Геномная дактилоскопия

Введение

В течение последних нескольких лет в геноме человека были обнаружены. Такие гипервариабельные районы содержат набор коротких, обычно GC-богатых тандемно повторенных единиц. Именно в этих тандемно повторяющихся структурах заключена молекулярная основа вариабельности указанных последовательностей. Длинные гомологичные участки повторяющейся последовательности подвержены рекомбинации, происходящей, как правило, вследствие неравного обмена при мейозе или митозе или вследствие ошибок при репликации ДНК. Рекомбинационные события обусловливают аллельные различия в числе тандемно повторяющихся единиц, имеющихся в данном локусе ГВР, и вследствие этого являются причиной наблюдаемого полиморфизма длины всего тандемного блока. Степень гетерозиготности по локусам ГВР высока, и, следовательно, они могут быть использованы в качестве маркеров при картировании генов.

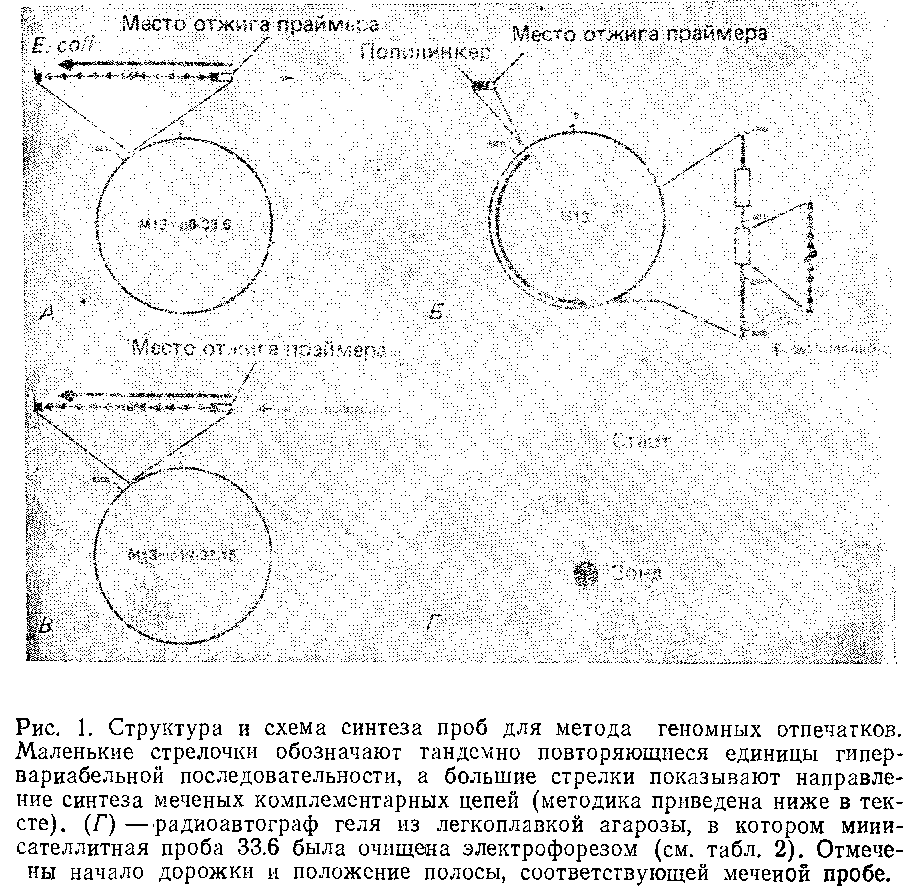
Недавно появились данные, свидетельствующие о том, что ГВР существуют в виде разбросанных по всему геному семейств, родство членов которых определяется гомологией с той или иной корпоследовательностью – основной единицей соответствующих тандемных повторов. Клонированные участки ГВР были использованы для приготовления проб, которые при гибридизации с геномной ДНК в мягких условиях выявляют множественные аллели ГВР. Получающийся сложный набор гипервариабельных полос и заключает в себе специфичную для данного индивидуума картину геномного «дактоотпечатка»; аллели, формирующие эту картину, обладают соматической стабильностью, стабильны в клетках зародышевого пути и наследуются в соответствии с правилами Менделя.

Таким образом, используя один зонд, можно одновременно наблюдать за наследованием большого количества аллелей. Эволюционная нестабильность, вследствие которой эти последовательности гипервариабельны, не настолько велика, чтобы затруднить сегрегационный анализ, поэтому метод геномной дактилоскопии может быть применен лри изучении генетического сцепления. Индивидуальный характер гибридизационной картины позволяет использовать метод в судебной медицине и применять его в качестве инструмента, с помощью которого с высокой степенью достоверности могут быть уточнены структуры родословных.

**1.** Гибридизационные зонды для геномной дактилоскопии

**1.1 Зонды на основе минисателлитов**

Джеффризу с соавторами удалось впервые показать, что зонды на основе корпоследовательности ГВР могут быть использованы для одновременного выявления большого количества соответствующих генетических локусов. С помощью зонда, синтезированного путем тандемного лигирования 33 повторяющихся элементов ГВР из гена человеческого миоглобина, они обнаружили на Саузернблотах, несущих гидролизат геномной ДНК человека, картину из большого числа полос. Некоторые из них оказались полиморфными последовательностями. Для выделения этих родственных участков исследователи отобрали из геномной библиотеки соответствующие клоны, используя в качестве зонда при гибридизации в мягких условиях конкатенированные повторы из гена миоглобина. Когда же эти выделенные клоны, в свою очередь, выступили в роли зондов, с помощью некоторых из них в геномной ДНК удалось выявить сложную картину гипервариабельных полос. Клоны содержали последовательности, имеющие области гомологии с тандемными повторами миоглобиновых ГВР, которые и представляли собой общие для этой группы корпоследовательности. Последовательности этого семейства были названы минисателлитами. Для изучения доступны два из них, являющиеся вариантами основной корпоследовательности. Соответствующие зонды, названные 33.6 и 33.15, существуют в виде рекомбинантных форм векторов, полученных на основе бактериофага М13. Каждая из них выявляет около 15 высокополиморфных полос в диапазоне 4–20 т. п.н. Среднее значение гетерозиготности по этим полосам около 90%, при этом вариабельность пропорциональна длине фрагмента и наибольший фрагмент фактически на 100% гетерозиготен в популяции.



**1.1.1** Приготовление минисателлитных зондов

Одноцепочечные матрицы для синтеза зондов можно легко и быстро получить, используя стандартные методики.

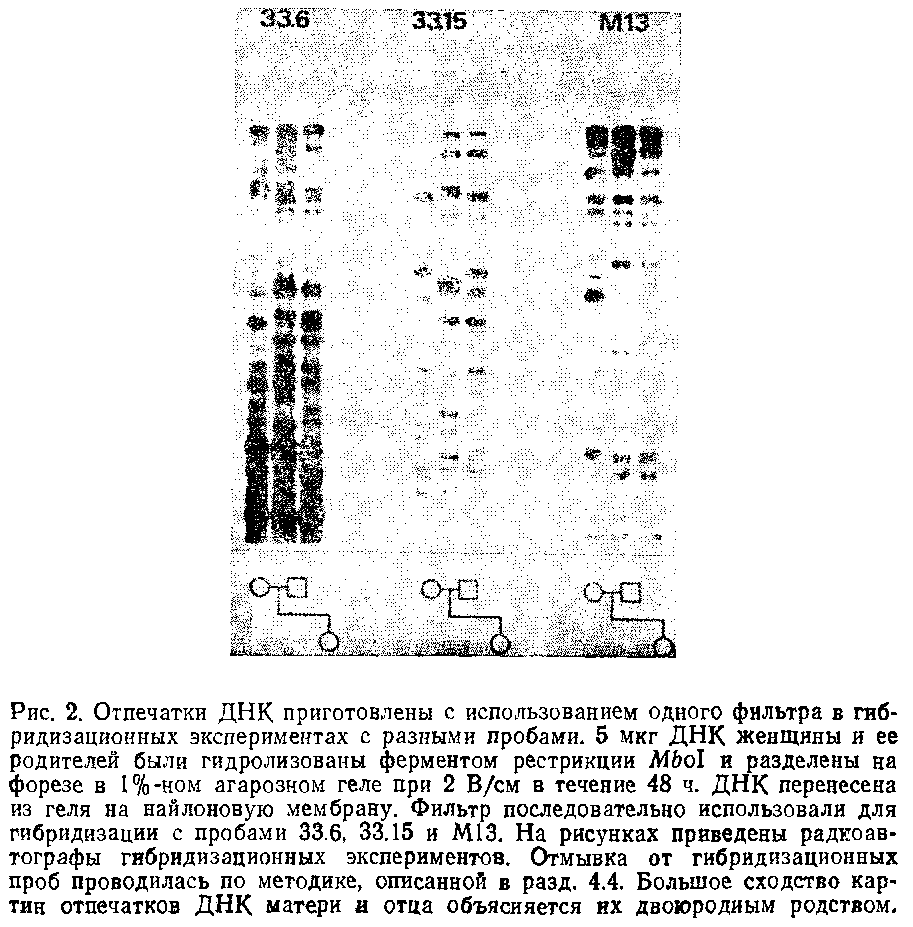
Матрицы, приготовленные по методике из табл. 1, обычно получаются в концентрации около 100 мг/мкл, однако лучше оценить концентрацию данной партии путем сравнения со стандартными растворами в геле, окрашенном бромидом этидия.

Учитывая, что молекула матрицы является одноцепочечной и кольцевой, ее нельзя пометить с помощью рассеянных затравок или методами никтрансляции. В этом случае следует воспользоваться техникой достройки праймера.

По этой методике синтетический олигонуклеотид, комплементарный последовательности, находящейся со стороны З-конца минисателлитной вставки, отжигается на матрице в растворе. В результате чего появляются З'-гидроксильные группы, к которым большой фрагмент ДНК-полимеразы I может присоединить меченые дезоксирибонуклеозиды. По истечении времени, необходимого для синтеза комплементарной цепи по всей длине вставки, можно использовать рестриктазу для расщепления новой двухцепочечной молекулы у б'-конца вставки.

Последовательность является частью фагового гена III, кодирующего белок, который участвует в прикреплении к бактериальным F-лилям в ходе инфекции и содержит два участка GC-богатых тандемных повторов. Эти последовательности отжигаются с конкурентной ДНК, содержащейся в большинстве гибридизационных буферов, поэтому для получения картины геномных «отпечатков пальцев» необходимо исключить этот компонент при использовании указанного зонда.

Семейство ГВР, выявляемое тандемным повтором из гена III бактериофага М13, отличается от семейств детектируемых с помощью минисателлитных зондов 33.6 и 33.15, но сходно с ними высокой степенью полиморфизма, а также тем, что частота аллелей уменьшается пропорционально их размерим. Средняя частота совпадения для одной полосы у индивидов европейского происхождения, не находящихся в родстве, составляет около 0,20 для фрагментов больше 2 т. п.н., причем для отдельного индивида таких полос выявляется 15–20. Таким образом, тандемный повтор М13 выявляет картину «отпечатков ДНК», сравнимую по своей сложности с картинами, получаемыми при использовании минисателлитов, и позволяет проанализировать 15 – 20 дополнительных гипервариабельных участков в ДНК человека и других млекопитающих. Так как повтор М13 не содержит сайтов узнавания для рестриктаз Hinll, Mbol, НаеШ и Alul и поскольку распределение детектируемых фрагментов сходно с таковым для отпечатков, полученных с помощью минисателлитов, представляется удобным и эффективным готовить блоты для гибридизации по очереди с каждой из трех, проб.



**1.2.1** Приготовление зонда на основе гена 111 фага **М13**

Тандемные повторы М13 локализуются в положениях 1700–1900 и 2300–2500 фагового генома. Можно использовать эти данные и свойства фагового вектора для синтеза гибридизационного зонда с высокой удельной активностью. После отжига 17-нуклеотидного секвенационного праймера на одноцепочечной ДНК-матрице любого из фагов М13 серии и достройки второй цепи фрагментом Клёнова с участием меченых дезоксинуклеозидтрифосфатов, реакцию останавливают добавлением соли. Если время инкубации выбрано таким образом, что удвоение цепи происходит не далее гена III, то вновь синтезированная цепь длиной до 4,5 т. п.н. содержит меченые нуклеотиды, а проксимальная часть генома фага М13 длиной 2,5 т. п.н., имеющая тандемные повторы, остается одноцепочечной. Это позволяет тандемному повтору гибридизоваться с комплементарными последовательностями ДНК, иммобилизованной на фильтре, и тем самым делает возможным использование меченой ДНК фага в качестве гибридизационной пробы. После освобождения от невключившихся меченых нуклеотидов можно измерить удельную активность ДНК-зонда, которая обычно варьирует в пределах 1–2ХЮ9 имп./мин/мкг ДНК матрицы. Однако, по-видимому, очистка зонда на колонке с сефадексом, снижает его гибридизационную активность, поэтому лучше избегать этого этапа и использовать в качестве зонда полученную реакционную смесь без дальнейшей ее очистки.

**2.** Гель-электрофорез **и** перенос ДНК

С обзором методов гель-злектрофореза можно познакомиться в работе. Однако полезно все же вкратце выделить специфические условия, повышающие эффективность анализа методом геномной дактилоскопии.

**2.1 Выбор фермента рестрикции**

Обсуждаемые молекулярные основы вариабельности членов семейств ГВР заключаются в разном количестве повторяющихся единиц в каждом ГВР-локусе. Чтобы выявить вариации подобного рода с возможно большим разрешением, необходимо гидролизовать ДНК ферментом рестрикции, который расщепляет достаточно часто, но не нарушает целостности повторяющихся единиц ГВР. Для обсуждаемых ГВР удобно использовать рестриктазы HinW, Mbol, AluA и Haelll.

В некоторых случаях встречаются фрагменты с такими близкими значениями Ш, что их нельзя разрешить на картине фореза. Тогда полезно повторить эксперимент с использованием другого фермента. При этом имеет место «перетасовка» рестриктов ГВР, что дает нам лишний шанс разделить сомнительные полосы. Когда используется фермент рестрикции, узнающий четырехнуклеотидную последовательность, эффект «перетасовки» обычно едва заметен, особенно для высокомолекулярных фрагментов, так как вклад в величину этих фрагментов фланкирующих последовательностей невелик, независимо от того, какой фермент используется. Тем не менее, некоторые полосы можно идентифицировать благодаря тому, что их подвижность по другим ферментам будет лишь немного изменена. Для достижения эффекта «перетасовки» наиболее подходящим считают фермент рестрикции Haelll.

**2.2 Приготовление образцов ДНК для электрофореза**

ДНК, получаемая обычным способом, обладает достаточно хорошим качеством для проведения анализа методом геномной дактилоскопии. Наилучшие результаты получают, если на дорожку приходится 1–5 мкг ДНК. Поскольку имеет значение не только присутствие, но и интенсивность отдельной полосы, важно достичь равномерного распределения материала между всеми дорожками геля. Хорошо известно, что измерения оптической плотности растворов ДНК дают ненадежные значения концентраций, которые могут рассматриваться только как грубые оценки. Для проверки полноты рестрикции и правильности нанесения необходимо поставить контрольный форез с аликвотами гидролизата ДНК и на основе этого подобрать оптимальное количество наносимой пробы.

Иногда для уменьшения наносимого объема бывает нужно осадить ДНК из рестрикционной смеси, но обычно в этой процедуре все же нет необходимости: хорошие результаты получаются и без этапа очистки. Если образец был осажден, очень важно промыть его 70%-ным этанолом для удаления избытка соли, которая нарушает подвижность фрагмента ДНК; важно также высушить осадок под вакуумом перед растворением» иначе присутствие следов этанола вызовет флотацию материала из лунок при нанесении.

2.3 **Условия электрофореза**

Оптимальные условия электрофореза для геля, используемого при геномной дактилоскопии, определяются требованиями к информативности получаемых отпечатков ДНК. Если данные необходимы для подтверждения структуры родословной или для окончательного уточнения картины в случае неудачно прошедшей гибридизации, или для сравнения образцов ДНК из разных тканей индивида, то желательно получить информацию о возможно большем числе полос, поддающихся разрешению. Для этих случаев наиболее удобно использовать 1%-ный агарозный гель и вести форез при 2 В/см в течение 48 ч с двумя сменами форезного буфера. Такие условия обеспечивают разрешение полос в диапазоне 2–25 т. п.н. Для сегрегационного анализа более важно достичь хорошего разделения наиболее высокомолекулярных гиперполиморфных фрагментов. В этом случае следует вести форез в 0,8%-ном геле при 1,5 В/см в течение 72 ч с пятью сменами форезного буфера. При этих условиях все фрагмент короче 3 т. п.н. выходят из геля, и в нем остаются оптимально разделенными более информативные длинные фрагменты, по которым наблюдается высокая степень гетероаиготности.

**2.4 Перенос по Саузерну**

После проведения электрофореза гель окрашивают бромидом этидия, и для оценки полноты проведения фореза и регистрации картины разделения фрагментов относительно маркеров размера гель фотографируют в УФ-свете.

Наиболее критическим моментом в приготовлении отпечатков ДНК с помощью зондов на основе ГВР является гибридизация и отмывка фильтров. Оптимальные условия варьируют в зависимости от характера зондов.

3. Использование метода геномной дактилоскопии в анализе генетического сцепления

**3.1 Обоснование метода**

Методический подход, получивший название «обратная генетика» и оказавшийся очень полезным при изучении наследственных заболеваний, в настоящее время взят на вооружение широким кругом исследователей. За последние годы наши представления о мышечной дистрофии Дюшенна, муковисцидозе и ретинобластоме значительно расширились. Если раньше усилия ученых были направлены на выяснение закономерностей наследоцания заболеваний, то сейчас на повестке дня стоит вопрос о локализации и молекулярной характеристике генов, ответственных за эти болезни.

Изучение большинства наследственных заболеваний шло именно таким путем. В некоторых случаях хромосомная локализация интересующего гена устанавливается достаточно легко. Это бывает, если у индивидов с аберрантным фенотипом наблюдаются устойчивые аномалии в определенной области хромосомы. К сожалению, для большинства наследственных заболеваний такая благоприятная в плане диагностики ситуация не характерна и для локализации гена необходимо проводить анализ генетического сцепления.

Применение ПДРФ в анализе сцепления стало обычной процедурой и значительно ускорило процесс картирования. Однако этот подход имеет и некоторые недостатки. В большинстве случаев ПДРФ представляет собой всего лишь диморфизм, и соответствующие локусы не могут поэтому характеризоваться более чем 50%-ной гетерозиготностью. Для того чтобы данное скрещивание дало какую-нибудь информацию о сцеплении, по крайней мере один из родителей должен быть гетерозиготным как по интересующей, так и по маркерной области.

Таким образом, большая часть семей, подвергнутых анализу на сцепление с применением ПДРФ, окажется совершенно неинформативной. Это пример неэффективного использования информационных ресурсов родословной. Как инструмент в анализе сцепления метод геномных отпечатков нашел применение в случаях тех заболеваний, биохимическая природа которых пока неизвестна или когда приуроченность к подозреваемому гену была исключена. Поскольку не существует простого метода определения, какая из полос «отпечатка» генома данного индивида представляет ту или иную полосу в картине другого индивида, обычная процедура получения данных о сцеплении по нескольким малым родословным невозможна. Сложность гибридизационных картин не позволяет достоверно проследить наследование аллельных пар, но в рамках большой родословной и, конечно же, в груше сибсов отдельный фрагмент данного размера может рассматриваться для простоты предварительного анализа как отдельный локус. По этой причине наличие одной большой родословной является обязательным требованием для традиционного сегрегационного анализа полос в геномном «отпечатке». Такие родословные редко когда можно получить для рецессивно наследуемых нарушений, однако для доминантных заболеваний – это реально.

Для сегрегационного анализа полос в «отпечатках» ДЬЩ в силу допущенного упрощения, сделанного в отношении их аллелизма, классический метод правдоподобия не может быть, применен. В данном случае допустима лишь оценка шансов на сцепление. Таким образом, метод следует рассматривать как быстрый предварительный анализ на сцепление. Если результат окажется обнадеживающим, т.е. если в отпечатках будет найдена полоса, наличие которой для данной родословной коррелирует с экспрессией аномального фенотипа, то эту полосу следует выделить и клонировать. Таким путем можно получить локус-специфичный зонд и с его помощью проанализировать другие родословные с целью подтвердить или отвергнуть наличие сцепления.

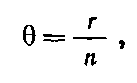
Недостаток больших родословных для рецессивно наследуемых признаков, конечно, ограничивает применение метода геномной дактилоскопии, но не исключает полезность его использования для анализа рецессивных болезней. Если в родословной имеются близкородственные браки, возможно также картирование по гомозиготности. Суть этого метода в следующем: если оба родственных индивида несут редкий рецессивный тену то вероятнее всего они унаследовали его от единого предка. Если индивиды находятся в достаточно далеком родстве, то общие для них последовательности будут составлять лишь малую долю генома. К тому же поскольку аллельные частоты фрагментов, образующих полосы в отпечатках, очень малы, то маловероятно, что в геноме родственников эти фрагменты совпадут, если они произошли от разных предков. Если при близкородственном скрещивании больные дети имеют общую для их геномных отпечатков полосу в удвоенном количестве, а здоровые дети наследуют одинарную дозу или вовсе не имеют этой полосы, то тем самым подтверждается гипотеза о физическом сцеплении между участком, ответственным за признак, и данной полосой. Условившись об аллелизме, можно подсчитать шансы на сцепление, но, как уже отмечалось, для подтверждения или опровержения наличия сцепления фрагмент необходимо клонировать.

**3.2 Статистические расчеты**

На блоте среднего качества с помощью одной пробы можно проследить до 10–20 полос, что в сумме для трех проб дает 30–60 маркеров. Генетические данные указывают на то, что не существует заметной кластеризации ГВР, поэтому есть основания предположить их независимое расположение. Вероятность Р того, что по крайней мере один из 45 независимо расположенных маркеров находится на генетическом расстоянии в 15% от сайта, ответственного за признак, может быть оценена с помощью алгоритма Элстона и Ланга равной приблизительно 0,33. Таким образом, при наличии соответствующей родословной этот метод позволяет успешно обнаружить сцепление.

Необходимо определить критическое значение вероятности сцепления с тем, чтобы после подсчета этой вероятности на основе предварительного анализа отпечатков можно было сделать заключение о необходимости следующей стадии – клонирования «подозрительной» полосы. В классических методах правдоподобия критическое значение лод-балла составляет 3: г=3; лод-балл, превышающий 3, рассматривается как свидетельство в пользу сцепления. Учитывая, что затраты на клонирование фрагмента достаточно велики, для перехода к этому этапу оправданным представляется требование уверенного превышения критической величины лод-балла на этапе предварительного анализа. Как же рассчитываются шансы на сцепление?

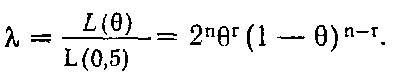
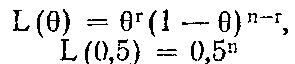
Легко получить численную оценку рекомбинации между маркерной областью и участком, ответственным за признак:



где г – количество рекомбинантов, п – общее количество возможных рекомбинаций.

Исходя из этого, обычным путем можно получить соотношение для оценки вероятности сцепления:

и



Это соотношение не учитывает поправку на использование множественных маркеров. Заметим, однако, что существуют некоторые сомнения в необходимости коррекции результатов в тех случаях, когда используется менее 100 маркеров. По мнению Отта, при увеличении количества маркеров с ростом вероятности ошибочного выявления сцепления растет и вероятность того, что хотя бы один из маркеров действительно сцеплен с изучаемым признаком, при этом последняя вероятность растет быстрее, чем первая, поэтому доля истинных сцеплений среди всех выявленных не уменьшается. Являясь математически убедительным, этот вывод противоречит интуиции. Высокий уровень ошибочных позитивных ответов на предварительном этапе сделал бы метод отпечатков малоинформативным. В связи с этим представляется целесообразным сделать условие перехода ко второй стадии анализа более строгим, подняв значение порога. В будущем при наличии более убедительной модели, описывающей эту задачу, можно было бы/его снизить.

3.3 **Практические аспекты сегрегационного анализа**

3.3.1 Минимальный размер родословных

Для того чтобы сегрегационный анализ полос в отпечатках ДНК сибсов был информативным, необходимо изучить не менее 10 мейозов. Такое количество сибсов далеко не всегда доступно для изучения, и поэтому приходится проводить сегрегационный анализ по большой родословной.

Для фрагментов короче 3 т. п.н. частота встречаемости одного и того же фрагмента у индивидов, не находящихся в родстве, очень мала. Таким образом, необязательно иметь для ист следований образцы ДНК обоих родителей близнецов; при наличии отпечатков ДНК одного из родителей отпечатки второго можно восстановить с высокой степенью достоверности. Если есть ДНК обоих родителей, то образцы следует разгонять на форезе в соседних дорожках геля. В этом случае фрагменты в геномных отпечатках потомков можно с легкостью отнести к тому или иному родителю.

3.3.2 Пример с сегрегационного анализа, проводимого с помощью метода геномных отпечатков

Ниже приводится пример анализа наследования полос в геномных отпечатках отдельной широкой сибсовой ветви.

Из-за несовершенства агарозного геля при электрофорезе лучше ие использовать для считывания данных отпечатков компьютерную технику, а «прочитать» их «вручную». Это легко проделать, если отдельно исследовать каждого из сибсов, параллельно проводя сегрегационный анализ полос в отпечатках каждого из родителей. Удобно присваивать полосам одного родителя цифры, а полосам другого – буквы.

После регистрации и отсчета полос для каждого из сибсов можно проанализировать полученные результаты на наличие сцепления. Если рассматривается гипотеза сцепления локуса, ответственного за изучаемый наследственный дефект, с какой-либо из полос в отпечатках ДНК больного родителя, то можно сосчитать число рекомбинантных и нерекомбинантных индивидов, численно оценить долю рекомбинантов и рассчитать вероятность. Эта процедура проделывается для каждой из полос в отпечатках ДНК больного родителя.