

КАРТИРОВАНИЕ ГЕНОВ, ДЕТЕРМИНИРУЮЩИХ РАСПРОСТРАНЕННЫЕ
БОЛЕЗНИ ЧЕЛОВЕКА

Аксенович Т.И.

Институт цитологии и генетики Сибирского отделения Российской академии наук,

Новосибирск, 630090, пр. Лаврентьева, 10, факс: 3833331257,

e_mail:aks@bionet.nsc.ru

АННОТАЦИЯ

Медленный прогресс в локализации и идентификации генов, контролирующих распространенные болезни человека, обычно объясняют сложностью их генетического контроля и неспособностью существующих методов учитывать специфику этого класса болезней. В данной работе мы проанализировали возможности и ограничения существующих методов картирования и показали, что сейчас мы умеем картировать лишь те гены, индивидуальный вклад которых в контроль болезни значим. Но даже для этих генов результаты часто неоднозначны и противоречивы, и это, в частности, объясняется неправильным формированием выборки.

КЛЮЧЕВЫЕ СЛОВА: картирование, сцепление, неравновесие по сцеплению, ассоциации, генетическая гетерогенность

В последние десятилетия был достигнут значительный прогресс в понимании генетической природы редких моногенных болезней [5, 18]. Этому в немалой степени способствовало создание новых методов молекулярной генетики, которые дали в руки исследователей огромное число генетических маркеров, позволили секвенировать участки генома, выявляя патогенные мутации, и анализировать экспрессию различных генов. Однако, несмотря на прогрессирующие успехи молекулярной генетики, приведшие к секвенированию генома, мы до сих пор не можем заметно продвинуться в понимании механизмов генетического контроля другой группы патологий, имеющих сложную генетическую детерминацию и называемых комплексными. К этой группе относятся все распространенные болезни человека. Не удивительно, что идентификация генов, детерминирующих предрасположенность к этим болезням, является одной из самых важных задач современной генетики.

МОНОГЕННЫЕ И КОМПЛЕКСНЫЕ БОЛЕЗНИ

Чем же комплексные болезни так сильно отличаются от моногенных, что методы, прекрасно зарекомендовавшие себя при изучении последних, оказываются здесь бессильными?

Моногенные болезни, называемые еще менделевскими, характеризуются простотой генетического контроля. Для них выполняется правило классической генетики: один ген - один белок - один признак (болезнь). Другими словами, причиной каждой моногенной болезни является повреждение одного определенного гена (генетическая однородность болезни), и проявление мутации этого гена слабо

модифицируется другими генами и средовыми воздействиями (полная пенетрантность генотипов). Именно на этих свойствах моногенных болезней были основаны методы картирования и идентификации контролирующих их генов.

Для большинства комплексных болезней изменение структуры белка не является ни необходимым, ни достаточным условием развития патологии. В их контроле принимает участие большое число генов, значительная часть которых является регуляторными. Вследствие этого основной причиной болезни становится не нарушение структуры белка, а изменение его концентрации в определенном месте и в определенное время. Комплексные болезни имеют более разнообразное фенотипическое проявление, чем моногенные, и не всегда можно однозначно поставить диагноз. Манифестация болезни существенно зависит от возраста и внешних условий (неполная пенетрантность генотипов). Это подтверждается результатами близнецового анализа: конкордантность по распространенным болезням у монозиготных близнецов всегда ниже 100% [20, 39]. Часто в семьях встречается сочетание нескольких патологий (синтропия). По-видимому, для комплексных болезней, в отличие от менделеевских, уместно правило: много генов – много вариантов их экспрессии – много фенотипических проявлений. Картирование генов, контролирующих столь сложные признаки, является чрезвычайно трудной задачей, которая в общем виде вряд ли может быть решена в настоящее время.

Обычно считается, что комплексные признаки по типу контроля являются мультифакториальными, т.е. вклады отдельных генетических и средовых факторов пренебрежимо малы, и предрасположенность или устойчивость к болезни определяется сочетанием большого числа малых эффектов [14]. В настоящее время

мы не умеем ни картировать, ни идентифицировать такие гены. Однако, даже если в среднем по популяции эффекты отдельных генов малы, то в разных семьях, у разных больных решающим в развитии патологии может оказаться небольшой набор генов значимого эффекта. В тех ситуациях, когда вклад гена оказывается значимым, этот ген становится доступным для анализа. Опыт показывает, что практически все идентифицированные гены, отвечающие за предрасположенность к болезням, были обнаружены при анализе семей или ограниченных популяций, в которых они обладали мажорным эффектом.

Существует несколько подходов к идентификации таких генов. Основной из них, называемый позиционным, базируется на сканировании генома. Каждый участок генома последовательно тестируют на причастность к болезни, оценивая вероятность локализации в нем гена, контролирующего болезнь. Участки, характеризующиеся максимальными вероятностями, рассматривают как места возможной локализации искомым генов. Хотя для комплексных и моногенных болезней методы картирования генов значительно различаются, их объединяет общая идеология – на каждом этапе анализа тестируется один локус генома. Теоретически в этом локусе может находиться несколько генов, отвечающих за предрасположенность к болезни, но на практике, как правило, в каждом локусе удается идентифицировать только один такой ген. Методы, основанные на тестировании отдельных локусов не позволяют учитывать эффекты взаимодействия генов и для анализа комплексных болезней являются слишком упрощенными. Однако, в настоящее время нет корректных методов одновременного поиска нескольких взаимодействующих локусов значимого эффекта. Предлагается сначала

локализовать отдельные гены, а потом анализировать эффекты их комбинирования и взаимодействия [16].

СКАНИРОВАНИЕ ГЕНОМА: АНАЛИЗ СЦЕПЛЕНИЯ И АНАЛИЗ АССОЦИАЦИЙ

В основе картирования генов лежат хорошо известные биологические явления: сцепление генов, их рекомбинация во время мейоза и полиморфность генома.

Благодаря сцеплению, мутация, детерминирующая болезнь, передается потомкам вместе с блоком окружающих ее аллелей соседних локусов. Рекомбинация со временем уменьшает размер этих блоков. Чем ближе расположены два локуса, тем дольше их аллели сохраняются в одном блоке. Идентификация блоков, полученных от различных родителей, обеспечивается полиморфностью генома, многие локусы которого имеют не один, а несколько вариантов нуклеотидных последовательностей. Такие локусы служат генетическими маркерами. Для того, чтобы картировать ген, вызывающий болезнь, достаточно доказать совместную сегрегацию болезни и блока маркерных аллелей.

Существует два методических подхода к картированию генов, контролирующих сложные болезни человека: анализ сцепления и анализ ассоциаций.

Основная идея анализа сцепления, или рекомбинационного анализа заключается в поиске блока маркеров, которые передаются от больного родителя преимущественно больным потомкам и не передаются здоровым. В разных семьях аллельный состав таких блоков может различаться, но их позиция в геноме должна быть одинакова. Информативными для анализа сцепления являются только те локусы, которые помечены гетерозиготными маркерами. Поэтому

предпочтительными для анализа являются полиморфные маркеры, имеющие много аллелей. Материалом для анализа сцепления всегда служат родственники: это могут быть пары больных сибов или расширенные родословные. Анализ сцепления позволяет локализовать ген на участке в 5-50 сМ. Это происходит потому, что доступными для генотипирования являются представители не более 3-5 поколений. За это время реализуется немного рекомбинационных событий, и блоки передаваемых генов велики. Идентификация косегрегирующих с болезнью блоков осуществляется с помощью различных методов статистического анализа. Самыми мощными считаются методы, базирующиеся на известной модели наследования признака [35], установление которой для комплексных болезней – достаточно сложная задача. Поскольку искажение модели наследования приводит к потере мощности, более популярными являются статистические методы, свободные от модели наследования. В их основе лежит анализ идентичности по происхождению маркерных аллелей у пар больных родственников [17].

Второй метод картирования – анализ ассоциаций или неравновесия по сцеплению. Неравновесие по сцеплению между двумя аллелями разных локусов выражается в том, что частота их совместной встречи в популяции отличается от ожидаемой при случайной независимой встрече. Одной из основных, хотя и не единственной, причиной существования неравновесия по сцеплению в популяции является тесное сцепление. Например, если в момент возникновения мутации, вызывающей болезнь, рядом находился определенный маркерный аллель, то в течение многих поколений этот аллель будет передаваться вместе с мутацией. Рекомбинация постепенно разрушает ассоциацию и происходит это тем быстрее, чем

дальше друг от друга расположены локусы. Для тесно сцепленных локусов неравновесие по сцеплению сохраняется десятки поколений [26].

Основная идея картирования с помощью анализа ассоциаций заключается в следующем. Если у большинства больных в популяции мутантный аллель имеет общее происхождение, окружающие маркеры находятся с ним в неравновесии по сцеплению. Для локализации гена, контролирующего болезнь, надо найти такой маркер, один из аллелей которого преобладает у больных. В отличие от анализа сцепления, здесь предполагается, что у больных из разных семей этот маркер имеет не только одинаковую локализацию в геноме, но и содержит один и тот же аллель. Поэтому при анализе ассоциаций не надо исследовать родословные, материалом для этого анализа могут служить независимые группы больных и здоровых людей. Тем не менее, предположение об общности мутации у большинства больных означает наличие общего предка, существовавшего много поколений назад. За это время, необходимое для распространения болезни в популяции, произошло много рекомбинационных событий, и неравновесие по сцеплению могло сохраниться только между мутацией и аллелем тесно сцепленного маркера. Поэтому с помощью анализа неравновесия по сцеплению удастся локализовать ген на участке менее 1 сМ. Маркеры должны плотно покрывать генетическую карту и число аллелей не должно быть слишком большим. Идеальными маркерами для анализа неравновесия по сцеплению являются SNP-маркеры, характеризующиеся полиморфизмом единичных нуклеотидов.

Как видно, анализ ассоциаций обладает рядом преимуществ, а именно, он может осуществляться на популяционных данных, для него подходят SNP-маркеры, он обладает высокой разрешающей способностью. Вместе с тем, анализ ассоциаций имеет ряд недостатков. Как показывает опыт, воспроизводимость результатов, полученных этим методом, крайне низка [7, 13]. Так Хиршхорн с соавторами обнаружили, что из 166 повторно тестируемых локусов только шесть продемонстрировали ассоциацию во всех повторах [13]. В чем здесь причина? Почему результаты, полученные с помощью анализа ассоциаций, часто не подтверждаются на других выборках или в других популяциях? Почему в локусах, указанных этим методом, часто не находят генов, контролирующих болезнь?

АНАЛИЗ АССОЦИАЦИЙ: ОГРАНИЧЕНИЯ И ПЕРСПЕКТИВЫ

Основной причиной получения ложно положительных результатов является то, что тесное сцепление генов - не единственная причина возникновения неравновесия по сцеплению. Оно может появиться в выборке из-за кластеризации данных или подразделенности популяции [2]. Кластеризация выражается в том, что при наследственной патологии в группу больных часто попадают близкие родственники. Если в семье таких людей с большой частотой встречается какой-то аллель любого локуса, то его частота будет повышенной в группе больных, независимо от локализации маркера. В этом случае на основании анализа ассоциаций будет указан локус, не содержащий искомого гена, и дальнейшие исследования будут направлены по ложному пути. Проблема кластеризации данных является достаточно серьезной, и сейчас разработан ряд пакетов программ, позволяющих установить степень родства

двух людей на основании информации об их маркерных генотипах [1]. Без такой предварительной проверки положительный результат анализа ассоциаций нельзя интерпретировать как указание на сцепление маркера с искомым геном.

Вторая причина, когда может быть получен ложно-положительный результат – подразделенность популяций. Популяции человека не являются панмиксными. В них всегда присутствует подразделенность, основанная на этнических, религиозных, социальных, культурных особенностях [10]. Если в одной субпопуляции одновременно наблюдается повышенная частота болезни и повышенная частота какого-то аллеля маркера, то ассоциация болезни с этим аллелем будет установлена независимо от его положения. Поскольку контролировать подразделенность человеческой популяции чрезвычайно сложно [21], было предложено использовать родительский контроль [30]. В этом случае вместо группы больных и здоровых людей набирается только группа больных, но у каждого из них определяются генотипы родителей. Из четырех родительских аллелей маркерного гена только два передаются больному потомку. Два других составляют контрольную группу. Если частота какого-то аллеля в группе переданных оказывается выше, чем в группе переданных аллелей, то ассоциация считается установленной. Эта ассоциация уже однозначно интерпретируется как сцепление и указывает на расположение искомого гена в геноме. С помощью родительского контроля решается также и проблема кластеризации данных.

Другое ограничение метода связано с тем, что с его помощью не всегда можно обнаружить сцепление, устанавливаемое с помощью рекомбинационного анализа (низкая мощность). Одной из причин этого служит постепенное разрушение

неравновесия по сцеплению между мутацией и аллелем соседнего локуса, возникшее в результате тесного сцепления. Чаще всего это происходит из-за рекомбинаций между геном, контролирующим болезнь, и маркерным локусом. В результате этого рядом с мутировавшим геном оказывается другой маркерный аллель. Поэтому при картировании с помощью анализа ассоциаций следует использовать очень плотные генетические карты, чтобы для каждой возможной позиции искомого гена существовал близко расположенный маркер. Другая причина разрушения неравновесия по сцеплению – мутации в маркерных локусах. К счастью, они возникают достаточно редко, особенно, если в качестве маркеров используются SNP [19].

Помимо разрушения неравновесия по сцеплению, к снижению мощности приводит разное происхождение мутации в гене, контролирующем болезнь. Для распространенных болезней человека предположение об общности происхождения мутации в большой открытой популяции, на котором основано использование неравновесия по сцеплению для картирования генов, кажется весьма спорным. Разное происхождение мутаций в одном и том же гене подтверждается тем, что для многих генов, участвующих в детерминации болезни, обнаружены различные патогенные мутации, приводящие к нарушению одного и того же метаболического процесса и имеющие одинаковое фенотипическое проявление (см., например, [11]). Очевидно, что такие мутации впервые возникали у разных людей, и что у каждого из них рядом с мутировавшим геном могли находиться разные маркерные аллели.

Все перечисленные факторы снижают мощность анализа ассоциаций, но не анализа сцепления, поскольку здесь нет привязки к какому-то определенному аллелю

маркера. Оптимальной является стратегия, когда сначала с помощью анализа сцепления выявляется крупный блок, содержащий картируемый ген, а затем с помощью анализа ассоциаций этот блок сужается [2].

Еще одна ситуация, когда анализ ассоциаций не выявляет ген, специфически влияющий на предрасположенность к болезни, возникает, когда максимальный вклад в предрасположенность к болезни в разных семьях обеспечивается разными генами и в то же время у большинства больных повышена частота аллеля, оказывающего слабый неспецифический эффект на предрасположенность к любым болезням. Тогда при анализе группы больных из разных семей эффект этого неспецифического аллеля может оказаться больше, чем любого из специфичных генов, представленных лишь у отдельных больных, и этот неспецифический локус будет картирован в первую очередь. Не удивительно поэтому, что большую часть вариабельности количественных признаков, связанных с той или иной патологией, можно объяснить генотипами локуса HLA [7, 36].

ГЕНЕТИЧЕСКАЯ ГЕТЕРОГЕННОСТЬ: ПОДХОДЫ К РЕШЕНИЮ ПРОБЛЕМЫ

Одной из основных проблем генетического анализа распространенных болезней человека является их генетическая гетерогенность, выражающаяся в том, что в разных семьях и у разных людей в контроле болезни принимают участие разные наборы генов. Эта гетерогенность затрудняет картирование независимо от того, каким методом оно осуществляется и требует для анализа выборки огромного размера. Чтобы решить эту проблему, необходимо снизить генетическую

гетерогенность в анализируемой выборке. Сделать это можно различными способами.

Первый из них заключается в анализе отдельных форм болезни. Для многих болезней с поздним возрастом проявления известны случаи, когда болезнь проявляется относительно рано. Было показано, что частота больных среди родственников пробанда особенно высока в семьях с ранним проявлением болезни [23]. Это свидетельствует о высоком вкладе генетической компоненты в контроль ранних форм болезни и позволяет думать, что в их детерминации участвует небольшое число генов. Перспективность использования этого подхода продемонстрирована на примере ряда болезней: рака молочной железы [31, 41], диабета второго типа [6], рака толстого кишечника [27], болезни Альцгеймера [9] и др. Однако решить все задачи картирования с помощью этого подхода вряд ли удастся – не для каждой болезни существуют рано проявляющиеся формы, и при их анализе выявляются только те гены, которые специфичны для данной формы болезни.

Второй подход заключается в изучении изолированных популяций человека [8, 28, 29, 32]. В таких популяциях велик эффект основателя, в результате чего снижена генетическая гетерогенность распространенных там болезней. Сейчас созданы проекты по изучению ряда изолятов, и уже получены первые результаты [2, 4, 24, 37]. Изолированные популяции дают совершенно уникальный материал для генетического анализа распространенных заболеваний, но вряд ли можно рассчитывать, что с их помощью будут выявлены все гены значимого эффекта, встречающиеся в открытых популяциях.

В этом смысле более перспективным кажется анализ расширенных родословных [3, 12, 34]. В каждой из таких родословных значительно сужен набор генов, участвующих в контроле болезни. Поэтому если размер родословной велик, на ее материале можно картировать гены, оказывающие в этой семье большой вклад в предрасположенность к болезни (см., например, [22]). Наличие родословных позволяет осуществлять картирование, используя как анализ сцепления, так и анализ ассоциаций, а для анализа ассоциаций обеспечивается родительский контроль. Кроме того, имея набор расширенных родословных, можно картировать несколько генов, доминирующих в каждой из семей. И наконец, как правило, в любой большой родословной встречаются разные болезни из списка наиболее распространенных, и появляется возможность одновременно анализировать несколько патологий и изучать плеiotропное действие генов.

Таким образом, анализ основных принципов картирования генов, контролирующих распространенные болезни человека, выявляет ряд ограничений методов, основанных на тестировании неравновесия по сцеплению. Многих проблем удастся избежать при правильном формировании выборки. Например, частота ложно-положительных результатов снижается, если вместо сравнения групп больных и здоровых используется родительский или родственный контроль. Проблемы, связанные с генетической гетерогенностью распространенных болезней, решаются подбором таких стратегий сбора материала, которые уменьшают эту гетерогенность в выборке. Наиболее перспективной кажется стратегия, основанная на сборе больших родословных. Она позволяет использовать различные методы анализа,

выявлять широкий спектр генов, контролирующих болезнь, учитывать плеiotропные эффекты генов.

Сегодня ни у кого не вызывает сомнения, что идентификация генов, контролирующих распространенные болезни человека, является чрезвычайно сложной задачей. Для ее решения необходимо рассматривать полиморфизм многих сайтов, расположенных по всему геному. Именно поэтому современные методы картирования базируются на сканировании генома с помощью анализа сцепления или анализа ассоциаций. На первый взгляд кажется, что такое рассмотрение учитывает многие особенности комплексных болезней: участие большого числа генов, их взаимодействие. Однако, несмотря на вовлечение в анализ всего генома, на каждом отдельном этапе анализа рассматривается только один локус и тестируется его причастность к контролю болезни. Тестируя шаг за шагом все регионы генома, мы можем идентифицировать те его сайты, полиморфизм которых ассоциирован с болезнью. Сегодня не существует корректных методов, позволяющих одновременно тестировать эффекты нескольких локусов, учитывая их взаимодействие [16]. Разработанные методы проверяют совместные эффекты только тех локусов, значимость вклада которых доказана при их индивидуальном тестировании [15, 16, 25]. Таким образом получается, что общая идеология анализа комплексных болезней мало чем отличается от идеологии моногенных болезней – в поле зрения исследователей оказываются лишь те гены, индивидуальный вклад которых в контроль признака значим [33, 34, 39]. Конечно, такой подход существенно ограничивает наши возможности, и сейчас ведется интенсивная работа по созданию новых методов и подходов, учитывающих специфику комплексных болезней [14, 25,

38]. Тем не менее, существует много ситуаций, где традиционный подход оказывается эффективным. Именно ему мы обязаны теми знаниями о генетической природе комплексных болезней, которыми уже располагаем, именно с помощью этого подхода выполняются все работы по картированию генов, публикуемые в каждом номере ведущих генетических журналов. Это позволяет надеяться, что возможности существующих методов еще не исчерпаны, и с их помощью можно получать новую информацию о генетике распространенных болезней.

Работа выполнена при поддержке грантов Российского фонда фундаментальных исследований (04-04-48074) и NWO (047-016-009).

Аксенович Татьяна Иосифовна
Новосибирск, 630090, пр.Лаврентьева, 10,
Институт цитологии и генетики СО РАН
(383)333-28-40
aks@bionet.nsc.ru

ЛИТЕРАТУРА

1. Abecasis G.R., Cherny S.S., Cookson W.O., Cardon L.R. GRR: graphical representation of relationship errors // *Bioinformatics*. 2001. Vol.17, № 8, P.742-743
2. Abecasis G.R., Ghosh D., Nichols T.E. Linkage disequilibrium: ancient history drives the new genetics // *Hum. Hered.* 2005. Vol.59, № 2, P.118-124
3. Almasy L., Blangero J. Challenges for genetic analysis in the 21st century: localizing and characterizing genes for common complex diseases and their quantitative risk factors // *Gene Screen*. 2000. Vol.1, P.113-116
4. Aulchenko Y.S., Vaessen N., Heutink P., Pullen J., Snijders P.J., Hofman A., Sandkuijl L.A., Houwing-Duistermaat J.J., Edwards M., Bennett S., Oostra B.A., van Duijn C.M. A genome-wide search for genes involved in type 2 diabetes in a recently genetically isolated population from the Netherlands // *Diabetes*. 2003. Vol.52, № 12, P.3001-3004
5. Botstein D., Risch N. Discovering genotypes underlying human phenotypes: past successes for mendelian disease, future approaches for complex disease // *Nat. Genet.* 2003. Vol.33 Suppl., P.228-237
6. Bowden D.W., Akots G., Rothschild C.B., Falls K.F., Sheehy M.J., Hayward C., Mackie A., Baird J., Brock D., Antonarakis S.E., et al. Linkage analysis of maturity-onset diabetes of the young (MODY): genetic heterogeneity and nonpenetrance // *Am. J. Hum. Genet.* 1992. Vol.50, № 3, P.607-618
7. Cardon L.R., Bell J.I. Association study designs for complex diseases // *Nat. Rev. Genet.* 2001. Vol.2, № 2, P.91-99
8. Chapman N.H., Thompson E.A. Linkage disequilibrium mapping: the role of population history, size, and structure // *Adv. Genet.* 2001. Vol.42, P.413-437

9. Chartier-Harlin M.C., Araria-Goumidi L., Lambert J.C. Genetic complexity of Alzheimer's disease // *Rev. Neurol. (Paris)*. 2004. Vol.160, № 2, P.251-255
10. Freedman M.L., Reich D., Penney K.L., McDonald G.J., Mignault A.A., Patterson N., Gabriel S.B., Topol E.J., Smoller J.W., Pato C.N., Pato M.T., Petryshen T.L., Kolonel L.N., Lander E.S., Sklar P., Henderson B., Hirschhorn J.N., Altshuler D. Assessing the impact of population stratification on genetic association studies // *Nat. Genet.* 2004. Vol.36, № 4, P.388-393
11. Gent J., Braakman I. Low-density lipoprotein receptor structure and folding // *Cell Mol. Life Sci.* 2004. Vol.61, №19-20, P.2461-2470
12. Gulcher J.R., Kong A., Stefansson K. The role of linkage studies for common diseases // *Curr. Opin. Genet. Dev.* 2001. Vol.11, №3, P.264-267
13. Hirschhorn J.N., Lohmueller K., Byrne E., Hirschhorn K. A comprehensive review of genetic association studies // *Genet. Med.* 2002. Vol.4, №2, P.45-61
14. Hirschhorn J.N., Daly M.J. Genome-wide association studies for common diseases and complex traits // *Nat. Rev. Genet.* 2005. Vol.6, №2, P.95-108
15. Hoh J., Wille A., Zee R., Cheng S., Reynolds R., Lindpaintner K., Ott J. Selecting SNPs in two-stage analysis of disease association data: a model-free approach // *Ann. Hum. Genet.* 2000. Vol.64, №5, P.413-417
16. Hoh J., Ott J. Mathematical multi-locus approaches to localizing complex human trait genes // *Nat. Rev. Genet.* 2003. Vol.4, №9, P.701-709
17. Holmans P. Nonparametric Linkage // in *Handbook of Statistical Genetics*, ed. by D.J. Balding et al, John Wiley & Sons, Ltd, 2001, P. 487-505

18. Jimenez-Sanchez G., Childs B., Valle D. Human disease genes // *Nature*. 2001. Vol.409, №6822, P.853-855
19. Johnson G.C., Todd J.A. Strategies in complex disease mapping // *Curr. Opin. Genet. Dev.* 2000. Vol.10, №3, P.330-334
20. Kato T., Iwamoto K., Kakiuchi C., Kuratomi G., Okazaki Y. Genetic or epigenetic difference causing discordance between monozygotic twins as a clue to molecular basis of mental disorders // *Mol. Psychiatry*. 2005. doi:10.1038/sj.mp.4001662
21. Marchini J., Cardon L.R., Phillips M.S., Donnelly P. The effects of human population structure on large genetic association studies // *Nat. Genet.* 2004. Vol.36, №5, P.512-517
22. Martin L.J., Comuzzie A.G., Dupont S., Vionnet N., Dina C., Gallina S., Houari M., Blangero J., Froguel P. A quantitative trait locus influencing type 2 diabetes susceptibility maps to a region on 5q in an extended French family // *Diabetes*. 2002. Vol.51, P.3568-3572
23. Murff H.J., Byrne D., Syngal S. Cancer risk assessment: quality and impact of the family history interview // *Am. J. Prev. Med.* 2004. Vol.27, №3, P.239-245
24. Njajou O.T., Vaessen N., Joosse M., Berghuis B., van Dongen J.W., Breuning M.H., Snijders P.J., Rutten W.P., Sandkuijl L.A., Oostra B.A., van Duijn C.M., Heutink P. A mutation in SLC11A3 is associated with autosomal dominant hemochromatosis // *Nat. Genet.* 2001. Vol.28, №3, P.213-214
25. Nelson M.R., Kardia S.L., Ferrell R.E., Sing C.F. A combinatorial partitioning method to identify multilocus genotypic partitions that predict quantitative trait variation // *Genome Res.* 2001. Vol.11, №3, P.458-470

26. Ott J. Analysis of human genetic linkage // Baltimore and London, The Johns Hopkins University Press. 1999. -382 p
27. Peltomaki P., Sistonen P., Mecklin J.P., Pylkkanen L., Jarvinen H., Simons J.W., Cho K.R., Vogelstein B., de la Chapelle A. Evidence supporting exclusion of the DCC gene and a portion of chromosome 18q as the locus for susceptibility to hereditary nonpolyposis colorectal carcinoma in five kindreds // Cancer Res. 1991. Vol.51, №16, P.4135-4140
28. Peltonen L., Palotie A., Lange K. Use of population isolates for mapping complex traits // Nat. Rev. Genet. 2000. Vol.1, №3, P.182-190
29. Rannala B. Finding genes influencing susceptibility to complex diseases in the post-genome era // Am. J. Pharmacogenomics. 2001. Vol.1, №3, P.203-221
30. Spielman R.S., Ewens W.J. The TDT and other family-based tests for linkage disequilibrium and association // Am. J. Hum. Genet. 1996. Vol.59, №5, P.983-989
31. Spurr N.K., Kelsell D.P., Black D.M., Murday V.A., Turner G., Crockford G.P., Solomon E., Cartwright R.A., Bishop D.T. Linkage analysis of early-onset breast and ovarian cancer families, with markers on the long arm of chromosome 17 // Am. J. Hum. Genet. 1993. Vol.52, №4, P.777-785
32. Terwilliger J.D., Zollner S., Laan M., Paabo S. Mapping genes through the use of linkage disequilibrium generated by genetic drift: 'drift mapping' in small populations with no demographic expansion // Hum. Hered. 1998. Vol.48, №3, P.138-154
33. Terwilliger J.D. On the resolution and feasibility of genome scanning approaches // Adv. Genet. 2001. Vol.42, P.351-391

34. Terwilliger J.D., Goring H.H. Gene mapping in the 20th and 21st centuries: statistical methods, data analysis, and experimental design // Hum. Biol. 2000. Vol.72, №1, P.63-132
35. Thompson E.A. Linkage analysis // in Handbook of Statistical Genetics, ed. by D.J. Balding et al, John Wiley & Sons, Ltd. 2001. P. 541- 563
36. Thomson G., Esposito M.S. The genetics of complex diseases // Trends Cell Biol. 1999. Vol.9, №12, P.M17-20
37. Vaessen N., Heutink P., Houwing-Duistermaat J.J., Snijders P.J., Rademaker T., Testers L., Batstra M.R., Sandkuijl L.A., van Duijn C.M, Oostra B.A. A genome-wide search for linkage-disequilibrium with type 1 diabetes in a recent genetically isolated population from the Netherlands // Diabetes. 2002. Vol.51, №3, P.856-859
38. Wang W.Y., Barratt B.J., Clayton D.G., Todd J.A. Genome-wide association studies: theoretical and practical concerns // Nat. Rev. Genet. 2005. Vol.6, №2, P.109-118
39. Weiss K.M., Terwilliger J.D. How many diseases does it take to map a gene with SNPs? // Nat. Genet. 2000. Vol.26, №2, P.151-157
40. Wong A.H., Gottesman I.I., Petronis A. Phenotypic differences in genetically identical organisms: the epigenetic perspective // Hum. Mol. Genet. 2005. Vol.14, Spec №1, P.R11-18.
41. Wooster R., Neuhausen S.L., Mangion J., Quirk Y., Ford D., Collins N., Nguyen K., Seal S., Tran T., Averill D., et al. Localization of a breast cancer susceptibility gene, BRCA2, to chromosome 13q12-13 // Science. 1994. Vol.265, №5181, P.2088-2090

GENETIC MAPPING OF COMMON HUMAN DISEASES

Axenovich T.I.

Institute of Cytology and Genetics of Siberian Division of Russian Academy of Sciences,
Novosibirsk, 630090, Lavrentiev ave., 10, Fax: 3833331257, e_mail:aks@bionet.nsc.ru

SUMMARY

Slow progress in detection and mapping of genes controlling common diseases is usually ascribed to a complexity of genetic control of the diseases and inability of current methods to account their specificity. In this paper we analyze capabilities and restrictions of various mapping methods and demonstrate that they are able to map only the genes of large effects. However, even for these genes the results of mapping are often ambiguous and contradictive due to incorrect data sampling.