

УДК 575.162

«ПРЯМАЯ» И «ОБРАТНАЯ» ГЕНЕТИКА. ГЕНЕТИКА КОЛИЧЕСТВЕННЫХ ПРИЗНАКОВ

© 2014 г. Н.М. Белоногова

Федеральное государственное бюджетное учреждение науки Институт цитологии и генетики
Сибирского отделения Российской академии наук, Новосибирск, Россия,
e-mail: belon@bionet.nsc.ru

Поступила в редакцию 13 мая 2013 г. Принята к публикации 1 февраля 2014 г.

ВВЕДЕНИЕ

Классическая генетика, начиная с опытов Менделя и его последователей, занималась изучением закономерностей наследования наблюдаемых признаков (фенотипов) живых организмов. Опираясь на анализ и количественный учет фенотипических классов потомства, полученного от скрещивания особей с разными признаками, Альфред Стёртевант еще в 1913 г. составил первую генетическую карту. Лишь спустя 15 лет был впервые выявлен «трансформирующий фактор», идентифицированный впоследствии как ДНК (эксперимент Фредерика Гриффита в 1928 г.). Еще позже была расшифрована структура ДНК (1953), разгадан генетический код (1966), секвенирован первый геном (1977), появились возможности манипуляции с ДНК и конструирования трансгенных организмов.

В 80-х годах XX в. эти возможности открыли новый путь исследований – анализировать не фенотип и его генетический контроль, а саму последовательность ДНК и эффекты ее мутирования. Такой подход стали называть «обратной» генетикой, а традиционный путь исследования от фенотипа к генотипу – «прямой» генетикой (рис. 1).

«Прямая» и «обратная» генетика дополняют друг друга в исследовании закономерностей формирования фенотипа на основе генотипа. Хотя на этом пути был достигнут значительный прогресс, до сих пор для большинства фенотипов генетический контроль полностью

не расшифрован, не понятен характер действия генов при формировании признаков.

Сложную генетическую детерминацию имеют фенотипы, измеряемые количественно: рост, вес, другие антропометрические характеристики, физиологические и психологические показатели, а также предрасположенность к таким распространенным заболеваниям человека, как диабет, ишемическая болезнь сердца, гипертония, различные формы рака и т. д. Вызовом современной генетики стало исследование генетического контроля таких фенотипов, контролируемых большим числом генов малого эффекта наряду с факторами внешней среды.

«ПРЯМАЯ» ГЕНЕТИКА

«Прямой» генетикой называют исследования, в которых исходный интерес представляет фенотип, а эксперимент производится для обнаружения генетических факторов, влияющих на проявление фенотипа.

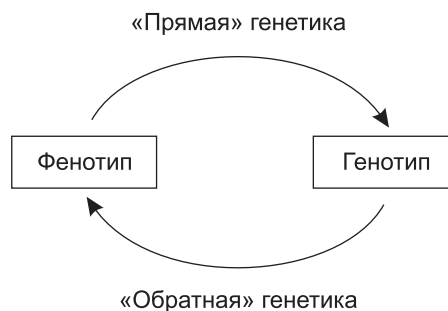


Рис. 1. «Прямая» и «обратная» генетика.

Методы современной «прямой» генетики представляют собой результат естественного развития методов классической генетики, которая прошла путь от наблюдаемого фенотипа к его материальным генетическим детерминантам. На современном этапе, когда известна полная последовательность геномной ДНК человека и большого числа модельных организмов, задача «прямой» генетики состоит в *картировании* генов, контролирующих различные признаки, т. е. в локализации конкретных последовательностей ДНК, отвечающих за тот или иной фенотип. Сделать это можно, выявив связь между фенотипическим разнообразием и разнообразием по генотипу в определенной позиции генома, — связь, указывающую на совместное наследование фенотипа и генотипа в ряду поколений.

Источником фенотипического и генотипического разнообразия может служить природная популяция, в которой оно поддерживается естественными процессами. Кроме того, у модельных организмов изменчивость может быть создана в лабораторных условиях путем индуцированного мутагенеза, при этом отбор мутантных особей для эксперимента производится исходя из их фенотипа.

Существует два подхода к картированию генов: тестирование генов-кандидатов и полногеномное сканирование. В обоих случаях в выборке должны быть определены и протестированы генотипы. При анализе генов-кандидатов тестируются аллели тех генов, для которых можно заранее предположить участие в контроле признака. При полногеномном сканировании проверяются тысячи позиций в геноме, по которым существует разнообразие в популяции и генотипы в которых могут быть достаточно легко определены у особей в выборке. Такие позиции называют *генетическими маркерами*. Маловероятно, что среди них окажется искомая мутация, но маркер может быть физически *сцеплен* с мутацией, если располагается очень близко на той же хромосоме. В этом случае аллели маркера преимущественно наследуются вместе с мутацией и неслучайно распределяются у особей с мутантным фенотипом (рис. 2).

Совместное наследование фенотипа и аллелей генетического маркера может быть протестировано с помощью *анализа сцепления* на выборке родственных особей. На модельных

объектах исследователь производит скрещивания по определенной схеме, что существенно упрощает последующий анализ. В генетике человека используется информация о реальных родословных.

В *параметрическом анализе сцепления* вероятность совместного наследования признака и маркерного генотипа оценивается исходя из модели наследования признака. Если модель наследования не известна, применяются менее мощные методы *непараметрического анализа сцепления*.

Анализ сцепления позволяет локализовать мутацию на участке 5–50 сМ, поскольку число рекомбинантных событий в нескольких поколениях родословной невелико. Для дальнейшего уточнения позиции используются другие подходы, в том числе анализ неравновесия по сцеплению, или *анализ ассоциаций*. В его основе лежит предположение о предковой хромосоме с мутацией, которая настолько тесно сцеплена с аллелем маркера, что за много поколений рекомбинация не разрушила этого сцепления (рис. 3).

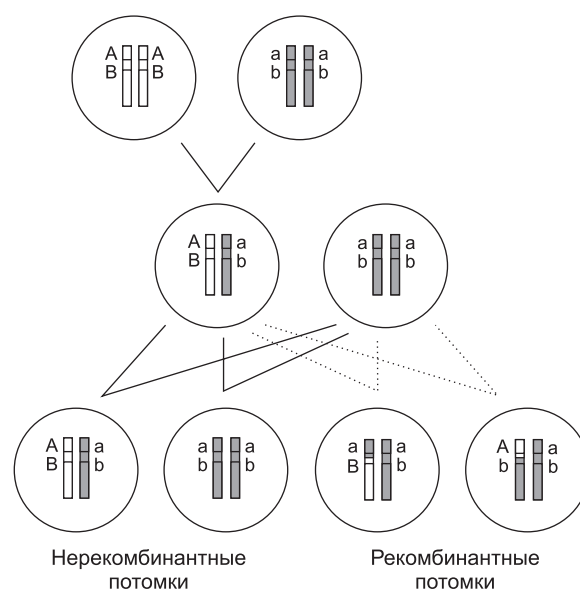


Рис. 2. Принцип генетического сцепления на примере потомства от анализирующего скрещивания двойной гетерозиготы AaBb с двойной гомозиготой по рецессивным аллелям тех же генов aabb.

При тесном сцеплении доля рекомбинантов в потомстве стремится к нулю, поскольку очень низка вероятность кроссинговера на коротком участке хромосомы между генами. Для несцепленных генов (например, размещающихся на разных хромосомах) доля рекомбинантов была бы равна 50 %.

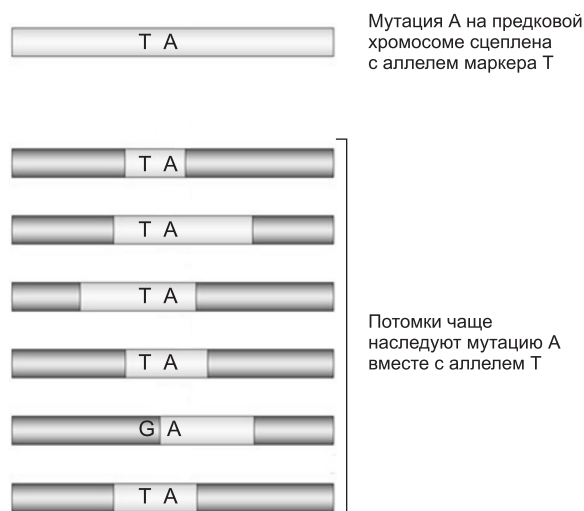


Рис. 3. Неравновесие по сцеплению.

Анализ ассоциаций может производиться на выборке неродственных особей и позволяет сузить интервал картирования до размеров < 1 сМ. Однако результат анализа ассоциаций указывает лишь на *неравновесие по сцеплению*, которое может быть обусловлено не только физическим сцеплением, но и другими причинами.

Анализ неравновесия по сцеплению оказался пригодным не только для уточнения результатов анализа сцепления: *полногеномный анализ ассоциаций* стал самостоятельным рабочим инструментом. В этом типе анализа требуется знать генотипы огромного числа генетических маркеров. Технология ДНК-микрочипов позволяет быстро считывать генотипы так называемых SNP маркеров (Single Nucleotide Polymorphism – разнообразие по одному нуклеотиду). Для каждого маркера точно известна его позиция в геноме, и благодаря неравновесию по сцеплению набор из нескольких сотен тысяч SNP маркеров информативен также и в отношении остальной последовательности геномной ДНК.

Полногеномный анализ ассоциаций позволяет с высокой точностью локализовать генетические варианты, влияющие на признак. И все же в интервал картирования часто попадает сразу несколько генов. В этом случае выбор гена-кандидата производится с учетом данных «обратной» генетики: фенотипические последствия мутирования уже изучены для многих генов или их гомологов на модельных объектах. Эта же информация может быть

использована в самом полногеномном анализе для придания большего веса потенциальным генам-кандидатам.

Таким образом подходы «прямой» генетики являются естественным продолжением классической генетики: путь от наблюдаемого фенотипа к определяющей его последовательности ДНК. Какой бы ни была стратегия исследования генетического контроля признака, на каждом этапе для повышения эффективности может быть привлечена информация из «обратной» генетики.

«ОБРАТНАЯ» ГЕНЕТИКА

«Обратная» генетика – область исследований, где отправной точкой служит некоторая последовательность ДНК и эксперименты направлены на выявление ее фенотипических эффектов.

Возможности для «обратных» генетических исследований появились сравнительно недавно, с развитием методов молекулярной биологии. Важную роль здесь играют манипуляции с ДНК и продуктами транскрипции. ДНК может быть встроена в геном другого организма, удалена или изменена. Мутации индуцируются либо направленно (с использованием гомологичной рекомбинации), либо случайным образом, с последующим отбором мутаций в определенном гене (техника TILLING, «генные ловушки»). На уровне транскрипта работают, например, некоторые процедуры генного сайленсинга.

Популярными объектами «обратной» генетики являются модельные организмы – вирусы, дрожжи, арабидопсис, дрозофила, *C. elegans*, *Danio rerio*, домовая мышь и др. Полученные результаты важны для общей генетики, генетики животных и растений, генной инженерии. Особую ценность имеет экстраполяция результатов, полученных путем «обратной» генетики на модельных животных, на гомологичные гены человека. Самый крупный проект, позволяющий производить такого рода экстраполяцию, – исследования *нокаутных мышей* (рис. 4).

Нокаут гена – его выключение в результате искусственного целенаправленного изменения ДНК. Например, часть кодирующей последовательности гена удаляется или заменяется некодирующей, либо транскрипция гена искусствен-

но нарушается вставкой («генные ловушки»). Если животные с мутацией жизнеспособны, но фенотип изменен, то по характеру изменений можно сделать вывод о функции гена и о том, в какие процессы вовлечен его продукт. Тогда и гомологичный ген человека может рассматриваться как кандидатный, потенциально вовлеченный в контроль аналогичного фенотипа.

Однажды полученные линии нокаутных мышей коммерчески доступны и могут использоваться для прояснения молекулярных механизмов действия соответствующих генов.

Результаты исследования нокаутных мышей оказались полезны для изучения целого ряда заболеваний и патологических состояний человека, включая рак, диабет, ожирение,

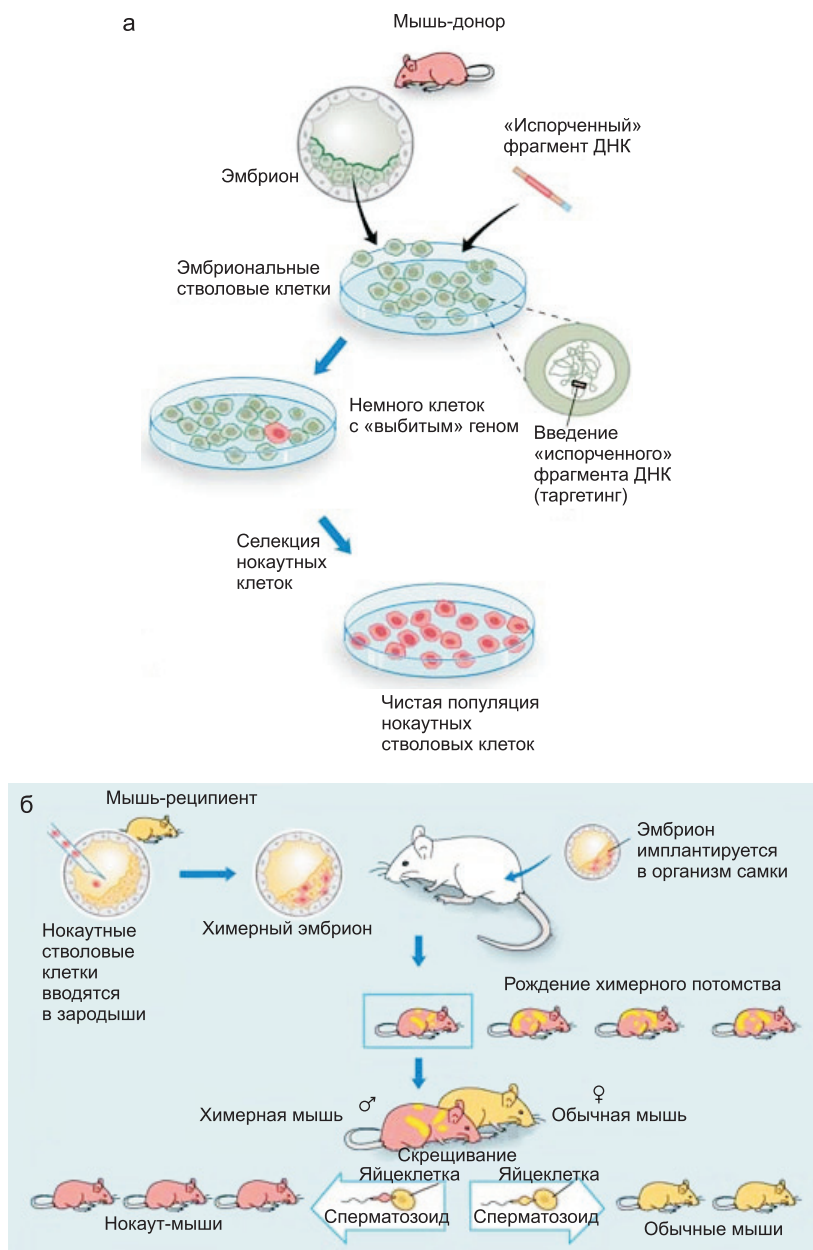


Рис. 4. Создание нокаутной линии мышей.

а – стволовые клетки выделяют из эмбриона донорской мыши (мышь розового цвета) на ранней стадии развития и культивируют в лаборатории; б – генетически модифицированные нокаутные стволовые клетки вводятся в зародыш другой мыши (мышь желтого цвета), который имплантируется в организм суррогатной мамы – самки мыши (из: (Белоконева, 2007. <http://www.nkj.ru/archive/articles/12348/>)).

сердечно-сосудистые заболевания. В 2007 г. Нобелевскую премию по физиологии и медицине получили создатели технологии генного нокаута Марио Капекки, Оливер Смитис и Мартин Эванс. Первая нокаутная мышь была создана ими еще в 1989 г.

Созданы нокаутные линии мышей для 2 400 генов, почти 800 из них описаны фенотипически. Еще для 12 тыс. генов существуют линии нокаутных эмбриональных стволовых клеток. В 2003 г. появилась также возможность создания нокаутных крыс.

По оценкам Barbañic с соавт. (2007), у 10–15 % мышинных нокаутов исследователи не находят никаких фенотипических отклонений. Одна из причин – способность генома компенсировать нарушения в одних генах за счет изменения активности других. В такой ситуации могут быть информативны двойные, тройные и т. д. нокауты (организмы, у которых выключена функция двух, трех и более генов).

С другой стороны, около 15 % нокаутов не доживают до стадии взрослого организма. Исследование функций таких генов может быть произведено на *условных нокаутах*, у которых функция гена выключается лишь в определенных тканях или органах и в определенное время.

Другой вариант – не выключать ген полностью, но снизить его экспрессию. Такую процедуру называют *нокдауном*. Нокдаун гена можно осуществить путем изменения самой последовательности ДНК (этим достигается постоянное снижение экспрессии). Еще один вариант – использование техник генного сайленсинга, когда матричная РНК связывается гомологичными малыми РНК (РНК-интерференция) или морфолиновыми олигонуклеотидами. Этим путем достигается временное снижение экспрессии гена (*транзистентный нокдаун*). Блокирование тех или иных функциональных сайтов матричной РНК позволяет не только выключать трансляцию, но и модифицировать сплайсинг. Это дает возможность исследовать назначение отдельных экзонов, отдельных белковых доменов.

Генные *нокины* – организмы, полученные в результате вставки кодирующей последовательности ДНК в нужное место генома путем сайт-специфического инсерционного мутагенеза. Может быть вставлена последователь-

ность мутантного гена и созданы условия для его оверэкспрессии. Присутствие мутантных вариантов продукта изучаемого гена наряду с его нормальными формами может дать информативный фенотип.

Наконец, если известно достаточно много о гене и его белке, ДНК гена может быть отредактирована так, чтобы изменились регуляторные сайты его белкового продукта (например сайты фосфорилирования).

Методы «обратной» генетики сейчас все чаще используются для подтверждения результатов генетического картирования. В качестве примера можно привести опубликованную недавно работу Manzini с соавт. (2012), в которой исследователи установили причинную связь между мутациями в гене *GTDC2* человека и развитием синдрома Уокера–Варбург. Это тяжелое наследственное заболевание заканчивается летальным исходом в первые месяцы жизни в результате поражения головного мозга, часто сопровождающегося дисплазией сетчатки и мышечной дистрофией.

Обнаружив высокую частоту мутаций в гене *GTDC2* у больных синдромом Уокера–Варбург, авторы работы исследовали фенотип транзистентного нокдауна по гомологичному гену *gtdc2* у рыбы *Danio rerio* (рис. 5). Характер нарушений фенотипа подтвердил гипотезу о том, что мутации в гене *GTDC2* приводят к развитию синдрома Уокера–Варбург. Более того, фенотип нокдауна удавалось «спасти» (восстановить до нормального состояния) путем введения в эмбрион рыбы матричной РНК гена *GTDC2* человека.

Таким образом, «обратная» генетика идет навстречу «прямой»: она начинает с последовательности ДНК и исследует ее влияние на фенотип. По сути, «обратной» генетикой являются любые направленные изменения последовательности ДНК или генной экспрессии, для которых исследуется фенотипический эффект, в том числе создание трансгенных организмов и ряд подходов функциональной геномики. «Обратная» генетика дает ценную информацию о том, как генотип особи реализуется в фенотипе. Эта информация постоянно востребована в «прямой» генетике при поиске кандидатных генов и для подтверждения результатов генетического картирования.

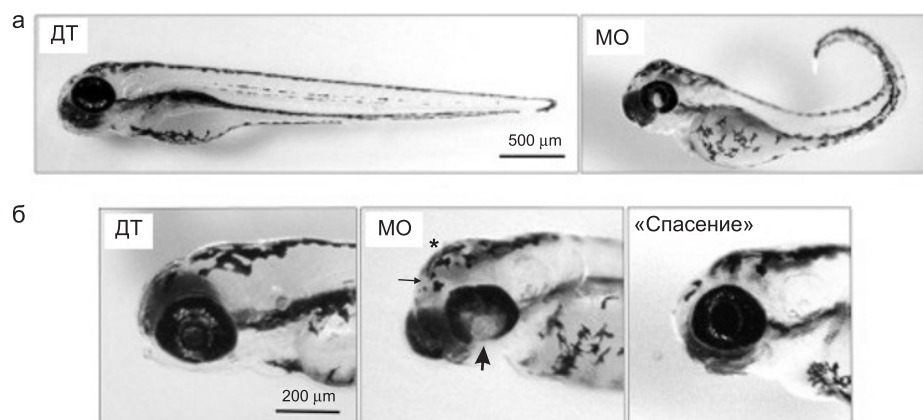


Рис. 5. Трехдневные эмбрионы *D. rerio*.

а – эмбрион дикого типа (ДТ) и эмбрион, развившийся после нокдауна гена *gtdc2* путем введения в оплодотворенную яйцеклетку антисмысловых морфолиновых олигонуклеотидов (МО). Видны множественные пороки развития: сокращение длины, изогнутый хвост, нарушения в развитии сетчатки, изменение размеров и формы головы; б – «спасение» фенотипа нокдауна введением мРНК гена *GTDC2* человека. При этом устраняются признаки тяжелой патологии, характерные для нокдауна (в центре): гидроцефалия (звездочка), кровоизлияния (тонкая стрелка), нарушения в развитии сетчатки (толстая стрелка). Фенотип «спасенного» нокдауна (справа) лишен этих дефектов и близок к дикому типу (слева) (из: (Manzini *et al.*, 2012. <http://www.sciencedirect.com/science/article/pii/S0002929712003667>)).

ГЕНЕТИКА КОЛИЧЕСТВЕННЫХ ПРИЗНАКОВ

Планируя «Опыты над растительными гибридами» (1865), Грегор Мендель сознательно отверг признаки, различия в которых основываются «на трудно подчас определяемом “более или менее”». Он рассматривал лишь *качественные* признаки гороха, строго принимающие одно из двух возможных значений: морщинистая или гладкая поверхность семени, желтая или зеленая окраска семядолей и т. д.

Примерно в то же время Френсис Гальтон начал изучать закономерности наследования признаков, поддающихся *количественной* оценке: рост, вес, сила мышц и другие антропометрические показатели, ряд свойств характера и даже предрасположенность к заболеваниям. Гальтону удалось обнаружить закономерности наследования количественных признаков и вывести свои законы, не похожие на менделевские.

В начале XX в. Рональд Фишер примирил обе концепции, предположив *полигенный* характер детерминации количественных признаков. Десятки и сотни аллелей по всему геному вносят вклад в разнообразие по количественным признакам, причем индивидуальный эффект

каждого аллеля может быть ничтожно мал и подвержен влиянию других генов и факторов внешней среды. Вызовом для генетики стали поиск генных детерминант количественных признаков, а также оценка их эффектов и характера проявления.

Актуальность задачи связана прежде всего с тем, что генетическая предрасположенность ко многим распространенным заболеваниям человека тоже может рассматриваться как количественный признак со сложным характером наследования. Более того, знание природы заболевания позволяет выявлять его *эндофенотипы* – наследуемые физиологические характеристики, связанные с риском развития болезни. Например, уровни триглицеридов и холестерина липопротеидов низкой и высокой плотности в сыворотке крови являются эндофенотипами сердечно-сосудистых заболеваний, индексы метаболизма и инсулинорезистентности – эндофенотипами диабета. Основная часть эндофенотипов – количественные признаки, генетическая природа которых слабо изучена.

Поскольку подавляющее число генов, контролирующих количественные признаки, имеют слабые индивидуальные эффекты, их влияние невозможно установить без статистического анализа больших выборок. Если анализируется

популяционная выборка и известно, что она однородна, никак не структурирована и особи в ней не связаны родством, то допустимо моделировать признак с помощью простой линейной регрессии:

$$y \sim \mu + \beta g.$$

В этой модели y – вектор значений признака из n элементов, где n – число особей в выборке. Значения признака имеют некоторое распределение, для признаков с большой долей полигенной дисперсии оно очень близко к нормальному; μ – математическое ожидание признака; g – вектор генотипов по тестируемому генетическому маркеру. Вектор имеет n значений и кодирован в соответствии с моделью наследования (рецессивная, доминантная, аддитивная, сверхдоминирование). В простейшем случае, когда проверяется аддитивная модель, значение g у особи соответствует числу редких аллелей маркера; β – размер эффекта тестируемого генетического варианта. Аппарат линейной регрессии позволяет легко оценить его величину и статистическую значимость отличия от нуля.

Если существует зависимость между признаком и генотипом по маркеру, то средние значения признака будут разными в разных генотипических классах (рис. 6).

Чем больше отличаются средние (т. е. чем больше β) и чем меньше стандартные ошибки

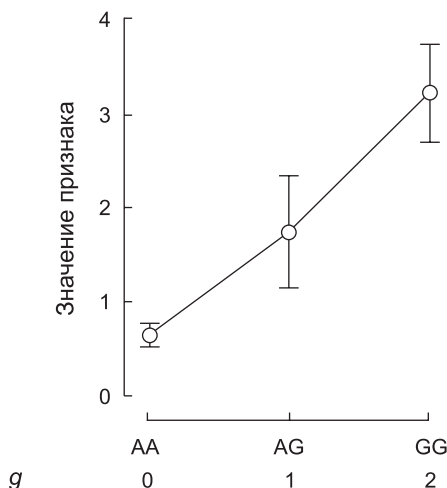


Рис. 6. Пример зависимости значений количественного признака от числа аллелей G генетического маркера.

Показаны средние значения признака и стандартные ошибки для трех генотипических классов.

средних, тем выше статистическая значимость зависимости. На размер β исследователь повлиять не в силах. Зато он может принять меры для уменьшения стандартной ошибки: прежде всего, увеличить объем выборки и этим обеспечить высокую численность всех генотипических классов, даже если речь идет о редких аллелях.

Как правило, объем выборки неродственных особей должен быть очень большим для успешного картирования генетических вариантов с малым вкладом в дисперсию признака. Современные успехи в картировании количественных признаков человека были достигнуты на объединенных выборках в сотни тысяч человек. Однако необходимая статистическая мощность может быть достигнута и другими путями, один из которых – анализ родственных выборок, особенно собранных в небольших изолированных популяциях.

При анализе родственной выборки необходимо дополнительно учитывать структуру генетической дисперсии. Если этого не сделать, число ложноположительных результатов резко возрастет. Модель, учитывающая структуру родства, усложняется до вида:

$$y = \mu + \beta g + G + e.$$

Здесь G – полигенная компонента, моделирующая аддитивные вклады большого числа генов в разнообразие признака. Предполагается, что она имеет многомерное нормальное распределение (сумма большого числа независимых эффектов) со структурой дисперсии $R\sigma_G^2$, где σ_G^2 – дисперсия полигенной компоненты и R – матрица родства особей в выборке.

Матрица родства R может быть построена на основе родословной, если родословная известна. Она также может быть посчитана исходя из генотипов особей по большому числу генетических маркеров (как минимум несколько тысяч) – это так называемое *геномное родство*. Если маркеров очень много (сотни тысяч), то геномное родство может оказаться точнее, чем родство, посчитанное по родословной. Оценка геномного родства бывает полезна и при анализе случайной выборки из популяции: она помогает обнаружить и учесть генетическую неоднородность выборки. Слагаемое e – остальные случайные эффекты. Для них тоже предполагается многомерное нормальное распределение со

структурой дисперсии $I\sigma_e^2$, где σ_e^2 – остаточная дисперсия, а I – единичная матрица.

Оценка параметров μ , β , σ_G^2 , σ_e^2 производится путем численной максимизации функции правдоподобия выборки. Обычно этот вычислительно емкий этап не может быть выполнен для каждого из сотен тысяч маркеров в отдельности, но существует ряд алгоритмов и аппроксимаций, позволяющих существенно упростить вычисления почти без потери статистической мощности.

Модель наследования признака может быть расширена эффектами взаимодействия генотипа и факторов среды, эффектами межгенного взаимодействия, материнским эффектом и т. д. В простейшем случае, когда тестируется аддитивная модель, а ковариации между отдельными компонентами считаются равными нулю, дисперсия признака может быть представлена в виде:

$$\sigma_y^2 = \sigma_a^2 + \sigma_G^2 + \sigma_e^2,$$

где σ_a^2 – дисперсия, объясняемая аддитивным эффектом тестируемого генетического маркера. Она равна $2p(1-p)\beta^2$, где p – частота редкого аллеля маркера. Величина $h^2 = (\sigma_a^2 + \sigma_G^2) / \sigma_y^2$ показывает долю аддитивной генетической дисперсии в общей дисперсии признака и называется наследуемостью в узком смысле, или просто *наследуемостью*. Эта величина имеет большое значение в генетике количественных признаков и селекции. Наследуемостью в широком смысле (H^2) называют долю всей генетической дисперсии, включающую наряду с аддитивными эффектами эффекты доминирования, межгенного взаимодействия, родительские эффекты и т. п. H^2 не может быть меньше, чем h^2 .

В естественных популяциях наследуемость для разных признаков различна. Примером признака с высокой h^2 может служить человеческий рост. Разнообразие по этому признаку (после учета эффектов пола и возраста) примерно на 80 % обусловлено генотипом. Были предприняты значительные усилия в направлении картирования генов роста человека, и на сегодняшний день ассоциация при полногеномном уровне статистической значимости показана уже для 180 генетических вариантов. Однако их генотипы, взятые вместе, объясняют всего около 10 % дисперсии роста.

В 2010 г. появилась идея проверить, какую долю дисперсии роста могут объяснить все ге-

нетические маркеры, тестируемые при картировании, независимо от статистической значимости их эффекта. Оказалось, – всего 45 % (Yang *et al.*, 2010). То есть генетические варианты, объясняющие еще 35 % (= 45 – 10) дисперсии, не были картированы из-за низкой статистической мощности исследований, не позволяющей обнаружить аллели с очень малыми индивидуальными эффектами при заданном уровне значимости. А остальные 35 % (= 80 – 45) не могли быть обнаружены, вероятно, потому, что соответствующие генетические варианты не находятся в неравновесии по сцеплению с генотипированными SNP маркерами: даже если их аллели достаточно сильно изменяют значе-

«Прямая» генетика – область исследований, в которых осуществляется поиск генетических факторов, влияющих на проявление заданного фенотипа. Поиск производится в рамках одной из двух основных стратегий: полногеномное сканирование и тестирование **генов-кандидатов** (для них есть основания заранее предполагать связь с исследуемым фенотипом). Во всех случаях в том или ином виде используется принцип **генетического сцепления** (совместного наследования из-за близкого физического расположения) предполагаемой мутации с тестируемыми **генетическими маркерами** – точками разнообразия в геноме, локализация которых известна.

«Обратная» генетика изучает фенотипические эффекты заданной последовательности ДНК. При этом исследователь производит манипуляции на молекулярном уровне: изменяет последовательность ДНК, удаляет ее из генома (**нокаут**), помещает ее в новое окружение (**нокин**) или снижает экспрессию гена (**нокдаун**).

Эндофенотипы – наследуемые характеристики, занимающие промежуточное положение между генотипом и окончательным фенотипом. Эндофенотипами заболеваний являются многие величины, измеряемые в клинической диагностике (например уровень холестерина в сыворотке крови).

Наследуемость – доля дисперсии признака, определяемая генетическим разнообразием особей в популяции. Часто под наследуемостью (h^2) имеют в виду долю аддитивной генетической дисперсии.

ние признака у своих носителей, для анализа они останутся невидимыми, если частота их в популяции меньше 1 %.

На рис. 7 для нескольких количественных признаков человека показана доля генетической дисперсии, объясняемая генотипами тех генетических вариантов, для которых вовлеченность в контроль этих признаков уже установлена. Видно, что доля «разгаданной» генетической дисперсии пропорциональна доле аддитивной генетической дисперсии (наследуемости) признака. Однако основная часть генов детерминант до сих пор не локализована для этих и большинства других количественных признаков человека.

Современная генетика количественных признаков стоит перед задачей картирования генетических вариантов, эффекты которых измеряются в лучшем случае десятими долями процента от общей дисперсии. В зависимости от их природы могут быть использованы разные пути увеличения мощности (табл.).

Ресурсом увеличения мощности классических «прямых» исследований картирования количественных признаков могут служить и

данные «обратной» генетики. На основе независимой внешней информации о функционировании и фенотипических эффектах тех или иных генов, получаемой из «обратной» генетики, можно присваивать тестируемым маркерам априорные веса, выбирать гены-кандидаты.

У нокаутных мышей может быть измерен большой спектр количественных фенотипов, аналогичных фенотипам человека. В лабораторных исследованиях модельных организмов эффект отдельной мутации будет заметнее на фоне сниженной полигенной дисперсии инбредных линий и в условиях контролируемой внешней среды. Тем не менее для обнаружения слабых эффектов могут потребоваться большие выборки животных. Отчасти этим может объясняться то, что у 10–15 % линий нокаутных мышей не наблюдалось фенотипических отклонений: слабые количественные изменения могут быть найдены лишь при анализе больших выборок.

Все же эффект редкого аллеля может быть легче обнаружен у мутантной линии модельных животных, чем при популяционных исследованиях. Согласно упомянутой выше работе (Yang *et al.*, 2010), значительная доля дисперсии

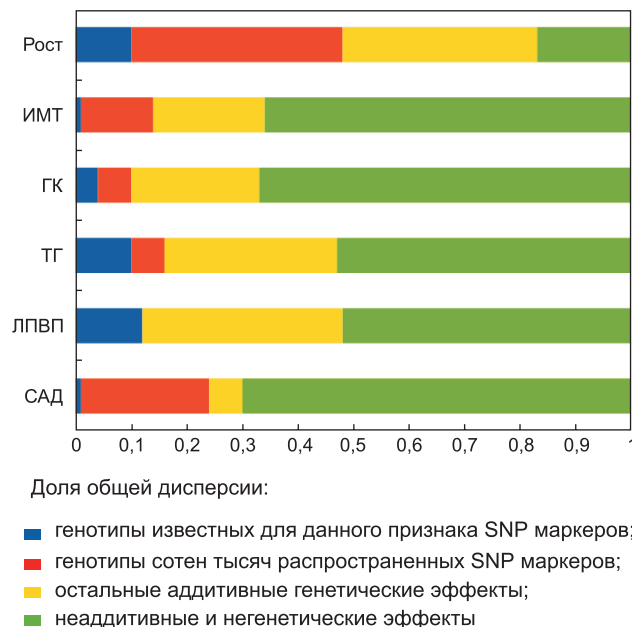


Рис. 7. Результаты исследований генетической архитектуры на примере 6 признаков человека: рост, индекс массы тела (ИМТ), концентрация глюкозы (ГК), триглицеридов (ТГ), холестерина липопротеидов высокой плотности (ЛПВП) в сыворотке крови натощак, систолическое артериальное давление (САД). По материалам публикации (Vattikuti *et al.*, 2012).

Таблица

Пути увеличения мощности статистического анализа

Тип искомых детерминант	Способ увеличения статистической мощности
Распространенные аллели малого эффекта	• Дальнейшее наращивание объемов исследуемых выборок, в том числе с использованием техники мета-анализа (объединение результатов из разных выборок)
Редкие мутации	• Одновременный анализ нескольких генетических маркеров в пределах одного гена • Глубокое исследование генетического разнообразия с целью поиска редких вариантов (например, секвенирование экзомов – последовательности ДНК экзонов всех генов) • Использование родственных выборок, особенно выборок из генетических изолятов, а также комбинация методов анализа сцепления и ассоциаций
Генетические детерминанты неаддитивной дисперсии	• Учет эффектов доминирования, импринтинга, межгенных взаимодействий, материнского эффекта

полигенных признаков может определяться именно редкими аллелями, которые находятся в слабом неравновесии по сцеплению с тестируемыми маркерами (частота аллелей тестируемых маркеров редко бывает ниже 5 %). Пусть мутация объясняет 0,5 % дисперсии признака и распространена в популяции с частотой 1 %. При стандартных ошибках I и II рода (5 и 20 % соответственно) и с учетом потери мощности из-за неполного неравновесия по сцеплению для обнаружения такой мутации путем анализа ассоциаций потребуется выборка объемом 7 524 особи. Чтобы зафиксировать смещение матожидания признака у гомозигот по такой мутации (например, в линии нокаутных животных), потребовалось бы всего 17 гомозигот и столько же контрольных особей.

При всех проблемах экстраполяции результатов с модельных объектов на человека исследования нокаутных мышей и другие подходы «обратной» генетики, несомненно, могут быть полезны для выявления группы генов, потенциально значимых для контроля количественных признаков человека. Если генетическому разнообразию этих генов дать больший вес в «прямом» анализе картирования, это может повысить его итоговую эффективность.

Таким образом, количественные признаки, как правило, имеют сложную детерминацию: они формируются под влиянием большого числа генов с малыми индивидуальными вкладами и при участии разнообразных факторов

внешней среды. На современном этапе картирование генов малого эффекта представляет собой основную трудность в расшифровке генетической архитектуры количественных признаков. Среди путей решения проблемы – дальнейшее наращивание объемов выборок, совершенствование методов анализа для учета специфики искомых эффектов, объединение информации из различных источников.

ЗАКЛЮЧЕНИЕ

Как правило, непосредственный практический интерес представляет именно фенотип, а детерминирующая его последовательность ДНК интересна не сама по себе, а из-за своих фенотипических эффектов. Однако, чтобы удовлетворить односторонний интерес, часто требуются разносторонние подходы. «Прямая» и «обратная» генетика исследуют один и тот же процесс реализации генетической информации в фенотипе с противоположных сторон, и благодаря этому причинно-следственные связи вскрываются более эффективно.

Промежуточным звеном на пути от гена к признаку является эндофенотип – биохимическая или физиологическая характеристика, генетический контроль которой может быть проще, чем для сложного признака. Картирование генов эндофенотипов, большинство которых измеряется в количественной шкале, – пробный камень, на котором будет проверена эффектив-

ность взаимодействия различных подходов. В случае успеха эта работа позволит прояснить механизмы формирования предрасположенности человека к ряду распространенных заболеваний, что будет способствовать их лечению, ранней диагностике и профилактике, а также откроет новые перспективы для генной инженерии, повысит эффективность создания трансгенных организмов с желаемыми свойствами.

ЛИТЕРАТУРА

- Аксенович Т.И. Статистические методы генетического анализа признаков человека. Уч. пособие. 2-е изд., перераб. и доп. Новосибирск: НГУ, 2003. 160 с.
- Аксенович Т.И., Белоногова Н.М. Картирование генов с помощью неравновесия по сцеплению или аллельных ассоциаций. Уч. пособие. Новосибирск: НГУ, 2008. 97 с.
- Белоконев О. Нобелевские премии 2007 года. Гены под прицелом // Наука и жизнь. № 12. 2007. <http://www.nkj.ru/archive/articles/12348/>
- Barbaric I., Miller G., Dear T.N. Appearances can be deceiving: phenotypes of knockout mice // Brief Funct. Genomic Proteomic. 2007. V. 6. P. 91–103.
- Manzini M.C., Tambunan D.E., Hill R.S. *et al.* Exome sequencing and functional validation in zebrafish identify *GTDC2* mutations as a cause of Walker-Warburg syndrome // Am. J. Hum. Genet. 2012. V. 91. P. 541–547.
- Yang J., Benyamin B., McEvoy B.P. *et al.* Common SNPs explain a large proportion of the heritability for human height // Nat. Genet. 2010. V. 42. P. 565–569.
- Vattikuti S., Guo J., Chow C.C. Heritability and genetic correlations explained by common SNPs for metabolic syndrome traits // PLoS Genetics. 2012. V. 8. e1002637.