-разному, в зависимости от того, находится ли это основание в метилированном состоянии или нет. В частности, рестриктазы - Msp1 и Hpa11, узнают последовательность CCGG, но в отличие от Msp1, Hpa11 не расщепляет ДНК в тех сайтах, где внутренний CpG динуклеотид метилирован. Некоторые сегменты генома, особенно это относится к повторяющимся последовательностям, полностью метилированы в местах 5'-CCGG-3' и частично метилированы в 5'-GCGC-3' - сайтах рестрикции для Hpa1. В других сегментах наблюдается характерный рисунок частичного метилирования в 5'-CCGG-3' последовательностях (Behn-Krappa et al., 1991). Различные индивидуумы, независимо от их этнического происхождения, практически не различаются по характеру метилирования ДНК в одних и тех же типах тканей, тогда как в процессе онтогенетической дифференцировки происходят значительные изменения рисунков метилирования. В перевиваемых культурах клеток опухолевого происхождения число метилированных сайтов резко уменьшено.

Высказано предположение о наличии прямой связи между метилированием ДНК и состоянием генетической активности в клетках. Существует класс белков, которые специфическим образом связываются с метилированными участками ДНК, делая их недоступными для действия ряда ферментов, в том числе, воз-

можно, и для полимераз. Получено много прямых экспериментальных доказательств роли метилирования ДНК в инактивации эукариотических промоторов, а, значит, и в регуляции активности генов. Напротив, гипометилирование промоторной области генов, в особенности CpG островков, как правило, свидетельствует о функциональной активности генов. Показано, что необычные структуры в молекуле ДНК, также как экзогенная ДНК, инкорпорированная в процессе генетической трансформации, нередко подвергаются метилированию. Известно, что метилирование играет важную роль в инактивации X хромосомы у самок, в регуляции экспрессии генов в процессе развития, а также непосредственно вовлечено в феномен хромосомного (геномного) импринтинга, связанного с различиями пенетрантности некоторых аллелей в зависимости от их происхождения, то есть прохождения через материнский или отцовский гаметогенез (Баранов, 1991).

В GC-богатых изохорах локализовано большое количество CpG островков - последовательностей от 500 до 2000 п.о., характеризующихся очень высоким содержанием гуанина и цитозина ($G+C > 60\%$), представленных в виде кластеров неметилированных CpG дуплетов и, так называемых, G/C боксов - локусов, родственных сайту узнавания для одного из транскрипционных факторов Sp1 - (G)4C(G)4C (Lindsay, Bird, 1987; Bird, 1986; Aissani, Bernardi, 1991). CpG острова содержат много сайтов узнавания для чувствительной к метилированию эндонуклеазы HpaII, а также сайты для редкощепящих рестриктаз, узнающих неметилированные CpG дуплеты. В частности, более 80% Nor1-сайтов связано с CpG-богатыми островками. Как правило, CpG островки локализованы в 5'-фланкирующих последовательностях, 5'-экзонах и 5'-интронах всех изученных хаузкипинг-генов и 40% тканеспецифических генов. CpG островки являются характерной особенностью транскрибируемых участков генома. Их идентификация в клонированных последовательностях геномных библиотек существенно облегчает поиск конкретных структурных генов (см. раздел 2.4). Наибольшая плотность CpG островков наблюдается в теломерных участках хромосом 1, 9, 15, 16, 17, 19, 20, 22 (Antonarakis, 1994). Точные молекулярные методы регистрации CpG островков показали, что их число в геноме человека приближается к 45000 (Antequera, Bird, 1993).

Можно также отметить существование в геноме человека сайтов, гиперчувствительных к действию ДНК-азы 1 и структурно отличающихся от основной массы хроматина. Присутствие таких сайтов показано для многих генов млекопитающих и, по-видимому, это необходимое, но не достаточное условие их экспрессии. Локализация гиперчувствительных сайтов может меняться в процессе развития и под действием гормонов. В некоторых случаях эти участки маркируют положение транскрипционных регуляторных элементов генома, действующих как в положительном, так и в отрицательном направлениях. В других случаях это области функционально активных генов, находящихся в деспирализованном состоянии и имеющих одностороннюю структуру. Именно такие односторонние участки ДНК особенно высокочувствительны к ДНК-азе 1. На этом их свойстве основан метод нуклеотид-трансляции *in situ*, позволяющий непосредственно на хромосомных препаратах визуализировать функционально активные районы хромосом. С этой целью хромосомные препараты обрабатывают ДНК-азой 1, после чего непосредственно на них с помощью ДНК-полимеразы проводят синтез ДНК в присутствии меченых нуклеотидов. При этом метка включается преимущественно только в те участки хромосом, где находятся функционально ак-

тивные гены (Verma, Babu, 1989).

ГЛАВА VIII.

БИОЛОГИЧЕСКИЕ МОДЕЛИ НАСЛЕДСТВЕННЫХ БОЛЕЗНЕЙ ЧЕЛОВЕКА.

Раздел 8.1. Генетические линии животных.

Большая роль в исследовании проблем генетики человека и медицинской генетики принадлежит мутантным генетическим линиям животных и, в особенности, генетическим линиям мышей (Конюхов, 1969, 1980; Корочкин, 1978). Высокий процент сходства по нуклеотидными последовательностям между кодирующими, регуляторными и даже интронными областями гомологичных генов млекопитающих и человека, а также наличие большого числа консервативных групп сцепления с идентичным расположением генов наряду с возможностями использования очень мощных экспериментальных подходов для идентификации и клонирования генов линейных животных позволяют проводить параллельные исследования, значительно ускоряющие эффективность поиска и молекулярного анализа индивидуальных генов человека.

Для многих моногенных заболеваний человека животные, несущие мутации в гомологичных генах, являются лучшими, а зачастую и единственными моделями для исследования молекулярных основ патогенеза и отработки оптимальных схем лечения, в том числе и с применением методов генной терапии (см.Главу IX). Поиск таких биологических моделей, прежде всего, ведется, среди уже существующих генетических линий животных с установленным типом наследования определенных аномальных признаков. Наиболее трудным при этом является доказательство идентичности мутантных генов и, соответственно первичных биохимических дефектов, у человека и у линейных животных.

В различных питомниках мира, в том числе и в России, созданы и поддерживаются коллекции, насчитывающие от десятков до несколько сотен генетических линий различных экспериментальных животных - мышей, крыс, кроликов, собак и др. (Конюхов, 1969; 1980; Staat, 1969; Hogan et al, 1989; Бландова и др., 1990). Среди них генетические линии мышей наиболее многочисленны в первую очередь из-за высокой плодовитости, удобства содержания, относительной легкости экспериментального манипулирования и целого ряда других причин. Некоторые из этих линий представляют собой случайные находки, другие, а их большинство, получены в результате действия различных мутагенных факторов. Так, значительное число биологических моделей было получено путем биохимической селекции потомства мышей самцов, обработанных сильными мутагенами - этилнитрозмочевинной, триэтиленмеламином или облученных Рентгеном. Так были смоделированы на мышах альфа-талассемия, полицитемия, почечный ацидоз (Erickson, 1988). Однако, такой способ получения животных-моделей, хотя и более эффективен, чем поиск спонтанно мутировавших особей, также основан на чистой случайности и не позволяет направленно менять структуру нужного гена.

Процесс создания подобных генетических линий обычно включает отбор особей с фенотипическими отклонениями; анализ наследования этих фенотипических признаков; длительное близ-

кородственное разведение отселектированных особей. При моногенном наследовании такие линии могут либо целиком состоять из мутантных гомозигот, либо поддерживаться через гетерозиготных особей в случае сниженной жизнеспособности и нарушения плодовитости у гомозигот.

На первом этапе поиска адекватной модели какого-либо моногенного наследственного заболевания руководствуются сходством клинических проявлений течения болезни и фенотипом мутантных животных. Однако, одного этого сходства недостаточно (Конюхов, 1969). Необходимо доказать гомологичность генотипической природы наблюдаемых нарушений, то есть доказать, что у человека и у животных (мышей) фенотипические изменения обусловлены мутациями в гомологичных генах. Огромная мировая генетическая коллекция мышей насчитывает несколько сотен линий, в каждой из которых различные дефекты наследуются по моногенному типу. Спонтанные биологические модели наследственных болезней известны и достаточно полно изучены для многих других экспериментальных и домашних животных. Представляется удивительным, что, несмотря на большое сходство геномов млекопитающих и наличие близких по первичной структуре и тождественных по функциям структурных генов, для значительной части наследственных болезней человека генетические аналоги среди животных до сих пор не найдены (Конюхов, 1969).

Это ограничение в настоящее время может быть преодолено путем целенаправленного конструирования генетических модельных линий животных. Экспериментальные основы такого подхода уже хорошо разработаны (Erickson, 1988; Аллен и др., 1990; Melton, 1993; Stewart et al., 1994). Для этого используют технику культивирования и трансфекции эмбриональных стволовых клеток (см. ниже), отбор *in vitro* клонов с нужными генетическими изменениями и пересадку их в зародыши или в соматические ткани животных. Для анализа экспрессии мутантных генов *in vivo* и оценки их биологического действия особенно удобными оказались трансгенные животные.

Раздел 8.2. Трансгенные животные.

Трансгенных животных получают в результате искусственного введения - трансгеноза, чужеродного генетического материала, представляющего из себя фрагмент гена или иную последовательности ДНК, в оплодотворенную яйцеклетку или в ранние зародыши млекопитающих. Подобные модели являются идеальными экспериментальными системами для исследования молекулярно-генетических основ онтогенеза, для изучения функции чужеродного гена, оценки его биологического действия на организм, а также для производства различных манипуляций со специфическими клеточными клонами *in vivo*. Разработано несколько способов получения трансгенных животных. Исторически более ранним и широко применяемым до настоящего времени является микроинъекция чужеродной ДНК в пронуклеус - ядро оплодотворенной яйцеклетки. Существуют детальные описания этого метода (Аллен и др., 1990; Hogan et al, 1989). Суть метода состоит в том, что под контролем микроскопа при помощи микроманипулятора в мужской пронуклеус оплодотворенной яйцеклетки тонкой иглой (до 1 микрона) вводят около 2 пиколитров раствора ДНК. Чужеродная ДНК, вначале свободно лежащая в нуклеоплазме, в течение нескольких последующих делений дробления случайным образом интегрирует в один из сайтов какой-либо хромосомы, то есть встраивается в ДНК-реципиента. При этом, как показали эксперименты с меченой ДНК, в различ-

ных бластомерах одного и того же дробящегося зародыша интеграция может происходить в разные хромосомные сайты и число интегрированных копий ДНК в каждом из этих сайтов может значительно варьировать. Тем не менее, поскольку сам эмбрион развивается, по-сути, из одного бластомера, во всех клетках такой особи после рождения чужеродная ДНК обычно находится только в одном каком-нибудь хромосомном сайте, хотя у разных особей она интегрируется по-разному и в разные сайты. После введения чужеродной ДНК в пронуклеус яйцеклетку трансплантируют самке-реципиенту. Доля трансгенных животных в потомстве таких самок варьирует от 10% до 30%. Это означает, что подобный механический вариант трансфекции чужеродных генов на ранней стадии эмбриогенеза является чрезвычайно эффективным. Идентификацию трансгенных животных производят путем анализа геномной ДНК на наличие экзогенных последовательностей, используя при этом методы блот-гибридизации или ПЦР. Экспрессию введенного гена анализируют путем идентификации специфических мРНК и/или соответствующих белковых продуктов в различных тканях трансгенного животного.

Другой, более прогрессивный способ получения трансгенных животных основан на том, что трансфекции подвергается не зигота, а тотипотентные эмбриональные стволовые клетки (см. ниже), которые затем трансплантируют в полость бластоцисты (Gardner, 1978). Этот метод и его решающие преимущества в плане генетического моделирования подробно рассмотрены в разделе 8.4.

Как правило, инъекцированная ДНК при встраивании в хромосому образует блок из множества tandemно расположенных копий, при этом число единиц повтора в блоке у разных особей может варьировать от единицы до нескольких сотен. После интеграции введенной ДНК в хромосому различные генетические конструкции устойчивы и стабильно передаются потомству в соответствии с законами Менделя. Встраивание введенной ДНК в функционально значимые области генома может приводить к их дестабилизации и сопровождаться появлением мутаций, спектр которых очень разнообразен. Таким образом, животные, полученные при введении одного и того же гена, будут различаться как по сайтам интеграции, так и по количеству копий встроеной чужеродной ДНК, а в некоторых случаях, по уровню мутабельности и по типам индуцированных мутаций. Таким образом, каждое трансгенное животное в этом смысле уникально.

Трансгенные животные являются чрезвычайно удобным объектом для анализа роли отдельных элементов гена в регуляции его работы. Так, сопоставление характера экспрессии введенного гена у животных, различающихся по длине фланкирующих последовательностей инъекцированной ДНК, дает возможность обнаружить элементы гена, контролирующие его работу в разных типах тканей. Для облегчения анализа регуляторных последовательностей гена часто вводят генетические конструкции, сочетающие эти элементы с геном-репортером, экспрессия которого выражается в появлении известной и легко определяемой ферментативной активности. Использование для трансгеноза рекомбинантных молекул ДНК, представляющих собой различные комбинации регуляторных элементов и кодирующих последовательностей, ведет к более глубокому пониманию молекулярных механизмов активации генов в разных типах тканей.

Как уже указывалось, случайный характер интеграции чужеродной ДНК нередко индуцирует мутации и нарушает экспрессию нормальных генов реципиента. В ряде случаев наблюдаемые отклонения в развитии оказываются аналогичными или сходными с

уже известными наследственными нарушениями у человека и подобные животные также могут использоваться в качестве генетических моделей заболеваний. Этот подход был применен для получения моделей таких заболеваний, в патогенезе которых решающую роль играет эффект дозы генов. В частности, путем трансфекции зиготы мышей генами бета-глобина, коллагена, ренина, антигенов гистосовместимости удалось получить биологические модели таких заболеваний, как бета-талассемия, несовершенный остеогенез, гипертония и диабет, соответственно (Erickson, 1988). Во всех перечисленных случаях введение дополнительной дозы экспрессирующего гена приводило к нарушению баланса белковых генопродуктов в клетках и, как следствие этого, было причиной патологических процессов.

Раздел 8.3. Экспериментальное моделирование.

Другой вариант биологического моделирования основан на получении животных с определенными очень специфическими, но ненаследственными изменениями. Эти животные также могут быть использованы для анализа молекулярных основ патогенеза и разработки методов адекватного лечения. Рассмотрим несколько примеров подобного экспериментального моделирования.

Описанная технология трансгеноза (введение генов в пронуклеус) может быть использована, в частности, для направленного получения животных с избирательными дефектами (уродствами) тех или иных тканей и органов. Метод заключается в возможности селективной элиминации тех специфических типов клеток, которые отсутствуют или дефектны у больных с моделируемым типом заболевания. Такие животные могут быть получены при инъекции в зародыш рекомбинантной ДНК, содержащей какой-либо цитотоксический ген, например, ген дифтерийного токсина, находящийся под контролем работающих в определенных типах клеток регуляторных элементов ДНК. При активации этих контролируемых элементов на определенной стадии развития экспрессия токсического гена приводит к избирательной гибели всей специфической популяции клеток, то есть такая система действует как очень точный скальпель.

Дальнейшая модификация метода заключается в использовании для трансгеноза условно летального гена, каким является, например, ген тимидинкиназы вируса Герпеса. Клетки, экспрессирующие этот ген, функционируют совершенно нормально. Однако, на любой стадии онтогенетического развития можно вызвать их селективную гибель при введении животному ганцикловира – противогерпесного препарата. Эта система дает больше возможностей для экспериментального анализа роли специфических клонов клеток в процессе нормального развития, а также для изучения патологических процессов, связанных с гибелью этих клеток. Подобная методология используется также при разработке генотерапевтических подходов для лечения некоторых ненаследственных, в частности онкологических заболеваний (см Главу IX).

Весьма многообещающим методом моделирования представляется направленное выключение работы определенных генов путем введения в доимплантационные зародыши антисмысловых мРНК. Такой подход был применен, в частности, при попытке моделирования болезни Гоше – лизосомного заболевания, обусловленного дефицитом бета-глюкуронидазы (Bevilacqua et al., 1988). Естественно, что в этом случае выключение экспрессии гена носит транзитный характер, то есть моделью, по-сути, является само животное – реципиент антисмысловой мРНК матрицы.

Другой пример экспериментального моделирования основан

на пересадке тканей или клеток атимусным иммунодефицитным мышам nu/nu. У мышей этой линии в связи с отсутствием тимуса и выраженным врожденным иммунодефицитом не происходит отторжение трансплантированных чужеродных тканей. Более того, у таких животных может происходить дифференцировка трансплантированных подкожно эмбриональных зачатков и регенерация пересаженных кусочков тканей из различных органов других видов животных и человека. Так например, кусочки трахеи крысы с нанесенными на них клетками бронхогенного эпителия человека, имплантированные подкожно атимусным мышам, формируют структуру поверхностного эпителия, сходную с той, которая имеется в бронхах человека. Именно таким путем мыши nu/nu были активно использованы для анализа экспрессии мутантных вариантов гена муковисцидоза человека, а также для испытания эффективности коррекции этого генетического дефекта с помощью методов генотерапии. В последнем случае мутантные эпителиальные клетки пациентов с муковисцидозом вначале подвергали трансфекции ретровирусными или аденовирусными векторами, несущими, наряду с геном - репортером, полноразмерную кДНК нормального гена муковисцидоза. Относительная простота подобных моделей и возможность генетического манипулирования с клетками человека до их трансплантации атимусным мышам делают этот подход весьма привлекательным для решения многих экспериментальных вопросов. Основные недостатки таких моделей связаны с трудностями содержания и разведения атимусных мышей и их низкой жизнеспособностью. Генетические линии животных в этом отношении имеют значительные преимущества.

Раздел 8.4. Конструирование модельных генетических линий животных.

Современный уровень экспериментальной эмбриологии млекопитающих и современные достижения молекулярной генетики позволяют осуществлять направленное получение генетических моделей наследственных болезней путем введения сайт-специфических модификаций в геном млекопитающих. Такой значительный качественный прорыв в генетическом моделировании стал возможен благодаря появлению принципиально новой технологии манипулирования с ранними зародышами млекопитающих. Особенно важными в этом отношении оказались два новых методических подхода: получение зародышей-химер, состоящих из клеточных клонов разных зигот, путем введения тотипотентных клеток в полость бластоцисты (Gardner, 1978) и разработка технологии культивирования клеточных векторов, так называемых эмбриональных стволовых клеток (Evans, Kaufman, 1981). С другой стороны, появились методы сайт-специфического переноса клонированных последовательностей ДНК в геном эукариот, основанные на отборе клеточных клонов, в которых после трансфекции происходит инсерция экзогенной ДНК в гомологичном сайте геномной ДНК без какого-либо нарушения последовательности ДНК в месте встраивания.

Конструированию генетических моделей должны предшествовать идентификация и сравнительный анализ двух гетерологичных генов - гена человека, вследствие нарушения работы которого развивается моделируемое заболевание, и его гомолога у выбранного для моделирования животного. При выборе объекта моделирования, в первую очередь, руководствуются методическими возможностями экспериментального манипулирования с животными. Важное значение имеет сходство кодирующих областей гетерологичных генов по нуклеотидным последовательностям. В большинстве случаев мыши представляются наиболее удобным

объектом для моделирования. Современный алгоритм формирования генетической линии животных с мутациями в заданном гене предполагает: (1) наличие культур тотипотентных, то есть способных к неограниченному развитию и дифференцировке, эмбриональных стволовых клеток; (2) создание на базе рекомбинантных ДНК генно-инженерных конструкций для направленного переноса генов; (3) трансфекцию этих конструкций в культуры эмбриональных стволовых клеток последующим скринингом и отбором клонов со специфическими генетическими модификациями; (4) введение отобранных модифицированных клеток в зародыш на стадии бластоцисты по методу Гарднера с целью получения химерных трансгенных животных; (5) отбор химерных особей, несущих модифицированные гены в различных тканях и органах; (6) селекцию особей, гетерозиготных по данной мутации; (7) инбредное разведение и селекцию гомозигот (Рис.8.1).

Как упоминалось ранее, идеальной системой для направленного переноса мутаций в геном млекопитающих являются эмбриональные стволовые клетки – ЭСК (Evans, Kaufman, 1981; Erickson, 1988; Labosky et al., 1994). Первичные культуры этих клеток получают из клеток бластоцисты (внутренней клеточной массы) или из первичных половых клеток ранних постимплантационных зародышей. При выращивании на питательном слое из эмбриональных фибробластов ЭСК сохраняются в недифференцированном состоянии от трех месяцев до года. При этом они могут быть несколько раз заморожены и оттаяны без потери способности к дифференцировке. ЭСК, введенные в бластоцель (полость бластоцисты), сохраняют свою тотипотентность и могут участвовать в формировании, практически, всех эмбриональных зачатков и органов развивающегося зародыша. В результате образуется животное – химера, состоящее из клеточных клонов двух разных типов: клеток исходного родительского генотипа и ЭСК. Если эти клетки различаются, например, по генам окраски шерсти, животное – химера будет иметь поперечную или пятнистую окрасенность. При этом все животные, независимо образом полученные в результате введения в одинаковые по генотипу зародыши одной и той же линии клеток, будут отличаться друг от друга по характеру пятнистости, так как все химеры различны по набору клеточных клонов, развившихся и дифференцировавшихся из введенных в зародыш ЭСК. Химерные животные, у которых ЭСК дифференцировались в половые клетки и дали начало полноценным зрелым гаметам будут устойчиво передавать своим потомкам генетическую информацию, содержащуюся в ЭСК. Таких животных иногда называют зародышевыми трансмиттерами. При скрещивании их с мышами дикого типа часть потомков будет уже гетерозиготна по мутантным генам ЭСК, то есть будут нести мутацию в гаплоидном состоянии в каждом типе клеток. Это в равной степени относится и к мутациям, искусственно введенным предварительно в ЭСК. Скрещивая таких гетерозигот, можно получить животных, гомозиготных по заданной мутации. Естественно, последнее достижимо только в том случае, если мутация не окажется летальной в гомозиготном состоянии у животных этого вида.

Возможность вести селекцию нужных мутантных или трансгенных клонов ЭСК и лишь затем их использовать в качестве клеточных векторов нашло широкое применение в генетическом моделировании. Первоначально для этой цели ЭСК обрабатывали различными мутагенами (этилнитрозомочевинной) отбирали клоны клеток, несущих мутацию в нужном гене, и затем использовали их для создания инъекционных химер по Гарднеру. Таким способом на мышах была получена модель болезни Леш-Нихана – мутация гена гипоксантин-фосфорибозил-трансферазы

(Hooper et al., 1987). С разработкой технологии адресной доставки чужеродной ДНК в гены-мишени этот способ генетического моделирования стал особенно эффективным. Сайт-специфическая модификация генов ЭСК достигается за счет гомологичной рекомбинации между экзогенной и хромосомной ДНК. При трансфекции большая часть проникших в ядра молекул рекомбинантной ДНК сохраняется там в течение двух-трех дней в виде кольцевых эписом и в дальнейшем теряется либо происходит интеграция трансфецирующей плазмиды в геном клетки-хозяина путем негомологичной рекомбинации, то есть в случайные сайты хромосомной ДНК. В таких клетках экспрессия введенных генов устойчиво сохраняется. Частота интеграции экзогенной ДНК может быть повышена при использовании линейных плазмид и специальных, преимущественно, ретровирусных векторов экзогенной ДНК (см. Главу IX). Случаи стабильной интеграции экзогенной ДНК могут быть легко выявлены, если трансфецирующие плазмиды или вектора содержат селективируемый маркерный ген. Чаще всего в качестве маркера используют прокариотический ген *neo*, сообщаящий клеткам устойчивость к неомицину. Клетки, в которых произошла интеграция такой плазмиды в хромосомную ДНК, будут образовывать устойчивые клоны при выращивании на среде G418, содержащей неомицин, в то время как все другие клоны клеток будут в этих условиях деградировать.

Раздел 8.5. Методы направленного переноса генов.

Наиболее важным шагом на пути искусственного получения мутантной линии животных является отбор клонов ЭСК с сайт-специфической модификацией определенного гена. Однако, случаи инсерции экзогенной ДНК в ген-мишень очень редки, их общая частота, обычно, не превышает 10^{-6} . Предпринимаются попытки генетической модификации ЭСК с тем, чтобы повысить в них частоту гомологичной рекомбинации. Идентифицированы некоторые гены, контролируемые этот процесс у мышей и у человека. Однако, в любом случае схемы направленной модификации генов должны включать селекцию нужных клонов клеток. Впервые направленная сайт-специфическая модификация была выполнена также на гене гипоксантин-фосфорибозил-трансферазы (см. выше) и была получена еще одна генетическая линия мышей, моделирующая вызванную дефектом в *HPRT*-гене болезнь Леш-Нихана у человека (Thomas, Capresci, 1987). Успех этих исследований, в первую очередь, обусловлен существованием простых схем отбора клеток с функционирующим и нефункционирующим *HPRT*-геном на селективных средах. Важно также, что этот ген локализован в X-хромосоме и в XY-клетках он представлен одной копией. При этих условиях случаи модификации гена, вызванные инсерцией экзогенной ДНК в правильное положение, легко идентифицируются - Рис.8.1 (см. Главу X).

В настоящее время предложено несколько вариантов для направленного переноса неселективируемых генов за счет дополнительной инсерции в трансфецирующую плазмиду селективируемого маркерного гена в таком положении, при котором его экспрессия происходит преимущественно при правильном встраивании векторной последовательности в ген-мишень (рис.8.2). Так, маркерный ген *neo*, помещенный в инсертируемую область ДНК плазмиды без собственного промотора, может экспрессироваться только находясь под контролем какого-либо другого промотора хромосомной ДНК. Для этого инсерция экзогенной ДНК должна произойти в область гена-мишени без сдвига рамки считывания. При случайной интеграции экспрессии маркерного гена не будет. Таким образом, отбор неомицин-устойчивых клеток приведет к

резкому увеличению частоты клонов, в которых произошла гомологичная рекомбинация между экзогенной и геномной ДНК. На этом же принципе основано использование генетических конструкций с геном *neo*, не содержащим поли-А последовательности в 3' области. Дальнейший поиск гомологичных рекомбинантов среди G418-устойчивых клеток проводят путем блот-гибридизации, используя в качестве ДНК-зонда фрагмент векторной последовательности, расположенный вне направленно переносимого участка экзогенной ДНК.

Особенно перспективным на сегодняшний день представляется метод позитивно-негативной селекции (Melton, 1994). Метод сочетает отбор клеток, в которых произошла интеграция экзогенной ДНК, с селективной элиминацией тех из них, где встраивание произошло за счет негомологичной рекомбинации. Для этого маркерный селектируемый ген *neo* с регуляторными последовательностями инсертируют в переносимую область ДНК плазмиды, а вне этой области встраивают условно летальный вирусный ген тимидинкиназы герпеса (HSV-tk). При интеграции такого вектора в геномную ДНК путем гомологичной рекомбинации HSV-tk ген не инкорпорируется в хромосому, тогда как при негомологичной рекомбинации этот ген будет присутствовать в неомицин-устойчивых клетках. Обработка таких клеток противогерпесным агентом - ганцикловиром, будет сопровождаться гибелью всех клонов, экспрессирующих вирусную тимидинкиназу (Рис.8.3).

Отбор клеток с модифицированным геном также может производиться с помощью ПЦР. При этом не используют какие-либо маркерные гены и/или селектируемые среды. Олигопраймеры для амплификации выбирают таким образом, что один из них гомологичен соседней с сайтом интеграции последовательности модифицируемого гена, а другой соответствует участку инсертируемой экзогенной ДНК (Рис.8.4). Метод позволяет обнаруживать присутствие пяти правильно модифицированных клеток среди 50 000. После трансфекции клетки разделяются на группы, в каждой из которых проводят тестирование с помощью ПЦР. При положительном ответе группу клеток разбивают на подгруппы и процедуру повторяют до тех пор, пока не удастся изолировать нужные клоны.

Направленное выключение генов-мишеней может быть достигнуто несколькими способами. Так называемые, нулевые мутации могут быть получены путем встраивания плазмиды, содержащей, наряду с экзонными последовательностями модифицируемого гена и селектируемым маркерным геном, сильные транскрипционные и трансляционные стоп-сигналы. При этом в разрушенном за счет инсерции экзоне транскрипция прекращается, в результате чего образуется укороченный белок, незащищенный от действия клеточных протеаз.

Более совершенной является разработанная недавно техника двойной замены гена. Для этого используют ЭСК, дефицитные по ферменту HPRT - HM1 (Melton, 1994). На первом этапе ген-мишень инактивируют путем замены одного из экзонов и прилежащих последовательностей на HPRT мини-ген. При этом в нокаутирующем векторе HPRT маркер фланкируется ДНК последовательностями, гомологичными месту вставки в ДНК гена-мишени. В этот же вектор включен и ген вирусной тимидинкиназы (Рис.8.5). После трансфекции отбираются клетки позитивные по HPRT и негативные по вирусной тимидин-киназе. Именно в таких клетках с высокой степенью вероятности произошла гомологичная рекомбинация с заменой одного из экзонов на инсертированный мини-ген HPRT. Факт такого встраивания доказывается при помощи ПЦР. На следующем этапе инсертированный HPRT ми-

ни-ген заменяют на отсутствующий фрагмент гена-мишени, в который предварительно вносят интересующие исследователя мутации. При этом альтернативный вектор несет те же фланкирующие ДНК-последовательности гена-мишени, что и первый (нокаутующий) вектор. Клетки HPRT минус на этом, 2-м этапе с большой вероятностью будут нести гомологичную рекомбинацию встроенной конструкции мини-HPRT гена и альтернативного фрагмента исходного гена. Факт такой рекомбинации контролируется с помощью ПЦР. Таким образом, вместо обычного выключения функции гена, что достигается уже на 1-м этапе, данная технология позволяет вносить в структуру гена дикого типа различные, заранее спланированные изменения, в том числе и специфические мутации, аналогичные таковым при наследственных болезнях у человека. Следовательно, данный подход позволяет проводить более тонкое генетическое моделирование и исследовать особенности функции мутантного гена *in vivo*.

Для введения специфических мутаций в определенные экзоны гена используют, так называемые "hit & run" векторы (Nasty et al., 1991). Перспективным также представляется использование дрожжевых YAC-векторов, несущих полноразмерные кДНК-овые последовательности гена. Так как уровень гомологичной рекомбинации у дрожжей достаточно высок, в такие конструкции легко вводить специфические мутации и затем использовать их для трансфекции ЭСК и получения трансгенных животных 4.

Отбор клонов эмбриональных стволовых клеток, в которых произошла направленная модификация гена-мишени, в значительной степени, предопределяет успех всего комплекса работ по созданию модельной генетической линии. Однако, и дальнейшие этапы этой программы, включающие получение химерных трансгенных животных, идентификацию зародышевых трансмиттеров (химер, продуцирующих трансфицированные половые клетки) и селекцию гетерозиготных, а затем гомозиготных мутантных особей, требуют большой квалификации, труда и времени. Осложняющим обстоятельством является то, что химерные животные нередко имеют сниженную жизнеспособность и плодовитость. То же может быть справедливо и в отношении гетерозиготных мутантных особей. В гомозиготном состоянии инсертированные мутации могут не только снижать жизнеспособность и плодовитость, но и обладать летальным или полуметальным эффектом уже в пренатальном периоде. В таком случае линия поддерживается путем отбора и скрещивания гетерозигот.

Несмотря на огромные методические сложности и высокую стоимость, направленное получение моделей наследственных болезней оправдывает затраченные усилия. Мутантные животные представляют уникальную возможность исследовать патофизиологические процессы, развивающиеся в организме вследствие нарушений работы определенного гена, анализировать влияние специфических мутаций на фенотип, тестировать новые лекарственные препараты и испытывать различные терапевтические подходы. Велика также роль генетических линий в разработке методов генной терапии (см. Главу IX).

ГЛАВА IV.

ТИПЫ И НОМЕНКЛАТУРА МУТАЦИЙ. МЕТОДЫ ДНК- ДИАГНОСТИКИ.

Раздел 4.1 Мутантные аллели, характеристика и типы му-

таций.

Каждый генетический локус характеризуется определенным уровнем изменчивости, то есть присутствием различных аллелей или вариантов последовательностей ДНК у разных индивидуумов. Применительно к гену, аллели разделяются на две группы – нормальные, или аллели дикого типа, при которых функция гена не нарушена, и мутантные, приводящие к нарушению работы гена. В любых популяциях и для любых генов аллели дикого типа являются преобладающими. Под мутацией понимают все изменения в последовательности ДНК, независимо от их локализации и влияния на жизнеспособность особи. Таким образом, понятие мутации является более широким по сравнению с понятием мутантного аллеля. Уместно, однако, заметить, что в научной литературе сравнительно часто встречающиеся в популяциях варианты последовательностей генов, не приводящие к заметным нарушениям функций, обычно рассматриваются как нейтральные мутации или полиморфизмы, тогда как понятия "мутация" и "мутантный аллель" зачастую употребляются как синонимы.

Как упоминалось ранее, различные изменения в нуклеотидной последовательности транскрибируемых областей ДНК могут по-разному проявляться в фенотипе. Часть из них не оказывает никакого влияния на структуру и функцию соответствующего белка. Примером могут служить замены нуклеотидов, не приводящие к замене аминокислот в силу вырожденности генетического кода. Мутантные аллели, в свою очередь, могут быть подразделены на три класса: (1) мутации, ведущие к полной потере функции (loss-of-function), (2) мутации, сопровождающиеся количественными изменениями соответствующих мРНК и первичных белковых продуктов и (3) доминантно-негативные мутации, изменяющие свойства белковых субъединиц таким образом, что они оказывают повреждающий эффект на жизнеспособность или функционирование экспрессирующих типов клеток (gain-of-function мутации). Наибольшим повреждающим действием обладают мутации, приводящие либо к образованию бессмысленного белка, либо к преждевременному окончанию его синтеза, то есть делеции или инсерции, не кратные трем нуклеотидам и потому вызывающие сдвиг рамки считывания, а также нонсенс мутации – замены нуклеотидов, при которых образуются терминирующие стоп-кодоны. Проявление таких мутаций зависит от их внутригенной локализации. Чем ближе мутации к 5' концу гена, то есть к началу транскрипции, тем короче их белковые продукты. Такие абортивные (truncated) белки неспособны к модификациям и быстро деградируют.

Фенотипическое проявление замен нуклеотидов в кодонах, так называемых миссенс мутаций, зависит от природы соответствующих аминокислотных замен в белке и от функциональной значимости того домена, в котором это произошло. Так, замены аминокислот в активных центрах белков могут сопровождаться полной потерей его функциональной активности, тогда как даже значительно более серьезные нарушения в других частях белка часто оказывают существенно меньшее влияние на фенотип. Мутации на стыке экзонов и интронов (так называемые сплайсинговые мутации) часто нарушают процессинг первичного РНК-транскрипта, в результате чего происходит либо неправильное вырезание соответствующей интронной области и трансляция бессмысленного удлинённого белка, не защищенного от протеолитического действия внутриклеточных ферментов, либо вырезание экзонов и образование делетированного белка. В обоих случаях сплайсинговые мутации, как правило, обуславливают тяжелое течение болезни. Нарушения в регуляторных об-

ластях генов сопровождаются количественными изменениями соответствующего продукта и не затрагивают структуры и функциональной активности белка. Проявление таких мутаций определяется, в конечном счете, пороговым уровнем концентрации белка, при котором его функция еще сохраняется. Как правило, регуляторные мутации менее серьезны и обладают более выраженным плейотропным (множественным) эффектом по сравнению с мутациями структурных генов.

Относительно недавно выявлен новый класс так называемых динамических мутаций, или мутаций экспансии, связанных с нестабильностью числа тринуклеотидных повторов в функционально значимых частях генов. Многие тринуклеотидные повторы, локализованные в транскрибируемых или регуляторных областях генов, характеризуются высоким уровнем популяционной изменчивости, в пределах которого не наблюдается фенотипических нарушений (Willems, 1994). Болезнь развивается лишь тогда, когда число повторов в этих сайтах превосходит определенный критический уровень. Наследование таких мутаций, как правило, отличается от классического Менделевского типа. Для них характерны: различная пенетрантность в сочетании с неполным доминированием; геномный импринтинг (различия фенотипических проявлений в зависимости от того, получена мутация от матери или от отца) и феномен антиципации - нарастание тяжести проявления заболевания в последующих поколениях (Willems, 1994).

Классическим примером мутаций экспансии является синдром ломкой X-хромосомы (FraXA), обусловленный присутствием удлинённых CCG повторов в 5'-нетранслируемой регуляторной области FMR1-гена (Xq27.3). Аналогичные нестабильные повторы обнаружены еще в трех ломких сайтах, причем два из них (FraXE и FraXF) расположены на очень небольшом расстоянии дистальнее FraXA. Во всех четырех случаях CCG-повторы локализованы вблизи от CpG островков, при этом увеличение числа копий триплетов выше определенного порогового уровня сопровождается гиперметилированием всей регуляторной GC-богатой области, вследствие чего и происходит резкое снижение и полное выключение транскрипционной активности - мутации по типу " утраты функции" (loss-of-functions). Таким образом, область CCG-повторов в этих локусах можно рассматривать, как своеобразный cis-действующий элемент транскрипции (Willems, 1994, Mandel, 1994).

Другой тип динамических мутаций описан для 6-ти различных тяжелых аутосомно-доминантных нейродегенеративных расстройств (см. Главу X). Для всех этих заболеваний обнаружено присутствие удлинённых CAG-повторов в открытой рамке считывания (ORF). Эти повторы транслируются в протяженные полиглутаминовые треки, предположительно локализованные в ДНК- связывающих доменах соответствующих белковых продуктов. В результате белковые молекулы приобретают новые свойства, нарушающие нормальные метаболические связи. Таким образом, нестабильные CAG-повторы можно рассматривать, как gain-of-function - мутации. Интенсивно обсуждается также возможность участия амплификации CAG-повторов в формировании предрасположенности к таким частым расстройствам центральной нервной системы, как шизофрения и маниакально-депрессивный психоз. Примером третьей группы болезней экспансии служит миотоническая дистрофия. При этом заболевании огромные CTG (или CAG) повторы локализованы в 3'-нетранслируемой области гена. Они также рассматриваются, как факторы, нарушающие нуклеосомную организацию гена и подавляющие его транскрипцию. Более подробно болезни экспансии рассмотрены в Главе X.

Раздел 4.2. Генетическая гетерогенность наследственных заболеваний.

Одним из важных обобщающих итогов молекулярно-генетических исследований моногенных болезней явилось доказательство их генетической гетерогенности. Последняя может быть вызвана разными причинами. Прежде всего, оказалось, что один и тот же биохимический эффект (фенотип) может быть обусловлен мутациями в разных генах. С другой стороны, мутации одного и того же гена, как установлено, могут приводить к совершенно разным клиническим проявлениям. Например, мутации гена адренорецептора, сцепленного с X-хромосомой, могут быть причиной нейродегенеративного заболевания - болезни Кеннеди, если они захватывают область тринуклеотидных повторов (Глава X), и в то же время приводить к синдрому тестикулярной феминизации, то есть нарушениям половой дифференцировки, если они затрагивают другие последовательности этого же гена. Крайним выражением такой гетерогенности может служить пример с геном рецептора тирозинкиназы -RET, различные мутации которого могут приводить к 4-м совершенно различным наследственным синдромам, таким как семейная медуллярная карцинома щитовидной железы, болезнь Гиршпрунга, множественная эндокринная неоплазия тип 2А (МЭН-2А) и тип 2В (МЭН-2В) (Hauningen, 1994). Подобные фенотипические разнообразия проявлений мутаций одного и того же гена получили название аллельных серий. Термин используется уже около 20 лет для описания групп из нескольких моногенных наследственных заболеваний, клинические проявления которых позволяют предполагать их связь с разными генами, в то время как биохимические и/или генетические исследования доказывают их аллельную природу, то есть в основе их патогенеза лежат разные мутации одного и того же гена.

В настоящее время известно более 100 таких болезней (Romeo, McKusick, 1994). Для каждого заболевания из подобной серии аллелизм мутаций уже доказан на молекулярном уровне. Причины подобного фенотипического разнообразия могут быть различными: (1) локализация мутантных аллелей в функционально разных доменах белка; (2) принципиально разный механизм действия мутаций (loss-of-function, gain-of-function); (3) присутствие в том же гене модифицирующего мутантного аллеля или полиморфизма и (4) влияние генетического окружения на проявление мутантного аллеля, то есть его взаимодействие с определенными аллелями гена-модификатора или даже несколькими такими генами. Углубленный молекулярно-генетический анализ практически каждого наследственного заболевания указывает на его значительную генетическую гетерогенность, связанную с различными мутациями гена. Некоторые примеры аллельных серий и генетической гетерогенности заболеваний будут рассмотрены более подробно в Главе X.

Раздел 4.3 Номенклатура мутаций.

Для практических целей и, главным образом, для чтения научной литературы, важно знать, как записываются мутации. До недавнего времени единой номенклатуры записи мутаций не существовало. В 1992 г. двумя американскими учеными Артуром Боде и Лап-Чи Тсуи была предложена универсальная стандартная система для обозначения разных мутаций (Beudet, Lap-Chee Tsui, 1993). Она рассчитана как на запись аминокислотных замен в белках, так и на нуклеотидные замены и перестановки в

ДНК. В первом случае, каждой аминокислоте соответствует однобуквенный символ (Табл.4.1), слева записывается нормальный вариант аминокислоты, справа - мутантный, а расположенный в центре номер соответствует месту замены в цепочке первичного продукта трансляции. Например, запись D44G означает замену аспарагина на глицин в 44-м положении полипептидной цепи, а A655E - аланина на глутамин в положении 655 белкового продукта. Так записываются различные варианты аминокислотных замен при миссенс мутациях. Буквой X обозначается место остановки синтеза полипептидной цепи при нонсенс мутациях. Например, Q39X означает замену глицина на стоп сигнал в 39-м кодоне, а W1282X - триптофан-триплета на стоп-кодон в положении 1282. Отсутствие одной или нескольких аминокислот обозначают значком Δ . Так, наиболее частая мутация, приводящая к муковисцидозу - Δ F508 - означает отсутствие фенилаланина в 508 положении трансмембранного регуляторного белка муковисцидоза. Полиморфизмы, связанные с равноценной по функциональной значимости заменой аминокислот, записывают через черточку. Например, M/V470 - метионин или валин в положении 470.

Таблица 4.1. Символы аминокислот.

| Аминокислоты | 1 | 0 | Трехбуквенный символ | 1 | 0 | Однобуквенный символ | 1 |
|-----------------------|---|---|----------------------|---|---|----------------------|---|
| Аланин | 1 | 0 | Ala | 1 | 0 | A | 1 |
| Аргинин | 1 | 0 | Arg | 1 | 0 | R | 1 |
| Аспарагин | 1 | 0 | Asn | 1 | 0 | N | 1 |
| Аспарагиновая кислота | 1 | 0 | Asp | 1 | 0 | D | 1 |
| Asn и/или Asp | 1 | 0 | Asx | 1 | 0 | B | 1 |
| Цистеин | 1 | 0 | Cys | 1 | 0 | C | 1 |
| Глутамин | 1 | 0 | Gln | 1 | 0 | Q | 1 |
| Глутаминовая кислота | 1 | 0 | Glu | 1 | 0 | E | 1 |
| Gln и/или Glu | 1 | 0 | Glx | 1 | 0 | Z | 1 |
| Глицин | 1 | 0 | Gly | 1 | 0 | G | 1 |
| Гистидин | 1 | 0 | His | 1 | 0 | H | 1 |
| Изолейцин | 1 | 0 | Ile | 1 | 0 | I | 1 |
| Лейцин | 1 | 0 | Leu | 1 | 0 | L | 1 |
| Лизин | 1 | 0 | Lys | 1 | 0 | K | 1 |
| Метионин | 1 | 0 | Met | 1 | 0 | M | 1 |
| Фенилаланин | 1 | 0 | Phe | 1 | 0 | F | 1 |
| Пролин | 1 | 0 | Pro | 1 | 0 | P | 1 |
| Серин | 1 | 0 | Ser | 1 | 0 | S | 1 |
| Треонин | 1 | 0 | Thr | 1 | 0 | T | 1 |
| Триптофан | 1 | 0 | Trp | 1 | 0 | W | 1 |
| Тирозин | 1 | 0 | Tyr | 1 | 0 | Y | 1 |
| Валин | 1 | 0 | Val | 1 | 0 | V | 1 |

Принципиальная схема записи и нумерации нуклеотидов приведена на Рис.4.1. Отсчет нуклеотидов в молекуле ДНК начинается с первого смыслового кодона, так что нуклеотид под номером +1 соответствует первому нуклеотиду в молекуле кДНК. Вверх по течению (или справа налево от 3' к 5'-концу) от первого кодона нуклеотиды записывают со знаком "-", вниз по течению (от 5' к 3') - со знаком "+". Для многих генов отсутствие точных данных о положении иницирующего сайта и наличие нескольких мест инициации транскрипции существенно затрудняют нумерацию нуклеотидов. Нуклеотиды экзонов обозначают заглавными буквами, интронов - прописными.

В Табл.4.2. даны примеры обозначения различных мутаций с использованием как аминокислотной, так и нуклеотидной нумерации. Нуклеотидная система записи особенно важна для обозначения делеций, инсерций, сплайсинговых мутаций и полиморфизмов, не связанных с заменами аминокислот или происходящими в нетранслируемых частях гена. В случае делеции или инсерции одного или двух нуклеотидов приводится их буквенное обозначение. Например, 441delA, 485insTA. При делеции или инсерции трех и более нуклеотидов указывается только их число. Так, 852del22 означает делецию 22 нуклеотидов, начиная с 852-го нуклеотида, а 3320ins7 обозначает вставку 7 пар оснований после нуклеотида 3320. В случае больших вставок или делеций их размеры указываются в килобазах, например 2115ins13kb, или обозначаются соответствующие инsertированные/делетированные структурные элементы генома. Так, 2115insAlu означает инсерцию Alu-повтора после нуклеотида 2115. При обозначении сплайсинговых мутаций записывают номер крайнего нуклеотида в ближайшем к мутации экзоне, число нуклеотидов (со знаком "+" в случае 3' конца экзона и со знаком "-" в случае 5' конца) и характер нуклеотидной замены. (Рис.4.1, Таб.4.2). Например, 711+5G-T обозначает замену G на T в 5-м основании интрона, следующего за экзоном, заканчивающимся 711 нуклеотидом.

Раздел 4.4. Идентификация структурных мутаций, изоляция мутантных ДНК.

Идентификация мутантных аллелей, то есть обнаружение нарушений в первичной нуклеотидной последовательности ДНК, является самым прямым методом молекулярной диагностики наследственных заболеваний. Большие перестройки - делеции, дупликации, инверсии, транслокации - размером более 1 Мб, затрагивающие целые гены или даже несколько генов, могут быть обнаружены на цитологических препаратах с использованием техники высокого разрешения хромосомного анализа (анализ дифференциально окрашенных прометафазных хромосом). Еще большая разрешающая способность (до 50 kb) может быть достигнута при работе на специальном образом приготовленных (растянутых) интерфазных хромосомах с использованием техники гибридизации *in situ* (FISH - Глава II).

Мутации, изменяющие длину рестриционных фрагментов, могут быть выявлены путем блот-гибридизации рестрицированной геномной ДНК с соответствующими ДНК-зондами на стадиях генетического анализа, предшествующих молекулярному клонированию гена. К числу таких мутаций относятся достаточно протяженные, но не идентифицируемые цитогенетически, внутригенные делеции, инсерции и дупликации, а также точечные мутации, локализованные в сайтах рестрикции. Непременным условием реализации метода блот-гибридизации для поиска подобных мутаций является наличие ДНК-зонда, являющегося либо частью гена, либо тесно сцепленной с этим геном клонированной ДНК-последовательностью. При поиске таких мутаций геномную ДНК от здорового донора и больного обрабатывают часто щепящими рестриктазами, подвергают электрофорезу, блот-гибридизации с меченым ДНК-зондом по стандартной схеме (Глава I) и проводят сравнительный анализ расположения бэндов на радиоавтографе. Особенно информативными обычно оказываются рестриктазы Msp1 и Taq1, которые узнают сайты CGGG и TGGA, соответственно. Благодаря наличию CpG последовательностей эти сайты особенно часто подвергаются спонтанному мутированию (см.раздел 4.5). Наличие протяженных делеций, либо точечных

мутаций в сайтах рестрикции приводит к изменениям размеров рестрикционных фрагментов. Целенаправленный поиск других точечных мутаций (не затрагивающих сайты рестрикции) и небольших структурных аномалий возможен только для клонированных генов с известной нуклеотидной последовательностью смысловых участков ДНК. Методы идентификации подобных мутаций основаны, главным образом, на использовании полимеразной цепной реакции в её различных модификациях (см. Главу I, а также разделы 4.5 и 4.6).

Для анализа мутантных аллелей, прежде всего, необходимо иметь изолированные последовательности мутантного гена, которые могут быть получены либо путем клонирования или амплификации мутантной кДНК, либо за счет специфической амплификации отдельных экзонов, их частей и регуляторных областей гена с использованием в качестве матрицы геномной ДНК пациентов (Рис. 4.2). В первом случае отбирают клонированные к-ДНК-последовательности мутантного гена, проводя скрининг кДНК-овых библиотек, сконструированных из специфических тканей или культур клеток больного. При этом в качестве зондов используют кДНК-овые последовательности нормального гена. Другим источником кодирующих последовательностей мутантного гена может служить мРНК, изолированная из экспрессирующих тканей или клеток больного. Мутантную кДНК получают путем специфической амплификации перекрывающихся последовательностей кодирующих областей гена, используя в качестве матрицы тотальную кДНК, полученную при обратной транскрипции изолированной мРНК. Преимуществом этого подхода является то, что праймеры для амплификации выбирают из экзонных областей, нуклеотидные последовательности которых, как правило, становятся известны вскоре после идентификации и клонирования гена. В ряде случаев изоляция мутантной мРНК затруднена в связи с недоступностью образцов тканей или органов, в которых происходит экспрессия нужного гена (мозг, печень и др.). Однако, обнаружение следовых количеств, так называемой, незаконной или эктопической мРНК во многих клетках и тканях, в том числе в клетках крови, позволяет преодолеть и эти трудности (Kaplan et al., 1992). Успех подобной процедуры получения мутантной кДНК связан, в первую очередь, с разработкой эффективных методов выделения и обратной транскрипции мРНК с сохранением всех типов кДНК, включая те, для которых соответствующие мРНК присутствуют в ничтожных концентрациях (например, мРНК дистрофина при мышечной дистрофии Дюшенна (см. Главу X). Единственным принципиальным ограничением методов детекции мутаций в кДНК-овых последовательностях является невозможность выявления мутаций в регуляторных и интронных частях гена. Подобные мутации могут быть выявлены только при анализе геномной ДНК пациента.

Получение геномной мутантной ДНК обычно не представляет сложностей, так как она может быть изолирована из любых клеток или тканей больного независимо от характера экспрессии исследуемого гена. Однако, амплификация целых экзонов возможна только при знании нуклеотидных последовательностей фланкирующих интронных областей, из которых и производят подбор специфических олигопраймеров. Секвенирование интронов представляет собой достаточно трудоемкую задачу, решенную далеко не для всех клонированных генов. Таким образом, к ограничениям этого подхода следует отнести необходимость достаточно полной информации о структуре гена и о его первичной нуклеотидной последовательности. Кроме того, объектом тестирования могут быть лишь сравнительно небольшие области гена, отсюда для получения более полной информации необходи-

ма амплификация многих экзонов.

Стратегия идентификации мутаций может быть различной и, в конечном счете, определяется тем, имеем ли мы дело с ранее неизвестными мутациями, либо целью анализа является скринирование уже известных мутаций. В первом случае объектом исследования чаще всего служат клонированные или амплифицированные кДНК-овые последовательности, тогда как при молекулярной диагностике известных мутаций, как правило, анализируют амплифицированные фрагменты геномной ДНК.

Раздел 4.5. Первичная идентификация точечных мутаций.

Любые типы мутаций могут быть обнаружены путем прямого секвенирования мутантной кДНК или отдельных экзонов и часто первичный поиск нарушений в кодирующих областях гена осуществляют именно таким образом. Сам метод секвенирования уже был рассмотрен ранее (см.Главу I, раздел 1.6). Для некоторых генов, имеющих небольшие размеры, метод прямого секвенирования с успехом применяется как основной метод сканирования мутаций. Так, в частности, особенно удобным оказалось его применение для детекции мутаций в сравнительно небольших по размеру генах, таких, например, как ген фактора IX свертывания крови (гемофилия В). Использование эктопической мРНК для получения амплифицированных кДНК-овых фрагментов открывает особенно широкие возможности для применения метода прямого секвенирования.

Разработанные в последние годы модификации методов ПЦР значительно облегчили секвенирование амплифицированных фрагментов и повысили его эффективность. Так, в частности, предложен вариант ассиметричной ПЦР, когда при амплификации концентрация одного из олигопраймеров в несколько десятков раз превосходит концентрацию другого праймера, в результате чего синтезируется преимущественно только одна, нужная для секвенирования цепочка ДНК. Для этой же цели (получения одноцепочечной ДНК) предложено использование магнитных частиц с фиксированным на их поверхности стрептавидином. При этом один из праймеров для проведения ПЦР метится биотином. Затем к продуктам амплификации добавляют магнитные частицы с пришитым стрептавидином. Благодаря прочному связыванию биотин - стрептавидин, меченая биотином последовательность ДНК фиксируется на магнитных частицах. С помощью щелочного лизиса с частиц удаляют вторую немеченую комплементарную последовательность ДНК, которую и используют для секвенирования. Еще в одном варианте амплификацию проводят в присутствии праймеров, несущих сайт узнавания для фермента T7 - РНК полимеразы. После амплификации в системе *in vitro* проводят транскрипцию амплификата с помощью T7-РНК полимеразы и образовавшуюся одноцепочечную РНК используют для секвенирования - метод GAWTS (genome amplification with transcript sequences).

Однако, в общем случае секвенирование полноразмерной кДНК или всех экзонов для генотипирования мутаций у отдельных пациентов достаточно трудоемко, дорого и требует много времени. Поэтому на практике чаще проводят предварительный отбор более простыми методами амплифицированных, а иногда клонированных фрагментов ДНК, предположительно содержащих мутации, а затем секвенируют только эти участки ДНК. Методы поиска фрагментов ДНК, предположительно содержащих мутации, основаны на сравнительном анализе мутантных и нормальных последовательностей по целому ряду физических и химических характеристик, которые в значительной степени варьируют в

зависимости от типа мутационного повреждения. Следует подчеркнуть, что независимо от метода детекции мутации и практически независимо от её природы (замены нуклеотидов, делеции, дупликации и пр.) точные молекулярные характеристики каждой мутации могут быть получены только путем прямого секвенирования. При наличии в амплифицированном фрагменте известных сайтов рестрикции положение мутации может быть предварительно уточнено. Для этого продукты амплификации разрезают соответствующей эндонуклеазой и исследуют более короткие фрагменты.

Наиболее просто обнаруживаются мутации, изменяющие длину амплифицированных фрагментов, так как подобные нарушения легко выявляются при электрофоретическом анализе. Так, протяженные делеции, захватывающие целые экзоны, могут быть выявлены по изменению длины рестриционных фрагментов, гибридизующихся со специфическими ДНК-зондами. Более простая и эффективная методика выявления таких мутаций в генах, сцепленных с полом, основана на одновременной амплификации различных экзонов, наиболее часто вовлекаемых в подобные перестройки, так называемый мультиплексный вариант ПЦР. Разница в размерах и числе амплифицированных фрагментов позволяет легко идентифицировать такие мутации на электрофореze (Рис.4.2). Особенно широко этот метод используется для идентификации делеций в гене дистрофина, на долю которых приходится около 60% всех мутаций, приводящих к миодистрофии Дюшенна (см.Главу X). При отсутствии делеций все амплифицированные фрагменты после электрофоретического разделения и окрашивания можно наблюдать в виде отдельных полос. Если в исследуемой ДНК какие-то из экзонов делетированы, будут отсутствовать и соответствующие им полосы на электрофореграмме (Рис.4. 2). Выбирая специфические участки гена для амплификации, можно оценить размер делеции с точностью до отдельных экзонов, а также определить ее внутригенную локализацию.

Метод этот, однако, не обнаруживает подобные делеции, находящиеся в гетерозиготном состоянии или локализованные в аутосомных генах, так как нормальная гомологичная последовательность геномной ДНК может служить матрицей для амплификации любых фрагментов. Данный подход применим к анализу делеций в аутосомных генах только в тех случаях, когда возможно дополнить ПЦР количественной оценкой результатов амплификации – так называемая количественная ПЦР. Оригинальный метод идентификации подобных делеций у гетерозигот основан на использовании в качестве матричной ДНК для ПЦР кДНК, полученной путем обратной транскрипции из эктопической мРНК или из мРНК, изолированной из экспрессирующих данный ген тканей или культур клеток пациента. В отличие от нормального гомолога, в мутантной молекуле кДНК граничащие с делецией экзоны сближены. Если в качестве олигопраймеров для ПЦР будут выбраны последовательности из этих областей гена, только мутантная кДНК будет служить матрицей для амплификации небольшого участка между праймерами из фланкирующих делецию экзонов. В нормальной последовательности кДНК этот участок может быть слишком велик, для того чтобы прошла амплификация. Практически, для обнаружения гетерозигот по протяженным внутригенным делециям проводят мультиплексную ПЦР с использованием системы олигопраймеров, обеспечивающих амплификацию фрагментов, полностью перекрывающих всю молекулу кДНК. Наличие делеции регистрируют по появлению продуктов амплификации необычного размера.

Небольшие делеции и вставки нуклеотидов не приводят к

отсутствию амплифицированных фрагментов ДНК, но изменяют их размеры. Эти изменения могут быть зарегистрированы при электрофорезе продуктов амплификации в полиакриламидном или агарозном гелях (Рис.4.3). Именно этот метод используется для детекции наиболее часто встречающейся мутации в гене муковисцидоза – делеции трех нуклеотидов $\Delta F508$. После выявления различий между нормальной и мутантной ДНК по длине рестрикционных или амплифицированных фрагментов гена необходимо провести секвенирование необычного фрагмента, с целью определения изменений в первичной структуре мутантной ДНК последовательности по сравнению с нормальной.

При мутациях гена, представляющих собой замену одного или нескольких нуклеотидов, длины амплифицированных фрагментов остаются постоянными, однако, некоторые физико-химические свойства мутантных молекул ДНК меняются. Так, например, при гибридизации однонитевых ДНК, комплементарных нормальной и мутантной нитям ДНК, возникают структурные нарушения в месте негомологичного спаривания. С учетом этих особенностей разработаны различные варианты поиска мутантных фрагментов ДНК и идентификации в них точечных мутаций. Ведущими из этих методов являются: метод анализа конформационного полиморфизма однонитевой ДНК – SSCP, денатурирующий градиентный гель-электрофорез – DGGE, метод химического расщепления некомплементарных сайтов (CMC), метод гетеродуплексного анализа (НА) и, наконец, собственно метод секвенирования. Основные характеристики этих методов приведены в Табл.4.3

Таблица 4.3. Преимущества и недостатки основных методов первичной идентификации мутаций.

| метод | размер фрагмента (п.о.) | %% детекции мутаций | точность картирования мутации | сканирование экзонов | сканирование кДНК |
|--------|-------------------------|---------------------|-------------------------------|----------------------|-------------------|
| SSCP | 250 | 80% | нет | +++ | + |
| DGGE | 600 | 95% | нет | ++ | ++ |
| CMC | 1700 | >95% | да | + | +++ |
| PCR DS | 500 | >99% | да | ++ | ++ |
| НА | 300 | 80% | нет | ++ | + |

"+" – применимость метода для сканирования геномной ДНК и кДНК

SSCP (Single Strand Conformation Polymorphism) – метод анализа конформационного полиморфизма однонитевой ДНК, предложенный (Orita et al.1989, Clavac, Dean, 1993) основан на регистрации различий в электрофоретической подвижности однонитевых ДНК, одинаковых по величине, но различающихся вследствие нуклеотидных замен по пространственной организации молекул (Рис 4.4). Скручивание или конформация небольших однонитевых ДНК существенно зависит от их нуклеотидной последовательности, так что замена даже одного основания в молекулах одинакового размера приводит к изменению их пространственной структуры. Метод включает амплификацию специфических сегментов ДНК размером от 50 до 300 пар оснований, обычно в присутствии меченых трифосфатов, денатурацию образовавшихся продуктов ПЦР и нативный высокоразрешающий электрофорез в полиакриламидном геле. Иногда амплификацию прово-

дят без использования метки, но тогда для лучшего разделения ДНК и однозначной идентификации бэндов на электрофореграмме используют специальные гели - Hydrolink либо MDE (AT Biochem, USA), а также более чувствительные по сравнению с этидиумом бромидом методы окрашивания, такие как окраска серебром. На процесс конформации оказывают влияние различные внешние факторы - температура, концентрация акриламида и глицерина в геле, ионная сила буферных растворов (Clavac, Dean, 1993). Оптимальный подбор этих параметров позволяет эффективно разделять амплифицированные фрагменты ДНК, различающиеся даже всего на один нуклеотид. Конформационный метод выявления точечных мутаций быстро получил широкое распространение благодаря своей простоте и возможности обнаруживать любые типы замен. Эффективность детекции мутаций при размерах амплифицируемого фрагмента менее 200 п.о. составляет 70-95%, но при длине фрагмента, превышающей 400 п.о., вероятность обнаружения мутаций уменьшается до 50%.

DGGE (Denaturation Gradient Gel Electrophoresis) - метод денатурирующего градиентного гель-электрофореза, основан на зависимости свойств плавления (или денатурации) небольших двухнитевых молекул ДНК от их нуклеотидной последовательности, а точнее от соотношения А-Т и G-С пар в исследуемых фрагментах (Майерс и др., 1990; Fodde, Losekoot, 1994). Объясняется это тем, что G-С связь более прочная по сравнению со связью между нуклеотидами А и Т. Подобные различия в динамике плавления могут быть выявлены путем сравнения подвижности нормальных и мутантных двухнитевых фрагментов ДНК при их электрофорезе в денатурирующих условиях (Рис.4.5). Градиент денатурации достигается разницей температур, различной концентрации мочевины или формальдегида в гелях. При этих условиях одинаковые по величине двухнитевые молекулы ДНК, отличающиеся по нуклеотидной последовательности, денатурируют по-разному. Разработан компьютерный алгоритм, позволяющий предсказывать характер плавления в зависимости от нуклеотидной последовательности (Lerman, Silverstein, 1987). При электрофорезе амплифицированных двухнитевых фрагментов ДНК в геле с линейно возрастающим градиентом концентраций денатурирующих агентов плавление нитей ДНК происходит в строго специфичной для данной последовательности области, эквивалентной температуре плавления - t_m , то есть такой температуре, при которой каждая пара оснований с 50%-ой вероятностью может соединиться или разойтись. После начала плавления продвижение двухнитевого фрагмента ДНК в геле резко замедляется вследствие сложной пространственной конфигурации молекул, причем эта задержка будет длиться до тех пор, пока не наступит полная денатурация ДНК. В результате происходит разделение фрагментов ДНК, различающихся по нуклеотидному составу. Таким способом удается идентифицировать лишь около 50% однонуклеотидных замен в фрагментах ДНК длиной от 50 до нескольких сотен нуклеотидов. Связано это с тем, что при прохождении ДНК через гель может начаться частичная денатурация концов молекул еще до достижения оптимальной области плавления. Поэтому мутации, локализованные вблизи концов амплифицированных участков ДНК, оказывают меньшее влияние на процесс плавления и поэтому могут не выявляться. Эффективность обнаружения мутаций с помощью градиентного денатурирующего электрофореза может быть существенно повышена за счет присоединения к концам амплифицированной геномной ДНК синтетических фрагментов GC-нуклеотидов, длиной в несколько десятков пар оснований. Такие монотонные тугоплавкие концы выполняют роль своеобразных зажимов и резко увеличивают

шансы обнаружения для всех точечных мутаций, независимо от их локализации внутри исследуемого фрагмента ДНК. Эта модификация делает метод очень чувствительным (см. Таб. 4.2). В отличие от SSCP он пригоден для более крупных амплифицированных фрагментов ДНК. При исследовании фрагментов до 600 п.о. эффективность выявления мутаций этим методом достигает 95%. Чаще всего DGGE – метод применяется для скрининга мутаций в амплифицированных экзонах, при этом в качестве матрицы используют геномную ДНК. Этот метод может быть с успехом применен также для анализа индуцированных мутаций, так как позволяет улавливать точечные мутации, возникшие даже в одной из 100 обработанных мутагеном клеток. К недостаткам метода следует отнести техническую сложность получения равномерного градиента денатурирующего агента в полиакриламидном геле, а также высокую стоимость искусственно синтезированных GC-концов.

НА (Heteroduplex analysis) – гетеродуплексный анализ позволяет идентифицировать мутации, находящиеся в компаунде или в гетерозиготном состоянии. Следует заметить, что у подавляющего большинства пациентов с генными болезнями, наследуемыми по аутосомно-рецессивному типу, мутации в гомологичных хромосомах находятся в компаунде, то есть в каждом из гомологичных генов имеются функционально значимые нарушения, но их молекулярная природа и внутригенная локализация различны. Исключения составляют лишь мажорные мутации, частота которых в популяции достигает десятков процентов. Примерами таких мутаций являются $\Delta F508$ в гене муковисцидоза или R408W мутация в гене фенилкетонурии. Принцип НА-метода заключается в том, что при амплификации относительно небольших фрагментов генов гетерозигот или гомозиготных компаундов мутация может быть локализована лишь в одной из гомологичных нитей матричной ДНК. Поэтому в амплификационной смеси наряду с двумя типами гомодуплексов образуются гетеродуплексы между нормальной и мутантной цепочками ДНК. Такие гетеродуплексные молекулы ДНК имеют иную электрофоретическую подвижность по сравнению с гомодуплексами (не отличающимися между собой по подвижности) за счет конформационных особенностей в местах несовпадения нуклеотидов (mismatch) (Рис. 4.6). Эти различия могут быть обнаружены при электрофорезе в обычном полиакриламидном геле. Значительно более эффективное разделение гомо- и гетеродуплексов может быть достигнуто при использовании новых вариантов гелей – Hydrolink либо MDE. Вероятность идентификации точечных мутаций этим способом на фрагментах ДНК менее 300 п.о. достигает 80-90%. Детекция мутаций осуществляется как изотопным, так и неизотопными методами (Grompe, 1993).

CMS (Chemical Mismatch Cleavage) – метод химического расщепления некомплементарных сайтов, основан на способности некоторых химических агентов специфически разрывать нить ДНК в месте локализации неспаренного основания (Рис. 4.7). Так, цитозин чувствителен к действию гидроксилamina, а тимин – к действию тетроксид осмия. Некоторые модификации метода используют чувствительность тимина и гуанина к карбодиимиду. Последующая обработка пиперидином приводит к полному разрыву молекулы ДНК в модифицированном сайте. Выявление мутаций осуществляют с помощью меченых ДНК-зондов, соответствующих, как правило, нормальному варианту последовательности ДНК. Такими зондами могут быть синтезированные олигонуклеотиды, клонированные последовательности ДНК или амплифицированные фрагменты (Cotton, 1990; Cotton, Malcolm, 1991).

При проведении исследования эталонную меченую ДНК сме-

шивают с избытком тестируемой ДНК (или РНК). Тестируемыми образцами ДНК могут служить клонированные ДНК, обработанные соответствующими эндонуклеазами, либо амплифицированные фрагменты. Смесь нагревают до полной денатурации двухнитевых молекул и затем охлаждают, чтобы создать условия для образования дуплексов. При наличии мутаций в тестируемых образцах ДНК в гетеродуплексах, возникших в результате гибридизации между однонитевыми молекулами эталонной и тестируемой ДНК, образуются места негомологичного спаривания. После обработки соответствующими химическими агентами идентификация и локализация мутантных сайтов в исследуемых участках ДНК проводится путем электрофореза и авторадииграфии. Появление укороченных фрагментов ДНК на электрофореграмме (а точнее необычных бэндов в нижней части геля) свидетельствует о наличии мутантного сайта, а определение размера укороченных фрагментов однозначно определяет локализацию этого сайта в исходной тестируемой молекуле ДНК. Современные модификации метода SMC позволяют идентифицировать до 95-100% мутаций (Grompe, 1993). Большими преимуществами этого метода являются: (1) возможность исследовать протяженные участки ДНК - до 2 кб, (2) способность одновременно выявить и локализовать несколько мутаций в одном фрагменте ДНК и (3) возможность одновременно использовать несколько ДНК-зондов для поиска мутаций - мультиплексный вариант методики. К числу недостатков можно отнести высокую токсичность используемых химических реактивов. Последняя может быть частично ослаблена использованием карбодимидов для идентификации GT гетеродуплексов.

Весьма близким по принципу к SMC-методу является метод расщепления гетеродуплексов РНКазой А. С этой целью создаются условия для образования гетеродуплексов между тестируемой ДНК и комплементарной ей радиоактивно меченой РНК-пробой. При обработке РНКазой А происходит разрезание молекул РНК в местах нарушения спаривания оснований. Места точечных мутаций определяются как и в случае SMC, по размерам образовавшихся фрагментов после электрофореза и авторадииграфии. Необходимость использования радиоактивно меченой РНК-пробы и возможность детекции только около 50% точечных мутаций ограничивают широкое применение метода (Grompe, 1993).

Первичная идентификация мутаций может быть осуществлена путем анализа нарушений не в нуклеотидной последовательности гена, а в аминокислотной последовательности соответствующего полипептидного продукта. Для этого выделяют тотальную мРНК из лейкоцитов крови, проводят обратную транскрипцию, амплифицируют специфические экзоны кДНК (метод RT-PCR), встраивают амплифицированную область ДНК в экспрессионную систему и анализируют образовавшийся продукт. Этот метод особенно эффективен при детекции мутаций в протяженных генах, содержащих большое число экзонов, таких как ген миопатии Дюшенна или ген нейрофиброматоза 1.

Раздел 4.6. Молекулярное сканирование известных мутаций.

Рассмотренные выше методы обнаружения мутаций предполагают обязательное секвенирование содержащих их сегментов ДНК с целью точной идентификации нуклеотидных нарушений, оценки их фенотипического проявления и определения причастности к развитию болезни. Поэтому они редко используются в практической диагностике и при популяционном скрининге гетерозигот. После описания мутации появляется возможность ее анализа более простыми способами, не требующими секвенирования.

Как упоминалось выше, мутации, изменяющие длину амплифицированных фрагментов, могут быть выявлены с помощью нативного электрофореза в полиакриламидном или агарозном гелях.

Из миссенс мутаций наиболее просто диагностируются те замены нуклеотидов, которые приводят к исчезновению или образованию сайта узнавания для какой-нибудь из рестриктаз. Они выявляются по изменению длины амплифицированного фрагмента ДНК после его обработки соответствующей эндонуклеазой. Поэтому сразу после идентификации мутации проводится компьютерный поиск возможных сайтов рестрикции в месте локализации замены основания. Вероятность такого события довольно велика, так как для каждой из нескольких сотен известных в настоящее время рестрикционных эндонуклеаз сайтом узнавания служит своя специфическая последовательность ДНК, средние размеры которой составляют 5 - 6 нуклеотидов.

Если естественных рестрикционных сайтов в месте мутации найти не удастся, то такие сайты могут быть созданы искусственно. В частности, разработана методика создания с помощью ПЦР новых сайтов рестрикции в мутантных аллелях, но не в аллелях дикого типа - метод ПЦР-опосредованного сайт-направленного мутагенеза (Ng et al., 1991; Eiken et al., 1991). Для этого амплифицируемый участок ДНК выбирают таким образом, чтобы 3'-конец одного из праймеров непосредственно примыкал к мутантному сайту (Рис.4.8). Именно этот праймер неполностью комплементарен матричной ДНК. В нем изменяют один из нуклеотидов с 3'-конца так, чтобы в сочетании с нуклеотидом мутантного, но не нормального сайта в этом месте образовывался сайт рестрикции для какой-нибудь из эндонуклеаз. Тогда после рестрикции и электрофореза продуктов амплификации геномной ДНК у индивидуумов, не содержащих данную мутацию, на электрофореграмме будет присутствовать один нерестрицированный фрагмент, у гетерозигот появится два дополнительных фрагмента, соответствующих по длине рестрицированным сегментам ДНК, и у гомозигот по мутации будут присутствовать только эти два фрагмента.

Концептуально близким к этому варианту является метод получивший название "амплификация рефрактерной мутационной системы"- amplification refractory mutation system - ARMS. В основе метода лежит неспособность TaqI термофильной полимеразы к амплификации фрагмента при наличии несоответствия (mismatch) между матричной ДНК и 3'-концом одного из олигопраймеров (Newton et al., 1989; Bottema et al., 1990). Суть метода заключается в одновременном проведении двух ПЦР, для каждой из которых одним из праймеров служит аллель-специфическая мутантная или нормальная олигонуклеотидная последовательность, соответственно. При этом в качестве второго праймера для проведения двух реакций выбирают одну и ту же олигонуклеотидную последовательность, так что в обоих случаях могут амплифицироваться участки ДНК одинаковой протяженности. Мутантный сайт в аллель-специфических праймерах расположен не в центре, а ближе к 3'-концу, и чаще всего занимает предпоследнюю позицию. При определенных условиях, важнейшим из которых является концентрация ионов магния в растворе, наличие сайта негомологичного спаривания в 3'-области праймера препятствует началу синтеза ДНК. Таким образом, при наличии мутации в исследуемой матричной ДНК амплифицированные фрагменты образуются только в том случае, если в качестве аллель-специфического праймера выбирается мутантная последовательность, тогда как при использовании нормального олигонуклеотидного праймера ПЦР блокируется (Рис.4.9.). Метод нашел широкое применение для детекции мутаций при фенилкетону-

рии, бета-талассемии, муковисцидозе, при типировании генов HLA системы. Однако, сложности в подборе праймеров и в выборе оптимального режима ПЦР ограничивают широкое применение этого метода. Вместе с тем, его несомненным преимуществом является возможность применения полностью автоматического сканирования.

Таким же преимуществом обладают и методы детекции состояния гена, основанные на лигировании синтетических олигонуклеотидных зондов- OLA (oligonucleotide ligation assay). При проведении этих реакций специфические ДНК или РНК последовательности исследуют путем использования их в качестве матрицы для ковалентного связывания двух пар олигонуклеотидных зондов (Landegren, 1993). ДНК-зонды для лигирования подбирают таким образом, чтобы они были полностью комплементарны нормальному фрагменту ДНК в области локализации мутации, причем сама нуклеотидная замена должна находиться на стыке двух праймеров. Обычно в один из зондов вводят радиоактивную или флюоресцентную метку, а другой - метят биотином. После гибридизации при строго стандартных условиях синтезированные олигонуклеотидные последовательности сшивают ДНК-лигазами из термофильных микроорганизмов. Такие ферменты работают при высоких температурах и сохраняют свою активность в условиях, необходимых для проведения денатурации молекул ДНК. При наличии мутации в тестируемой молекуле ДНК на конце одного из зондов образуется сайт некомплементарного спаривания, непосредственно примыкающий к месту лигирования. Наличие терминального неспаренного основания в смежно расположенных последовательностях ДНК-зондов резко снижает скорость лигирования и при определенных условиях проведения реакции сшивки между зондами в этом случае не происходит. Метод включает несколько последовательных циклов гибридизации, лигирования и денатурации. Начиная со второго цикла, матричной ДНК для гибридизации зондов наряду с тестируемой пробой служат также лигированные последовательности. В дальнейшем проводят электрофоретический анализ меченых одностранных фрагментов ДНК. Система успешно апробирована на мутациях глобиновых генов при серповидно-клеточной анемии и на мутации delF508 при муковисцидозе.

Универсальным методом детекции замен оснований является метод аллель-специфических олигонуклеотидов - ASO, который включает амплификацию фрагментов ДНК и последующую дот- или слот-гибридизацию с мечеными аллель-специфическими олигонуклеотидами (Reiss, 1991). Для этого синтезируют два типа олигонуклеотидных последовательностей обычно размером 19 пар оснований, в которых мутантный сайт занимает центральное положение. Каждый из этих олигонуклеотидных зондов комплементарен нормальному или мутантному вариантам ДНК, соответственно. Условия гибридизации подбирают таким образом, чтобы дуплексы образовывались только при полной комплементарности гибридных пар. В этих условиях амплифицированные фрагменты ДНК без мутации будут гибридизоваться только с нормальным зондом, ДНК гомозигот по мутации - только с мутантным и ДНК гетерозигот - с обоими олигонуклеотидами (Рис.4.10). Для уменьшения перекрестной аллель-специфической гибридизации в реакционную смесь добавляют 30-кратный избыток конкурентного немеченого олигонуклеотидного ДНК-зонда. Разработаны удобные модификации метода ASO с использованием аллель-специфических ДНК-зондов, меченых биотином или пероксидазой хрена (Лебедева и др., 1994).

Наиболее быстрым, экономичным и удобным методом сканирования точечных мутаций является модифицированный вариант

ASO, так называемая гибридизационная система обратного дот-блота (reverse dot-blot hybridisation system) (Saiki et.al., 1989). Метод позволяет одновременно скринировать сразу много точечных мутаций и доступен автоматизации (Chebab, 1993). В этом случае проводят гибридизацию меченых продуктов ПЦР, обычно представляющих собой отдельные экзоны, с фиксированными на нейлоновых фильтрах аллель-специфическими олигонуклеотидными зондами (ASO). Предварительную иммобилизацию мутантных и нормальных ASO-зондов на мембранах осуществляют за счет присоединения гомополимерных Т-хвостов с дезоксирибонуклеотид-трансферазой на конце. При этом олигонуклеотидные последовательности остаются свободными и могут участвовать в гибридизации с мечеными амплифицированными фрагментами ДНК. После отмывки несвязавшихся молекул ДНК радиоавтографические или цветные пятна на фильтрах становятся заметными только в местах локализации олигонуклеотидов, полностью комплементарных тестируемой геномной ДНК. Реакцию, обычно проводят в присутствии ионов тетра-алкиламмония, уменьшающих зависимость температуры плавления от композиции оснований. Это позволяет использовать одинаковые условия гибридизации для различных олигонуклеотидов, то есть вести поиск сразу нескольких типов мутаций, локализованных в одном и том же экзоне гена. Данный метод положен в основу разработки специальных систем, предназначенных для одновременной детекции наиболее распространенных мутаций в исследуемом гене. Система представляет собой ленточный фильтр (стрип) с нанесенными пятнами олигопраймеров, каждый из которых соответствует определенной мутации. Стрип помещают в раствор со смесью тех меченых амплифицированных экзонов, которые могут содержать тестируемые мутантные аллели и создают условия для аллель-специфической гибридизации. Таким способом сканируют одновременно 42 мутации, ответственные за серповидноклеточную анемию, 34 мутации при муковисцидозе (Chebab, 1993)

Очень простой метод обнаружения и идентификации описанных ранее мутаций в амплифицированных фрагментах ДНК основан на анализе характера электрофоретического разделения продуктов ПЦР в MDE-гидросвязывающих гелях в присутствии высоких концентраций мочевины. Эти гели способствуют формированию гетеродуплексов в процессе электрофореза, причем расположение дуплексов очень специфично для различных мутантных аллелей, локализованных в одном и том же амплифицированном фрагменте. Это позволяет не только выявлять, но с высокой степенью вероятности идентифицировать известные мутации. В мультиплексном варианте методики возможен одновременный поиск мутаций в нескольких амплифицированных фрагментах. Простота и высокая скорость анализа способствуют разработке на основе разделения в MDE-гелях схем максимальной автоматизации процесса поиска и идентификации известных мутаций у пациентов и гетерозиготных носителей мутаций

Итак, методы выявления мутаций довольно разнообразны и постоянно совершенствуются. В первую очередь, молекулярному анализу подвергают те гены, повреждения которых сопровождаются развитием наиболее частых заболеваний. Исследования проводят либо в семьях высокого риска, где уже имеются больные с тем или иным моногенным заболеванием с целью выявления гетерозиготных носителей, либо непосредственно в популяциях, где имеется большая выборка больных и можно предполагать высокую частоту гетерозиготных носителей мутаций. При этом выбор конкретных схем идентификации мутантных аллелей, зачастую, определяется диагностическими возможностями лабораторий и стоимостью анализов. Особое внимание уделяют поиску

тех аллелей, которые встречаются в популяциях с высокой частотой. Именно для таких мутаций разрабатывают более простые и эффективные методы молекулярной диагностики, позволяющие тестировать больных и проводить скрининг гетерозигот среди их родственников или среди определенных групп населения.

Важнейшей характеристикой мутантного аллеля является его корреляция с тяжестью течения заболевания, то есть фенотипическое выражение мутации в гомозиготном и в гетерозиготном состоянии, а также в компаунде с другими мутантными аллелями того же гена. Для многих моногенных болезней обнаружено большое число аллелей, которые по своему клиническому проявлению могут быть подразделены на различные группы, от очень мягких, оказывающих незначительный повреждающий эффект, до летальных или полублетальных, обуславливающих смерть пациентов в раннем возрасте. Таким образом, молекулярная диагностика мутаций может иметь определяющее значение для прогнозирования развития заболевания и выбора адекватной тактики лечения. Подобные скринирующие программы в настоящее время разработаны и успешно осуществляются во многих странах для таких распространенных заболеваний как серповидноклеточная анемия, муковисцидоз (Rene et al., 1994). Реализация этих программ наряду с большой медицинской пользой приносит и массу социальных проблем, некоторые из которых будут рассмотрены ниже.

ВВЕДЕНИЕ В МОЛЕКУЛЯРНУЮ ДИАГНОСТИКУ И ГЕНОТЕРАПИЮ НАСЛЕДСТВЕННЫХ БОЛЕЗНЕЙ.

В.Н.Горбунова, В.С.Баранов

ОГЛАВЛЕНИЕ

ВВЕДЕНИЕ.

ГЛАВА I. СТРУКТУРА И МЕТОДЫ АНАЛИЗА ДНК.

Раздел 1.1 Общие представления, центральная догма, генетический код.

Раздел 1.2 Выделение ДНК, ее синтез и рестрикция.

Раздел 1.3 Блот-гибридизация по Саузерну, гибридизация *in situ*.

Раздел 1.4 ДНК-зонды, клонирование, векторные системы.

Раздел 1.5 Геномные и к-ДНК-овые библиотеки генов, их скрининг.

Раздел 1.6 Секвенирование последовательностей ДНК.

Раздел 1.7 Полимеразная цепная реакция.

ГЛАВА II. ГЕНОМ ЧЕЛОВЕКА, СТРУКТУРА ГЕНОВ.

Раздел 2.1 Определение генома и его основных элементов.

Раздел 2.2 Повторяющиеся последовательности ДНК.

Раздел 2.3 Мультигенные семейства, псевдогены, онкогены.

Раздел 2.4 Современное определение понятия "ген",

- транскрипция, регуляторные элементы генов.
- Раздел 2.5 Изменчивость генома, полиморфные сайты рестрикции, ПДРФ-анализ.
- Раздел 2.6 Вариабильные микро- и минисателлитные ДНК.
- Раздел 2.7 Мобильность генома, облигатные и факультативные элементы генома.
- Раздел 2.8 Изохоры, метилирование, гиперчувствительные сайты.
- ГЛАВА III. ГЕНЕТИЧЕСКИЕ КАРТЫ, ПОЗИЦИОННОЕ КЛОНИРОВАНИЕ.
- Раздел 3.1 Классификация генетических карт, оценка сцепления.
- Раздел 3.2 Соматическая гибридизация, цитогенетический анализ, картирование анонимных последовательностей ДНК.
- Раздел 3.3 Генетические индексные маркеры.
- Раздел 3.4 Хромосом-специфические библиотеки генов, пульсирующий гель-электрофорез.
- Раздел 3.5 Позиционное клонирование, прогулка и прыжки по хромосоме, идентификация и изоляция генов.
- Раздел 3.6 Каталог генов и генных болезней МакКьюсика. Международная программа "Геном человека".
- ГЛАВА IV. ТИПЫ И НОМЕНКЛАТУРА МУТАЦИЙ. МЕТОДЫ ДНК-ДИАГНОСТИКИ.
- Раздел 4.1 Мутантные аллели, характеристика и типы мутаций.
- Раздел 4.2 Генетическая гетерогенность наследственных заболеваний.
- Раздел 4.3 Номенклатура мутаций.
- Раздел 4.4 Идентификация структурных мутаций, изоляция мутантных ДНК.
- Раздел 4.5 Первичная идентификация точечных мутаций.
- Раздел 4.6 Молекулярное сканирование известных мутаций.
- ГЛАВА V. ПОПУЛЯЦИОННЫЙ АНАЛИЗ МУТАЦИЙ. ЭНДОГЕННЫЕ МЕХАНИЗМЫ СПОНТАННОГО МУТАГЕНЕЗА.
- Раздел 5.1 Популяционный анализ мутаций, полиморфизм, неравновесность по сцеплению.
- Раздел 5.2 Частоты спонтанного мутагенеза.
- Раздел 5.3 Эндогенные механизмы возникновения мутаций.
- Раздел 5.4 Механизмы поддержания и распространения мутаций в популяциях.
- ГЛАВА VI. ЭКСПЕРИМЕНТАЛЬНЫЙ АНАЛИЗ ЭКСПРЕССИИ ГЕНОВ.
- Раздел 6.1 Дифференциальная активность генов, выбор адекватных биологических моделей.
- Раздел 6.2 Анализ регуляторных элементов гена, изоляция и исследование мРНК, искусственные транскрипционные системы.
- Раздел 6.3 Анализ трансляции, ДНК-экспрессионные системы.
- ГЛАВА VII. МОЛЕКУЛЯРНО-ГЕНЕТИЧЕСКИЕ АСПЕКТЫ ПРЕНАТАЛЬНОЙ ДИАГНОСТИКИ НАСЛЕДСТВЕННЫХ БОЛЕЗНЕЙ.

- Раздел 7.1 Прямые и косвенные методы молекулярной диагностики.
- Раздел 7.2 ДНК-диагностика при различных типах наследования.
- Раздел 7.3 Группы риска, поиск гетерозиготных носителей мутаций.
- Раздел 7.4 Особенности применения молекулярных методов в пренатальной диагностике моногенных болезней.
- Раздел 7.5 Доимплантационная диагностика, общая схема пренатальной диагностики, точность прогнозирования.

ГЛАВА VIII. БИОЛОГИЧЕСКИЕ МОДЕЛИ НАСЛЕДСТВЕННЫХ БОЛЕЗНЕЙ ЧЕЛОВЕКА.

- Раздел 8.1 Генетические линии животных.
- Раздел 8.2 Трансгенные животные.
- Раздел 8.3 Экспериментальное моделирование.
- Раздел 8.4 Конструирование модельных генетических линий животных.
- Раздел 8.5 Методы направленного переноса генов.

ГЛАВА IX. ГЕННАЯ ТЕРАПИЯ.

- Раздел 9.1 Определение, историческая справка, программы генной терапии.
- Раздел 9.2 Типы генотерапевтических вмешательств, выбор клеток-мишеней.
- Раздел 9.3 Методы генетической трансфекции в генной терапии.
- Раздел 9.4 Конструирование векторных систем и совершенствование методов трансформации клеток человека.
 - 9.4.1 Основные векторные системы.
 - 9.4.2 Методы физического переноса чужеродной ДНК в клетки эукариот.
 - 9.4.3 Липосомный метод трансфекции.
 - 9.4.4 Рекомбинантные вирусы.
 - 9.4.5 Перспективы создания "идеальных" векторных систем.
- Раздел 9.5 Генотерапия моногенных наследственных заболеваний.
- Раздел 9.6 Генотерапия ненаследственных заболеваний: опухоли, инфекции.
- Раздел 9.7 Некоторые этические и социальные проблемы генной терапии.

ГЛАВА X. МОЛЕКУЛЯРНО-ГЕНЕТИЧЕСКАЯ ХАРАКТЕРИСТИКА НЕКОТОРЫХ МОНОГЕННЫХ ЗАБОЛЕВАНИЙ.

- Раздел 10.1 Хромосомная локализация и принципы классификации генов наследственных болезней.
- Раздел 10.2 Метаболические дефекты лизосомных ферментов. Болезни накопления.
- Раздел 10.3 Болезни экспансии, вызванные "динамическими" мутациями.
- Раздел 10.4 Моногенные наследственные болезни, диагностируемые молекулярными методами в России.
 - 10.4.1 Муковисцидоз.

- 10.4.2 Миодистрофия Дюшенна.
- 10.4.3 Гемофилия А.
- 10.4.4 Гемофилия В.
- 10.4.5 Болезнь Виллебранда.
- 10.4.6 Фенилкетонурия.
- 10.4.7 Синдром Леш-Нихана.
- 10.4.8 Болезнь Вильсона-Коновалова.
- 10.4.9 Адено-генитальный синдром.
- 10.4.10 Спинальная мышечная атрофия.
- 10.4.11 Атаксия Фридрейха.

ЗАКЛЮЧЕНИЕ.

ГЛАВА V.

ПОПУЛЯЦИОННЫЙ АНАЛИЗ МУТАЦИЙ. МЕХАНИЗМЫ СПОНТАННОГО МУТАГЕНЕЗА.

Раздел 5.1 Полиморфизм, неравновесность по сцеплению.

Молекулярная идентификация мутантных аллелей и разработка эффективных методов их диагностики у больных и у гетерозиготных носителей являются той экспериментальной базой, которая позволяет исследовать распространенность мутаций в различных этнических группах и популяциях. Одновременно возможен анализ сцепления мутантных аллелей с другими генетическими маркерами и оценка на этой основе предполагаемых механизмов возникновения и поддержания в популяциях наблюдаемого уровня изменчивости. Но прежде чем перейти к описанию основных целей и методов популяционного анализа мутаций определим более точно два основных понятия, часто используемых при проведении подобных исследований - это понятия полиморфизма и неравновесности по сцеплению.

Генетическая изменчивость локуса в определенной популяции измеряется уровнем полиморфизма, и количественным выражением этой меры служат частоты аллелей. В однолокусной двухаллельной системе частоты двух вариантов аллелей - A1 и A2, обозначаются буквами p и q, соответственно, и $(p+q=1)$. Когда мы говорим, что в популяции существует полиморфизм по мутации, это значит, что ее частота выше определенного выбранного в соответствии с какими-то причинами условного уровня, чаще всего выше 5%. Высокополиморфными считаются локусы

с близкими значениями частот аллелей. В больших популяциях со случайным характером скрещивания, то есть в панмиктических популяциях, действует закон Харди-Вайнберга, согласно которому частоты гомозиготных особей в популяции равны p^2 и q^2 , соответственно, а доля гетерозигот равна $2pq$. Таким образом, в двухаллельной модели частота гетерозигот в популяции по полиморфным локусам составляет от 10% и более, а по высокополиморфным локусам может достигать 50%. Если число аллелей в локусе больше двух, общая частота гетерозигот в популяции равна $1 - (p_1^2 + p_2^2 + \dots + p_n^2)$, где p_k - частота аллеля A_k , $k = 1, 2, \dots, n$ и $p_1 + p_2 + \dots + p_n = 1$. При этом доля гетерозигот в популяции возрастает с увеличением числа аллелей и достигает максимального предела в каждой группе из n аллелей при равенстве их частот. Таким образом, при увеличении числа аллелей, встречающихся с равными вероятностями, частота гетерозигот в популяции будет стремиться к единице, то есть подовляющее число особей в популяции будут гетерозиготными по данному локусу. Такая ситуация характерна для высокополиморфных микросателлитных повторов, в частности STR, что и делает их наряду с другими преимуществами наиболее удобными генетическими маркерами.

Как мы уже упоминали, частыми моногенными наследственными заболеваниями считаются такие, при которых один больной ребенок встречается среди 2000 - 10000 новорожденных, то есть q^2 колеблется в пределах $1/2000 - 1/10000$. Соответственно, частота мутаций - q , в этих генах составляет от 0.5% до 2.5% и доля гетерозиготных носителей - $2pq$, достигает 1% - 5%. Для более редких заболеваний частоты гетерозигот не превышают десятых долей процента. Тем не менее, если учесть, что общее количество различных моногенных заболеваний составляет около 5000 нозологий, оказывается, что, в среднем, каждый человек является носителем мутантных аллелей около десяти генов. В общем случае, набор этих генов различен у разных людей, и только в тех семьях, где оба родителя являются носителями мутаций одного и того же гена появляется 25%-ый риск рождения больного ребенка.

До сих пор мы рассматривали изменчивость в одном локусе - А. Понятие неравновесности по сцеплению определяет отношения между изменчивостью в двух локусах - А и В. Ситуация здесь может быть двоякой. Если изменчивость в этих локусах независима, то аллели двух локусов встречаются в популяции в случайных комбинациях. Однако, если аллели различных локусов в одних комбинациях встречаются чаще, чем в других, то говорят, что существует неравновесность по сцеплению. Количественно эта связь оценивается с помощью детерминанта неравновесности по сцеплению - d . Рассмотрим, чему равен детерминант неравновесности по сцеплению в двухлокусной двухаллельной системе. Пусть P_{nm} , где $n=1,2$ и $m=1,2$, частоты четырех возможных в этом случае типов гамет - $A_1B_1, A_2B_2, A_1B_2, A_2B_1$, а P_n и R_m - частоты аллелей локусов А и В, соответственно. Если сочетания аллелей в гаметах случайны, то теоретически ожидаемые частоты гамет четырех типов равны произведению частот входящих в эти гаметы аллелей, то есть $P_{nm} = P_n \cdot R_m$. В этом случае произведение частот двух типов гамет (A_1B_1 и A_2B_2), находящихся в состоянии "притяжения", равно произведению частот гамет в состоянии "отталкивания" (A_1B_2 и A_2B_1), то есть $P_{11} \cdot P_{22} = P_1 \cdot P_2 \cdot R_1 \cdot R_2 = P_{12} \cdot P_{21}$. Однако, если сочетания аллелей в гаметах неслучайны, то эти произведения различны. Их разность служит мерой неравновесности по сцеплению - d . Итак $d = |P_{11} \cdot P_{22} - P_{12} \cdot P_{21}|$. При отсутствии неравновесности по сцеплению $d = 0$. Верхняя гра-

ница $d = 0.25$. Максимальная неравновесность по сцеплению достигается при полном отсутствии двух типов гамет, находящихся либо в состоянии "притяжения", либо в состоянии "отталкивания", при одновременном равенстве частот оставшихся двух типов гамет.

Мутации в различных локусах возникают, как правило, независимо друг от друга и наличие неравновесности по сцеплению, по-видимому, отражает тот факт, что мутация в одном из локусов возникла в хромосоме, несущей определенный аллель другого локуса, и далее эта хромосома получила распространение в популяции. Неравновесность по сцеплению является очень важной популяционной характеристикой, позволяющей судить о порядке и примерном времени возникновения различных мутаций, а также оценивать возможные механизмы их поддержания в популяции. Обнаружение сильной неравновесности по сцеплению между специфическими мутациями гена и определенными аллелями маркерных локусов имеет важное практическое значение. Часто наблюдается неравновесность по сцеплению между мутантными аллелями гена и одновременно несколькими маркерными локусами. В этом случае анализ маркерных гаплотипов, то есть наборов аллелей различных локусов, локализованных в одной хромосоме, дает возможность с высокой степенью вероятности оценивать характер мутационного повреждения и проследивать его наследование в семьях больного.

Раздел 5.3 Частоты спонтанного мутагенеза.

Считается, что средняя частота спонтанного возникновения мутаций в структурных локусах человека колеблется в пределах от $10(-5)$ до $10(-6)$ на одну гамету за каждое поколение. Однако, эта величина может значительно варьировать для разных генов, меняясь в пределах от $10(-4)$ для высокомутабильных локусов до $10(-11)$ в наиболее устойчивых частях генома. Эти различия зависят от многих факторов и, в первую очередь, от характера мутационного повреждения, от механизма возникновения мутации и локализации нарушения. Большое значение также имеет сам ген, протяженность его кодирующих областей и те функции, которые выполняют контролируемые им молекулы в обеспечении жизнедеятельности клеток и всего организма, в целом. Так например, нарушение работы генов, продукция которых необходима на ранних стадиях эмбриогенеза, может приводить к гибели плода. Такие мутации трудны для диагностики и в практической медицине мы чаще всего имеем дело только с теми мутациями, которые не проявляют летального эффекта на ранних стадиях эмбрионального развития. Тем ни менее, не исключено, что ранние эмбриональные летали составляют немалый процент среди мутантных аллелей различных генов и вносят определенный вклад в снижение репродуктивной функции.

Особого внимания заслуживает проблема мутирования в STR-сайтах и в VNTR- локусах, часто используемых в настоящее время для геномной дактилоскопии и молекулярной диагностики. Эти участки генома, изменчивость в которых обусловлена различиями в числе тандемных повторов, формально можно отнести к разряду высокомутабильных. Заметим сразу, однако, что прямое сопоставление темпов мутирования в кодирующих областях генома и в мини- и микросателлитных последовательностях ДНК, по-видимому, некорректно, так как физическая основа изменчивости, наблюдаемой в этих функционально и структурно различающихся локусах совершенно различна. Мутабельность в STR-сайтах обусловлена нестабильностью числа повторов, причем возникающие мутации затрагивают, как правило, лишь весь-

ма ограниченное число кластеризованных копий, чаще всего 1 или 2. Подобные тандемные повторы редко наблюдаются в смысловых последовательностях ДНК. Исключение составляют лишь мутации экспансии, но и для них характерно (1) существование большого числа аллелей дикого типа, отличающихся друг от друга небольшим числом копий, и (2) значительные различия по длине повторяющегося участка между нормальными и мутантными вариантами гена. Кроме того, изменчивость в STR-сайтах в основе своей носит нейтральный характер и поэтому темп мутирования в этих локусах не подвержен жесткому контролю со стороны естественного отбора.

Прямые исследования показали, что в большинстве случаев частота возникновения спонтанных мутаций в микросателлитных STR-сайтах варьирует в пределах от 0.1% до 0.45% на гамету, что должно учитываться при использовании этих маркеров в практической медицине. Частота мутирования в вариабильных (C-A)_n-повторах, используемых в качестве индексных маркеров в Genethon - картах сцепления составляет 0.05% на гамету. Показано, что для ряда VNTR-локусов (MSB2) частота мутирования достигает 0.7% на гамету. Для других локусов (M17) обнаружено достоверное превышение скорости мутагенеза в зародышевых клетках по сравнению с соматическими. Для многих достаточно стабильных STR-локусов обнаружены существенные межпопуляционные различия по частоте аллелей, что позволяет использовать эту изменчивость для генетической характеристики отдельных популяций. В то же время другие STR-сайты значительно чаще подвергаются спонтанному мутированию, что приводит к уравниванию паттерна аллелей и делает эти маркеры малопригодными для популяционных исследований. Имеются данные, что в наиболее вариабильных STR-локусах частота спонтанных мутаций может достигать 5% на гамету за поколение (Jeffreys et al., 1988).

Раздел 5.3. Эндогенные механизмы возникновения мутаций.

Основную часть мутаций, ведущих к наследственным болезням, составляют точечные мутации, делеции и в меньшей степени инсерции и дупликации. При этом, как показывает детальный сравнительный анализ, частота, тип и локализация этих мутаций отнюдь неслучайны и зависят от многих, пока еще невыясненных эндогенных механизмов мутагенеза. В пользу такого вывода свидетельствует уникальный характер спектра мутаций для каждого из многих десятков генов, первичная структура ДНК и типы мутаций которых уже хорошо изучены. Так, для многих структурных генов доминирующими по спектру и частоте являются точечные мутации (ген трансмембранного регуляторного белка муковисцидоза, ген фенилаланингидроксилазы, ген фактора IX свертывания крови, бета-глобиновый ген и мн др.), тогда как для других - достаточно протяженные структурные перестройки типа делеций, дупликаций и инсерций (гены дистрофина, фактора VIII свертывания крови, 21-гидроксилазы) (см. Главу X). К структурным факторам, определяющим эндогенные механизмы мутагенеза, можно отнести (1) наличие прямых и обратных повторов и симметричных элементов вблизи места перестройки; (2) высокую концентрацию CpG динуклеотидов; (3) наличие внегенных последовательностей ДНК, гомологичных фрагментам структурного гена; (4) мобильные элементы генома. Естественно, что реализация этих структурных факторов в те или иные типы мутаций возможна лишь в процессе репликации (1,2) и рекомбинации (3) ДНК хромосом.

Как подробно рассмотрено в серии работ D.N Cooper и

М. Krawczak (1990, 1991), наличие в первичной структуре ДНК прямых повторов, идентичных повторяющихся последовательностей, инвертированных повторов, шпилечных структур, квази-палиндромных последовательностей и симметричных элементов генома (например CTGAAGTC) нередко ведет к образованию петель при репликации ДНК вследствие скользящего нарушения спаривания (slipping mispairing) родительской и дочерней цепей ДНК. Эти новые структурные элементы ДНК либо уничтожаются ферментами системы репарации, что ведет к делециям, либо сохраняются и дублируются, что приводит к дупликациям и инсерциям, при этом возникшие изменения закрепляются при последующих раундах репликации. Авторы приходят к следующим выводам: (1) возникновение подобных мутаций происходит неслучайно, но зависит от особенностей первичной структуры ДНК в месте перестройки; (2) в основе структурных перестроек ДНК лежат ошибки репликации; (3) принципиально сходные модели эндогенного мутагенеза характерны как для делеций, так и для инсерций. Считается, что именно механизм скользящего нарушения спаривания ответственен за мутации экспансии, приводящие к быстрому увеличению числа тринуклеотидных повторов и к нарушению работы соответствующих генов, а также за высокую изменчивость, наблюдаемую во многих местах локализации мини- и микросателлитных tandemных повторов.

Недавно показано, что повышенной эндогенной мутагенностью обладают вообще все последовательности ДНК, находящиеся в определенном конформационном состоянии, а именно в состоянии изгиба (bent DNA) (Milot et al., 1992). Известно, что такая конформационная структура ДНК свойственна промоторным частям генов, местам начала репликации (origins of replication), местам контакта хромосом с ядерным матриксом. Именно эти участки ДНК являются местами посадки ферментов топоизомераз, вовлеченных в процессы репликации, транскрипции, рекомбинации, в том числе, как оказалось, и в процесс негомологичной (незаконной -illegitimate) рекомбинации. Установлено, что именно негомологичная рекомбинация может приводить не только к внутригенным делециям, дупликациям и другим мутациям на молекулярном уровне, но и является одной из основных причин крупных структурных хромосомных перестроек типа транслокаций, инверсий и других.

Замены или утраты отдельных оснований в геномной ДНК могут возникать в результате нарушения процессов репликации и репарации. Ошибки в ДНК матрице, вызванные действием повреждающих внешних агентов, либо спонтанной деградацией оснований закрепляются в последующих циклах репликации. Основные типы спонтанной деградации включают потерю оснований и процесс дезаминирования. Особенно чувствительны к дезаминированию цитозиновые остатки. Установлено, что у позвоночных почти половина всех цитозиновых остатков в ДНК метилирована в 5-ом положении. Процесс метилирования особенно часто захватывает области повторов 5' CpG 3', расположенные как внутри генов, так и в их промоторных частях. При дезаминировании 5-метилцитозин превращается в тимин. В цикле последующей репликации, возникший в результате дезаминирования мисмэтч (T-G) может либо коррегироваться в нормальный вариант (C-G), либо приводить к мутациям типа трансцизий (T-G) или (C-A). Естественно, что гены, имеющие в своей структуре большой процент CpG оснований, особенно часто подвергаются спонтанному мутированию типа трансцизий. В частности, преобладание подобных точечных мутаций известно для генов факторов IX и VIII свертывания крови, для гена фенилкетонурии и других. Так, из 76 мутаций гена фактора IX в 21 случае найдены

транскрипции CpG - TpG или CpA (Green et al., 1990). Преобладание таких мутаций отмечено и в 22 CpG дуплетах гена фенилаланингидроксилазы у больных фенилкетонурией (Abadie et al., 1989).

Другим важным фактором эндогенного мутагенеза является наличие тесно сцепленных с генами гомологичных последовательностей ДНК (псевдогенов). В мейозе такая ситуация нередко приводит к неравной гомологичной рекомбинации и, как следствие этого, к генной конверсии, сопровождающейся структурными перестройками типа делеций, дупликаций и т.п. Подобный механизм мутаций, как оказалось, является доминирующим для гена 21-гидроксилазы (Morel, Miller, 1991), а также для гена фактора VIII свертывания крови (Lakich et al., 1993). Важная роль ошибок рекомбинации в этиологии структурных поломок гена особенно очевидна при анализе гена дистрофина, мутации которого ведут к миопатии Дюшенна. Известно, что в 60% случаев мутации этого гена представляют собой делеции, захватывающие один или несколько соседних экзонов. Известно также, что подавляющее большинство делеций сосредоточено в двух "горячих" районах этого гигантского по размерам гена (2,2 Мб), и при этом частота внутригенных рекомбинаций почти в 4 раза больше, чем можно предполагать, исходя из его размеров (Oudet et al., 1992). Любопытно отметить, что в одной из этих горячих точек (интрон 7) недавно обнаружен кластер транспозоноподобных повторяющихся последовательностей. Удельный вес мобильных (транспозоноподобных) элементов типа Alu и Line повторов (см. Главу 2) в возникновении генных мутаций до конца не выяснен. Имеются единичные наблюдения о реальном перемещении этих элементов по типу конверсии и их интеграции в структурные гены аденозиндезаминазы, аполиipoproteина С, факторов VIII и IX свертывания крови, кальмодулина (Vidaud et al., 1993).

Раздел 5.4 Механизмы поддержания и распространения мутаций в популяциях.

Частоты и характер распределения мутаций в популяциях зависят от многих факторов, главными из которых являются частоты мутагенеза и давление естественного отбора. Значительное влияние на этот процесс оказывают также структурные особенности популяций, такие как размеры, степень географической и этнической изолированности, величина инбридинга, характер миграции населения.

Для всех мутаций, возникающих за счет повышенного уровня спонтанного мутагенеза, характерны следующие особенности - случайный характер внутригенной локализации мутаций, сходство типов нарушений при отсутствии полной молекулярной идентичности между ними. В отличие от спонтанных мутаций, вызванных эндогенными причинами, для мутаций, индуцированных действием неблагоприятных факторов внешней среды, промышленными и сельскохозяйственными вредностями, ионизирующим облучением, химическими агентами и прочим, специфика в типах мутаций и в характере их локализации не наблюдается. В популяциях, находящихся в области действия таких неблагоприятных факторов, будет повышена частота мутаций в различных генах, однако спектр индуцированных мутаций будет достаточно разнообразным.

Рассмотрим теперь влияние отбора на процесс поддержания и распространения мутаций в популяциях. Многие гены мутационно наследственных заболеваний рецессивны, то есть мутации в них в гетерозиготном состоянии не оказывают вредного

влияния на жизнеспособность. Поэтому после возникновения мутация может распространяться в популяции до определенной концентрации, практически не подвергаясь элиминирующему действию естественного отбора. В дальнейшем частота этой мутации достигнет равновесного состояния и не будет повышаться за счет выщепления гомозиготных особей, жизнеспособность и репродуктивные качества которых резко снижены. При этом скорость элиминации мутации из популяции резко замедляется при снижении ее частоты и, практически, после возникновения мутация может сохраняться в популяции на протяжении многих десятков и даже сотен поколений. Различные мутации могут случайным образом получить большее распространение в изолированных популяциях или среди групп населения, отличающихся повышенным уровнем инбридинга. В целом, при отсутствии давления отбора по отношению к гетерозиготным особям общая концентрация мутантных аллелей в популяции определяется частотой их спонтанного возникновения, при этом пул мутаций будет состоять из большого количества разнообразных аллелей, каждый из которых будет представлен редкими или даже единичными случаями в различных популяциях.

Однако, специфические мутации могут получить значительно более широкое распространение в тех случаях, когда гетерозиготные особи имеют какие-либо селективные преимущества. Таким эффектом может обладать сама мутация, но более вероятно возможность неравновесности по сцеплению между этой мутацией и селективными аллелями другого локуса. Гетерозиготы могут получить преимущество при изменении условий окружающей среды, в каких-то экстремальных ситуациях или среди определенных групп населения. Так например, мутации, повышающие устойчивость организма к действию инфекционных агентов, могут получить широкое распространение в период массовых эпидемий. Одновременно повысится частота всех аллелей других локусов, находящихся в неравновесности по сцеплению с данной мутацией. Мутантные аллели, обеспечивающие селективное преимущество гетерозигот, становятся преобладающими во многих популяциях, не полностью изолированных друг от друга. При этом наибольшая частота таких аллелей будет наблюдаться в районах, где влияние поддерживающего отбора было максимальным (например, в эпицентре эпидемии). По мере удаления от этого района концентрация таких мутантных аллелей будет уменьшаться, причем их распределение в разных популяциях будет коррелировать с характером миграции населения. Подобный характер распределения определенного мутантного аллеля в частично изолированных популяциях принято связывать с так называемым эффектом основателя или родоначальника.

Исследование спектров распределения мутаций в различных популяциях позволяет делать предположения относительно возможного происхождения таких повреждений и тех механизмов, которые лежат в основе их распространения среди населения.

Рассмотрим наиболее вероятные интерпретации различных вариантов распределения аллелей в популяциях. Мутации, представленные у единичных больных или в группе родственных индивидуумов и не имеющие специфической внутригенной локализации, по-видимому, являются следствием естественного мутационного процесса. Если в каких-то популяциях концентрация мутаций в различных генах повышена, вероятно, они находятся в зоне действия внешних неблагоприятных факторов, индуцирующих возникновение нарушений в структуре ДНК. В тех случаях, когда локализация и типы мутаций носят специфический характер можно предполагать наличие особых молекулярных механизмов контроля повышенного уровня мутагенеза в определенных

районах генома. Распространение специфических мутаций в изолированных популяциях происходит за счет их ограниченного размера и повышенного уровня инбридинга (эффект родоначальника). И, наконец, обнаружение градиентного распределения мутаций, преобладающих в различных, частично изолированных популяциях позволяет предполагать селективное преимущество гетерозиготных носителей мутаций на определенных этапах эволюционного развития.

Таким образом, сопоставляя спектры распределения однотипных мутаций у жителей разных континентов, разных стран, у людей, принадлежащих к различным расам и национальностям можно определить степень генетической близости между всеми этими группами и реконструировать их филогенетические отношения (Cavalli-Sforza, Piazza, 1993). Одним из практических следствий этих исследований является возможность прогнозировать наиболее вероятные мутации в различных генах у пациентов разного этнического происхождения, что приводит к сужению спектра поиска специфических мутаций. Особый интерес в этом смысле представляют наиболее распространенные мутации (например delF508 при муковисцидозе; R408W - при фенилкетонурии и многие другие). Для профилактики наследственных заболеваний необходима разработка эффективных и простых методов молекулярной диагностики таких мутаций как у больных, так и у гетерозиготных носителей с целью проведения скринирующих программ среди населения и выявления максимально возможного числа семей с повышенным риском рождения больного ребенка.

ГЛАВА VII.

МОЛЕКУЛЯРНО-ГЕНЕТИЧЕСКИЕ АСПЕКТЫ ПРЕНАТАЛЬНОЙ ДИАГНОСТИКИ НАСЛЕДСТВЕННЫХ ЗАБОЛЕВАНИЙ.

Раздел 7.1 Прямые и косвенные методы молекулярной диаг-

ностики.

Локализация и клонирование кДНК-овых последовательностей генов открывают принципиально новые возможности диагностики наследственных заболеваний, основанные на исследовании мутантных аллелей у пациентов, членов их семей или у предполагаемых гетерозиготных носителей патологических мутаций. Это в равной мере относится и к пренатальной диагностике, которая может быть проведена с использованием молекулярных методов анализа на самых ранних стадиях развития плода (см.7.5). Эти же подходы вполне приемлемы для диагностики до появления каких-либо клинических или биохимических симптомов болезни (досимптоматическая диагностика), что позволяет выработать и начать рациональную тактику лечения (упреждающая терапия), а также эффективно выявлять гетерозиготных носителей в семьях высокого риска, что является важным фактором профилактики наследственных болезней. Решающими преимуществами молекулярной диагностики являются её универсальность, возможность использовать для анализа любые ДН-содержащие клетки или ткани, причем анализ может быть произведен на любых стадиях онтогенеза, начиная со стадии зиготы.

Принципиально различают прямую и непрямую ДНК-диагностику моногенных наследственных болезней. В общем случае, использование прямых методов диагностики возможно лишь для клонированных генов с известной нуклеотидной последовательностью полноразмерной кДНК, при этом необходимо предварительное генотипирование мутантных аллелей у родителей. В случае прямой диагностики объектом молекулярного анализа является сам ген, точнее мутации этого гена, идентификация которых и составляет основную задачу исследования. Такой подход особенно эффективен при наличии точной информации о природе, частоте и локализации наиболее распространенных (доминирующих по частоте) мутаций соответствующих генов, а также о наличии в них особенно легко мутирующих "горячих" точек. К таковым относятся мутация delF508 и ряд других мутаций при муковисцидозе, делеционные мутации при миодистрофии Дюшенна, мутация R408W при фенилкетонурии, инверсионная мутация при гемофилии А, протяженная делеция при адено-генитальном синдроме, экспансии триплетных повторов в случае "динамических" мутаций при синдроме ломкой X-хромосомы и при ряде других нейродегенеративных заболеваний (см. Главы IV и X). Методы, используемые для направленного поиска этих мутаций, подробно рассмотрены в Главе IV. В ряде случаев (муковисцидоз, фенилкетонурия, серповидно-клеточная анемия) эти методы удалось автоматизировать, что позволяет одновременно тестировать сразу несколько (до 30 и более) различных мутаций. При этом появляется реальная возможность выявлять свыше 95-98% мутантных хромосом, что делает целесообразным и экономически оправданным скринирование всей популяции отдельных стран на выявление мутантных особей для последующей организации эффективных профилактических мероприятий, направленных на предупреждение рождения больных детей. Подобные программы по муковисцидозу уже успешно проводятся в ряде стран Западной Европы (Великобритания, Дания, Франция) и Северной Америки.

Главное преимущество прямого метода - это высокая, достигающая до 100%, точность диагностики и отсутствие необходимости анализа всей семьи на предмет её информативности (см.ниже). Последнее обстоятельство особенно важно для проведения пренатальной диагностики тяжелых, зачастую летальных наследственных болезней (муковисцидоз, миодистрофия Дюшен-

на, гемофилия А и др). Такие семьи нередко обращаются за медико-генетической помощью уже после того как больной ребенок умер. Так, по нашим наблюдениям до 80% семей с высоким риском муковисцидоза обращаются по поводу необходимости дородовой диагностики уже после смерти больного ребенка (Baranov et al., 1992).

Однако, существует огромное количество наследственных болезней, для которых мутации не описаны либо не найдено мажорных мутаций в исследуемых популяциях. И даже во всех тех случаях, когда имеются мажорные мутации, наряду с ними, описаны многочисленные редко встречающиеся (вплоть до единичных случаев), так называемые минорные мутации. Кроме того, всегда сохраняется возможность присутствия у пробанда неизвестных мутаций, а клонирование гена больного человека для целей прямого секвенирования, даже ограниченного только смысловой его частью – кДНК, далеко не всегда возможно в силу очевидных финансовых и временных ограничений такого подхода. Эти трудности успешно преодолеваются благодаря наличию не прямых (косвенных) методов молекулярной диагностики.

Этот исторически более ранний подход основан на использовании сцепленных с геном полиморфных маркеров, с помощью которых проводится идентификация мутантных хромосом (точнее хромосом, несущих мутантный ген) в семьях высокого риска, то есть у родителей больного и его ближайших родственников. В настоящее время косвенные методы молекулярной диагностики принципиально возможны практически для всех моногенных заболеваний с известной локализацией контролирующего гена, для каждого из которых уже разработана удобная система вне- и внутригенных полиморфных индексных маркеров (см. Главу III).

Более того, косвенные методы молекулярной диагностики пригодны даже для тех болезней, гены которых еще не идентифицированы и мутации не известны. Единственным и не переменным условием этого является наличие полиморфных сайтов рестрикции либо коротких tandemных повторов типа STR, находящихся в непосредственной близости от мутантного гена или, что еще лучше, внутри него (чаще всего в интронах). При помощи этих полиморфных сайтов удастся маркировать мутантные аллели гена и проследить их передачу потомству (см. Главу 111). Ранние исследования непрямым методом проводились почти исключительно с использованием полиморфных сайтов рестрикции – двухаллельной системы, информативная емкость которой не превышает 50%, а реальное число гетерозигот по данному признаку в популяции, зачастую, оказывается существенно ниже 0.5. Следовательно, в лучшем случае только половина гетерозиготных носителей наследственного заболевания, вызванного рецессивной мутацией какого-нибудь гена, может быть доступна не прямой молекулярной диагностике с использованием одного полиморфного сайта рестрикции. Повышение информативности в случае ПДРФ-анализа могло быть достигнуто только путем увеличения числа полиморфных сайтов. Как правило, для диагностики необходимо не один, а 3-4 полиморфных сайта, что далеко не во всех случаях возможно. Многие из полиморфных сайтов локализованы вне генов на расстояниях, при которых кроссинговер может в заметном проценте случаев исказить результаты диагностики. Кроме того, практическое использование рестрикционных сайтов зачастую затруднено из-за отсутствия или большой стоимостью соответствующих эндонуклеазных рестриктаз.

Все эти недостатки могут быть устранены при использовании в качестве молекулярных маркеров высокополиморфных tandemно повторяющихся коротких три- и тетрамерных повторов (STR) (см. Главу III). Для многих генов найдены уникальные

паттерны полиморфных аллелей, отличающихся по числу "коротких" единичных повторяющихся последовательностей нуклеотидов. Такие "количественные" полиморфизмы, как уже упоминалось (см. Главу III), очень широко распространены по всему геному и присутствуют в интронных и фланкирующих областях многих генов. Появление в конце 1994г 0.7-сантиморганной карты генома человека, построенной на базе высокополиморфных динуклеотидных (С-А)_n повторов, сделало реальным маркирование, практически, любого картированного гена. Особенную диагностическую ценность представляют внутригенные маркеры, для которых резко снижена вероятность кроссинговера с мутантными аллелями гена, а следовательно, особенно высока точность диагностики. Кроме того, как оказалось, многие внутригенные полиморфные короткие tandemные повторы обнаруживают сильное неравновесие по сцеплению с определенными мутантными аллелями гена, что значительно облегчает их идентификацию в отягощенных семьях. Индекс гетерозиготности таких полиаллельных мини- и микросателлитных систем нередко превышает 0.8. Их применение позволяет маркировать, то есть сделать информативными (см. 7.4) для ДНК-диагностики практически все семьи высокого риска при условии наличия больного ребенка или доступности для молекулярного анализа его патанатомического материала.

Изучение полиморфных маркеров у больного пробанда и его родителей и выяснение аллельной природы молекулярного маркера, так называемое "определение фазы сцепления" (см. раздел 7.4), составляет основу для дальнейшей диагностики косвенными методами (Евграфов, Макаров, 1987). Применение косвенных методов молекулярной диагностики предусматривает также в качестве обязательного предварительного этапа исследование частоты аллелей соответствующих полиморфных сайтов в анализируемых популяциях, среди больных и гетерозиготных носителей мутаций, а также определение вероятности рекомбинации и неравновесности по сцеплению между маркерными сайтами и мутантными аллелями гена. Такие исследования проводятся с помощью методов блот-гибридизации по Саузерну, либо ПЦР (см. разделы 1.3.; 1.7; 2,5). В первом случае в распоряжении исследователя должны быть соответствующие ДНК-зонды, во втором - должны быть известны нуклеотидные последовательности районов ДНК, включающие соответствующие полиморфные сайты, для выбора олигопраймеров (см. Главы II, IV).

Раздел 7.2. ДНК-диагностика при различных типах наследования.

Напомним, что значительное число моногенных заболеваний наследуется по рецессивному типу. Это значит, что при аутосомной локализации соответствующего гена болеют только гомозиготные носители мутаций. Гетерозиготы клинически здоровы, но с равной вероятностью передают своим детям мутантный или нормальный варианты гена. Таким образом, на протяжении длительного времени мутация в скрытом виде может передаваться из поколения в поколение. При аутосомно-рецессивном типе наследования в родословных тяжелых больных, которые либо не доживают до репродуктивного возраста, либо имеют резко сниженные потенции к размножению, редко удается выявить больных родственников, особенно, по восходящей линии. Исключения составляют семьи с повышенным уровнем инбридинга, который возникает либо за счет высокой частоты близкородственных браков, либо за счет вступления в брак людей из одинаковых изолированных популяций ограниченной численности. Больные

дети с вероятностью 25 % рождаются в тех семьях, где оба родителя являются гетерозиготными носителями мутаций одного и того же гена. Возможно также рождение больного ребенка в такой семье, где только один из супругов несет мутацию, а вторая мутация возникла в гамете его партнера в момент, предшествующий оплодотворению. Доля таких семей в общей группе риска относительно невелика, а риск повторного рождения в них больного ребенка не превышает общей частоты спонтанного возникновения мутаций в данном гене. Для болезней, сцепленных с полом, то есть контролируемых генами, локализованными в X-хромосоме, характерно то, что болеют преимущественно мальчики, тогда как носителями являются девочки.

Y-хромосома содержит очень мало генов, большинство из которых (около 10) картировано в так называемой псевдоаутосомной области короткого плеча, гомологичной таковой на коротком плече X-хромосомы. Важнейшим истинным геном Y-хромосомы, то есть геном, представленным только на этой хромосоме, является ген SRY (Yp21.1), детерминирующий развитие пола по мужскому типу. Мутации этого регуляторного гена приводят к нарушениям половой дифференцировки (XY-женщины), а его перенос вследствие ошибок рекомбинации псевдоаутосомных районов на короткое плечо X-хромосомы обуславливает синдром реверсии пола - Sex reverse (XX-мужчины) (McElreavey et al., 1993). Практически важно, что присутствие даже небольшого околоцентромерного фрагмента длинного плеча Y-хромосомы при кариотипе XX является безусловным показанием для удаления зачатков гонад у таких индивидуумов в связи с высокой вероятностью экспрессии гена Gba, ведущего к их злокачественному перерождению - гонадобластоме (Giquel et al., 1992). В отличие от Y-хромосомы, большая по размеру X-хромосома человека несет до 5% всех структурных генов, многие из которых уже идентифицированы (см. Главу III).

Все рецессивные аллели X-хромосомы у мальчиков проявляются, так как находятся в гемизиготном состоянии. Девочки могут болеть в том случае, если они гомозиготны по мутации. Такая возможность может осуществиться в семье, где болен их отец, а мать является носителем мутации и передала дочери свой мутантный аллель. Если отец больной девочки здоров, можно предполагать, что мутация возникла в той его гамете, которая участвовала в оплодотворении. X-сцепленные заболевания у девочек (миодистрофия Дюшенна, гемофилия А) могут быть следствием сочетанного проявления мутации этих генов у одного из родителей и делеции соответствующих фрагментов хромосом у другого. В этих редких случаях у девочек, как и у мальчиков, мутации X-сцепленных генов будут находиться в гемизиготном состоянии.

При доминантном наследовании для развития болезни достаточно одного мутантного аллеля. Такие больные с вероятностью 50% рождаются в семьях, где один из родителей болен. Очень редко больные дети с доминантным типом наследования могут родиться и у здоровых родителей в результате мутирования одной из гамет. Однако, вероятность повторного рождения больного ребенка в такой семье такая же, как и для популяции в целом. Пренатальная диагностика доминантных болезней проводится достаточно редко по ряду причин. Во-первых, такие болезни составляют относительно небольшой процент среди всех моногенных заболеваний. Во-вторых, тяжелые болезни, сопровождающиеся летальным исходом в раннем возрасте или приводящие к бесплодию, не передаются по наследству, а появляются каждый раз заново вследствие мутации при созревании гамет. Необходимость же предотвращения рождения больных, которые не

только доживают до репродуктивного возраста, но и способны оставить потомство, является вопросом дискуссионным. Проведение пренатальной диагностики таких заболеваний принимается с учетом многих конкретных обстоятельств. Актуальность этой проблемы стала особенно очевидной в последние годы, когда была открыта многочисленная группа доминантных нейродегенеративных заболеваний (см. Главу IV), которые проявляются сравнительно поздно, нередко уже в репродуктивном возрасте, тяжело протекают и, по-сути, не имеют сколько-нибудь реальной терапии. Для таких заболеваний первостепенное значение приобретают методы досимптоматической диагностики с использованием методов ДНК-анализа (см. Главу X).

Значительно больше распространены моногенные болезни с частичным доминированием и неполной пенетрантностью. Для подобных заболеваний риск рождения больного ребенка в отягощенной семье зависит от конкретных значений этих параметров, которые, в свою очередь, определяются молекулярными механизмами, лежащими в основе формирования такого рода отклонений от Менделевского типа наследования.

Разработка молекулярных методов диагностики болезней, вызванных мутациями в митохондриальных генах, как показывают исследования последних лет (McKusick, 1994), приобретает особое значение. Как правило, в основе различных митохондриальных болезней лежат нарушения в системе окислительного фосфорилирования. Поскольку некоторые ткани обладают повышенной чувствительностью к подобным нарушениям, сходная клиническая картина заболеваний может наблюдаться вследствие мутаций разных митохондриальных генов. Наследование таких болезней значительно отличается от Менделевского типа, в первую очередь, из-за материнского характера наследования митохондрий, наличия в зиготе и соответственно во всех клетках организма большого числа копий митохондриальных хромосом, из-за особенностей сегрегации этих хромосом при делении клеток и, наконец, из-за тех количественных и качественных изменений в митохондриальной ДНК, которые сопровождают процессы онтогенетической дифференцировки клеток и старения организма.

Для всех "митохондриальных болезней" характерен материнский тип наследования и присутствие мутантных аллелей только в части хромосом (так называемая гетероплазмия), доля которых может варьировать в разных тканях. Как правило, существует определенное пороговое значение доли мутантных хромосом, превышение которого ведет к появлению и прогрессированию заболевания. Сокращение числа митохондрий, происходящее в норме при старении, может способствовать увеличению доли мутантных хромосом в определенных тканях и тем самым быть причиной болезни. Неслучайно поэтому, что для многих "митохондриальных болезней", характерно позднее начало. В силу функциональных особенностей митохондриальных генов часто наблюдается кумулятивный эффект двух и более мутаций, локализованных в разных генах. Важным представляется также то обстоятельство, что в митохондриальных генах OXPHOS-комплекса частота возникновения мутаций выше, чем в ядерных генах, кодирующих субъединицы окислительного фосфорилирования (McKusick, 1994).

Раздел 7.3. Группы риска, поиск гетерозиготных носителей мутаций.

Как следует из вышеизложенного, точная диагностика в сочетании с детальным анализом типа наследования того или

иною заболеванием имеет определяющее значение для формирования групп риска, то есть отбора семей, в которых вероятность рождения больных детей повышена. Прежде всего, это те семьи, где уже есть или был ребенок, страдающий каким-либо моногенным наследственным заболеванием. Для аутосомно-рецессивных болезней с большой долей вероятности можно считать, что оба родителя этого ребенка являются гетерозиготными носителями мутантных аллелей соответствующего гена и риск повторного рождения больного ребенка в такой семье составляет 25%, независимо от исхода предыдущих родов. Поэтому в таких случаях рекомендуется обязательная пренатальная диагностика плода при каждой последующей беременности.

Для сцепленных с полом заболеваний важной практической задачей является выявление случаев спонтанного возникновения мутаций в родительском поколении. Для таких распространенных заболеваний, как миодистрофия Дюшенна и гемофилия А, почти 1/3 всех случаев имеет спонтанное происхождение. При этом мутации гена дистрофина, как правило, возникают в оогенезе, то есть у матери, а мутации гена фактора Y111, обычно, появляются во время сперматогенеза у деда больного ребенка (см. Главу X). В отличие от тех семей, в которых мать гетерозиготна, вероятность повторного рождения больного ребенка в семьях со спонтанной мутацией не превышает среднепопуляционную частоту, и потому нет необходимости проводить пренатальную диагностику плода при последующих беременностях. Специального рассмотрения в этой связи заслуживает, однако, вопрос гонадного мозаицизма, то есть наличия в гонадах матери генетически нормальных и мутантных ооцитов. Особенно велик риск такого состояния в случае миодистрофии Дюшенна. Гонадный мозаицизм в силу своей органной специфики достаточно трудно доказать или опровергнуть. Между тем, считается что 6,7% спорадических случаев иодистрофии Дюшенна обусловлена гонадным мозаицизмом у матери (Essen et al., 1992).

Подробное медико-генетическое консультирование семей, в которых зарегистрированы спонтанные случаи рождения детей с X-сцепленными заболеваниями в сочетании с соответствующими лабораторными, в том числе и молекулярными исследованиями, как правило, позволяют ответить на вопрос о происхождении мутации. Так, гетерозиготное носительство у матери может быть заподозрено, в частности, по содержанию соответствующих белковых продуктов (например, фактора Y111 свертывания крови при гемофилии А; дистрофина в мышцах и креатинкиназы в сыворотке крови при миодистрофии Дюшенна) либо при помощи специальных ДНК-методов, позволяющих идентифицировать мутантный аллель у матери. Если дифференцировка этих случаев невозможна или доказано, что мутация у больного ребенка не является спонтанной, следует предполагать, что мать является гетерозиготной носительницей и с 50%-ой вероятностью будет передавать болезнь своим сыновьям. В этом случае пренатальная диагностика обязательна и должна сопровождаться определением пола плода. Следует, однако, подчеркнуть, что установление мужского пола плода на сегодняшний день отнюдь не является показанием для прерывания беременности, поскольку в 50% случаев они получают от матери X-хромосому с нормальным аллелем гена и являются вполне здоровыми. Определить, какую именно X-хромосому (с нормальным или мутантным аллелем) получил будущий мальчик, и является задачей молекулярной диагностики. С помощью прямых и непрямых методов ДНК-диагностики эта задача уже практически решается для очень многих сцепленных с полом заболеваний (см. Главу X).

Наиболее эффективной мерой профилактики наследственных

заболеваний является выявление гетерозиготных носителей мутаций, так как при этом удастся предотвратить рождение первого больного ребенка в семьях высокого риска. Родственники больного с большой вероятностью могут быть гетерозиготными носителями мутантных аллелей, поэтому в тех случаях, когда это возможно, они подлежат обследованию в первую очередь. Для болезней, сцепленных с полом, это родственники по женской линии – сестры, дочери и тетки пробанда. Их диагностика особенно важна, так как вероятность рождения больных сыновей в потомстве носительниц мутаций очень высока и не зависит от генотипа супруга. При аутосомно-рецессивных заболеваниях половина сибсов родителей и 2/3 здоровых сибсов больного будут гетерозиготными носителями мутации. Поэтому в тех семьях, где принципиально возможна молекулярная идентификация мутантных аллелей, необходимо обследовать максимальное число родственников больного пробанда для выявления гетерозиготных носителей. Иногда в больших семьях с разветвленными родословными удается проследить наследование неидентифицируемых мутаций с помощью косвенных методов молекулярной диагностики.

Для заболеваний, распространенных в определенных популяциях или в каких-то этнических группах и обусловленных присутствием одного или нескольких преобладающих и легко идентифицируемых мутантных аллелей, возможно проведение тотального скрининга на гетерозиготное носительство этих мутаций среди определенных групп населения, например, среди беременных женщин или среди новорожденных. Считается, что подобный скрининг экономически оправдан в том случае, если при проведении процедуры выявляются аллели, составляющие не менее 90 – 95 % всех мутаций данного гена в исследуемой популяции. Выявленные при подобных обследованиях носители мутаций также составляют группу риска, и в последующем должны быть аналогичным образом протестированы их супруги. Однако, даже в том случае, если мутация найдена только у одного из родителей, вероятность рождения больного ребенка несколько выше популяционной частоты, но, конечно, значительно меньше 25%. Конкретное значение этого риска зависит от общей частоты мутаций соответствующего гена в популяции. В таких семьях по желанию родителей также может быть проведена пренатальная диагностика и прослежено наследование мутантного аллеля. При отсутствии этой мутации у плода прогноз считается благоприятным, независимо от того, какие аллели ребенок получит от второго супруга.

Раздел 7.4 Особенности применения молекулярных методов в пренатальной диагностике моногенных болезней.

Следующим шагом после отбора нуждающейся в пренатальной диагностике семьи является комплексное молекулярное обследование ее членов – родителей и, если есть такая возможность, больного ребенка. Эти исследования могут быть достаточно длительными и поэтому желательно их проводить до наступления беременности. Результаты молекулярного обследования семьи служат основой для назначения и выбора способов проведения пренатальной диагностики. После этого семья планирует беременность, и в определенные сроки женщина поступает в клинику для проведения процедуры, обеспечивающей забор необходимых для диагностики тканей плодного происхождения. При этом следует учитывать, что определение конкретных сроков этой процедуры зависит от многих медицинских показаний (в первую

очередь, акушерских), обеспечивающих максимальную безопасность подобного инвазивного вмешательства как для матери, так и для будущего ребенка (Баранов и др., 1994; Баранов, 1994).

Наиболее объективная информация о наличии моногенного наследственного заболевания у плода может быть получена при анализе его ДНК и обнаружении мутационных изменений в кодирующих или регуляторных последовательностях соответствующих генов. Практическая диагностика мутантных аллелей в семьях высокого риска проводится для заболеваний с известным спектром наиболее часто встречающихся мутаций. При этом для каждой болезни разрабатываются относительно простые и наиболее эффективные методы идентификации мутантных аллелей. В настоящее время насчитывается уже несколько сотен таких заболеваний и количество их быстро увеличивается (Baranov, 1993; Баранов, 1994; см. Главу X).

Для генотипирования мутаций, то есть для выяснения природы мутантных аллелей у больного и его родителей используются различные стандартные методы, подробно изложенные в главе IV. Выбор конкретного метода зависит от типа предполагаемых мутаций и от методических возможностей диагностических центров. При отсутствии у пробанда наиболее частых и ранее описанных мутаций его ДНК может быть направлена в специализированные молекулярно-генетические лаборатории для более тщательного анализа с использованием всего комплекса методов идентификации мутаций вплоть до получения и секвенирования мутантных кДНК-вых последовательностей гена. Однако, подобные исследования очень дороги, требуют много труда и времени и потому в обычной клинической практике используются достаточно редко. В этих случаях чаще прибегают к косвенным методам молекулярной диагностики (см. раздел 7.1). Для многих заболеваний, эти методы все еще остаются единственно возможными (см. Главу X). Однако, как уже указывалось, они требуют обязательного обследования полной семьи, включая больного ребенка. При его отсутствии молекулярное маркирование мутантных генов у гетерозиготных родителей становится невозможным, а, следовательно, невозможна и пренатальная диагностика.

Идентификацию мутантных генов осуществляют путем сравнения маркерных генотипов родителей и больного ребенка. Если не произошел кроссинговер между маркерным локусом и геном, все больные дети в семье должны иметь одинаковый маркерный генотип. Поэтому при проведении пренатальной диагностики сходство генотипов плода и пробанда является необходимым, однако в ряде случаев, вовсе не достаточным условием для подтверждения наличия заболевания. Так, например, у гомозиготных по маркеру родителей все дети будут иметь одинаковый генотип по этому локусу независимо от присутствия мутантных аллелей гена. Необходимым условием дифференцировки нормальных и мутантных генов у носителей мутаций является их сцепление с разными аллелями маркерного локуса. Поэтому, дискриминация мутантных генов у родителей возможна только с помощью гетерозиготных маркеров. При этих условиях мутантные гены родителей могут быть определены по тем маркерным аллелям, которые присутствуют у больного ребенка. Однако, идентификация по маркерным аллелям обоих мутантных генов возможна лишь в тех случаях, когда родители гетерозиготны, а пробанд (больной) гомозиготен по маркерному локусу.

Возможность идентификации мутантных генов родителей на основе анализа маркерных генотипов определяет информативность семьи по отношению к данному маркеру. На Рис.7.1

представлены возможные маркерные генотипы в полностью и частично информативных, а также в неинформативных семьях при определении их путем блот-гибридизации с ДНК-зондом. Аналогичные схемы могут быть составлены и в тех случаях, когда маркерные генотипы определяются путем анализа содержащих полиморфные локусы амплифицированных фрагментов ДНК. В информативных семьях маркерный генотип больного ребенка отличается от маркерных генотипов обоих родителей и здорового сибса, поэтому можно проследить наследование мутантных генов при каждой беременности. В этом случае сходство маркерных генотипов пробанда и плода достаточно для неблагоприятного прогноза, а носители мутации будут иметь родительский генотип. В частично информативных семьях идентифицируется только один из мутантных генов, так как маркерный генотип больного ребенка совпадает с генотипом одного или даже обоих родителей. В этом случае при сходстве маркерных генотипов плода и пробанда вероятность болезни будущего ребенка составляет 50%. Все дети с другим маркерным генотипом будут здоровы, но половина из них будет нести один из мутантных аллелей. Неинформативный маркер не пригоден для идентификации мутантных генов, а следовательно, и для проведения молекулярной диагностики в данной семье. При отсутствии полной информативности необходимо исследовать другие сцепленные с геном полиморфные локусы и выбрать в качестве маркеров те из них, которые у родителей пробанда находятся в гетерозиготном состоянии. В частично информативных семьях можно проводить пренатальную диагностику с такой же степенью достоверности, как и в полностью информативных семьях, при одновременном использовании двух полиморфных локусов, каждый из которых маркирует разные мутантные гены обоих родителей. Информативными оказываются также те семьи, в которых один из мутантных аллелей может быть идентифицирован прямым анализом соответствующего участка гена, а наличие другого мутантного аллеля доказывается с помощью маркерного локуса. Для многих моногенных наследственных заболеваний количество идентифицированных высокополиморфных локусов, тесно сцепленных с контролирующим геном, достаточно для того, чтобы более 90% семей высокого риска оказались полностью информативными, а значит, пригодными для проведения в них пренатальной диагностики с использованием молекулярных методов тестирования состояния плода.

Таким образом, при отсутствии возможности прямой идентификации соответствующих мутантных аллелей, анализ информативности семьи следует начинать с наиболее полиморфных маркерных локусов. так как при этом больше вероятность того, что родители больного ребенка окажутся гетерозиготами по выбранному маркеру. При последующем выборе маркеров важно также учитывать наличие между ними неравновесности по сцеплению, так как чаще всего информативность семей в отношении маркеров, находящихся в сильном неравновесии по сцеплению, будет одинакова. Действительно, неслучайный характер цис- и транс-расположения аллелей в неравновесных по сцеплению локусах обуславливает повышенную вероятность таких событий, при которых гомозиготы по одному из маркеров оказываются гомозиготными и по другому. То же самое справедливо и в отношении гетерозигот. Поэтому при анализе информативности семьи, в первую очередь, следует выбирать маркерные локусы, находящиеся между собой в равновесии по сцеплению.

Следует также учитывать, что определенные аллели полиморфных сайтов, расположенных близко к мутации, возникшей однократно много лет назад и распротранившейся в популяции,

будут оставаться в очень тесном неравновесии по сцеплению. Примером может служить неравновесность между delF508 (ориентировочный возраст возникновения 30-50 000 лет) и 6-ти членным тетрамерным повтором TAGG в интроне 6 гена муковисцидоза (Chebab et al., 1993; Агбанглы и др., 1994). Поэтому в семьях высокого риска с большой долей вероятности, конкретное значение которой определяется детерминантом неравновесности по сцеплению, гомозиготы по этому маркерному аллелю, окажутся также гомозиготами и по мутации, то есть будут больны. Конечно, пренатальный диагноз не может быть основан только на этой информации, однако, её следует учитывать при выработке комплексного заключения относительно здоровья будущего ребенка.

Во всех случаях при использовании косвенных методов молекулярной диагностики необходимо также помнить, что маркерный генотип плода определяет наличие мутантных генов с точностью до вероятности кроссинговера между мутантным аллелем и маркерным локусом. Поэтому, чем ближе расположен маркер по отношению к мутантному аллелю, тем меньше вероятность ошибки при проведении пренатальной диагностики. Для того чтобы полностью избежать или, по крайней мере, резко снизить вероятность такой ошибки, желательно использовать внутригенные маркеры или проводить одновременное тестирование ДНК плода с помощью двух маркерных локусов, фланкирующих мутантный аллель. Последний подход особенно оправдан при работе с очень протяженными генами (например геном дистрофина длиной около 2,2 миллионов пар оснований), где вероятность внутригенного кроссинговера по некоторым данным может достигать 2-2,5%.

Раздел 7.5 Доимплантационная диагностика, точность прогнозирования.

Следует подчеркнуть, что общий подход к молекулярному тестированию, практически, не зависит от формы нозологии, так как для анализа генов и мутантных аллелей используется один и те же методы (см. Глава IV). При наличии у больного или у его родителей идентифицируемых мутаций или при нахождении информативных для данной семьи ДНК-маркеров проведение пренатальной диагностики становится возможным на любой стадии развития и при любом сроке беременности.

ДНК-диагностика на начальных этапах развития человека, безусловно, имеет свои методические особенности и тесно сопряжена с приемами экстракорпорального оплодотворения и трансплантации зародышей. Эти обстоятельства позволяют выделить ее в самостоятельное научно-практическое направление - доимплантационную диагностику генных и хромосомных болезней (Verlinsky, Kuliev, 1993). Несмотря на очевидные успехи в этой области, достигнутые в ряде зарубежных лабораторий (доимплантационная диагностика муковисцидоза, некоторых X-сцепленных заболеваний - миодистрофии Дюшенна, синдрома ломкой X-хромосомы, гемофилии), данное направление исследований все еще находится на уровне научных разработок и не имеет широкого применения. Объектом ДНК-диагностики на этих стадиях развития могут служить полярные тельца овулировавших яйцеклеток и отдельные бластомеры зародыша, полученные методом микрургии (Рис.7.2). Недостатком этого подхода является сравнительно невысокий процент (до 30%) приживаемости оперированных зародышей после их искусственной трансплантации в матку. Подробно с методами доимплантационной диагностики, её современным состоянием и перспективами можно ознакомиться в уже цитированной монографии (Verlinsky, Kuliev, 1993). Есть

определенные основания считать, что в связи с быстрым ростом числа центров экстракорпорального оплодотворения в России, заметным увеличением эффективности их работы, оцениваемой по числу наступивших беременностей и родов, методы доимплантационной диагностики наследственных болезней будут привлекать к себе все большее число исследователей. Однако, практическая значимость такого подхода в виду его высокой стоимости и недостаточной надежности еще сравнительно долго будет оставаться проблематичной, по крайней мере, в нашей стране.

Обобщенный опыт различных молекулярных диагностических центров мира свидетельствует о том, что оптимальным для пренатальной диагностики является первый триместр беременности (10-12- недели), поскольку при неблагоприятном прогнозе беременность может быть прервана обычным медицинским абортom. Зачастую молекулярную диагностику проводят и во втором триместре беременности (обычно на 17-21 неделях). Необходимые для анализа образцы ДНК плода выделяют из биоптатов хориона (плаценты), клеток амниотической жидкости (амниоцитов) или из лимфоцитов пуповинной крови при помощи соответствующих инвазивных процедур - хорионбиопсия, плацентобиопсия, амниоцентез или кордоцентез, которые проводятся под контролем ультразвука (Рис. 7.3). Подробнее об этих методах, их преимуществах и недостатках можно узнать в соответствующих монографиях и обзорах (Papp, 1992; Баранов, 1994; Баранов и др., 1994). Диагностика во 2-м триместре беременности нередко показана для неинформативных семей в случае тех нозологий, пренатальная диагностика которых принципиально возможна путем биохимического тестирования первичных нарушений (например муковисцидоз) или оценки каких-либо иных проявлений болезни у плода.

Основным источником ДНК для диагностики в постнатальном периоде являются лимфоциты крови. Реже для этих целей используют другие ткани и биологические жидкости, содержащие клеточные элементы (слюна, кости, моча). Если семья частично информативна (идентифицируется либо доступен молекулярной маркировке лишь один мутантный аллель), также рекомендуется диагностика в 1-м триместре беременности, так как она позволяет отвергнуть диагноз у 50% всех плодов при аутосомно-рецессивных заболеваниях, что доказывается отсутствием идентифицируемого мутантного аллеля. В противном случае дальнейшая тактика определяется исходя из пожеланий женщины и возможностей для уточнения диагноза на более поздних сроках развития (II-й триместр беременности). Так, в случае пренатальной диагностики муковисцидоза результаты молекулярных исследований могут быть дополнены биохимическим тестированием ферментов амниотической жидкости на 18-20-й неделях беременности (Горбунова и др., 1991; Baranov et al., 1992); в случае гемофилии А - прямым серологическим исследованием активности фактора Y11 свертывания крови (Aseev et al., 1994); в случае миодистрофии Дюшенна - иммуноцитохимическим анализом дистрофина в биоптатах мышц плода (Hoffman et al., 1992) и т.д. Следует однако подчеркнуть, что современный уровень молекулярных знаний о природе гена, его мутациях и полиморфизмах позволяет в идеале проводить молекулярную диагностику наиболее частых моногенных заболеваний, практически, во всех случаях. Возможности молекулярной диагностики в России пока весьма ограничены (Baranov, 1993; см. Глава X). Кроме того, ситуация осложняется сравнительно поздним (часто, уже во время беременности) обращением семей высокого риска на пренатальную диагностику, что не позволяет провести молекуляр-

ные исследования ее информативности в полном объеме. Связано это, в первую очередь, с плохой информированностью населения о работе соответствующих медико-генетических служб.

Из всех существующих способов пренатальной диагностики методы, основанные на анализе ДНК, являются наиболее точными, так как они выявляют первичные нарушения в структуре генов. Это прежде всего относится к тем заболеваниям, диагностика которых осуществляется путем прямой идентификации мутантных аллелей. Тем ни менее, даже в этом случае окончательное заключение о здоровье будущего ребенка хотя и приближается к абсолютному, но все же носит вероятностный характер. Основным источником ошибок при проведении пренатальной диагностики служит возможная контаминация материала плода, используемого для постановки диагноза, другими тканями, в первую очередь, материнского происхождения. Поэтому сразу после взятия диагностического материала путем биопсии хориона, амниоцентеза или кордоцентеза проводят тщательную его оценку с использованием разнообразных лабораторных методов. Отличие молекулярных методов тестирования от биохимических или любых других заключается в их повышенной чувствительности и огромной разрешающей способности. Загрязнение материала особенно опасно в тех случаях, когда прогноз дается на основе анализа амплифицированных фрагментов ДНК. Попавшие в тестируемые образцы клетки чужеродного происхождения могут послужить источником донорской ДНК для ПЦР и тогда может быть получено несоответствующее действительности заключение о том, что ребенок будет здоров. Риск подобных ошибок может быть значительно уменьшен при работе в стерильных условиях и при постановке анализов в нескольких параллельных пробах. Конечно, в любом диагностическом центре возможны лабораторные ошибки чисто технического порядка, не зависящие от используемых методов тестирования. Однако, достоверность прогнозирования, основанного на прямом анализе мутантных аллелей плода, обычно, превышает 99%. При использовании косвенных методов молекулярной диагностики появляется дополнительный риск ошибки, связанный с возможностью рекомбинации между мутантным аллелем и маркерным локусом (особенно при больших размерах гена, как, например, в случае гена дистрофина). Конкретное значение этого риска зависит от взаимного расположения используемых для диагностики маркеров и соответствующих мутантных аллелей. Обычно, для молекулярной диагностики используют маркеры, частота рекомбинации которых с мутантными аллелями гена не превышает десятых, а иногда и сотых долей процента.

В случае неблагоприятного прогноза в отношении здоровья будущего ребенка решение о прерывании или продолжении беременности принимают родители на основании предоставленного в их распоряжение общего заключения. После рождения ребенка или прерывания беременности желательно проводить верификацию диагноза всеми доступными для этого методами, включая и те, которые были использованы для постановки пренатального диагноза.

ЗАКЛЮЧЕНИЕ

При всем разнообразии тем, затронутых в монографии, весь изложенный в ней материал, по-сути, касается трех

основных проблем : 4(1) Генетическое картирование и геном человека, 4(2) Молекулярная диагностика генных болезней, 4(3) генокоррекция наследственных дефектов и генотерапия. В решении каждой из названных проблем достигнуты серьезные успехи. Напомним некоторые из них.

Известно, что первый ген человека (ген цветной слепоты-дальтонизма) картирован на X-хромосоме человека в 1911г. Первый аутосомный ген - только в 1968г. К 1973 на всех хромосомах человека было картировано всего 64 гена, а к 1994г. на генетических картах уже локализовано свыше 60 000 маркерных ДНК последовательностей (главным образом, фрагментов кДНК экспрессирующихся генов-EST см.Главу III.), а также более 5 000 полноразмерных структурных генов. Благодаря многочисленным полиморфным сайтам и, главным образом, микросателлитным молекулярным маркерам, созданы подробные (1,5-2 сантиморганные) генетические карты для каждой хромосомы человека. Это позволило перейти от функционального к позиционному картированию, т.е. от картированию новых генов непосредственно на физической карте ДНК целого генома.

По мнению авторитетных специалистов по генетическому картированию таких как Питер Гудфеллоу, Поль Вайссенбах и др. других, дальнейшее наращивание плотности молекулярных маркеров на хромосомах человека уже лишено смысла, тем более, что в процессе идентификации все новых и новых генов методом EST, новые полиморфные сайты все равно будут найдены. Не менее оптимистично обстоят дела и с секвенированием, т.е. с выяснением первичной нуклеотидной последовательности всей двухметровой молекулы ДНК человека. Достаточно заметить, что первоначальная стоимость секвенирования одной пары нуклеотидов оценивалась в 1 \$, сейчас она составляет уже около 40 центов и продолжает снижаться. Причина этого - широкая автоматизация монотонного процесса секвенирования. Так, 10 роботов фирмы Applied Biosystems за одну неделю секвенируют более 30 000 000 пар оснований. Дальнейшее совершенствование технологии секвенирования, созданы принципиально новых подходов (метод "чипов"-Мирзабеков А.Д., Ed.Southern), повышающих в десятки раз степень автоматизации этого процесса в высокоспециализированных центрах позволяет надеяться, что секвенирование всего генома человека будет завершено в 2006 году, а вполне вероятно и к 2000 году!

Однако, само по себе завершение гигантского по замыслу и грандиозного по реализации научного проекта "Геном человека" отнюдь не означает, что процесс познания генома завершен. Уже сейчас становится очевидным, что не существует какого-то усредненного генома человека, каждый геном как и каждый человек сугубо индивидуален. Эта индивидуальность генома проявляется не только на уровне отдельной личности, но и на уровне этнических групп, наций, отдельных популяций и рас (Cavalli-Sforza, 1994). Геном человека как система динамичная, очень разнообразна. Анализ этого разнообразия (diversity) - одно из важнейших продолжений программы Геном человека. Еще более актуальным является выяснение "функциональной карты генома". Секвенирование позволит расшифровать порядок всех 3 4.5 x 10⁹ нуклеотидов. Но ведь это только начало. Определить границы генов, выяснить положение многочисленных регуляторных элементов, их интронно-экзонную структуру и, наконец, функциональное назначение каждого из 60 000 генов, назначение которых пока неизвестно - вот следующая поистине глобальная задача молекулярной генетики. Вполне вероятно, что именно на этом пути удастся решить загадку "избыточной ДНК", понять эволюцию (филогенез) генома

человека и, возможно, расшифровать партитуру симфонии жизни – т.е. последовательность включения и выключения генов в ходе онтогенеза.

На генетические карты человека уже в 1994 г. нанесено 933 гена, мутации которых приводят к различным наследственным заболеваниям, причем более 400 из них проклонированы, т.е. выделены в чистом виде и размножены вне организма человека в составе ДНК фагов, вирусов, дрожжей и бактерий. Для многих из этих генов, в особенности, сопряженных с наиболее частыми, социально значимыми заболеваниями, подробно изучены спектры мутаций, охарактеризованы аллельные полиморфизмы и разработаны схемы молекулярной диагностики. Причем, если в 1993 г. таких заболеваний было около 130 (Bob Williamson), то в 1994 – более 600. По-сути, уже сегодня каждый наследственный недуг, ген которого картирован, доступен молекулярной диагностике прямыми или косвенными методами.

Помимо моногенных болезней, проблемы молекулярной диагностики которых в значительной степени уже решены, все больше внимания сегодня уделяют мультифакториальным заболеваниям. На повестке дня молекулярная диагностика предрасположенности к таким широко распространенным недугам как атеросклероз, ишемия сердца, онкологические, психиатрические заболевания, диабет и многие другие. Выяснение генетической природы этих болезней, равно как досимптоматическая диагностика многих моногенных болезней с поздней манифестацией, ставит перед исследователями, от молекулярными биологами

и

врачами, проблему целесообразности такой досимптоматической диагностики, права личности на исключительность знаний о собственном геноме, точнее о тех мутациях и генетических предрасположенностях, которые закодированы в нем еще до рождения. Для некоторых заболеваний (муковисцидоз, фенилкетонурия) такая ранняя диагностика, безусловно, целесообразна, так как позволяет начать лечение до начала заболевания. Для тех же нозологий, где реальной терапии пока нет (хорея Гентингтона, другие болезни "экспансии", миодистрофия Дюшенна и др.) целесообразность такой диагностики и, главное, конфиденциальность полученной информации широко обсуждаются.

Да, уже сейчас вполне реально говорить о "генетическом паспорте новорожденных", т.е. о том, что уже вскоре после рождения с помощью автоматизированной системы удастся проанализировать весь спектр наиболее частых мутаций широко распространенных заболеваний как моногенной так и мультифакториальной природы, в том числе и генов, мутации которых с высокой вероятностью могут привести к раку молочной железы, толстого кишечника, к атеросклерозу, диабету и многим другим тяжелейшим недугам. Жить, зная, в чем себя не ограничивая, засунув как страус голову в песок, либо, зная, что у вас мутация в гене глутатионтрансферазы, а следовательно, высока вероятность болезней легких (особенно рака) воздержаться от курения? Что лучше, добровольные ограничения и периодические осмотры или состояние счастливого неведения, грозящее неминуемой катастрофой? А сколько мультифакториальных заболеваний можно избежать, зная о слабых и сильных сторонах своего генома! В настоящее время в США проводятся массовые опросы населения, цель которых выяснить целесообразность досимптоматической диагностики в семьях высокого риска, доступность (конфиденциальность) этой информации для членов семьи, нанимателей, страховых компаний и пр. Иными словами, практически овладев плодом Древа Познания – собственным геномом – человечество поставлено перед дилеммой как сделать так, чтобы

польза от него оказалась много весомее, чем потенциальный вред. Неслучайно, сегодня уже на новом, молекулярном уровне всплывают идеи "улучшения человеческой породы" Фрэнсиса Гальтона, правда, неимеющие ничего общего с примитивной евгеникой прошлого. Обрели реальность и казавшиеся невозможными в недалеком прошлом идеи патентования отдельных генов и их фрагментов, потенциально особенно перспективных для молекулярной диагностики и биотехнологий. Несмотря на протесты руководителей программы "Геном человека" Фрэнсиса Коллинса, Томаса Каски и широкой научной общественности патентноспособность генома человека, по крайней мере, его отдельных фрагментов, генов (например, гена BRCA-1-мутации которого резко увеличивают вероятность рака молочной железы) получила одобрение законодательных комитетов США.

Еще больше этических и моральных вопросов порождает генотерапия. Однако, и в этой области наметилась определенная эволюция взглядов как специалистов, так и широкой общественности. От полной неприемлемости такого подхода в 70-80-х годах 20-го века до признания безопасности (при соблюдении необходимых правил) генноинженерных манипуляций на соматических клетках. Между тем, логика подсказывает, что по мере того как все больше соматических мутаций удастся исправить с помощью генной терапии, тем значительней будет их вклад на уровне половых клеток. Тем больше будет шанс того, что при вступлении в брак гомозигот по аутосомно-рецессивным заболеваниям (а вероятность такого события будет возрастать по мере совершенствования методов лечения генетических болезней), все дети будут здоровы, хотя и получают от своих родителей мутантные гены. Это станет особенно реально в связи с успешной разработкой методологии доимплантационной диагностики наследственных болезней. Поэтому в научной литературе все настойчивее звучит тема генокоррекции на уровне половых клеток или ранних зародышей человека. Современный методический уровень пока еще не позволяет с абсолютной уверенностью осуществлять направленный перенос гена в половые клетки и клетки дробящихся зародышей. Пока этого удалось достичь только в экспериментах с эмбриональными стволовыми клетками, причем на 1-м этапе возникший организм, в действительности, является мозаиком по введенному гену (см. Главу V 4I 0II). Естественно, это не означает, что проблема гомологичной рекомбинации нормального и мутантного гена на уровне половых клеток и ранних зародышей принципиально неразрешима. Однако, потребуется еще много времени, возможно не одно десятилетие XXI-го века, прежде, чем эта задача будет успешно решена. В настоящее время мнение специалистов и широкой общественности - максимально предотвратить возможность попадания чужеродного генетического материала в половые клетки, с тем, чтобы избежать непредсказуемых, а, скорее всего, весьма печальных, последствий для человечества такого трансгеноза.

Нет, однако, сомнений в том, что, если конец XX-го века ознаменован расшифровкой молекулярной структуры генома человека, то век XXI прославится выяснением его функций на уровне отдельных генов, генных сообществ, хромосом, также всего генома в целом в контроле процессов онто- и филогенеза.

Молодым людям, вступающим в XXI век, необходимо знать, на каком этапе находится наука о структуре и функциях генома человека. Это важно не только с сугубо утилитарных, медицинских позиций (хотя и этот аспект крайне важен), но и с позиций общечеловеческих. Ибо всякое эпохальное открытие на-

уки (а именно таковым и является расшифровка генома человека) до недавнего времени использовалось не только во благо, но и во вред человечеству (4п 0ечальный пример тому - открытие расщепления ядра урана, породившее атомную бомбу). Неразумные эксперименты с геномом человека могут привести к еще более страшным последствиям. Уберечь генофонд человечества, всячески оберегая его от рискованных вмешательств, и при этом извлечь максимальную выгоду из уже полученной бесценной информации в плане диагностики, профилактики и лечения многих тысяч наследственно обусловленных недугов - вот задача, которую необходимо решать уже сегодня 4и 0с которой мы придем в новый XXI век !

ГЛАВА IX.

ГЕННАЯ ТЕРАПИЯ.

Раздел 9.1 Определение, историческая справка, программы генной терапии.

В широком смысле слова генная терапия означает лечение путем введения в ткани или в клетки пациента смысловых последовательностей ДНК. Первоначально генная терапия рассматривалась как возможность исправления дефекта в гене. Считалось, что основным объектом для такого лечения будут служить моногенные наследственные заболевания человека, причем теоретически представлялась вероятной коррекция генного дефекта как на соматическом уровне, так и на уровне зародышевых (половых) клеток. Многочисленные эксперименты по созданию трансгенных животных, начатые после 1980 г., а также на культурах клеток внесли существенные коррективы в эти теоретические представления. Во-первых, оказалось значительно проще исправлять не сам дефект в гене, то есть заменять весь мутантный ген или его мутированный фрагмент на нормальный, а вести коррекцию путем введения в организм пациента полноценно работающего гена (обычно его кДНК). Во-вторых, несмотря на решающие успехи генной инженерии последних лет, исследования по генной терапии у человека осуществляются исключительно на соматических тканях, в которых в норме происходит экспрессия дефектного гена. Генная терапия на уровне половых и зародышевых клеток человека ввиду возможных серьезных последствий для генофонда человечества представляются весьма проблематичной и на данном этапе наших знаний - малореальной. И наконец, в-третьих, уже разработанная и применяемая на практике методология генной терапии оказалась пригодной для лечения не только моногенных наследственных заболеваний, но и таких широко распространенных болезней, какими являются злокачественные опухоли, многие виды тяжелых вирусных инфекций, включая спид, сердечно-сосудистые и другие заболевания. Учитывая эти обстоятельства, генную терапию на современном этапе можно определить как лечение наследственных, онкологических, некоторых инфекционных (вирусных) и других заболеваний путем введения генов в клетки пациентов с целью направленного изменения генных дефектов, либо придания клеткам новых функций (Culver, 1994). Первые клинические испытания методов генной терапии были предприняты 22 мая 1989г. с целью генетического маркирования опухоль-инфильтрующих лимфоцитов в случае прогрессирующей меланомы. Маркированные прокариотическим геном нео, Т-лимфоциты были устойчивы к неомицину и

могли быть легко отселектированы в культуре, что позволило детально проследить их судьбу в кровотоке и избирательное накопление в опухолях (подробней см. 9.5).

Первым моногенным наследственным заболеванием, в отношении которого были применены методы генной терапии, оказался наследственный иммунодефицит, обусловленный мутацией в гене аденозин-дезаминазы. 14 сентября 1990г. в Бетезде (США) 4-х летней девочке, страдающей этим достаточно редким заболеванием (1 : 100 000), были пересажены ее собственные лимфоциты, предварительно трансформированные *ex vivo* геном ADA (ген ADA + ген neo + ретровирусный вектор). Лечебный эффект наблюдался в течение нескольких месяцев, после чего процедура была повторена с интервалом 3-5 месяцев (Anderson, 1992; Culver, 1994). В течение 3-х лет терапии в общей сложности проведено 23 внутривенных трансфузии ADA-трансформированных Т-лимфоцитов без видимых неблагоприятных эффектов. В результате лечения состояние пациентки (Ашанти В. ДеСильва) настолько улучшилось, что она смогла вести нормальный образ жизни и не бояться случайных инфекций. Столь же успешным оказалось и лечение второй пациентки с этим заболеванием (подробней см. раздел 9.5). В настоящее время клинические испытания генной терапии этого заболевания проводятся в Италии, Франции, Великобритании и Японии.

Другие моногенные наследственные заболевания, в отношении которых уже имеются официально разрешенные протоколы и начаты клинические испытания, касаются семейной гиперхолестеринемии (1992); муковисцидоза (1993); гемофилии В (1992); болезни Гоше (1993). В отношении многих других заболеваний медицинские протоколы клинических испытаний находятся в стадии утверждения (см. раздел 9.5.). К 1993г. только в США к клиническим испытаниям генно-инженерных конструкций на человеке было допущено 53 проекта (Culver, 1994). К 1995г. в мире число таких проектов возросло до 100 и более 400 пациентов было непосредственно вовлечено в эти исследования (Hodgson, 1995). Подавляющее большинство таких проектов (86) касалось лечения онкологических заболеваний, а также СПИДа.

Таким образом, от опытов на животных и теоретических построений 80-х годов уже в 1990 году удалось приступить к реальному лечению моногенных заболеваний, число которых стремительно нарастает. Естественно, что подобные революционные перемены могли возникнуть только в результате решающих успехов молекулярной биологии в картировании генов, мутации которых приводят к наследственным заболеваниям (см. Главу III), выяснении молекулярной природы этих мутаций (см. Главу IV), успехов в секвенировании и клонировании генов (см. Главы I и II), создании генно-инженерных конструкций (см. Главу II), отработки и совершенствования методов их доставки (см. ниже). Следует также подчеркнуть, что качественный скачок в области генной терапии, когда сам ген стал рассматриваться как лекарственный препарат, стал возможен благодаря тому, что предшествующие экспериментальные и клинические исследования доказали безопасность генной терапии.

Вместе с тем, и в современных исследованиях по генной терапии необходимо учитывать, что последствия манипулирования генами или рекомбинантными ДНК *in vivo* изучены недостаточно. Следует помнить, что введение в организм человека последовательностей ДНК, не находящихся под контролем собственных им регуляторных элементов, может приводить к трудно предсказуемым изменениям метаболических процессов и сопровождаться функциональным дисбалансом. Современные представления о структуре генома и его взаимодействиях с экзогенными ДНК и

вирусными последовательностями, часто используемыми в качестве векторов для переноса генов (см. 9.2), могут оказаться недостаточными для прогнозирования возможных нежелательных или неконтролируемых последствий такого вмешательства. Поэтому при разработке программ генной терапии принципиальное значение имеют вопросы безопасности предлагаемых схем лечения как для самого пациента, так и для популяции в целом (Anderson, 1992; Miller, 1992). Важно, чтобы при проведении испытаний ожидаемый лечебный эффект или возможность получения дополнительной полезной информации превосходили потенциальный риск предлагаемой процедуры. Неслучайно, в странах с наиболее продвинутым уровнем исследований в этой области, особенно в США, медицинские протоколы с использованием смысловых последовательностей ДНК подвергаются обязательной экспертизе в соответствующих комитетах и комиссиях. Клинические испытания предложенной генотерапевтической процедуры возможны только после ее одобрения соответствующим законодательно утвержденным органом. В США таковыми являются: Консультативный Комитет по Рекомбинантным ДНК (Recombinant DNA Advisory Committee - RAC), Комитет по лекарствам и пищевым продуктам (Food and Drug Administration - FDA), с последующим обязательным утверждением проекта директором Национального Института Здоровья (National Institute of Health) (Miller, 1992; Anderson, 1992; Culver, 1994). В Европе такие протоколы составляются и утверждаются в соответствии с рекомендациями Европейской Рабочей Группы по Переносу Генов и Генной Терапии (European Working Group on Human Gene Transfer and Therapy) (Cohen-Haguenauer, 1995). Программы генной терапии для клинических испытаний должны включать следующие разделы: обоснование выбора нозологии для проведения курса генной терапии; определение типа клеток, подлежащих генетической модификации; схему конструирования экзогенной ДНК; обоснование биологической безопасности вводимой генной конструкции, включающая опыты на культурах клеток и на модельных (трансгенных) животных; разработку процедуры ее переноса в клетки пациента; методы анализа экспрессии введенных генов; оценку клинического (терапевтического) эффекта; возможные побочные последствия и способы их предупреждения (Culver, 1993; Cohen-Haguenauer, 1995).

Важнейшим элементом в программе генной терапии является анализ последствий проводимых процедур. Этот контроль проводят на всех этапах терапии, причем исследования выполняют на различных уровнях. Прежде всего, после переноса гена осуществляют поиск модифицированных клеток в организме пациента и следят за динамикой этих клеток в определенных тканях. Этот поиск может быть облегчен при наличии маркерного гена в конструкции. Присутствие последовательностей экзогенной ДНК в модифицированных клетках чаще всего идентифицируют с помощью ПЦР. На следующем этапе производят анализ экспрессии введенных генов путем идентификации и количественной оценки соответствующего РНК-транскрипта, либо белкового продукта гена. В тех случаях, когда это возможно, проводят анализ коррекции первичного биохимического дефекта. Затем, все полученные данные сопоставляют с результатами комплексного медицинского обследования и вносят необходимые исправления и добавления в проводимую схему лечения.

Раздел 9.2. Типы генотерапевтических вмешательств, выбор клеток-мишеней.

Рассмотрим наиболее общие принципы, лежащие в основе

построения программ генной терапии. Итак, генная терапия предполагает введение последовательностей ДНК в клетки-мишени. Она проводится либо с целью коррекции наследственной патологии, возникшей вследствие генетического дефекта, либо для придания этим клеткам новых функций, способствующих устранению патологических процессов. В первом случае, в организм больного вводят нормально работающий гомолог дефектного гена. Вторым подходом применяют при лечении, таких заболеваний, как опухоли или инфекции. В этих случаях вводят гены, обладающие условным цитотоксическим эффектом или способствующие формированию выраженного иммунного ответа. Мишенями для таких генов служат пораженные ткани, иммунные клетки, специфическим образом проникающие в эти ткани, либо предварительно трансформированные *in vitro* другие клетки. Таким образом, в зависимости от характера заболевания и предполагаемого генотерапевтического подхода объектом генетической трансфекции могут служить самые разные соматические клетки, как несущие дефектный ген, так и нормальные клетки, приобретающие терапевтические свойства после трансфекции. В зависимости от способа введения экзогенных ДНК в геном пациента генная терапия может проводиться либо в культуре клеток (*ex vivo*), либо непосредственно в организме (*in vivo*). Клеточная генная терапия или терапия *ex vivo* предполагает выделение и культивирование специфических типов клеток пациента, введение в них чужеродных генов, отбор трансфицированных клеток и реинфузию их тому же пациенту (Рис. 9.1). В настоящее время большинство допущенных к клиническим испытаниям программ генной терапии использует именно этот подход (Culver, 1994). Осуществление таких программ возможно лишь в крупных специализированных центрах, требует больших материальных затрат и высоких биотехнологий.

Генная терапия *in vivo* основана на прямом введении клонированных и определенным образом упакованных последовательностей ДНК в специфические ткани больного. При этом вводимые ДНК, как правило, интегрируются с молекулами, обеспечивающими их адресную доставку в клетки-мишени (см. 9.3). Этот очень перспективный подход, рассчитанный на массовое лечение широко распространенных заболеваний, пока реально апробирован только для лечения муковисцидоза (Crystal et al., 1994). Особенно перспективным для лечения генных болезней *in vivo* представляется введение генов с помощью аэрозольных или инъекционных вакцин. Аэрозольная генотерапия разрабатывается, как правило, для лечения пульмонологических заболеваний, таких как муковисцидоз, эмфизема, рак легких, при которых объектами генетической модификации являются специфические типы легочных клеток (Hoffman, 1991). Инъекционные вакцины могут использоваться для модификации различных типов клеток и со временем, по-видимому, станут наиболее распространенным и универсальным способом доставки чужеродного генетического материала в любые ткани.

Эффективность курса генной терапии в значительной степени зависит от правильного выбора типов соматических клеток, в которых должна быть проведена генетическая модификация. Так например, при лечении какого-либо наследственного заболевания, обусловленного дефектом секреторного белка, генетической коррекции, в принципе, могут быть подвергнуты любые клетки, тогда как для нерастворимых или мембран-связанных белков выбор ограничен теми клетками, где экспрессируется соответствующий ген (см. раздел 8.5). Разработке программы генной терапии предшествуют тщательный анализ тканеспецифической экспрессии соответствующего гена, идентификация пер-

вичного биохимического дефекта, исследование структуры, функции и внутриклеточного распределения его белкового продукта, а также биохимический анализ патологического процесса. Все эти данные учитываются при составлении соответствующего медицинского протокола. Кроме того, план генотерапевтических вмешательств определяется также доступностью клеток-мишеней, периодом их жизни и характером миграции в организме, эффективностью и специфичностью трансфекции клеток, длительностью экспрессии введенного гена.

Наиболее перспективной представляется возможность генетической модификации не самих уже дифференцированных клеток с наследственным дефектом, а их предшественников, то есть долго живущих стволовых клеток. В частности, многообещающей является трансформация тотипотентных эмбриональных стволовых клеток, которые при создании определенных микроусловий могут дифференцироваться, практически, в любые соматические клетки организма (Hodgson, 1995). Следует упомянуть в этой связи предложенный недавно эффективный метод получения стволовых клеток гемопоэтического ряда, перспективных для генотерапии наследственных заболеваний крови (Berardi et al., 1995).

Как правило, определение типа клеток, подлежащих генетической модификации, завершается оценкой результатов переноса гена в системе *in vitro* и проведения экспериментов на животных моделях в тех случаях, когда это возможно. Апробацию процедуры генокоррекции наследственного заболевания проводят на первичных культурах экспрессирующих клеток больного либо на перевиваемых культурах, полученных после предварительной трансформации первичных культур. На этих клеточных моделях оценивают эффективность выбранной системы переноса экзогенной ДНК, определяют экспрессию вводимой генетической конструкции, анализируют ее взаимодействие с геномом клетки, отрабатывают способы идентификации первичного дефекта и его коррекции на биохимическом уровне.

Однако, многие проблемы генной терапии не могут быть решены на уровне клеток. Важное значение имеет анализ влияния введенных ДНК-последовательностей на межклеточные взаимодействия, определяющие работу соответствующих тканей и органов. Такие исследования могут быть проведены только *in vivo*. Так, например, в культуре клеток можно определить количество синтезированного белка, необходимое для нормализации биохимического дефекта, но этих данных недостаточно для ответа на вопрос, какое количество клеток в организме должно быть модифицировано для восстановления нарушенной функции. Используя культуры клеток, можно разработать биохимическую систему адресной доставки рекомбинантных ДНК, однако, проверка надежности работы этой системы может быть осуществлена только на уровне целого организма. Показатели длительности и характера экспрессии введенного гена в культуре клеток могут использоваться лишь в качестве ориентировочных параметров для оценки необходимой периодичности повторения терапевтических процедур. Кроме того, многие побочные эффекты и, в первую очередь, возможные ошибки в регуляции экспрессии чужеродного гена и опасность вирусной контаминации в результате использования компетентного по репликации вектора (см. ниже), могут быть выявлены только *in vivo*. Поэтому такое внимание в программах по генной терапии уделяется экспериментам *in vivo* на естественных или искусственно полученных моделях соответствующих наследственных болезней у животных (см. Главу VIII). Успешная коррекция генетических дефектов у таких животных и отсутствие нежелательных побочных эффектов генной терапии является важнейшей предпосылкой для разрешения кли-

нических испытаний.

Раздел 9.3 Методы генетической трансфекции в генной терапии.

Решающим условием успешной генотерапии является обеспечение эффективной доставки, то есть трансфекции (в широком смысле) или трансдукции (при использовании вирусных векторов) чужеродного гена в клетки-мишени, обеспечение длительной персистенции его в этих клетках и создание условий для полноценной работы, то есть экспрессии. Трансфекция может проводиться с использованием (1) чистой ("голой"-naked) ДНК, лигированной в соответствующую плазмиду, либо (2) комплексированной ДНК - плазмидная ДНК комплексированная с солями, белками (трансферрином), органическими полимерами (DEAE - декстраном, полилизинном, липосомами или частицами золота), либо (3) ДНК в составе вирусных частиц, предварительно лишенных способности к репликации. Залогом длительной персистенции чужеродной ДНК в клетках-реципиентах является ее встраивание в геном, то есть в ДНК клетки-хозяина. Пребывание экзогенной ДНК в ядре в свободном состоянии (в виде, так называемых, эписом) с неизбежностью ведет к ее элиминации даже в неделящихся клетках и, соответственно, к транзитной экспрессии (обычно, в течение нескольких месяцев). Необходимой предпосылкой экспрессии чужеродной ДНК является наличие соответствующих промоторов, причем в случае наличия тканеспецифических промоторов можно добиться экспрессии введенного гена только в определенных тканях и клетках (см. ниже). Основные методы доставки чужеродных генов в клетки подразделяются на химические, физические и биологические. Эффективность трансфекции и интеграционная способность трансдуцированной чужеродной ДНК при различных способах трансфекции в ДНК-клетки мишени приведены в Табл.9.1.

Таблица 9.1. Основные характеристики генетической трансфекции in vitro (Culver, 1994).

| Методы | Трансдукция | Интеграция | Экспрессия |
|---|-------------|------------|------------|
| Химические: | | | |
| Са-фосфат преципитация | низкая | низкая | транзитная |
| Физические: | | | |
| Электропорация | низкая | низкая | транзитная |
| Микроинъекция | высокая | низкая | транзитная |
| "Бомбардировка" частицами золота | высокая | низкая | транзитная |
| Слияние: | | | |
| Липосомы | низкая | низкая | транзитная |
| Рецептор-опосредо- ванный эндоцитоз: | | | |
| ДНК-белковый комплекс | высокая | низкая | транзитная |
| ДНК-комплекс- вирусная капсида | высокая | низкая | транзитная |
| Рекомбинантные вирусы: | | | |

| | | | |
|--|---------|---------|--------------|
| Аденовирус | высокая | низкая | транзиторная |
| Адено-ассоцииро- ванный вирус (AAV) | высокая | низкая | длительная ? |
| Вирус герпеса (HSV) | низкая | низкая | слабая |
| Вирус иммуно- дефицита (HIV) | высокая | высокая | длительная ? |
| Вирус мышиной лейке- мии Молони (MoMLV) | высокая | высокая | длительная ? |
| Вирус ветряной оспы (Vaccinia) | высокая | низкая | слабая |

Как следует из представленных данных, введение чужеродных генов *in vitro* может быть весьма эффективно как при помощи некоторых физических способов доставки (электропоации, бомбардировки частицами золота), так и, практически, при всех вариантах биологической доставки, особенно, с помощью рекомбинантных вирусов. Однако, реально интеграция в геном клетки-реципиента может быть достигнута только в случае ретровирусных или адено-ассоциированных векторов, обладающих необходимыми для встраивания в эукариотическую ДНК свойствами. При этом отсутствие встраивания в геномную ДНК, как правило, коррелирует с транзиторной (временной) экспрессией гена. Следовательно, только вирусные векторы или генетические конструкции, включающие вирусные последовательности способны к активной трансфекции, а в ряде случаев и к длительной экспрессии чужеродных генов. Следует напомнить в этой связи, что из 100 уже одобренных проектов генотерапии 95 предполагают использовать вирусную трансдукцию и 86 из них основаны на применении ретровирусных векторов.

Несмотря на усилия многих генно-инженерных лабораторий, Центров, а в последнее время и фармацевтических фирм, отсутствие идеальных векторов, обладающих эффективной (100%) трансфекцией как *ex vivo* так и *in vivo*, в сочетании с высокой пакующей способностью (включение генетической конструкции от 1 до 1 000 тысяч п.о.), интегрирующих в геном или неинтегрирующих, но обеспечивающих длительную и, что особенно важно, регулируемую экспрессию, при отсутствии опасности онкогенных модификаций или иных нежелательных побочных эффектов, продолжает оставаться одним из серьезных препятствий на пути внедрения генотерапии (Hodgson, 1995).

Раздел 9.4. Конструирование векторных систем и совершенствование методов трансформации клеток человека.

9.4.1. Основные векторные системы.

В Табл.9.1 приведены основные типы векторных систем. Остановимся подробнее на способах их конструирования, преимуществах и недостатках, некоторые из которых уже упоминались в предыдущем разделе. Как правило, вводимая генетическая конструкция представляет собой полноразмерную кДНК-овую последовательность определенного гена, инсертированную в экспрессионный вектор, то есть находящуюся под действием сильного промотора. Выбор подходящего промотора зависит от многих параметров, главным из которых является необходимый уровень экспрессии гена в клетках-мишенях. Вектор часто содержит один из маркерных генов, таких как гены *neo*, *beta-Gal*, ген люциферазы и др., присутствие которых в трансду-

цированных клетках может быть легко обнаружено по наличию либо соответствующего белкового продукта (гистохимически для генов бета-Gal и люциферазы), либо маркерных последовательностей ДНК. Если в качестве маркера выбран селективируемый ген (neo), отбор клеток, трансфицированных *in vitro*, может производиться автоматически на соответствующих селективных средах. Существует два типа конструкций; один - на основе плазмидной ДНК, другой - на базе вирусов. Плазмидные конструкции удобны для клонирования, генно-инженерных манипуляций и получения большого количества рекомбинантной ДНК. Однако, бактериальные плазмиды, в отличие от вирусных конструкций, не способны самостоятельно проникать в эукариотические клетки. Для введения инсертированной в плазмиду экзогенной ДНК в клетки человека необходимо перенести ее в подходящий вирусный вектор или применить способ, облегчающий ее прохождение через клеточные мембраны.

9.4.2 Методы физического переноса чужеродной ДНК в клетки эукариот.

Уместно заметить, что чужеродная ДНК может спонтанно проникать в клетки эукариот, благодаря наличию на наружных клеточных мембранах белков, специфически связывающих ДНК. Путем эндоцитоза (впячивания внутрь клеточной мембраны) чужеродная ДНК попадает в цитоплазму в составе эндосом, где обычно быстро разрушается лизосомальными ферментами. Только небольшая часть экзогенной ДНК выходит из эндосом, попадает в ядро и, если не разрушается эндогенными нуклеазами, то может быть интегрирована в ДНК клетки. Такое, однако, случается достаточно редко. Известное исключение составляют мышцы, в которых благодаря низкой активности эндогенных нуклеаз и низкой пролиферативной активности введенная ДНК долго (до 1 года) может сохраняться и даже экспрессироваться в миофибриллах (Hansen et al., 1991).

Эффективная доставка чужеродной ДНК непосредственно в ядро клетки-мишени может быть достигнута путем микроинъекции (метод применяемый сегодня почти исключительно для создания трансгенных животных путем введения экзогенной ДНК в пронуклеус оплодотворенной яйцеклетки - см.Главу VIII); при помощи электропорации (кратковременного воздействия сильным электрическим полем); путем перфорации клеточных мембран золотыми или вольфрамовыми микрочастицами конъюгированными с чужеродными ДНК и разогнанными до высокой скорости (метод бомбардировки). Эти методы доставки применимы, главным образом, для клеток, культивируемых *in vitro*. Исключение составляет лишь метод бомбардировки, который при наличии специального генного "ружья" с успехом применяется и *in vivo* (Yang et al., 1990).

Для повышения эффективности переноса обычно используют системы доставки - соединения или группы соединений, взаимодействующие с ДНК с образованием компактных структур, облегчающих проникновение ДНК в клетки и защищающих ее от действия нуклеаз (Власова и др., 1994). Самой простой системой доставки является система кальций-фосфатной копреципитации, широко применяемая для трансфекции клеток *in vitro*. Более сложный и многообещающий вариант трансфекции представляет собой рецептор-опосредованный транспорт, предусматривающий создание достаточно сложной, обычно трехкомпонентной конструкции: ДНК-поликатион + лиганд + вирус (Рис. 9.2). В качестве лигандов используются специфические белки, такие как трансферрин, эритропоэтин, асиалоглюкопротеин, конъюгирован-

ный с альбумином инсулин и некоторые другие, взаимодействующие с клеточными рецепторами и обеспечивающие фиксацию генной конструкции на специфических клетках, то есть адресную доставку чужеродной ДНК в клетки определенного типа (например асиалогликопротеин - в клетки печени, трансферрин и эритропоэтин - в клетки крови и т.д). Лиганды ковалентно присоединяются к связывающим и компактизирующим чужеродную ДНК катионным носителям (полилизину, DEAE-Dextran и др.).

Важным компонентом системы является аденовирус или его N-концевой фрагмент, выступающие в качестве эффективных эндогенных агентов, обеспечивающих выход экзогенной ДНК из эндосом после попадания ее в цитоплазму клеток-мишеней. Адресная доставка и эффективная защита от лизосомальных ферментов обеспечивают высокую трансфекционную способность таких конструкций, их несомненную перспективность для генной терапии *in vivo* (Hodgson, 1995).

Мы уже упоминали о возможности сочетания векторного и физико-химического подхода при конструировании систем для переноса генов в клетки человека. Одна из таких систем основана на использовании филаментного фага fd для трансфекции эпителиальных клеток. Гены fd, кодирующие белки оболочки фага, экспрессируются на его поверхности. В один из них встраивают последовательность, кодирующую поли-L-лизин. Полилизиновые остатки в составе слитого белка связываются с плазмидной ДНК и удерживают ее на поверхности фага. В другой ген оболочки фага встраивают последовательность ДНК, кодирующую какой-либо агент, специфически связывающийся с апикальной поверхностью эпителиальных клеток и интернализирующий (обеспечивающий проникновение) фага внутрь клетки. С этой целью были апробированы гены белков патогенных бактерий, поражающих кишечный эпителий - интерналин и инвазин, а также последовательности ДНК, кодирующие пептидные фрагменты вариабельного района моноклональных антител Ab11. Было показано, что во всех трех случаях достигается адресная доставка и интернализация фага в клетки-мишени, то есть система успешно функционирует.

Направленный перенос генов во многие типы клеток, содержащие трансферриновые рецепторы, может быть осуществлен при комплексировании ДНК с трансферрином. Использование в этом комплексе аденовирусного вектора существенно облегчает прохождение ДНК через эндосомы и попадание её в ядро. Идеальными белковыми лигандами для специфических клеточных рецепторов могут служить моноклональные антитела или их фрагменты, направленные против тех элементов рецепторов, которые находятся на наружной поверхности клеточной мембраны. Подобная система разработана для рецептор-опосредованного генного переноса в эпителиальные клетки. Она основана на использовании против-секреторных SCFab-фрагментов антител для полимерного иммуноглобулинового рецептора pIgR. Этот рецептор транспортирует IgA и IgM в респираторные эпителиальные клетки, связывая иммуноглобулины и интернализируя их путем эндоцитоза. Показано, что в системе *in vitro* частота трансфекции эпителиальных клеток при использовании SCFab-поли-L-лизин-ДНК комплекса такая же, как и при введении экзогенной ДНК посредством трансферринового рецептора. Аналогичные подходы могут быть применены для введения генов и в другие типы клеток.

9.4.3 Липосомный метод трансфекции.

Эффективный внутриклеточный транспорт и защита от деп-

радации лизосомальными ферментами достигаются при использовании в качестве векторов липосом-липидных пузырьков, обладающих выраженными фузогенными свойствами - способностью сливаться с клеточными мембранами. Особенно перспективны в этом отношении липосомы, полученные на основе катионных липидов, обеспечивающих 100% связывание ДНК в конденсированные нуклеолипидные частицы. Положительный заряд на поверхности таких пузырьков обеспечивает их активное слияние с отрицательно заряженными клеточной мембранами и прямое попадание чужеродной ДНК в цитоплазму, минуя эндосомы и, соответственно, не подвергаясь действию лизосомных гидролаз. Очень эффективный перенос высокоочищенных ДНК или РНК-последовательностей в соматические, особенно, в мышечные ткани может быть осуществлен с помощью препаратов липофектина или липофектамина (Caplen et al., 1994). Гораздо более высокая частота трансфекции по сравнению с липосом-опосредованным переносом получена в экспериментах на культурах клеток при использовании ДНК-липидного комплекса с циклическим амфипатическим пептидом грамицидином S.

Особенно перспективными в последнее время представляются комплексы, в которых липосомы конъюгируют с мембранными антителами к определенным белкам-мишеням (иммунолипосомы) либо с белками-лигандами (см. выше). Эти конструкции обеспечивают эффективную адресную доставку чужеродной ДНК в клетки-мишени. Подобная схема была успешно апробирована для переноса гена сывороточного альбумина человека в гепатоциты линейных крыс Nagase с наследственной дисальбуминемией. Доказано присутствие и экспрессия введенного таким образом гена человека в клетках печени крыс. Аналогичные результаты получены в опытах на линейных кроликах Watanabe, дефектных по рецепторам липопротеинов низкой плотности - LDL. Эти животные моделируют одно из наиболее частых моногенных заболеваний человека - семейную гиперхолестеринемия. При внутривенной инъекции кроликам липидного асиалогликопротеинового комплекса с плазмидной ДНК, несущей нормальный LDL-ген, уровень холестерина в крови животных устойчиво снижался.

Важным преимуществом рецептор-опосредованных систем на основе липосом является их низкая иммунореактивность. Они лишены и многих других недостатков, свойственных вирусным векторным системам. Вместе с тем, до сих пор не решена проблема низкой частоты трансформации клеток при липосомном переносе. Это обстоятельство существенно ограничивает применение липосом в генной терапии (Crystal, 1995). Тем не менее, в настоящее время рецептор-опосредованный вариант передачи генетической информации в клетки эукариот с использованием в качестве лигандов специфических антител, рецепторных белков, а также вирусных последовательностей и липосом позволяет в одной системе совместить преимущества физико-химических методов переноса ДНК и вирусных векторов и потому представляет один из наиболее перспективных и быстро развивающихся направлений в трансфекции эукариотических клеток.

9.4.4 Рекомбинантные вирусы.

Конструирование векторов на базе вирусов представляет собой наиболее интересный и перспективный раздел генотерапии. Эволюционно сложившаяся система обеспечения эффективного проникновения в клетки-мишени, а в случае ретровирусов и система интеграции в клеточный геном, позволяет рассматривать вирусы как естественные векторы чужеродной ДНК для клеток млекопитающих. Действительно, только с помощью вирусных

векторов пока удается достичь такого уровня трансфекции клеток человека *in vitro* и *in vivo*, который необходим для проявления лечебного эффекта. Это доказывают многочисленные эксперименты на животных и первые клинические испытания утвержденных программ генотерапии (см. 9.2). Вместе с тем, нельзя недооценивать и вполне реальную опасность патологических процессов, связанных с использованием вирусных частиц. В качестве векторов применяют следующие рекомбинантные вирусы: ретровирусы, аденовирусы, аденоассоциированные вирусы, вирус герпеса, вирус СПИДа (HIV), вирус ветряной оспы и некоторые другие (Anderson, 1992; Culver, 1994; Lowenstein, 1994; Hodgson, 1995; Kay, Woo, 1992). Учитывая большую практическую значимость этих векторов, рассмотрим их более подробно.

Ретровирусные векторы. . Генные конструкции на основе ретровирусов (РНК-содержащих вирусов) особенно часто используются для трансдукции ДНК *ex vivo*. Наиболее популярный ретровирусный вектор - вирус мышиного лейкоза Molony (MoMLV). По сравнению с другими типами векторов ретровирусы обладают уникальной способностью эффективно переносить чужеродные гены и стабильно интегрировать их в геном делящихся соматических клеток. Для безопасности ретровирусные последовательности модифицируют таким образом, что в инфицированных ими клетках вирусные белки не производятся. Это достигается за счет удаления или инактивации всех кодирующих последовательностей вируса. Репликация вирусных векторов может происходить только в специальных "пакующих" клетках, в геном которых встроены все гены, производящие вирусные белки. При введении ретровирусных векторов в эти клетки образуются вирионы, несущие векторную РНК и способные лишь проникать в клетки-мишени, но не размножаться в них. Недостатком этой системы, также как и других векторных систем на основе вирусов, является возможность контаминации производящей клеточной линии нормальным ретровирусом и получения на этой основе компетентного по репликации вектора. Для предотвращения этого необходимо регулярное тестирование "пакующей" линии клеток. Возможна также контаминация ретровирусного вектора клеточными РНК, некоторые из них могут обратно транскрибироваться и встраиваться в геном трансдуцированных клеток. Последствия такого события могут быть выявлены в экспериментах на животных моделях. Другими серьезными недостатками ретровирусных векторов является: (1) их способность переносить генетический материал только в пролиферирующие клетки; (2) способность индуцировать мутации при случайной интеграции в геном; (3) возможность спонтанной активации онкогена; (4) небольшие размеры переносимой ДНК-вставки - до 8 тысяч п.о.; (5) сравнительно низкий титр 10^{6-7} /мл рекомбинантных вирусных частиц, получаемых для трансдукции; (6) необходимость конструирования соответствующих "пакующих" клонов клеток.

Аденовирусные векторы. . В отличие от ретровирусов аденовирусы активно инфицируют неделящиеся клетки, обладают большей потенциальной пакующей способностью (ДНК-вставка > 8 кб), имеют высокий титр 10^{11} /мл, однако, не обеспечивают встраивание чужеродной ДНК в геном трансформированной клетки (Hodgson, 1955). Использование их перспективно для генотерапии клеток верхних дыхательных путей, легких и других органов - мозга, печени, мышц, кожи и пр. Они эффективны при доставке аэрозольным способом, что было использовано в первых клинических испытаниях по генотерапии муковисцидоза (Crystal et al., 1994). В аденовирусные векторы также инсертируют маркерные гены - neo, CAT или бета-галактозидазный

ген (бета-Gal) для того, чтобы идентифицировать трансдуцированные клетки. Для конструирования векторов используют дефектные по репликации аденовирусы, которые получают путем вырезания из вирусной ДНК генов, кодирующих белки (E1a, E1b) – так называемые аденовирусные векторы 2-го поколения. В настоящее время создаются аденовирусные векторы 3-го поколения, в которых помимо генов E1a и E1b удаляют и регуляторный ген E4. Такая конструкция может поддерживаться только в присутствии клеток-хелперов (например, в культуре клеток почек человека). Удаление большого числа аденовирусных генов из векторов часто сопровождается их дестабилизацией. Это один из главных недостатков аденовирусных векторов, так как в ряде случаев остаются гены, трансдуцированные в клетки-мишени, способствуют формированию иммунного ответа. Именно выраженный иммунный ответ при повторных введениях аденовирусного вектора с инсерцией гена CFTR, оказался наиболее серьезным препятствием для успешной генотерапии муковисцидоза (Crystal et al., 1994). Некоторые аденовирусные белки способны оказывать цитотоксический эффект на высокоспециализированные клетки человека. Схема поддержания аденовирусных векторов сходна с той, которая используется для производства ретровирусных векторов. Велика опасность их контаминации хелперным реплицирующимся вирусом. Кроме того, аденовирусы редко интегрируются в геномную ДНК и потому экспрессия переносимых ими генов, как правило, носит временный характер (Табл. 9.1). Способность инфицировать, практически, любые клетки как *in vivo*, так и *in vitro* делает особенно актуальной адресную доставку таких конструкций и введение в их состав тканеспецифических промоторов. Например, промотор гена альфа-фетопротеина при необходимости экспрессии гена в клетках печени, либо промоторы генов сурфактантных белков В и С для экспрессии чужеродных генов в клетках легких.

Аденоассоциированные вирусы. Обладают значительно меньшей пакующей способностью (около 5 кб). Однако, в отличие от ранее рассмотренных вирусов не обладают онкогенной активностью, не патогенны, способны интегрироваться в геном, где пребывают в латентном состоянии. Уникальной особенностью AAV является их способность к стабильной неслучайной интеграции в один из районов хромосомы 19. Специфичность интеграции вируса определяется наличием в его геноме гена *her*. Близко родственные AAV, так называемые, парвовирусы (H1, MVM, Lu-III), обладают еще меньшей пакующей способностью – около 2 кб и не имеют специфичного встраивания, однако, они также рассматриваются как потенциально перспективные векторы.

Вирус простого герпеса (HSV). Крупный (152 кб) ДНК-содержащий вирус, при трансформации не интегрируется в геном, формируя в ядрах эпизомные структуры. Уникальная особенность HSV является его выраженная тропность к клеткам нервной системы. Отсюда его перспективность как векторной системы для лечения опухолей мозга, болезни Паркинсона и многих других. Его известное преимущество – достаточно большая пакующая способность (>30кб). Важным этапом в создании вектора на основе HSV является удаление из его генома области ICP22, ответственной за синтез литических белков, и индукция мутации 1814, вызывающей блок транскрипции вирусной ДНК. В последнее время на основе HSV стали получать искусственные производные вируса, так называемые ампликон-продукты, лишенные, практически, всех генов HSV, но способные к репликации.

Конструкции векторов, используемых для переноса экзогенных ДНК в клетки человека, постоянно совершенствуются в

зависимости от типа клеток-мишеней. Так, новый тип векторов, сконструированных на основе псевдо-аденовирусов, сочетает в себе все преимущества аденовирусных векторов, но при этом собственные вирусные гены, практически, не оказывают никакого повреждающего эффекта на трансфецированные клетки-мишени, так как содержат очень мало регуляторных элементов и последовательностей, ответственных за упаковку и репликацию аденовируса. Кроме того, псевдо-аденовирусные векторы с успехом переносят чужеродные последовательности ДНК как в делящиеся, так и в покоящиеся клетки. Изучаются также перспективы использования для генной терапии других вирусных систем, таких как SV40, вирус иммунодефицита (HIV), вирус ветряной оспы и многие другие. В частности, заслуживают внимания эписомные (неинтегрирующиеся в геном реципиента) векторы, полученные на основе очень крупного вируса Эпштейн-Барра, способного нести вставку размером от 60 до 330 кб (Sun et al., 1994).

9.4.5 Перспективы создания "идеальных" векторных систем.

Обзор существующих данных позволяет прийти к заключению, что, несмотря на усилия многих лабораторий мира, все уже известные и испытанные *in vivo* и *in vitro* векторные системы далеки от совершенства (Hodgson, 1995). Если проблема доставки чужеродной ДНК *in vitro* практически решена, а ее доставка в клетки-мишени разных тканей *in vivo* успешно решается (главным образом, путем создания комбинированных рецептор-опосредованных конструкций), то другие характеристики существующих векторной системы - стабильность интеграции, регулируемая экспрессия, безопасность - все еще нуждаются в серьезных доработках. Прежде всего, это касается стабильности экспрессии. Последняя может быть достигнута либо при интеграции чужеродной ДНК непосредственно в геном реципиента, либо путем обеспечения длительной персистенции экзогенной ДНК в ядре. До настоящего времени интеграция в геном достигалась только при использовании ретровирусных либо адено-ассоциированных векторов (Табл. 9.1). Случайная интеграция трансфектной ДНК в геном происходит достаточно редко, причем в случае ретровирусных векторов это происходит только в делящихся клетках. Повысить эффективность стабильной интеграции можно путем совершенствования генных конструкций типа рецептор-опосредованных систем (Рис. 9.2). Однако, эти векторные конструкции должны включать только часть вирусных генов, например, гены обратной транскриптазы, ретровирусной интегразы, некоторые транспозоновые гены, парвовирусные гер-гены (см. 9.4.4). Учитывая возможный мутагенный эффект случайной интеграции, весьма перспективным представляется создание достаточно стабильных эписомных векторов. В частности, в последнее время особое внимание уделяется созданию векторов на базе искусственных хромосом млекопитающих (Mammalian Artificial Chromosomes), которые могли бы достаточно автономно находиться в ядре, сохраняя способность к репликации и экспрессии. Удобными моделями для этого представляются автономно реплицирующиеся циркулярные микрохромосомы раковых клеток (Hodgson, 1995).

Особенно привлекательной в плане генной коррекции представляется возможность замены всего мутантного гена или его мутировавшей части (например, одного экзона) на нормальный аналог, что может быть достигнуто путем гомологичной рекомбинации. При этом в идеале можно ожидать не только длительную персистенцию введенного гена, но и сохранение нор-

мальной экспрессии. С этой целью в конструкции, используемые для переноса ДНК, включают агенты, повышающие частоту гомологичного спаривания, например, бактериальную рекомбиназу. Показано, что в этих условиях частота гомологичной рекомбинации может превышать $2.5 \cdot 10^{-4}$. Это достаточно для того, чтобы с помощью ПЦР отобрать нужные клоны клеток. Для направленного введения фрагментов гена в строго определенные локусы генома недавно разработана система двойной замены, основанная на использовании HPRT-зависимых эмбриональных стволовых клеток и векторной конструкции содержащей ген HPRT (гипоксантин-фосфорибзилтрансферазы) и ген тимидин-киназы вируса герпеса (HSV). Двойная селекция трансформантов позволяет отобрать клетки, в которых произошла гомологичная рекомбинация. Такой подход нашел широкое применение при создании искусственных моделей наследственных болезней у человека (см. подробнее Главу VIII). Однако, в клинической практике он еще не используется.

Раздел 9.5 Генотерапия моногенных наследственных заболеваний.

Вопросы генотерапии наследственных заболеваний подробно рассмотрены в многочисленных обзорах (Ledley, 1987; Anderson, 1992; Pyeritz, 1993; Breakefield, 1993; Lowenstein, 1994; Kay, Woo, 1994; Brown et al., 1994; Дризе, 1994; Crystal, 1995) и достаточно полно суммированы в недавно опубликованной монографии (Culver, 1994).

Уже через год после первого введения маркерного гена в организм человека была проведена успешная клеточная соматическая генотерапия наследственного заболевания, обусловленного дефицитом аденозиндезаминазы (ADA) (см. 9.1). При этом заболевании в крови пациентов накапливается в высокой концентрации 2'-дезоксаденозин, оказывающий токсическое действие на Т- и В- лимфоциты, в результате чего у больных развивается серьезный комбинированный иммунодефицит. Для поддержания жизни пациентов проводят периодические гетерологичные трансплантации клеток костного мозга, однако, лишь для трети больных могут быть подобраны совместимые доноры. Энзим-замещающая терапия также приводит к заметному улучшению состояния пациентов, но, как правило, успех этот носит временный характер. План генной терапии, разработанный сотрудниками Национального Института Здоровья США (NIH) и одобренный RAC, заключался в назначении больным аутологичных лимфоцитов, трансдуцированных нормальным ADA-геном. Осуществление этого плана потребовало выполнения следующих процедур: изоляции клеток из крови пациента; активации и иммуностимуляции роста Т-лимфоцитов в культуре; трансдукции их ретровирусным вектором, несущим нормальный ADA-ген и маркерный ген нео; отбора трансдуцированных клеток на селективной среде; внутривенной реинфузии модифицированных Т-лимфоцитов пациенту. Первой пациенткой, подвергшейся этой терапии, была 4-х летняя девочка (см. раздел 9.1). На протяжении 10.5 месяцев ей было сделано 8 аутологичных вливаний трансдуцированных Т-лимфоцитов и после полугодового перерыва программу реинфузий повторяли каждые 3-5 месяцев. Уже после первого цикла число Т-лимфоцитов нормализовалось, концентрация ADA в циркулирующих клетках крови увеличилась с 1% до 20 - 25% нормального уровня и резко улучшились основные иммунные характеристики. Вопреки многим прогнозам, на протяжении более, чем 6 месяцев после прекращения массивных вливаний в кровотоке пациентки устойчиво сохранялось высокое число кор-

ректированных Т-клеток, что позволило в дальнейшем снизить количество вводимых клеток и значительно увеличить промежутки между этими процедурами. Спустя три месяца после первых клинических испытаний была начата программа генной терапии ADA-дефицита у второй 9-летней пациентки. После 11 инфузий трансдуцированных аутологичных Т-клеток состояние этой девочки также заметно улучшилось и отмечалась полная нормализация соответствующих биохимических и иммунологических показателей. Таким образом, необходимо еще раз отметить, что при лечении обеих пациенток был достигнут очевидный клинический эффект (Anderson, 1992; Culver, 1994).

Однако, в обоих случаях не все иммунные функции восстановились полностью. По-видимому, это было связано с тем, что коррекция генетического дефекта проводилась в зрелых Т-лимфоцитах. В связи с этим предложены программы генной терапии с помощью реинфузии смешанной популяции трансдуцированных Т-лимфоцитов и периферических стволовых клеток крови. Возможность изоляции и трансдукции таких тотипатентных стволовых клеток показана в экспериментах на приматах.

Успех первых клинических испытаний явился мощным стимулом для ускорения развития новых генотерапевтических методов применительно к другим наследственным заболеваниям. В Табл. 9.2 представлен список болезней, для которых принципиально возможен генотерапевтический подход и генокоррекция наследственного дефекта с большой вероятностью будет осуществлена уже в обозримом будущем, а также те заболевания, для которых уже имеются официально утвержденные протоколы и которые находятся на разных стадиях клинических испытаний.

Таблица 9.2. Наследственные заболевания, генокоррекция которых находится на стадии клинических испытаний (КИ), экспериментальных разработок (ЭР) и принципиально возможна (ПВ). (Culver, 1994; Lowenstein, 1994)

| | Болезнь | Дефектный ген | Клетки-мишени | Стадия |
|----|------------------------------|---|-----------------------------|--------|
| 1 | Иммунодефицит | аденозиндезаминаза | лимфоциты | КИ |
| 2 | Иммунодефицит | пуриннуклеозид-фосфорилаза | лимфоциты | ПВ |
| 3 | Семейная гиперхолестеринемия | рецептор липопротеинов низкой плотности | гепатоциты | КИ |
| 4 | Гемофилия В | фактор 1X | фибробласты | КИ |
| 5 | Гемофилия А | фактор Y111 | миобласты, фибробласты | ЭР |
| 6 | Болезнь Гоше (сфинголипидоз) | в-глюкоцереброзидаза | макрофаги, стволовые клетки | КИ |
| 7 | Болезнь Хантера | идуронат-сульфатаза | макрофаги, стволовые клетки | ПВ |
| 8 | Синдром Гурлера | L-идуронидаза | макрофаги, стволовые клетки | ПВ |
| 9 | Эмфизема легких | альфа-1-антитрипсин | лимфоциты | ЭР |
| 10 | Муковисцидоз | CF-трансмембранный регулятор | эпителий бронхов | КИ |
| 11 | Фенилкетонурия | фенилаланингидроксилаза | гепатоциты | ЭР |
| 12 | Гипераммонемия | орнитинтранскарбамилаза | гепатоциты | ПВ |
| 13 | Цитрулинемия | аргиносукцинатсинтетаза | гепатоциты | ПВ |
| 14 | Мышечная дистрофия Дюшенна | дистрофин | миобласты, миофибриллы | ЭР |
| 15 | Талассемия | бета-глобин | эритробласты | ЭР |

| | | | | |
|----|--|---|--|----|
| 16 | Серповиднокле- точная анемия | бета-глобин | эритробласты | ЭР |
| 17 | Респираторный дистресс-синдром | сурфактант белок В | эпителий бронхов | ЭР |
| 18 | Хронический грануломатоз | NADPH-оксидаза | гранулоциты | ЭР |
| 19 | Болезнь Альцгеймера | белок-предшественник в-амилоида (ААР) | нервные клетки | ЭР |
| 20 | Болезнь Паркинсона | тирозин-гидроксилаза | миобласты, фибробласты | ЭР |
| 21 | Метахромати- ческая лейко- дистрофия | арилсульфатаза А | стволовые клетки крови, нервные клетки | ПВ |
| 22 | Синдром Леш- Нихана | гипоксантин-фосфо- рибозил трансфераза | нервные клетки | ПВ |

Как следует из данных Таблицы 9.2, на стадии клинических испытаний в 1994г. уже находились 5 моногенных заболеваний. Для 10 генных болезней проводились экспериментальные исследования и отработывались требования, необходимые для получения официального разрешения клинических испытаний (см.9.1). Исследования по остальным заболеваниям находятся на начальных этапах. Список таких заболеваний очень быстро увеличивается. Обращает на себя внимание, что первые программы по генной терапии связаны с модификацией гемопоэтических клеток (Wivel, Walters, 1993). Клетки крови наиболее доступны для генетических манипуляций. После изоляции различные типы клеток крови могут быть легко размножены, подвергнуты трансфекции *in vitro*, а затем возвращены пациенту. Генетической модификации могут быть подвергнуты не только зрелые клетки (лимфоциты, макрофаги), но и их предшественники - стволовые клетки. Важным обстоятельством, в этой связи, является то, что процедура трансплантации клеток костного мозга уже широко используется в клинике. Разработаны и достаточно эффективные методы выделения стволовых гемопоэтических клеток человека (Berardi et al., 1995). В экспериментах на животных показано, что модифицированные клетки как миелоидного, так и лимфоидного рядов могут сохраняться в кровотоке на протяжении более двух лет после аутологичной пересадки клеток костного мозга, трансдуцированных *in vitro*. Путем трансфекции клеток крови соответствующими генами можно лечить не только собственно заболевания крови, но и использовать их для лечения многих других заболеваний как моногенной природы (Табл. 9.2), так и различных опухолей и инфекций (см.ниже).

Другими достаточно универсальными реципиентами чужеродных генов могут быть фибробласты и мышечные клетки (миобласты, миофибриллы). Они могут быть использованы для тех заболеваний, где необходима коррекция генов, белковые продукты которых должны поступать в сыворотку крови или дифундировать в соседние клетки. Особенно удобны для целей генной терапии скелетные мышцы, в которых благодаря отсутствию эндонуклеазной активности (см.раздел 9.4.2) принципиально возможен перенос генов *in vivo* путем прямой инъекции экзогенной ДНК. Инъекционная в мышцы ДНК способна экспрессироваться в миофибриллах находясь в неинтегрированном, эксхромосомном состоянии. Белковые продукты экспрессии в течение длительного времени после трансдукции будут поступать в кровь. Продолжительность экспрессии значительно увеличивается, если генетическую модификацию производят в аутологичных миоблас-

тах, которые после этого инъецируют в зрелую мышцу. Эти особенности уже позволили начать эксперименты по генной терапии таких заболеваний как гемофилии А и В, дефицит антитрипсина, диабет, врожденный дефицит гормона роста и даже болезнь Паркинсона (Culver, 1994; Lowenstein, 1994). Достаточно удобными для генетических модификаций оказались и фибробласты кожи, в первую очередь, благодаря легкости генноинженерных манипуляций *ex vivo*.

Раздел 9.6. Генотерапия ненаследственных заболеваний: опухоли, инфекции.

Параллельно с развитием исследований в области генокоррекции наследственных дефектов успешными также оказались поиски методов терапевтического использования смысловых последовательностей ДНК для лечения ненаследственных заболеваний и, главным образом, злокачественных опухолей и вирусных инфекций. Существенно, что именно в этих разделах патологии поиски путей генокоррекции проводятся особенно интенсивно, а число уже одобренных протоколов клинических испытаний во много раз превышает число таковых для лечения моногенных болезней (см. Рис. 9.1). Такое положение дел, по-видимому, прежде всего объясняется широкой распространенностью онкологических заболеваний и отсутствием достаточно эффективной терапии. В Табл. 9.3 перечислены основные методологические подходы к генотерапии различных опухолей, разработанные и широко используемые уже на современном этапе. Многие из этих подходов вполне приложимы и для борьбы с наиболее серьезными инфекционными заболеваниями, например, со СПИДОМ.

Таблица 9.3. Основные методологические подходы в генокоррекции онкологических заболеваний.

| П Р И Н Ц И П | ВВОДИМЫЕ ГЕНЫ |
|---|---|
| 1. Повышение иммунореактивности опухоли | гены чужеродных антигенов, цитокинов |
| 2. Генетическая модификация иммунных клеток | гены цитокинов, ко-стимуляторов |
| 3. Инсерция генов "чувствительности" либо генов "самоубийц" | гены тимидин-киназы HSV, цитозин дезаминазы |
| 4. Блок экспрессии онкогенов | антисмысловые Ki-ras мРНК, гены внутриклеточных антител |
| 5. Инсерция генов-супрессоров опухолей | p53 |
| 6. Защита нормальных клеток от химиотерапии. | гены лекарственной устойчивости тип 1. |
| 7. Индукция синтеза противоопухолевых в-в нормальными клеткам | гены интерлейкина-2, интерферона |
| 8. Продукция противоопухолевых рекомбинантных вакцин. | вакцины типа ВЦЖ, экспрессирующей опухолевой антиген |
| 9. Локальная радиопротекция нормальных тканей с помощью антиоксидантов. | гены трансферазы, глутатион синтетазы |

Подробный анализ используемых при этом подходов и результаты первых клинических испытаний выходит за рамки нашего изложения. Однако, материал этот настолько интересный и многообещающий, что мы позволим себе на нескольких примерах

охарактеризовать основные принципы построения таких генотерапевтических программ.

Как упоминалось ранее (см. 9.1), перенос гена в организм человека был осуществлен в 1989 году в большей степени в исследовательских, а не в терапевтических целях. Это был маркерный прокариотический ген *neo*, сообщающий клеткам устойчивость к неомицину. Он был введен пациенту, страдающему злокачественной меланомой, в составе трансдуцированных TIL-клеток (Т-лимфоцитов, полученных из опухолевых тканей больного). В 1986г. вскоре после идентификации этого нового класса иммунных клеток, была предпринята попытка лечения меланомы путем аутологичной внутривенной инфузии TIL-клеток, предварительно выделенных из опухолей пациентов и интенсивно наращиваемых *in vitro* в присутствии ростового фактора IL-2. Примерно у трети пациентов лечение оказалось эффективным, хотя в последующем наблюдали значительное число рецидивов заболевания. Для анализа причин терапевтического эффекта TIL-клеток и совершенствования методики лечения меланомы необходимо было исследовать устойчивость вводимых Т-лимфоцитов и их миграцию в организме больного. С этой целью была произведена маркировка используемых для лечения TIL-клеток путем их трансдукции в культуре ретровирусным вектором, несущим ген *neo*, с последующим отбором неомицин-устойчивых клонов и выращиванием их на среде G418. Результаты исследований показали, что реинфузированные G418-устойчивые TIL-клетки действительно проникают в опухоль и могут быть обнаружены там в небольшом количестве даже спустя 9 недель после введения. Найдены отличия субпопуляции Т-лимфоцитов в опухоли от общей популяции инфузированных TIL-клеток.

После успешного испытания переноса маркерного гена *neo* в опухолевые ткани путем аутологичной реинфузии трансфицированных Т-лимфоцитов лечение меланомы было дополнено введением в вектор мышинного гена, контролирующего продукцию, так называемого фактора некроза опухоли - TNF. Преодолагалось, что локальная секреция этого токсичного для клеток белка в опухолевых тканях будет способствовать формированию иммунного ответа. Опасность данной терапевтической процедуры обусловлена возможностью разрушения TIL-клеток в печени, мозге и легких. Поэтому экспрессия TNF-гена под гетерологичным промотором может оказать сильный токсический эффект в этих органах. Первые клинические испытания описанной схемы лечения начаты в январе 1991 года в Национальном Институте Здоровья (NIH) США.

Другая программа генной терапии, предложенная для лечения меланом, основана на стимуляции противоопухолевого иммунитета, опосредованного Т-лимфоцитами. Для этого в изолированные опухолевые клетки пациента вводят TNF- или IL2-ген или какие-либо другие гены, секретирующие цитокины, и затем проводят иммунизацию пациента путем подкожного введения трансдуцированных клеток. Эта процедура сама по себе может привести к рассасыванию первичной опухоли или может быть использована для изоляции более эффективных TIL-клеток из лимфоузлов, вблизи от места инъекции. Подобная иммунизация может быть рекомендована для предотвращения рецидивов у пациентов, подвергавшихся другим курсам противоопухолевой терапии. Первая попытка прямого переноса гена в опухолевые клетки пациента без их предварительной изоляции также была предпринята с целью формирования иммунного ответа против злокачественной меланомы. Процедура включала прямую инъекцию в опухоль липосом-плазмидного комплекса с геном, контролирующим отсутствующий у пациента антиген гистосовместимости

HLA-B7. Другой тип модификации опухолевых клеток основан на введении в них гена тимидинкиназы Герпеса. Использованный в работе ретровирусный вектор обеспечивал включение генной конструкции только в активно пролиферирующие клетки, каковыми и являются клетки опухоли. Впервые эта схема была апробирована при лечении карциномы яичника. После интраперитонеальной аутологичной инъекции трансдуцированных клеток злокачественной карциномы пациентам назначали противогерпесный препарат – ганцикловир, избирательно убивающий клетки, экспрессирующие ген вирусной тимидин-киназы. Противоопухолевый эффект был обусловлен летальным действием токсина, образующегося в модифицированных клетках и последующей иммунной реакцией организма на опухолевые клетки.

Подходы, используемые для лечения вирусных инфекций путем введения в организм человека специфических нуклеиновых кислот, очень разнообразны и основаны на детальном исследовании молекулярных механизмов взаимодействия инфицирующих агентов с клетками-хозяина. Мы лишь коротко перечислим основные принципы, используемые при разработке соответствующих медицинских протоколов. Наибольшее количество противовирусных программ генной терапии предложено в рамках борьбы со СПИДом, хотя аналогичные методы разрабатываются для лечения гепатита, цитомегаловирусных, герпесных и иных вирусных инфекций. Одна из первых таких программ была направлена на разрушение регуляторных механизмов репликации вируса иммунодефицита – HIV, путем введения в Т-лимфоциты от 20 до 50 копий TAR-гена, кодирующего активирующий элемент, критический для переключения генетической программы клетки на вирусную репликацию. Другая программа включала введение в Т-лимфоциты гена CD4 вирусного антигена для специфического связывания HIV и выведения его в русло крови. Ряд программ основаны на введении в Т-клетки условно летальных генов, таких как ген вирусной тимидинкиназы, с тем, чтобы предотвратить нежелательные побочные эффекты в случае неконтролируемого размножения этих клеток или слишком сильного их действия на HIV-инфицированные клетки. Одним из направлений повышения эффективности терапевтического использования Т-лимфоцитов для лечения СПИДа является направленная модификация *ex vivo* генов главного комплекса гистосовместимости и конструирование на этой основе "универсальных донорских" клеток. Так, лишенные HLA-маркеров гетерологичные модифицированные Т-клетки могут быть трансплантированы пациентам без опасения иммунологической несовместимости. Подобный подход может оказаться эффективным при необходимости гетерологичной трансплантации в терапевтических целях любых типов клеток. Принципиально иным способом борьбы с вирусными инфекциями является введение в пораженные ткани антисмысловых последовательностей, способных гибридизоваться с вирусами и, таким образом, их нейтрализовывать (Cohen, Hogan, 1994; Wagner, 1994). Адресность доставки таких последовательностей может быть достигнута путем их комплексования с соответствующими белковыми лигандами (см. 9.4.2).

Раздел 9.7. Некоторые этические и социальные проблемы генной терапии.

Появление принципиально новых технологий, позволяющих активно манипулировать с генами и их фрагментами, обеспечивающими адресную доставку новых блоков генетической информации в заданные участки генома, совершило революцию в биологии и медицине. Как следует из вышеизложенного, сам ген все

чаще начинает выступать в качестве лекарства, применяемого для лечения не только моногенных, но и многих других, в том числе и значительно более распространенных недугов (опухоли, инфекции). Не за горами применение генотерапии и для борьбы с мультифакториальными заболеваниями (сердечно-сосудистые, психические, эндокринологические и многие другие). Уже сейчас, на современном уровне наших знаний о геноме человека теоретически вполне возможны такие его модификации путем генной трансфекции, которые могут быть предприняты с целью улучшения ряда физических (например, рост), психических и интеллектуальных параметров. Таким образом, современная наука о человеке на своем новом витке развития вернулась к идее "улучшения человеческой породы", когда-то постулированной выдающимся английским генетиком Фрэнсисом Гальтоном и развитой его учениками и последователями (Карл Пирсон, Лионель Пенроуз, Дж.Халдэй и др.). Дальнейший ход истории, как известно, полностью дискредитировал саму идею "улучшения" человеческой породы. Однако, грядущее "всевластие" человека над собственным геномом заставляет вновь и вновь возвращаться к этой теме, делают ее предметом постоянных оживленных дискуссий в широкой и научной печати (Ledley, 1987; Anderson, 1992; Wivel, Walters, 1993; Culver, 1994). Развернувшаяся в этой связи дискуссия позволяет подвести некоторые итоги и сделать определенные прогнозы.

Уже сейчас не вызывает сомнения, что первоначальные опасения, связанные с генной инженерией вообще и генной инженерией человека в частности были неоправданны. После многолетней дискуссии и всестороннего рассмотрения на разных уровнях было признанным целесообразным применение генной терапии для лечения многих заболеваний. Единственным и непременным ограничением, сохраняющим свою силу и в современных условиях, является то, что все генотерапевтические мероприятия должны быть направлены только на конкретного больного и касаться исключительно его соматических клеток.

По глубокому убеждению основных авторитетов генной терапии (Фр.Андерсон, Т.Каски, Фр.Коллинс, Дж.Вильсон и др.), а также согласно существующим регламентациям соответствующих "разрешительных" комитетов по генно-инженерным исследованиям (см. 9.1) современный уровень наших знаний не позволяет проводить коррекцию генных дефектов на уровне половых клеток и клеток ранних доимплантационных зародышей человека. Причина этого - реальная опасность засорения генофонда нежелательными искусственными генными конструкциями или внесение мутаций с непредсказуемыми последствиями для будущего человечества.

Вместе с тем, по мере совершенствования методов генной терапии, появления новых технологий, связанных с созданием более эффективных и безопасных векторных систем и более совершенных генетических конструкций, стремительным ростом объема информации о структуре генома, картировании новых генов в научной литературе все чаще и все настойчивее раздаются призывы к возобновлению дискуссии о целесообразности генокоррекции зародышевых и половых клеток человека (Wivel, Walters, 1993; Latchman, 1994).

Основным аргументом в пользу таких вмешательств является тот вполне очевидный факт, что по мере того как все большее число наследственных заболеваний будет доступно эффективной генной терапии, все большее число особей, гомозиготных по летальным мутантным генам, будет накапливаться в популяции. Соответственно, тем реальней будут ситуации, когда оба супруга окажутся гомозиготными носителями мутантного ге-

на. В этом случае получение здорового потомства потребует генетического вмешательства уже на ранних стадиях и, возможно, будет вполне безопасной и реальной трансфекция гамет или ранних зародышей.

Эксперименты на животных по созданию искусственных биологических моделей наследственных болезней (см.Главу VIII), а также первые клинические испытания по доимплантационной диагностике генных болезней (Verlinsky, Kuliev, 1993; см. Главу VI) убеждают в том, что такой генно-терапевтический подход может быть реальным уже в ближайшем будущем. Вполне естественно, что целесообразность его применения должна определяться не только генно-инженерными возможностями, но и его социальной значимостью и необходимостью. Вот только некоторые вопросы, которые должны быть решены в рамках предлагаемой генетиками широкой дискуссии:

Сможет ли в будущем генная терапия обеспечить столь полноценную генокоррекцию, которая не представит угрозы для потомства?

В какой мере полезность и необходимость генотерапевтической процедуры для одной супружеской четы перевесят риск такого вмешательства для всего человечества?

Сколь оправданы будут эти процедуры на фоне грядущего перенаселения планеты ?

Как будут соотноситься генно-инженерные мероприятия на человеке с проблемами гомеостаза общества и биосферы?

Таким образом, генетическая революция апофеозом которой явилась генотерапия не только предлагает реальные пути лечения тяжелых наследственных и ненаследственных недугов, но и в своем стремительном развитии ставит перед обществом новые проблемы, решение которых необходимо уже в ближайшем обозримом будущем.