

Генетика наследственных болезней человека

Эта статья посвящена молекулярной генетике наследственных болезней человека и обращена к читателю, не имеющему серьезного опыта общения с литературой подобного рода, поэтому ее можно рассматривать, как попытку рассказать по возможности простым языком о вещах, достаточно сложных. О достижениях современной молекулярной биологии, ставшей в последние десятилетия следующей после физики королевой наук, о генах и мутациях, о том, как работают молекулярные генетики в медицине, в частности, как они определяют носительство и проводят дородовую (пренатальную) диагностику наследственных заболеваний. Основное внимание в связи со спецификой данного издания будет уделено генам факторов VIII и IX системы свертывания крови и ДНК-диагностике гемофилии А и В.

К сожалению, совсем простым обыденным языком рассказать обо всем этом невозможно, и поэтому волей-неволей придется осваивать азы двух “иностранных” языков, языка нуклеиновых кислот и “птичьего” языка молекулярных биологов.

Всем (не только специалистам) хорошо известно, что дезоксирибонуклеиновая кислота (ДНК) занимает среди всех биополимеров самое важное место, поскольку именно она является носителем генетической информации. Как и все полимеры, ДНК состоит из соединенных друг с другом молекул мономеров. Мономеры, составляющие ДНК, называются *дезоксирибонуклеотидами* (чаще их называют просто нуклеотидами). Они содержат остаток фосфорной кислоты и сахар

дезоксирибозу с присоединенным к ней гетероциклическим азотистым основанием. Существуют четыре разновидности таких оснований – аденин (А), цитозин (Ц), гуанин (Г) и тимин (Т).

Вот эти четыре основания (буквы) и являются тем нехитрым алфавитом, при помощи которого написаны все генетические тексты. ДНК в клетках существует в виде двух цепей, прочно связанных друг с другом. Каждый нуклеотид одной цепи водородными связями соединен с нуклеотидом второй цепи, причем, благодаря высокому сродству между определенными основаниями, строго соблюдается следующее правило такого спаривания – напротив цитозина в одной цепи в другой всегда располагается гуанин, а напротив аденина – тимин. Именно между этими парами оснований наиболее эффективно образуются водородные связи. Только благодаря этому правилу, которое называется правилом *комплементарности*, и возможно точное копирование ДНК. Копирование ДНК в клетке осуществляется специальным ферментом ДНК-полимеразой. Используя одну из цепей ДНК в качестве матрицы, этот фермент синтезирует вторую цепь в строгом соответствии с изложенным выше принципом.

Сравнение последовательности нуклеотидов в ДНК с буквенным текстом отнюдь не случайно и использовано здесь не только для того, чтобы упростить изложение. На

самом деле любой молекулярный биолог или генетик в подавляющем большинстве случаев именно так и воспринимает ДНК, как текст, и анализирует ее тоже, как текст. И чаще всего анализ экспериментальных данных сводится к поиску в этом тексте каких-то определенных сочетаний букв или же к сравнению между собой разных текстов, которые подчас могут различаться всего одной буквой, как это бывает при поиске мутаций. Десять лет назад в биологической науке произошло событие, ставшее мировой сенсацией – был расшифрован полный геном человека, то есть вся совокупность нуклеотидных последовательностей ДНК, составляющих генетический аппарат человека. Между тем путь к этому сенсационному успеху был очень непростым и совсем не коротким.

Достаточно сказать, что основным методом, с помощью которого удалось прочитать весь геном человека, был разработан в Кембриджском университете замечательным английским ученым, дважды лауреатом Нобелевской премии Фредериком Сэнгером еще в конце 70-х годов прошлого века. Кстати, одну из этих двух премий он получил именно за разработку метода определения первичной структуры (последовательности нуклеотидов) ДНК. У биологов этот метод называется *секвенированием* (от английского слова *sequence* – последовательность). Для того, чтобы это достижение стало возможным, потребовалась долгая и кропотливая работа биологов, физиков, инженеров, математиков (биоинформатиков), придумавших новые методы синтеза и анализа, создавших высокоточные автоматические приборы и сложные компьютерные программы для обработки огромного массива экспериментальных данных. Так возник текст, который стал настольной книгой молекулярных биологов и генетиков. Его так и называли Книгой (как Библию), хотя при сравнении его объема с привычным книжным форматом становится понятно, что это на самом деле не книга, а целая библиотека. Для написания этого текста природе понадобилось 3.2-3.3 миллиарда букв (нуклеотидов). Именно таков размер человеческого генома. Для сравнения – один из самых масштабных шедевров мировой литературы, роман Льва Толстого “Война и мир” содержит приблизительно 2.5 миллиона букв. Если бы нашелся читатель, способный прочитывать по одной такой книге в день, то у него ушло бы на чтение геномной Книги около трех лет. Но, наверное, на этом стоит закончить с литературными фантазиями и ассоциациями и перейти собственно к генам, самые большие из которых составили бы главу в “Войне и мире”, а самые маленькие заняли бы всего полстраницы. Так, например, один из самых крупных генов, ген фактора VIII, имеет размер около 186 тысяч нуклеотидов, а один из самых мелких, ген бета-цепи белка гемоглобина, состоит всего из 1500 нуклеотидов. В заключение этой вводной части хотелось бы только сказать, что, получив в руки текст полного генома, молекулярные биологи, несомненно, серьезно продвинулись в понимании того, как он организован, и увидели много нового и часто неожиданного, но вместе с тем не только многие старые вопросы остались нерешенными, но возникла и масса других. В частности до сих пор остается открытым вопрос о том, сколько же всего генов у человека. Еще до того, как был расшифрован полный геном, многие ученые сходились на том, что их должно быть около ста тысяч. При первых же прочтениях их оказалось существенно меньше, всего 25-30 тысяч, что вызвало даже некоторое разочарование. Последующий более строгий анализ снизил это число до 21 тысячи. До этого казалось, что мы устроены сложнее. Хотя речь в данном случае идет, главным образом, о генах, кодирующих белки. Но уже и здесь все не очень просто, потому что многие из них являются гипотетическими, то есть они вычислены компьютером на основании анализа

определенных признаков, характерных для генных структур, но самих белков, которые кодируются этими предполагаемыми генами, ученые пока не видели, и никто не может сказать, существуют они на самом деле или нет. Кроме того, есть группа генов, кодирующих не белки, а РНК. Молекулы РНК (рибонуклеиновой кислоты)

состоят из рибонуклеотидов и отличаются от ДНК типом сахарного остатка (рибоза вместо дезоксирибозы) и одним гетероциклическим основанием (урацил вместо тимина). РНК выполняют в живой клетке много важнейших функций. Вся информация, содержащаяся в генах, точно копируется в молекулы РНК специальными ферментами – РНК-полимеразами. Этот процесс получил название *транскрипция*. РНК, считываемые с белок-кодирующих генов, называются информационными или матричными РНК (мРНК). После синтеза и созревания (об этом процессе речь пойдет немного позже) они транспортируются из ядра в цитоплазму, где на специальных органеллах, рибосомах, по содержащейся в них информации синтезируются белки. Процесс синтеза белков на рибосомах называется *трансляцией*. Рибосомы в свою очередь содержат в своем составе, кроме белков, молекулы рибосомных РНК (рРНК). Переводчиками генетической информации с языка нуклеиновых кислот на язык белков служат транспортные РНК (тРНК). Эти РНК, не кодирующие белков, а сами по себе являющиеся структурными и функциональными единицами, известны давно, они немногочисленны и их гены не вносят сколько-нибудь заметного вклада в общий счет генов человека. Но не так давно биологи открыли новый класс РНК, получивших название микроРНК. Это совсем маленькие молекулы, часто состоящие всего из двадцати нуклеотидов. Их предназначение – регулировать работу генов. Каждая такая микроРНК специфична к мРНК определенного гена. После образования комплекса с микроРНК мРНК теряет способность к трансляции. Таким способом регулируется активность едва ли не половины всех белок-кодирующих генов. Гены микроРНК имеют малые размеры и идентифицировать их в геноме не так просто, но они многочисленны и на общее число генов могут повлиять весьма существенным образом. Нельзя не сказать и о том, что гены составляют всего 10-20% от всего генома. Вопрос о том, что из себя представляет и зачем нужна остальная ДНК, которую называют то “эгоистической”, то “мусорной”, тоже по большому счету остается без определенного ответа. Возможно, она содержит заготовки для каких-то будущих генетических текстов или еще не идентифицированные регуляторные последовательности, или же какие-то структуры, обеспечивающие сверхплотную упаковку ДНК в клеточном ядре, а такая чисто техническая проблема тоже требует своего решения, ведь если представить себе человеческий геном в виде единой расправленной цепочки, то физическая длина ее составит около 2 метров, а уместить ее надо в клеточном ядре диаметром 10 микрометров.

Так что геномную Книгу еще читать и читать, и понимать. И в ее понимании большую помощь должны оказать новейшие методики и компьютерные программы, появившиеся за те десять лет, что прошли со времени расшифровки первого генома человека. Если на расшифровку первого генома ушли годы, то с использованием современных технологий стало возможным осуществлять такую работу за 1-2 месяца. Сравнительный анализ геномов разных людей, сравнение человеческого генома с геномами различных животных, а такие работы сейчас ведутся, безусловно, должны способствовать гораздо лучшему пониманию организации и функционирования генетического аппарата.

Перед тем, как окончательно перейти к разговору о генах и мутациях, осталось только напомнить, что ДНК в клеточном ядре не является единой молекулой, а упакована в отдельные хромосомы разного размера. В состав хромосом кроме ДНК входят также белки – *гистоны*, которые и обеспечивают компактность всей конструкции за счет образования специфических структурных единиц – *нуклеосом*. Все хромосомы, не являющиеся половыми, называются *аутосомами* и имеют свой порядковый номер от 1-го до 22-го. Номера присвоены хромосомам в соответствии с их “ростом”. Самая большая 1-я хромосома имеет размер около 250 млн. нуклеотидов и несет приблизительно 4 тысячи генов, а самая маленькая 22-я содержит около 50 млн. нуклеотидов и примерно 600 генов. Кроме аутосом есть еще две половые хромосомы, женская X-хромосома (около 150 млн. нуклеотидов и 1400 генов) и мужская Y-хромосома (около 50 млн. нуклеотидов и более 200 генов). В ядрах всех соматических (неполовых) клеток аутосомы представлены двумя копиями, а половые хромосомы распределяются следующим образом: у женщин – две X-хромосомы, а у мужчин – одна X-хромосома и одна Y-хромосома. Именно поэтому большинство наследственных болезней, связанных с нарушением генов, находящихся на X-хромосоме (такие заболевания называются X-сцепленными), затрагивают только мужчин. Почти все эти болезни относятся к *рецессивным*, то есть наличие нормальной копии гена вполне компенсирует дефектность второй копии. У мужчин же такой компенсации нет, поскольку все X-хромосомные гены представлены у них единственной копией. К наиболее тяжелым и распространенным X-сцепленным заболеваниям относятся гемофилия А и В, мышечная миодистрофия Дюшенна. Для женщин, являющихся носительницами таких заболеваний, безусловно, актуально проведение дородовой (пренатальной) диагностики, позволяющей исключить риск рождения больного ребенка. И, естественно, очень важной диагностической составляющей в этом случае является определение пола плода. Молекулярно-генетические методы анализа позволяют проводить пренатальную диагностику наследственных патологий на раннем этапе беременности (оптимальные сроки для биопсии хориона – 9-11 недель), когда определить пол плода с помощью ультразвуковых исследований не представляется возможным. С этой целью используют другие подходы – цитогенетические (хромосомный анализ) или молекулярно-генетические (анализ ДНК). Цитогенетический хромосомный анализ называют также кариотипированием (карио – ядро). Цитогенетики видят в световой микроскоп весь хромосомный набор, но для того, чтобы увидеть и хорошо различить хромосомы необходимо, чтобы в препарате были в достаточном количестве делящиеся клетки, если же их мало, то клетки подрашивают в специальных культуральных средах, и тогда вместо обычных 1-2 дней на определение пола может потребоваться неделя, а то и больше. Обычно результаты своих исследований цитогенетики записывают следующим образом – нормальный кариотип (46;X,X) или нормальный кариотип (46;X,Y). Нормальный кариотип означает, что все 46 хромосомы на месте, имеют стандартный вид, и нет ничего лишнего; X,X – пол женский, X,Y – пол мужской. Но иногда, хотя и в достаточно редких случаях, могут обнаружиться хромосомные аномалии, например, трисомия по 21-ой хромосоме. То есть в хромосомном наборе вместо положенных двух присутствуют три 21-х хромосомы, а такая “лишняя” 21-я хромосома ассоциируется с известным синдромом Дауна. Поэтому при проведении пренатальной диагностики цитогенетическое исследование оказывается полезным не только в плане определения пола. К вопросу о том, как определяют пол с помощью анализа ДНК, мы еще вернемся в процессе разговора о

молекулярно-генетической диагностике носительства и пренатальной ДНК-диагностике наследственных болезней. Но перед этим следует сказать несколько слов о том, как устроены белок-кодирующие гены высших организмов и, в том числе, человека, как с них считывается информация, и что из себя представляет генетический код, который служит для перевода этой информации с языка ДНК на язык белков. Белки, как известно, это биополимеры, состоящие из аминокислот. Природных аминокислот всего 20, то есть алфавит белкового языка существенно богаче, чем нуклеиновый. Каждой аминокислоте в ДНК соответствуют определенные последовательности из трех нуклеотидов, которые называют кодонами или триплетами. Всего, как нетрудно сосчитать, из 4-х букв нуклеинового алфавита можно составить 64 различных триплетов. Только две аминокислоты кодируются единственным триплетом, это метионин (АТГ) и триптофан (ТТГ).

Остальные аминокислоты имеют от 2-х до 6-ти кодонов. Так, например, важная аминокислота аргинин кодируется 6-ю триплетами – ЦГА, ЦГЦ, ЦГГ, ЦГТ, АГГ и АГА, а глицин 4-мя – ГГА, ГГЦ, ГГГ и ГГТ. Обращает на себя внимание тот факт, что основную смысловую нагрузку несут первые две буквы кодона, а третья зачастую никакого значения не имеет и может быть любой.

Это свойство генетического кода называют вырожденностью. Три кодона – ТГА, ТАГ и ТAA не кодируют никаких аминокислот. Они выполняют другую важную функцию – служат для рибосомы сигналом к окончанию синтеза белка, и поэтому всегда располагаются в самом конце кодирующей области гена и называются терминирующими или стоп-кодонами (иногда их еще называют антисмысловыми или нонсенс-кодонами). Есть еще одно неукоснительно выполняющееся правило - кодирующая область всех генов всегда начинается с одного и того же метионинового кодона АТГ (все белковые молекулы начинаются с метионина). Пока исследователям была известна общая организация только бактериальных генов, все было просто и понятно, поскольку в ДНК бактерий кодирующие белок нуклеотидные последовательности представляют собой единые непрерывные структуры. Но у высших организмов (эукариотов) все оказалось иначе. Почти все их гены устроены таким образом, что кодирующие участки в них, названные *экзонами*, разделены существенно большими по размеру не кодирующими последовательностями – *интронами*. Так, например, один из интересующих нас генов, ген фактора VIII, имеет общую длину 186 тыс. нуклеотидов и состоит из 26 экзонов, занимающих суммарно всего 7 тыс. нуклеотидов, а остальные 179 тыс. приходятся в основном на 25 его интронов, самый большой из которых, 22-ой, имеет длину 32 400 нуклеотидов, то есть в 4,5 раза больше, чем все экзоны, вместе взятые. РНК-полимераза в ходе транскрипции не в состоянии отличить кодирующие последовательности от не кодирующих и синтезирует полную копию всего гена, так называемую пре-мРНК. Для того, чтобы превращать ее в зрелую мРНК, с которой на рибосомах будет считываться белок, природе пришлось изобрести довольно сложный специальный механизм *сплайсинга* (сращивание), заключающийся в вырезании из пре-мРНК интронов и сшивании экзонов между собой в единую кодирующую молекулу. Этот процесс осуществляется особым комплексом, состоящим из малых ядерных РНК и нескольких специальных белков. Называется этот комплекс сплайсосомой. Он строго специфичен к небольшим участкам интронов, граничащим с экзонами, так называемым сайтам сплайсинга, чем и обеспечивается высокая точность процедуры вырезания/сшивки.

Любой интрон любого гена всегда начинается с динуклеотида ГТ и заканчивается динуклеотидом АГ.

Интересно, что интроны после вырезания переводятся в молекулы кольцевой формы. Это предотвращает их участие в процессе сшивания экзонов. Вопрос о том, почему природа так сложно устроила гены высших организмов, остается пока без определенного ответа, хотя на этот счет существует целый ряд гипотез. Согласно одной из них, в этом, по крайней мере, отчасти, заключен эволюционный потенциал эукариотов. Дело в том, что отдельные экзоны некоторых генов целиком кодируют участки белков, белковые *домены*, с определенными функциями и обмен такими экзонами между разными генами в процессе рекомбинации можно рассматривать, как один из способов создания белков с новым набором свойств. По другой гипотезе интроны нужны для регуляции активности генов. Действительно, некоторые интроны содержат специфические нуклеотидные последовательности, *энхансеры* (усилители), с которыми связываются различные регуляторные белки (транскрипционные факторы), способные обеспечивать более эффективную транскрипцию гена. Из важных структурных элементов гена следует еще упомянуть *промотор* и *сигнал полиаденилирования*. Промотор расположен перед кодирующей областью гена. Эту структуру узнает РНК-полимераза и связывается с ней в процессе инициации транскрипции. Расположенный с другой стороны гена, за кодирующей последовательностью, сигнал полиаденилирования (обычно он выглядит следующим образом – ААТААА) служит для навешивания на конец молекулы мРНК длинного (в среднем около 50 звеньев) “хвоста”, состоящего из одних адениновых оснований и обеспечивающего ее стабильность. Этот синтез осуществляет специальный фермент поли(А)-полимераза.

Теперь, когда нам известны общая организация и основные структурные элементы гена, можно приступить к рассмотрению различных типов мутаций, в той или иной степени нарушающих его функционирование. От типа мутации и ее позиции в гене в значительной мере зависит характер ее проявления или иными словами степень тяжести ассоциированного с ней заболевания. Наиболее распространенным типом мутаций являются единичные нуклеотидные замены. Если замена одной буквы на другую произойдет в кодирующей области гена, изменится смысл того кодона, который она затронула. Он может превратиться в кодон другой аминокислоты, то есть возникнет единичная аминокислотная замена в белке (такие генные нарушения называются миссенс-мутациями), или же поменяться на преждевременный терминирующий кодон (нонсенс-мутация), что довольно часто происходит с кодоном аргинина ЦГА. На этом обстоятельстве стоит остановиться особо, поскольку динуклеотиды ЦГ в эукариотических ДНК являются “горячими” точками мутаций. Частота мутирования в них примерно в 100 раз выше, чем в других участках генома. Дело в том, что цитозин в этих динуклеотидах часто представлен своим аналогом – 5-метилцитозином, который способен спонтанно превращаться в тимин. Если речь идет об уже упомянутом кодоне аргинина ЦГА, то он в таком случае превращается в стоп-кодон ТГА. Метилирование цитозина у эукариотов выполняет важную функцию. Оно является одним из способов регуляции активности генов. Гены с метилированными промоторами не активны. Хорошо известно явление полной инактивации одной из двух X-хромосом у женщин, получившее название *лайонизация* по имени Мэри Лайон, впервые высказавшей эту гипотезу. Инактивация носит

случайный характер, то есть в каждой клетке не активна либо одна, либо другая X-хромосома, и поэтому в целом у любой женщины находятся в рабочем состоянии гены обеих X-хромосом, причем уровень их работоспособности соответствует таковому у мужчин, имеющих единственную X-хромосому. Известно, однако, такое явление, как асимметричная инактивация X-хромосомы, когда клеток с одной инактивированной X-хромосомой оказывается существенно больше, чем с другой. В частности, подобной асимметрией иногда объясняются сами по себе очень редкие случаи проявления гемофилии у женщин, когда клеток с активной X-хромосомой, несущей мутантный ген фактора VIII или IX, значительно больше клеток с активной нормальной X-хромосомой. И раз уж речь зашла о проявлении гемофилии у женщин, то следует сказать, что второй его причиной может быть случайное совпадение – присутствие у одной женщины двух мутантных X-хромосом. Вероятность такого совпадения крайне низка. Частота встречаемости гемофилии А у мужчин в разных странах оценивается в пределах 1:5000/1:10000 (гемофилия В встречается в 5 раз реже). По всей вероятности, эти цифры приблизительно соответствуют средней частоте встречаемости “гена гемофилии”, как зачастую обобщенно называют мутантные гены факторов VIII и IX. В таком случае, вероятность наличия у женщины двух мутантных генов фактора VIII будет находиться в пределах 1:25/1:100 миллионов, а для гена фактора IX еще в 5 раз меньше. Но вернемся собственно к мутациям. Безусловно, нонсенс-мутации относятся к “тяжелым” генным нарушениям, поскольку они в конечном итоге приводят к образованию неполноценного белка, укороченного по сравнению с нормальным, так как его синтез на рибосоме будет прекращаться в том месте, где образовался новый стоп-кодон. И белок этот будет тем короче, чем ближе к началу кодирующей области гена возникла нонсенс-мутация. С миссенс-мутациями, приводящими к единичным аминокислотным заменам, дело обстоит несколько сложнее. Степень тяжести таких мутаций напрямую зависит от позиции замененной аминокислоты в молекуле белка и от того, насколько новая аминокислота по своим свойствам отличается от прежней. Особое значение имеют кислотно-основные свойства, гидрофобность или гидрофильность, общая структура боковой цепи аминокислоты (размеры, наличие атома серы). Если замена затронет функционально важный участок белка, например, активный центр фермента, сайты связывания с ДНК, с другими белками или низкомолекулярными веществами (субстратами, кофакторами), и при этом новая аминокислота будет принципиально отличаться от прежней, то такая мутация несомненно приведет к тяжелым последствиям. То же самое можно сказать и о заменах, серьезно нарушающих пространственную структуру белка, от которой существенно зависит не только его функциональная активность, но и стабильность. Многие миссенс-мутации, однако, не столь фатальны и позволяют белку в значительной степени сохранить свои функции. У большинства пациентов со средней и легкой формой гемофилии заболевание обусловлено именно такими “щадящими” миссенс-мутациями, сохраняющими фактору свертывания 5-20% активности. Кроме того, встречаются нуклеотидные замены, которые приводят к замещению аминокислоты, но при этом практически не влияют на свойства белка. В таких случаях говорят не о мутациях, а о *полиморфизмах*, то есть о нейтральных изменениях в ДНК, не имеющих фенотипического проявления. Хорошим примером таких полиморфизмов служат аномальные гемоглобины. Это варианты известного всем жизненно важного белка гемоглобина, работающего переносчиком кислорода, отличающиеся от нормального гемоглобина всего одной аминокислотой. Так вот среди аномальных гемоглобинов нередко встречаются такие, которые по своей стабильности и

способности связывать кислород ничем не хуже нормального гемоглобина, имеющегося у большинства людей. Единичные нуклеотидные замены могут затрагивать также сайты сплайсинга, о которых говорилось выше.

Эти мутации так и называются мутациями сплайсинга. Особенно тяжелыми такие нарушения бывают, когда замена происходит в канонических динуклеотидах ГТ или АГ. В измененных сайтах сплайсинга мРНК-предшественница (пре-мРНК) перестает расщепляться, и поэтому соседствующие с такими нарушенными сайтами экзоны удаляются из нее вместе с интронами. В итоге зрелая мРНК теряет целый экзон, а белок – весь участок, закодированный в этом экзоне, а это, как минимум, несколько десятков аминокислот. Кроме того, нуклеотидные замены могут

нарушить промотор или сигнал полиаденилирования. Но такие мутации, как правило, проявляются очень слабо. В большинстве случаев они приводят лишь к некоторому уменьшению количества мРНК и, следовательно, белка. Есть, однако, еще одна, довольно экзотическая разновидность мутаций – нуклеотидные замены во внутренних областях интронов, создающие аномальные сайты сплайсинга, которые вполне успешно конкурируют с нормальными, законными сайтами сплайсинга. В таких случаях наряду с нормальной мРНК, количество которой уменьшается, возникает химерная нефункциональная мРНК, содержащая кроме кодирующих последовательностей фрагмент интрона. Чаще всего подобные мутации фенотипически проявляют себя довольно слабо.

Другой тип мутаций – микроделеции и микроинсерции (в английском варианте они называются frameshift – сдвиг рамки), Микроделеции выражаются в том, что в ДНК теряется небольшой участок размером от одного до нескольких нуклеотидов. Нетрудно себе представить, что произойдет, если из какого-либо кодирующего участка гена убрать, к примеру, один нуклеотид. В кодирующей области гена (или мРНК) все очень жестко расписано и, начиная со стартового кодона АТГ, каждые следующие три нуклеотида означают конкретную аминокислоту, которая должна находиться в белке именно в данном месте. Последовательность кодонов называют обычно рамкой считывания белка. Делеция нарушит эту рамку (сдвиг рамки), и от точки делеции разбивка на кодоны станет совершенно иной. Кроме того, как правило, в новой рамке считывания возникают стоп-кодоны. В результате синтезируется укороченный белок с измененной концевой частью. Особый случай представляют собой микроделеции в 3, 6, 9 и т. д. нуклеотидов, то есть, когда число потерянных нуклеотидов кратно трем. Никаких стоп-кодонов и сдвига рамки при этом, естественно, не возникает, а просто из молекулы белка исчезают 1, 2, 3 ... аминокислоты. Несколько реже, чем микроделеции, встречаются противоположные им по смыслу мутации – микроинсерции, то есть вставки. В этом случае короткие фрагменты длиной от 1 до нескольких нуклеотидов не удаляются, а наоборот, вставляются в ДНК. Такое событие будет иметь те же последствия, что и микроделеция, с той только разницей, что в белке вместо потери окажутся лишние аминокислоты. К еще более редким событиям относятся протяженные делеции или макроделеции, приводящие к потере больших участков гена размером в несколько сотен или даже тысяч нуклеотидов (иногда ген может исчезнуть целиком). Последствия мутаций такого типа очевидны. И уже совсем редко встречаются протяженные инсерции, обусловленные, как правило, внедрением в ген так называемых ретропозонов, активных генетических элементов,

представленных в геноме многими тысячами копий и способных к дополнительному копированию. Считается, что сами по себе ретропозоны ничего функционального не кодируют, но некоторые их копии транскрипционно активны, то есть с них считывается РНК. Эта РНК может быть опять переведена в свою ДНК-копию при помощи содержащегося в клетках фермента – обратной транскриптазы. Обратная транскриптаза в отличие от ДНК-полимеразы, копирующей ДНК в ДНК (репликация), и РНК-полимеразы, синтезирующей РНК по матрице ДНК (транскрипция), способна к синтезу ДНК с РНК-матрицы (обратная транскрипция). Получающиеся таким образом свободные ДНК-копии ретротранспозона могут случайным образом встраиваться в геном и, попадая в какой-либо ген, нарушать его структуру. Кроме описанных выше основных типов мутаций, есть еще одна разновидность генных нарушений – инверсии. К инверсиям мы обязательно вернемся, поскольку они, являясь в целом весьма редким событием, тем не менее очень часто определяют возникновение тяжелой формы гемофилии А, что в известной степени является генетической особенностью этого заболевания.

Сейчас же, наверное, имеет смысл сказать несколько слов о том, как молекулярные генетики обнаруживают различные мутации и определяют таким способом носительство того или иного наследственного заболевания. Следует сразу оговориться, что молекулярно-генетическое исследование не является способом установления первичного диагноза. Оно вступает в силу и становится эффективным на уровне уже имеющегося предварительного диагноза, когда становится понятным, какой именно ген необходимо исследовать. Так, например, в случае гемофилии важно знать к какому типу она относится, А или В, потому что от этого зависит какой ген будет изучаться, ген фактора VIII или ген фактора IX. Первичный диагноз гемофилии устанавливается при помощи измерения активности факторов свертывания крови VIII и IX у больного из данной семьи. **Однако, достоверно определить таким способом носительство заболевания у женщины из семьи, затронутой гемофилией, практически невозможно, поскольку очень широк диапазон нормальных значений активности факторов (от 50 до 200%).** На практике достаточно часто встречаются случаи, когда женщины, имеющие активность фактора 100% и более, оказываются носительницами гемофилии, и наоборот, женщины с активностью фактора меньше 100% никакого отношения к заболеванию не имеют. **Корректный вывод о статусе носительства гемофилии у женщин можно сделать только на основании результатов генетического исследования.** Под генетическим исследованием в данном случае подразумевается либо определение мутации, приводящей к возникновению заболевания, либо семейный анализ с использованием нейтральных ДНК-маркеров (полиморфизмов).

Поиск мутаций осуществляется при помощи определения нуклеотидных последовательностей всех функционально важных участков гена, о которых мы говорили выше. Основное внимание при этом уделяется кодирующим областям (эксонам) и экзон-интронным сочленениям (сайтам сплайсинга). Процесс определения нуклеотидных последовательностей называется секвенированием (от англ. sequence – последовательность). Расшифрованные последовательности сравниваются с известной первичной структурой нормального гена, и если найдено отличие, то почти всегда можно сказать, является оно мутацией определенного типа или нейтральным полиморфизмом. Для того, чтобы определить нуклеотидную последовательность

интересующего нас фрагмента гена, необходимо наработать его в количествах, достаточных для проведения такого исследования. Обычно это количество измеряется десятками или сотыми долями микрограмма. Для того, чтобы нарабатывать (амплифицировать, то есть увеличивать число копий) нужные для исследования фрагменты генов в таких количествах, в конце 80-х годов прошлого века был разработан специальный метод, получивший название *полимеразная цепная реакция* (ПЦР). И вот уже двадцать с лишним лет ни одно молекулярно-генетическое исследование не обходится без использования этого метода. Суть его заключается в следующем. В небольшую пластиковую пробирку объемом 0,2 – 0,5 мл помещают все компоненты, необходимые для синтеза ДНК.

Это, прежде всего, ДНК-матрица, то есть геномная ДНК, выделенная, например, из ядерных клеток крови пациента, содержащая исходные копии изучаемого гена и ДНК-полимераза, фермент, который будет осуществлять копирование нужного фрагмента этого гена. Очень важной составляющей системы являются синтезированные химическим способом на специальных приборах, синтезаторах, искусственные кусочки одноцепочечной ДНК с заданной исследователем последовательностью нуклеотидов. Они называются праймерами или, в просторечии, затравками и обычно имеют длину около 20 нуклеотидов. Таких праймеров в системе должно быть два. Один из них комплементарен участку одной цепи ДНК и ограничивает размножаемый фрагмент гена с одной стороны, а второй комплементарен участку другой цепи и ограничивает фрагмент с другой стороны. Праймеры не только ограничивают нужный нам фрагмент гена, они абсолютно необходимы для работы ДНК-полимеразы, без таких затравок она просто не сможет начать синтез. Собственно, то же самое происходит и при репликации геномной ДНК в клетке. Только в этом случае затравками служат кусочки РНК, которые синтезируются РНК-полимеразой именно с этой целью. При завершении репликации они удаляются и заменяются на ДНК. Кроме перечисленных компонентов реакционная смесь содержит также все четыре дезоксирибонуклеозидтрифосфата, источник мономерных нуклеотидов, из которых будут строиться новые цепочки ДНК, и ионы магния, служащие при инициации реакции полимеризации ко-фактором, обеспечивающим создание прочного комплекса между ДНК-матрицей, ДНК-полимеразой и праймером. Синтез выбранного фрагмента гена в пробирке осуществляется в режиме повторяющихся циклов, причем каждый цикл состоит из трех стадий. На первой стадии пробирку с реакционной смесью выдерживают короткое время (обычно от 30 секунд до 1 минуты) при высокой температуре (94-95°C) для того, чтобы разъединить цепи ДНК-матрицы, ведь геномная ДНК находится в двухцепочечном состоянии, и освободить, таким образом, место для посадки праймеров. Эту стадию называют денатурацией или плавлением ДНК. На следующей стадии температуру снижают до 50-60°C и выдерживают смесь также от 30 секунд до 1 минуты, чтобы образовался комплекс между разошедшимися цепями матричной ДНК и праймерами. При этом, естественно, часть цепей ДНК вновь соединится в двухцепочечную структуру, но в основном все же они будут образовывать комплекс с праймерами, благодаря тому, что последние находятся в реакционной смеси в огромном избытке. Эта стадия называется реассоциацией или отжигом. И последняя стадия цикла, именуемая синтезом, проводится при 72°C обычно в течение одной или нескольких минут в зависимости от размера амплифицируемого фрагмента.

Температура 72°C выбрана не случайно. Поскольку весь процесс амплификации проходит в высокотемпературном режиме, необходимо, чтобы полимеризующий фермент выдержал эти жесткие условия и не потерял своей активности. Этому условию удовлетворяют ДНК-полимеразы, выделенные из термофильных бактерий, живущих в горячих источниках (чаще всего используют ДНК-полимеразу из *Thermusaquaticus*). Для этих ферментов температура 72°C является физиологическим оптимумом, при котором они работают в реальных природных условиях. На самом деле синтез новых цепочек ДНК начинается еще на стадии отжига, как только образуется комплекс матрица-фермент-праймер, только он проходит с меньшей скоростью, и стадия синтеза необходима для его полного завершения. Итак, что же произойдет в первом цикле амплификации? Здесь важно помнить, что цепи двунитевой ДНК-матрицы антипараллельны, то есть копируются они в противоположных направлениях, поэтому в нашем случае новые цепочки будут синтезироваться с двух имеющихся в системе праймеров навстречу друг другу. Вновь синтезированные цепи с одного конца будут ограничены праймером, с которого начался синтез, с другого же конца никаких принципиальных ограничений пока нет, и синтез остановится там, куда успеет дойти полимераза к моменту завершения цикла. Этот процесс будет повторяться во всех последующих циклах амплификации и, как не сложно догадаться, количество таких ограниченных с одной стороны цепочек будет возрастать в арифметической прогрессии, то есть в каждом новом цикле их будет синтезироваться столько же, сколько и в предыдущем, поскольку матрицей для них служит исходная геномная ДНК. Но второй цикл амплификации принципиально отличается от первого тем, что в нем начинают работать в качестве матриц вновь синтезированные цепочки ДНК, уже ограниченные с одной стороны одним из праймеров и содержащие внутри себя место посадки для второго праймера. Результатом синтеза на таких матрицах станут цепочки, ограниченные с обеих сторон заданными нами праймерами, и с каждым следующим циклом число их будет удваиваться, то есть расти уже не в арифметической, а в геометрической прогрессии.

Теперь можно произвести несложный расчет, решить простую арифметическую задачу. Предположим, нам нужно наработать фрагмент гена размером в 300 (3×10^2) нуклеотидов. В опыт мы взяли 0,1 микрограмм геномной ДНК (столько обычно и берут) и на выходе хотим получить такое же весовое количество нужного нам фрагмента. Приблизительный размер генома человека, как мы помним, составляет 3×10^9 нуклеотидов. Легко увидеть, что для достижения поставленной цели нам нужно амплифицировать искомый фрагмент в 10^7 (десять миллионов) раз. Десять циклов амплификации с геометрической прогрессией дадут нам увеличение количества нужного материала примерно в тысячу раз ($2^{10} = 1024$), двадцать циклов – в миллион раз, дополнительные 3-4 цикла дадут нам увеличение еще в 8-16 раз. Если прибавить сюда первые два цикла, на которых геометрической прогрессии еще нет, то в сумме получится 25-26 циклов. Учитывая то обстоятельство, что никакая реакция в реальной жизни не дает 100%-ного выхода,

число циклов амплификации обычно увеличивают до 30. Остается только сказать, что за разработку этого метода американский биохимик Кэрри Муллис в 1993 году получил Нобелевскую премию. После проведения ПЦР качество и приблизительное количество получившегося продукта оценивают при помощи электрофореза в агарозном или полиакриламидном геле. Эти гели представляют собой плотный мелкопористый

материал, в котором молекулы ДНК, имеющие отрицательный заряд, движутся под действием электрического поля от анода (-) к катоду (+). При этом они делятся по размеру: чем крупнее фрагменты, тем труднее они проходят сквозь поры геля, и тем медленнее движутся. По завершении электрофореза гель прокрашивают специальным красителем – бромистым этидием, после чего фрагменты ДНК можно увидеть в ультрафиолетовом свете (светящиеся полосы ярко оранжевого цвета). Параллельно с опытным образцом в геле разгоняют так называемый маркер длин, представляющий собой стандартный набор фрагментов ДНК известной длины (известно также приблизительное весовое количество маркерных фрагментов). По положению амплифицированного фрагмента относительно маркерных фрагментов и интенсивности его свечения можно оценить как его длину (если это ожидаемый продукт, то она должна соответствовать расчетной), так и приблизительное количество. Если амплификация прошла успешно, и получился нужный фрагмент в достаточном количестве, то остается только определить его нуклеотидную последовательность. Для этого традиционно используют упоминавшийся выше метод секвенирования по Сэнгеру. Он также, как и ПЦР, основан на полимеразной реакции и в его современных версиях также применяется термостабильная ДНК-полимераза. Однако, в отличие от реакции амплификации при секвенировании в системе присутствует только один праймер, комплементарный одному из концов изучаемого фрагмента ДНК, потому что в данном случае крайне важно, чтобы все синтезируемые цепочки имели один и тот же строго фиксированный конец. Реакцию секвенирования можно проводить в цикловом режиме. При этом, естественно, новые цепочки ДНК будут накапливаться не в геометрической прогрессии, как при ПЦР, а в арифметической. Основная же идея метода заключается в использовании в полимеразной реакции наряду с нормальными нуклеотидными мономерами их терминирующих производных

(терминаторов). Терминаторы или дидезоксинуклеотиды отличаются от нормальных дезоксинуклеотидов тем, что у них из остатка рибозы химическим путем удалена гидроксильная группа, по которой при синтезе ДНК происходит присоединение к растущей цепи следующего нуклеотида. Они нормально воспринимаются ДНК-полимеразой как субстраты и эффективно встраиваются в синтезируемую ДНК, но на этом синтез и завершается. Если в 4 пробирки поместить одну и ту же реакционную смесь и в каждую из них добавить один из терминаторов, то в пробирке с дидезоксиаденозином все синтезированные цепочки будут заканчиваться на А, в пробирке с дидезоксицитидином на С и т.д.. Цепочки при этом будут образовываться разной длины в зависимости от того, насколько успел продвинуться синтез к тому моменту, когда произошло присоединение терминирующего нуклеотида. В итоге суммарно в четырех пробирках образуется полный набор цепочек разной длины, каждая из которых оканчивается своей терминирующей буквой. Содержимое каждой из пробирок наносят затем на четыре смежные дорожки длинного полиакриламидного геля (50 см и более) и проводят электрофорез высокого разрешения, позволяющий разделять цепочки ДНК, различающиеся по длине на один нуклеотид. На результирующей электрофореграмме в каждой дорожке проявится свой набор полос, причем на каждом уровне будет располагаться только одна полоса, соответствующая одной из четырех букв. Двигаясь последовательно снизу вверх по ступенькам получившейся лесенки, мы и прочитаем интересующую нас нуклеотидную последовательность. Но так было раньше в эпоху так называемого “ручного” секвенирования. В современных автоматических секвенаторах все выглядит немного

по-другому. Реакцию проводят в одной пробирке, добавляя в нее все четыре терминатора, каждый из которых помечен своим флуоресцентным красителем, дающим излучение в определенном диапазоне длин волн. Электрофорез высокого разрешения осуществляют не в пластине, а в капилляре, рядом с нижним концом которого располагается детектор флуоресцентного излучения, регистрирующий последовательное прохождение через эту зону фрагментов, отличающихся по длине на один нуклеотид. Волновые диапазоны всех четырех красителей не пересекаются. Информация о последовательности прохождения сигналов четырех видов через детектор записывается компьютером и с помощью специальной программы преобразуется в графическую картинку, представляющую собой последовательность пиков четырех цветов (каждой из четырех букв соответствует свой цвет). Эту картинку легко можно перевести в нуклеотидную последовательность, прочитав ее глазами или воспользовавшись услугами соответствующей компьютерной программы. Описанным способом можно прочитывать фрагменты ДНК длиной примерно до 800 нуклеотидов. В качестве “примера из жизни” можно взять ситуацию с определением мутаций, приводящих к гемофилии В. Ген фактора IX, как уже упоминалось выше, занимает на X-хромосоме пространство в 34 тысячи нуклеотидов. Но в прочтении такой длинной последовательности целиком нет особой необходимости, поскольку интересующие нас мутации в подавляющем большинстве случаев сосредоточены в функционально важных участках гена, которые все вместе укладываются всего в 2,5-3 тысячи нуклеотидов. Это кодирующие области, разбитые на 8 экзонов (суммарно 1383 нуклеотидов), прилегающие к экзонам участки интронов, промотор и сигнал полиаденилирования. Все эти области можно амплифицировать в виде 7 отдельных фрагментов длиной от 200 до 700 нуклеотидов. При этом промотор войдет в один фрагмент с экзонам 1, сигнал полиаденилирования – в один фрагмент с последним, 8-ым экзонам, и также войдут в один фрагмент экзона 2 и 3, поскольку они расположены поблизости друг от друга. Если прочитать эти 7 фрагментов, то мы почти со 100%-ной вероятностью обнаружим мутацию, вызывающую гемофилию В. В том случае, когда анализируется ДНК больного гемофилией, чтение всегда будет однозначным (одна X-хромосома и один ген фактора IX), и мутацию мы найдем, сравнивая расшифрованные нуклеотидные последовательности с известной структурой нормального гена. При анализе ДНК женщины, являющейся носителем гемофилии, одновременно амплифицируются фрагменты обеих ее X-хромосом. Разделить их никак нельзя, и поэтому мутантный участок гена будет прочитываться вместе с нормальным. Если мутация представляет собой нуклеотидную замену, то на картинке, получаемой с секвенатора, в той позиции, где находится мутация, мы увидим двойной пик, то есть одновременно “нормальную” букву и “мутантную”. Если же мутация связана с потерей (делеция) или вставкой нуклеотидов (инсерция), то, начиная с этого места, удвоится вся последующая последовательность пиков. Поэтому при анализе женской ДНК не обязательно досконально прочитывать полученный текст, достаточно его бегло просмотреть и найти в нем соответствующее аномальное место.

В настоящее время известно свыше 1000 различных мутаций, ассоциированных с гемофилией В. Информация о них собрана в постоянно пополняющейся базе данных Haemophilia B Mutation Database (<http://www.kcl.ac.uk/ip/petergreen/haemBdatabase.html>).

В этой базе приводится не только характеристика мутаций, но и соответствующие им уровни активности фактора IX.

В России анализ мутаций в гене фактора IX для диагностики носительства и пренатальной (дородовой) диагностики гемофилии В проводится в лаборатории генной инженерии Гематологического научного центра (ФГБУ ГНЦ МЗ РФ, Москва) и Центре молекулярной генетики Медико-генетического научного центра (МГНЦ РАМН, Москва).

Что касается анализа мутаций, ассоциированных с гемофилией А, то большие размеры и достаточно сложная организация гена фактора VIII (только кодирующие области, разбитые на 26 экзонов, составляют около 7 тысяч нуклеотидов) долгое время существенно затрудняли их поиск и характеристику. Работы в этом направлении проводились лишь в немногих лабораториях наиболее развитых стран мира (Великобритания, Германия, Италия, США). Однако, в последние годы широкое внедрение методов молекулярной биологии в медицину, постоянно растущий интерес к популяционной генетике и совершенствование методов секвенирования ДНК привели к очевидной активизации исследований в области молекулярной генетики гемофилии А. Только в 2010-11 годах опубликовано несколько десятков работ, посвященных этой проблематике, причем к традиционным лидерам в этой области присоединился целый ряд стран Азии и Южной Америки. В настоящее время идентифицировано свыше 2000 различных генных дефектов, приводящих к гемофилии А. Наиболее полный их перечень содержится в международной базе данных HAMSTeRS (The Haemophilia A Mutation, Structure, Test, Resource Site; <http://europium.csc.mrc.ac.uk>). Во многих странах мира (Великобритания, Италия, Германия, Франция, Испания, Китай, Индия и др.) созданы аналогичные национальные и даже региональные регистры. **В России, к сожалению, работа по определению мутаций в гене фактора VIII в настоящее время только планируется.**

Большая часть выявленных нарушений гена фактора VIII приходится на миссенс- и нонсенс-мутации, но с достаточно высокой частотой встречаются также мутации в сайтах сплайсинга, микроделеции и микроинсерции. Значительно реже выявляются протяженные делеции и интронные мутации, приводящие к возникновению аномальных сайтов сплайсинга. В основном эти мутации уникальны, то есть, найдены у единичных пациентов, но есть два нарушения, которые во всех исследованных популяциях относятся к часто повторяющимся событиям. Это, инверсии, о которых упоминалось выше. И прежде всего, речь идет об инверсии в области интрона 22. Дело в том, что в интроне 22 гена фактора VIII есть участок в 9,5 тысяч нуклеотидов, который имеет еще две почти точные копии, но ориентированные в обратном направлении, которые находятся на расстоянии 300-400 тысяч нуклеотидов от начала гена. При пространственном сближении этого внутригенного участка с одной из своих внегенных копии происходит их взаимодействие, приводящее к тому, что ограниченная ими область гена, содержащая экзоны с 1-го по 22-ой, меняет свою ориентацию, т.е. попросту говоря, переворачивается на 180 градусов. Это и есть инверсия. В итоге получается такое производное гена фактора VIII, в котором сначала идут экзоны с 22-го по 1-ый, а далее следуют остальные экзоны с 23-го по 26-ой. Естественно, что

такая дикийная конструкция как ген работать не будет. **Инверсия с участием интрона 22 встречается в разных популяциях у 30-50% пациентов с тяжелой формой гемофилии А.** Это нарушение не выявляется при секвенировании, поскольку все функционально важные участки гена формально имеются в наличии и находятся в нормальном состоянии. Тестирование данной инверсии представляет собой отдельную и достаточно непростую задачу, поскольку для этого необходимо амплифицировать фрагменты ДНК длиной 10-12 тысяч нуклеотидов, в то время как стандартная ПЦР практически не дает возможности нарабатывать фрагменты длиннее 3 тысяч нуклеотидов. Поэтому для амплификации более протяженных последовательностей используют специальные смеси различных полимераз, и кроме того в таких случаях ужесточаются требования к качеству исходного препарата геномной ДНК, она должна быть очень высокомолекулярной. Такую же методику применяют для тестирования другой, более редкой инверсии в области интрона 1, происходящей по аналогичному механизму и встречающейся у 2-5% пациентов с тяжелой формой гемофилии А.

Особо следует отметить необычайно высокую частоту встречаемости при гемофилии А мутаций *de novo*, то есть вновь возникших (от 30 до 50% в разных популяциях). Имеются в виду случаи так называемой спорадической гемофилии, когда при отсутствии семейной истории заболевания в семье рождается больной мальчик. Этот факт, по всей вероятности, является следствием больших размеров и сложной организации гена фактора VIII, а также высокой повторяемости инверсионных мутаций, описанных выше. Для сравнения, частота встречаемости мутаций *de novo* при гемофилии В оценивается всего в 3%. Мутации при гемофилии, как и в случае любой другой наследственной патологии, почти всегда изначально возникают в единичных половых клетках в ходе сперматогенеза (наиболее распространенный вариант, более 90% случаев) или оогенеза и в очень редких случаях – в раннем эмбриогенезе. Появившийся на свет благодаря такой дефектной половой клетке ребенок становится первым больным, если это мальчик, или первым носителем заболевания, если это девочка, и именно с него начинается семейная история болезни. Вопрос с потомством мужчины, больного гемофилией, решается однозначно – сыновья всегда будут здоровы (речь, естественно, идет о гемофилии), а дочери всегда будут носительницами заболевания. У женщин, носительниц гемофилии, разумеется, есть альтернатива. Они могут с равной вероятностью передать в потомство как нормальный, так и мутантный ген. **Для исключения риска рождения ребенка, больного гемофилией, проводится дородовая (пренатальная) ДНК-диагностика.** Она дает возможность не только установить статус плода относительно гемофилии, но и не менее точно, чем это позволяют сделать методы хромосомного анализа, определить его пол. В зависимости от срока беременности используются 3 способа получения плодного материала: 1) взятие ворсин хориона или биопсия хориона (обычно на сроке 9-12 недель), 2) отбор амниотической жидкости (околоплодные воды) или амниоцентез (начиная с 13-14 недель) и 3) забор пуповинной крови или кордоцентез (после 22 недель). Предпочтительным, разумеется, является первый из этих способов, поскольку в случае неблагоприятного прогноза (больной мальчик) можно с наименьшими осложнениями прервать беременность на достаточно раннем сроке.

Для определения пола молекулярными генетиками была придумана очень простая и изящная методика, в основе которой лежит тесное эволюционное родство между X- и Y-хромосомами.

По существу Y-хромосома представляет собой вырожденную X-хромосому, утеревшую большую часть своих генов, и поэтому едва ли не все сохранившиеся на ней гены имеют на X-хромосоме свою пару. Кодирующие последовательности одноименных X- и Y-генов чрезвычайно консервативны, а вот не кодирующие интроны значительно отличаются не только по первичной структуре, но и, что самое существенное, по размеру (интроны X-генов обычно крупнее). Поэтому, для того, чтобы определить “мужская” ДНК или “женская”, амплифицируют такие интроны с помощью полимеразной цепной реакции, и, если ДНК “мужская” (X;Y), то в результате получаются два фрагмента разной длины, а если “женская” (X;X), то только X-фрагмент.

Для диагностики носительства и пренатальной ДНК-диагностики гемофилии А часто используется более доступная и быстрая, чем тестирование мутаций, методология, основанная на семейном анализе *полиморфизмов* гена фактора VIII и смежных с ним участков X-хромосомы, позволяющая осуществлять маркировку дефектного гена (для этого необходима ДНК больного из данной семьи) и проследить его судьбу в родословной. Под полиморфизмами в данном случае понимаются такие вариации в первичной структуре ДНК, которые фенотипически никак себя не проявляют. То есть в отличие от мутаций, они нейтральны. В структурном же плане они могут представлять собой то же, что и мутации. В частности, самым распространенным типом полиморфизмов ДНК являются одиночные нуклеотидные замены. В научной литературе их обозначают латинской аббревиатурой SNP (single nucleotide polymorphism). Большая часть полиморфизмов локализуется в функционально не значимых участках генома, в интронах или межгенных пространствах. Но они могут встречаться и в кодирующих областях. Например, в третьих вырожденных позициях кодонов, не влияющих на тип аминокислоты (так называемые синонимические замены). Понятно, что для ДНК-диагностики наследственной болезни могут быть использованы только полиморфизмы, локализованные внутри гена, с которым связано данное заболевание, или же в непосредственной близости от него. Тестировать такие полиморфизмы при помощи секвенирования неудобно, долго и достаточно дорого. Поэтому для их исследования используют специальные бактериальные ферменты, рестрикционные эндонуклеазы или, проще, рестриктазы. Рестриктазы узнают в ДНК строго определенные короткие нуклеотидные последовательности (сайты) обычно длиной в 4-6 нуклеотидов и разрезают ее в этом месте. Такой фермент, и иногда даже не один, есть в каждой бактерии. Это неотъемлемая составная часть ее “иммунной” системы. Другой ее частью является метилаза, узнающая тот же сайт, что и рестриктаза, и присоединяющая к его нуклеотидам метильные группы. Метилированный сайт рестриктазой уже не воспринимается. Таким способом хозяйская ДНК защищается от расщепления “своим” ферментом. А вот чужая ДНК, метилированная по другим сайтам или не метилированная вовсе, расщепляться будет. Так бактерия борется с проникшим в нее чужеродным генетическим материалом, в основном с бактериальными вирусами – бактериофагами. Рестриктазы из разных

бактерий узнают разные последовательности нуклеотидов, поэтому их ассортимент весьма широк, и почти всегда можно подобрать такой фермент, в участок узнавания которого попадает интересующая нас полиморфная нуклеотидная замена. В качестве живого примера можно привести полиморфизм, локализованный в последовательности АГЦТ интрона 1 гена фактора VIII. Это участок узнавания рестриктазы Alu I (название фермента составляется из первых букв латинского названия бактерии, из которой он получен, в данном случае - *Arthrobacter luteus*). Второй вариант этой последовательности - АГТТ из-за замены Ц на Т рестриктазой уже не узнается. Фрагмент гена, содержащий такую замену, амплифицируют с использованием ПЦР, обрабатывают рестриктазой и анализируют при помощи электрофореза. Если у консультируемой женщины оба гена по данному полиморфизму одинаковы, то на электрофореграмме мы увидим или только исходный неразрезанный фрагмент (генотип Т/Т), или только два укороченных фрагмента, получившихся после разрезания (генотип Ц/Ц). Такое состояние называется *гомозиготностью*. Если же гены разные, то мы одновременно увидим оба варианта (генотип Т/Ц). В этом случае мы говорим о *гетерозиготности*. Поскольку основная задача генетика при диагностике гемофилии А – найти такой полиморфизм, который позволит отличить нормальный ген фактора VIII от мутантного, то понятно, что именно состояние гетерозиготности является искомым. И чем больше в популяции число гетерозигот по данному полиморфизму, тем выше его диагностическая ценность (информативность). Для SNP максимально возможное значение гетерозиготности в популяции невелико. Оно наблюдается при одинаковой частоте встречаемости двух возможных вариантов и составляет 50%. Наибольшую информативность имеют полиморфизмы другого типа, так называемые микросателлитные повторы (STR – short tandem repeat), когда в ДНК много раз подряд повторяется один и тот же короткий мотив длиной обычно в 2, 3 или 4 нуклеотида. Например, ЦАГЦАГЦАГЦАГ... и так далее. Характеристикой конкретного варианта такого полиморфизма является число повторяющихся мотивов или единиц (в данном случае ЦАГ), и чем больше таких вариантов с разным числом мотивов, тем выше его гетерозиготность (информативность). Для некоторых полиморфизмов такого типа она превышает 90%. Несколько подобных генетических маркеров используются и для диагностики гемофилии А.

Эффективность и надежность описанного выше диагностического подхода напрямую зависят от количества применяемых полиморфных маркеров, их информативности (гетерозиготности) и относительного расположения в исследуемом гене. Анализ, проведенный при помощи нескольких информативных полиморфизмов, локализованных в разных частях гена, позволяет предельно снизить вероятность ошибочного диагноза, обусловленную возможной генетической рекомбинацией (*кроссинговером*). Дело в том, что в процессе деления половых клеток (мейоз), гомологичные (одноименные) хромосомы могут обмениваться своими частями. Это равный, честный обмен, без каких-либо потерь, но если такое событие затронет исследуемый ген, то оно может спутать все маркировки и привести к ошибочным выводам. Следует, однако, отметить, что кроссинговер, будучи распространенным явлением в целом, для данного конкретного гена – событие достаточно редкое. Даже в том случае, когда диагностика носительства гемофилии А проведена по единственному информативному полиморфному маркеру, вероятность ошибки не превышает 1%. Если же информативными оказались несколько маркеров из разных областей гена фактора VIII или соседних с ним участков X-хромосомы, то она и вовсе

близка к нулю. Основным недостатком данного подхода является невозможность точного установления носительства заболевания у женщин из семей со спорадической гемофилией (единственный больной в родословной, мутация *de novo*). В таких случаях достаточно надежно (с указанной выше малой вероятностью ошибки) определяется только отсутствие носительства. Альтернативный диагноз обозначается при этом не как достоверное носительство, а как высокий риск носительства (статистически около 90%). Кроме того, в отдельных ситуациях, для установления статуса консультируемой женщины относительно гемофилии необходимо бывает проведение достаточно обширного семейного исследования с привлечением значительного числа ее кровных родственников. И последнее, может быть, самое важное – в большинстве случаев успешная диагностика невозможна без наличия образца ДНК от больного из обследуемой семьи.

Молекулярно-генетическая диагностика гемофилии А, основанная на определении патологической мутации в гене фактора VIII, в значительной степени лишена этих недостатков и позволяет наиболее надежно определять статус пациентки относительно заболевания, в том числе и в семьях с мутациями *de novo*. Вместе с тем, значение мутационного анализа не ограничивается его практическим использованием в диагностических целях. Широкомасштабные исследования подобного рода, основанные на анализе биологического материала от сотен пациентов, позволяют выявить общий спектр мутаций, определить их частотные характеристики и выявить, таким образом, генетические нарушения, являющиеся особенностью изучаемой популяции. В ходе такого исследования всегда выявляется значительное число новых не встречавшихся ранее мутаций, среди которых особое значение имеют миссенс-мутации (и соответствующие им аминокислотные замены). Они позволяют получить новые сведения о структурно-функциональных особенностях белка фактора VIII, более подробно охарактеризовать участки, определяющие его внутримолекулярную стабильность, активацию тромбина и межмолекулярные взаимодействия с факторами IX и фон Виллебранда. Тестирование выявленных генных дефектов у родственников пациентов и сопоставление полученных результатов с данными семейного анализа полиморфных маркеров позволяют определить реальную частоту возникновения мутаций *de novo* в гене фактора VIII и установить характер их происхождения, герминальный (сперматогенез или оогенез) или эмбриональный. Значительный интерес представляет сопоставление результатов мутационного анализа с данными о наличии у пациентов ингибиторной формы гемофилии А, обусловленной развитием иммунного ответа на терапевтическое введение препаратов фактора VIII. Этому вопросу посвящено большое число работ. Согласно накопленным к настоящему времени данным синтез ингибиторных антител чаще всего провоцируется такими фатальными для FVIII мутациями, как инверсии, нонсенс-мутации, frameshift-мутации, протяженные делеции и мутации сплайсинга, но может быть опосредован и некоторыми миссенс-мутациями. Мутационный анализ дает возможность выявлять ответственные за ингибиторный эффект эпитопы FVIII (структурные элементы, распознаваемые ингибиторными антителами), что позволяет надеяться на создание в будущем его не иммуногенных терапевтических вариантов.