

КРИТЕРИИ ОТБОРА ГЕНЕТИЧЕСКИХ МАРКЕРОВ ДЛЯ АНАЛИЗА ПРЕДРАСПОЛОЖЕННОСТИ К МНОГОФАКТОРНЫМ ФЕНОТИПИЧЕСКИМ ОСОБЕННОСТЯМ

И. И. Низамутдинов¹✉, Д. О. Коростин¹, В. В. Ильинский¹, А. С. Ракитко^{1,2}

¹ ООО «Генотек», Москва

² Кафедра теории вероятностей, механико-математический факультет, Московский государственный университет имени М. В. Ломоносова, Москва

Принцип наследования многофакторных фенотипических особенностей (МФО) заключается в наличии значительно-го количества генетических маркеров, каждый из которых вносит некоторый вклад в вероятность развития данной особенности. В данной работе предложен алгоритм отбора ДНК-маркеров с целью разработки прогностических тест-систем для определения индивидуальной предрасположенности к МФО. Метод заключается в отборе генетических маркеров, показавших статистически достоверную ассоциацию с данной МФО, а также функционально значимые полиморфизмы, для которых описаны механизмы влияния на развитие МФО. Если функциональная значимость полиморфизма не описана, критерием его статистически достоверной ассоциации с МФО является достижение полногеномной значимости в одном из исследований и подтверждение данной ассоциации на независимой выборке. Научные публикации, используемые для оценки ассоциации генетических маркеров с МФО, в зависимости от типа исследования должны соответствовать приведенным критериям.

Ключевые слова: генетический маркер, мультифакториальная болезнь, однонуклеотидный полиморфизм, критерии отбора, полногеномное исследование, функциональная значимость, метаанализ, исследование случай–контроль

✉ **Для корреспонденции:** Низамутдинов Игорь Игоревич
пер. Наставнический, д. 17, стр. 1, корп. 15, г. Москва, 105120; igor@genotek.ru

Статья получена: 16.12.2016 **Статья принята к печати:** 21.12.2016

CRITERIA FOR THE SELECTION OF GENETIC MARKERS IN THE ASSESSMENT OF PREDISPOSITION TO MULTIFACTORIAL TRAITS

Nizamutdinov II¹✉, Korostin DO¹, Il'inskii VV¹, Rakitko AS^{1,2}

¹ Genotek Inc, Moscow, Russia

² Department of Probability Theory, Faculty of Mechanics and Mathematics, Lomonosov Moscow State University, Moscow, Russia

The principle of multifactorial traits (MTs) inheritance relies on the presence of a large number of genetic markers, with each marker contributing to the probability of developing those traits. This work proposes an algorithm for the selection of DNA markers that could be used to develop a prognostic test system for the assessment of individual predisposition to MTs. The method is based on the selection of genetic markers that have demonstrated a statistically significant association with an MT under consideration and have been described as functionally significant polymorphisms affecting MT development. If the functional significance of a polymorphism has not been described so far, then to be reliably associated with an MT, this polymorphism is expected to achieve genome-wide significance in one of the studies and such significance must be confirmed in an independent sample. Papers that are used to assess the association of genetic markers with MTs are expected to meet the proposed criteria depending on the study type.

Keywords: genetic marker, multifactorial disease, single nucleotide polymorphism, selection criteria, genome-wide association study, functional significance, meta-analysis, case-control study

✉ **Correspondence should be addressed:** Igor Nizamutdinov
per. Nastavnicheskiy, d. 17, str. 1, korp. 15, Moscow, Russia, 105120; igor@genotek.ru

Received: 16.12.2016 **Accepted:** 21.12.2016

Генетическое тестирование — важный элемент персонализированной медицины, позволяющий выявлять риски развития широко распространенных заболеваний и особенности функционирования организма. Информация о генетической предрасположенности может быть использована для подбора методов индивидуальной профилактики и ранней диагностики заболеваний, коррекции образа жизни, а также при составлении плана питания и физических упражнений.

Индивидуальная предрасположенность к сердечно-сосудистым заболеваниям, например, может быть вызвана нарушениями в большом количестве систем организма: изменениями коагуляционных свойств крови, дислипидемией, нарушениями в ренин-ангиотензиновой системе, повышением уровня гомоцистеина, рядом врожденных состояний, значительно повышающих риск сердечно-сосудистых патологий (болезни Фабри, Моямоя и пр.). Большинство таких нарушений обусловлено генетическими

факторами, что создает возможности для индивидуальной профилактики сердечно-сосудистых заболеваний, нацеленной на устранение факторов риска, свойственных конкретному индивиду. В случае отсутствия профилактики информация о генетической предрасположенности может быть полезна для ранней диагностики и более пристального наблюдения за симптомами болезни. Например, известно, что некоторые мутации в генах *MLH1*, *MSH2*, *MSH6*, *PMS2* или *EPCAM* могут повышать риск развития колоректального рака более чем в 10 раз [1]. Данная форма рака хорошо поддается лечению на ранних стадиях, с пятилетней выживаемостью пациентов около 94 %, и практически неизлечима на IV стадии. Помимо ранней диагностики и профилактики заболеваний информация о генетических факторах может быть использована для коррекции образа жизни. Например, известно, что полиморфизм гена *GC*, продукт которого отвечает за связывание и транспорт кальциферола и его метаболитов, может приводить к пониженной концентрации витамина D и его метаболитов в крови [2]. Увеличение в рационе количества продуктов, содержащих витамин D, может компенсировать данную генетическую особенность.

При занятиях фитнесом и спортом также важно учитывать генетические факторы. Например, в случае высокой генетической предрасположенности к варикозной болезни вен из программы тренировок необходимо исключить упражнения, сопровождающиеся сильным натуживанием.

Завершение проекта «Геном человека» привело к активному развитию направления индивидуального генетического тестирования. Прогностические тест-системы, выявляющие наличие ДНК-маркеров, ассоциированных с различными фенотипическими особенностями и заболеваниями, находят широкое применение во всем мире. Одним из существенных ограничений разработки таких тест-систем является проблема перехода от теории к практике (*from bench to bedside*) [3]: ассоциация генетических маркеров с фенотипическими особенностями демонстрируется в научных исследованиях путем изучения больших выборок людей, соответственно, возникает задача интерпретации наличия данных маркеров у конкретного человека.

Другая значительная проблема — это количество генетических маркеров, которое необходимо использовать при разработке тест-систем для определения индивидуальной предрасположенности к многофакторным фенотипам. Использование всех маркеров, для которых была показана ассоциация с исследуемой особенностью, неизбежно приводит к снижению специфичности тест-системы и повышению ее стоимости. В то же время небольшое количество используемых маркеров негативно сказывается на чувствительности и снижает прогностические свойства тест-системы. В настоящее время существует несколько подходов для решения указанных проблем.

Одним из подходов создания генетических тест-систем является учет небольшого количества ДНК-маркеров, статистически достоверная ассоциация которых была показана во многих исследованиях. Например, при оценке индивидуального риска инфаркта миокарда могут учитываться генетические маркеры в генах-кандидатах *eNOS* и *CX37* [4], при этом игнорируется возможность наличия у индивида других генетических маркеров, таких как мутация Лейдена, значительно повышающая риск развития этого острого состояния [5]. Подобный подход не позволяет полноценно оценить наличие всевозможных нарушений, приводящих к заболеванию, и, как следствие, имеет

невысокую чувствительность. С другой стороны, на сегодняшний день известно более 400 генетических маркеров, статистически значимая ассоциация которых с инфарктом миокарда была продемонстрирована в некоторых исследованиях типа случай-контроль. Учет всех известных генетических маркеров ограничивает набор лабораторных методов, которые можно использовать при анализе. Наиболее доступные для большинства лабораторий методы (ПЦР «в реальном времени», анализ полиморфизма длины рестриционных фрагментов и ряд других методов) при поточной работе подходят для тестирования небольшого количества генетических маркеров (нескольких десятков). Увеличение числа тестируемых маркеров приводит либо к увеличению времени анализа указанными выше методами, либо значительному повышению стоимости тест-системы из-за использования более дорогих технологий (ДНК-микрочипов).

Кроме того, использование большого количества генетических маркеров приводит к проблеме переобучения: хороших прогностических свойств алгоритма на обучающей выборке, но низкой чувствительности и специфичности на общей популяции.

В настоящей работе предлагается алгоритм отбора и оценки генетических маркеров с целью разработки прогностических тест-систем развития многофакторных фенотипических особенностей (МФО). Основная идея и особенность метода заключается в том, что отбор проходят не только те генетические маркеры, которые показали полногеномную ассоциацию, но и маркеры, не достигшие статистической значимости после поправки на множественность в полногеномных исследованиях, однако удовлетворяющие некоторым иным критериям (в частности, критерию функциональной значимости). Статья посвящена описанию и обсуждению данных критериев.

Выбор фенотипических особенностей

Фенотипические особенности (ФО) условно можно разделить на четыре типа: с низким или нулевым вкладом генетических факторов в формирование данной особенности, моногенные, полигенные и мультифакторные. В соответствии с настоящими критериями мы предлагаем разработывать генетические тест-системы только для ФО, наследуемость которых не ниже 30 %. Меньшее значение наследуемости свидетельствует о преобладании факторов внешней среды в формировании ФО, вероятность развития которой должна оцениваться исходя из образа жизни и условий окружающей среды.

За формирование моногенных особенностей отвечает мутация одного гена, примерами могут служить моногенные заболевания (муковисцидоз, фенилкетонурия). При разработке тест-систем для определения возможности развития моногенных ФО необходимо учитывать пенетрантность известных мутаций и процент случаев проявления фенотипа, за который они ответственны.

Полигенные особенности развиваются в результате полиморфизма большого количества генов, каждый из которых вносит некоторый вклад в их формирование. Примером полигенной особенности является цвет глаз, который почти полностью определяется генетическими факторами [6].

В данной работе мы не будем рассматривать подобные ФО, а сосредоточимся на мультифакторных признаках, в которые соизмеримый вклад вносят как генетические факторы, так и факторы внешней среды. При разработке

тест-систем для оценки вероятности развития у индивида мультифакторных особенностей необходимо оценивать статистическую значимость ассоциации генетических маркеров с исследуемой ФО и их функциональное влияние на проявление данной особенности.

Оценка статистической значимости ассоциации генетических маркеров

В настоящее время для широко распространенных МФО известно большое количество генетических маркеров. В исследованиях, посвященных изучению роли генетических факторов в развитии МФО, статистические параметры, характеризующие степень ассоциации генетических маркеров с МФО, и *p*-значение рассчитываются отдельно для каждого маркера. Степень ассоциации полиморфизма может быть охарактеризована различными статистическими параметрами, но в большинстве случаев это соотношение шансов (ОШ), относительный риск, бета-коэффициент или частоты аллелей среди больных и здоровых.

Параметром, определяющим достоверность полученных различий между людьми с МФО и без нее, является *p*-уровень значимости. Общепринятым критерием для заключения, что обнаруженные различия не случайны, является значение $p < 0,05$ [7]. При тестировании нескольких гипотез (исследовании нескольких полиморфных локусов) необходимо корректировать порог *p*-значения методом Бонферрони. В дальнейшем мы будем подразумевать необходимость наличия данной поправки при оценке статистической значимости генетического маркера.

В случае отсутствия функциональной значимости полиморфизма (см. раздел ниже) достоверно ассоциированными с МФО необходимо считать генетические маркеры, ассоциация которых была продемонстрирована в полногеномных исследованиях (GWAS), достигла полногеномной значимости ($p < 5 \cdot 10^{-8}$) и была подтверждена на независимой выборке [8].

Оценка функциональной значимости генетических маркеров

Функциональная значимость полиморфизма оценивается путем анализа его влияния на формирование исследуемой особенности. Например, при разработке тест-систем для оценки индивидуального риска развития заболевания необходимо учитывать генетические маркеры, вовлеченные в патогенез. Важность данного подхода обусловлена тем, что некоторые редко встречающиеся генетические маркеры не достигают полногеномной значимости и, соответственно, отсеиваются в ходе GWAS. Для заключения о наличии функциональной значимости маркера необходимо, чтобы он соответствовал одному из следующих критериев.

1) Известен механизм влияния генетического полиморфизма на формирование МФО

Данные полиморфизмы находятся в генах-кандидатах, для которых установлена вовлеченность в механизм формирования МФО. Например, продукт гена метилентетрагидрофолатредуктазы (*MTHFR*) играет важную роль в метаболизме витамина В9: катализирует процесс образования активной формы фолиевой кислоты, необходимой для превращения гомоцистеина в метионин. Полиморфизм

rs1801133 приводит к аминокислотной замене в белке *MTHFR*, что снижает его сродство к субстрату и, как следствие, приводит к снижению метаболизма гомоцистеина [9]. Нарушение метаболизма гомоцистеина является фактором риска развития гипергомоцистеинемии. При этом важно отметить, что наличие данного полиморфизма — не единственный фактор риска развития данного заболевания и в случае отсутствия других факторов не гарантирует его развитие.

2) Известен опосредованный механизм влияния генетического полиморфизма на формирование МФО

Например, полиморфизм *rs1799983* в гене эндотелиальной NO-синтазы является миссенс-мутацией, приводит к нарушению процессинга белка, что в итоге ведет к снижению активности фермента. Измененный белок синтезирует меньшее количество оксида азота (II), необходимого для вазодилатации сосудов. Это приводит к повышению давления и является фактором риска развития гипертонии [10]. Повышенное артериальное давление также является рисковым фактором ишемической болезни сердца (путем сужения просвета сосудов и развития эндотелиальной дисфункции). Данный полиморфизм может быть рассмотрен как генетический маркер, ассоциированный с риском развития ишемии.

Ассоциация всех функционально значимых маркеров с фенотипическими особенностями должна быть экспериментально подтверждена в исследованиях типа случай–контроль.

Критерии отбора научных публикаций

Научные публикации, посвященные анализу ассоциации генетического полиморфизма с фенотипическими особенностями, можно разделить на три типа: исследования случай–контроль и исследования количественных параметров, метаанализы и обзорные статьи.

Поскольку обзорные статьи не проводят статистического анализа ассоциации генетических маркеров с исследуемыми МФО, они должны быть исключены из рассмотрения при анализе целесообразности включения генетических маркеров в разрабатываемые тест-системы. В то же время данный тип публикаций может быть использован при составлении изначального списка генетических маркеров, которые в дальнейшем анализируются в соответствии с настоящими критериями.

В исследованиях случай–контроль анализируется ассоциация генетических маркеров с каким-либо заболеванием или физиологическим свойством организма путем сравнения частот аллелей среди людей с МФО и контрольной группой. Данные исследования делятся на два типа: полногеномный поиск ассоциаций (GWAS) и исследования ассоциации полиморфизма некоторых генов-кандидатов с МФО.

Полногеномный поиск ассоциаций — направление биологических исследований, связанных с поиском различий между геномами людей с определенными фенотипическими характеристиками и без них. В данных исследованиях проводится анализ ассоциации генетических маркеров, распределенных по всему геному с использованием ДНК-микрочипов высокой плотности.

В работах, посвященных исследованию ассоциации отдельных генов с МФО, ограничен набор исследуемых

генетических маркеров, и внимание сосредоточено на генах с известным или предполагаемым механизмом вовлеченности в развитие данных МФО.

Метаанализ — это исследование, в котором проводится анализ наличия ассоциации путем обобщения данных множества исследований. Основной принцип метаанализа — все исследования, включенные в анализ, должны проверять одну и ту же гипотезу.

В силу специфики различных типов исследований критерии, применяемые для отбора научных публикаций, несколько различаются.

Для уменьшения отбора генетических маркеров, для которых в ходе GWAS был получен ложноположительный результат, необходимо чтобы исследования соответствовали следующим критериям [11]:

1. В первичном полногеномном исследовании должно быть не менее 750 пациентов. Меньшие размеры выборки могут приводить к некорректной работе статистических тестов и, как следствие, увеличению количества ложноположительных и ложноотрицательных результатов.

2. Необходимо учитывать только генетические маркеры с $p < 0,01$.

3. Обнаруженные ассоциации должны быть подтверждены хотя бы в одном исследовании (в случае редких болезней подтверждения может и не быть). При этом необходимо, чтобы в этих исследованиях: $p < 0,01$; 95 % доверительные интервалы для ОШ должны пересекаться во всех исследованиях; публикации должны быть в журналах с импакт-фактором > 2 .

Публикации, посвященные исследованию влияния полиморфизма небольшого числа генов на развитие МФО, должны включать следующие условия:

1. Данные должны быть получены при исследовании тканей (прижизненная биопсия, материал аутопсии, послеоперационный материал) или биологических жидкостей участников.

2. Ассоциации должны быть экспериментально получены в рассматриваемой научной публикации. Публикации, в которых авторы цитируют выводы других авторов относительно исследуемых генетических маркеров, необходимо исключить из анализа.

3. $p < 0,05$.

4. Размеры выборки участников должны быть достаточно велики для возможности обнаружения ассоциации генетических маркеров с определенными частотами встречаемости [12].

5. Если есть несколько публикаций, в которых исследовали ассоциацию данного генетического маркера с риском развития заболевания, то для анализа выбирают: а) более раннюю (например, из двух статей 2009 и 2015 гг. — статью за 2015 г.); б) в которой исследование проводилось на большем количестве образцов.

При наличии метаанализа, посвященного исследованию ассоциации генетических маркеров с формированием МФО, полученные в нем данные являются приоритетными. При этом необходимо учитывать информацию, приведенную только в метаанализах высокого уровня, соответствующих следующим критериям [13]:

1. В настоящий момент не описаны четкие механизмы влияния исследованных в метаанализе генетических маркеров на патогенез заболевания. Если такие механизмы известны, то необходим анализ функциональной значимости полиморфизма.

2. В исследовании проведен сравнительный поиск литературы. Критерии отбора научных публикаций для ме-

таанализа должны подразумевать включение как публикаций, в которых ассоциация полиморфизма с заболеванием была подтверждена, так и тех, в которых ассоциация была опровергнута.

3. Указаны источники информации и ключевые слова, по которым проводился поиск.

4. Анализ целесообразности включения в метаанализ статей, список которых был получен автоматически, должен производиться вручную.

5. Должны быть указаны и обоснованы критерии включения и исключения публикаций (объем выборок, статьи только на английском языке, демографические характеристики участников и пр.).

6. Данные исследований должны быть комбинируемыми.

7. В исследовании должен быть проведен анализ предвзятости отдельных публикаций методом воронкообразного графика или анализа чувствительности.

8. Если в метаанализе рассматривали различные популяции и статистически значимая ассоциация была показана только для европейских популяций (Caucasian), то данный генетический маркер рассматривается в качестве ДНК-маркера, ассоциированного с фенотипической особенностью, при условии анализа нескольких публикаций, в которых исследовались европейские популяции.

9. В исследованиях, в которых была обнаружена значимая ассоциация полиморфизма, 95 % доверительные интервалы для ОШ (или других статистических параметров, характеризующих степень ассоциации) должны пересекаться.

Алгоритм оценки целесообразности включения генетического маркера в тест-систему

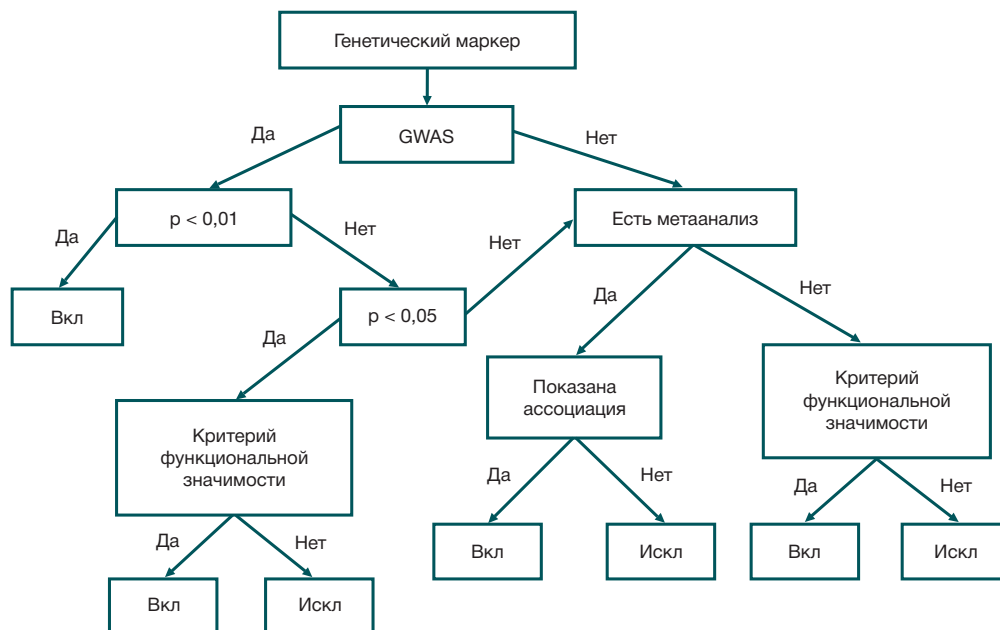
Последовательность действий при анализе необходимости включения генетических маркеров в тест-систему для определения вероятности развития МФО можно представить в виде следующей последовательности действий (рисунок).

Если ассоциацию генетического маркера исследовали в ходе GWAS и показали ее статистически достоверную значимость, а само исследование соответствовало описанным выше критериям, то такой маркер должен быть включен в тест-систему. Если же в ходе GWAS p -значение ассоциации генетического маркера оказалось больше 0,01, но меньше 0,05, то необходим анализ функциональной значимости данного маркера.

Если генетический маркер не исследовали в ходе GWAS либо отсеяли на первом этапе исследования, но был проведен метаанализ высокого уровня, который показал значимую ассоциацию с МФО, то данный генетический маркер также необходимо учитывать при разработке прогностических тест-систем. При наличии обучающей выборки необходимо убедиться, что добавление подобного маркера в тест-систему не приводит к увеличению эмпирического риска.

Если генетический маркер не исследовался в ходе GWAS и нет метаанализа его ассоциации, но при этом он функционально значим и есть исследования генов-кандидатов, подтверждающие его ассоциацию, то его также следует включать в прогностическую тест-систему.

После формирования списка генетических маркеров, которые целесообразно учитывать при разработке прогностической тест-системы, необходим анализ неравновесного сцепления генов.



Алгоритм отбора генетических маркеров для тест-систем, оценивающих индивидуальную предрасположенность к многофакторным фенотипическим особенностям

Пример составления списка генетических маркеров МФО

На сегодняшний день известно 6 генетических маркеров, ассоциация которых с риском развития ишемического инсульта достигает полногеномной значимости и была подтверждена на независимых выборках [14–16]. При этом в данный перечень не входят полиморфизмы генов *F5*, *F2*, *F7*, *F13B*, *MTHFR*, *ACE*, *APOE*, *GP1IIa*, *eNOS*, *PAI*, *GP1BA*, *ITGA2*, *ITGA2B*, *LPL*, *IL6* и *PON1*, ассоциация которых с риском развития инсульта была показана ранее при исследовании отдельных генов-кандидатов [17]. Полиморфизмы данных генов необходимо рассмотреть с позиции функциональной значимости и наличия метаанализов высокого уровня, в которых исследовалась их связь с ишемическим инсультом.

Коагуляционный фактор V (ген *F5*) является важным компонентом системы свертывания крови. Его функция заключается в активизации реакции образования тромбина из протромбина. Полиморфизм *rs6025* гена *F5*, известный также как мутация Лейдена, приводит к повышению устойчивости фермента к ингибиторам и, как следствие, к повышенной коагуляции крови. При этом был проведен метаанализ, в котором была подтверждена ассоциация данной мутации с риском развития инсульта [17]. Гиперкоагуляция является фактором риска развития сердечно-сосудистых заболеваний, включая инсульт, соответственно, данный полиморфизм также может быть рассмотрен и как функционально значимый.

Полиморфизм *rs179963* (*G20210A*) в гене протромбина (*F2*) локализован в 3'-нетранслируемой области и приводит к повышению уровня протромбина [18], что может приводить к гиперкоагуляции. Согласно результатам метаанализа [17], данный полиморфизм ассоциирован с риском развития ишемического инсульта.

Как было показано выше, полиморфизм *rs1801133* гена *MTHFR* ассоциирован с риском развития гипергомоцистемии. Повышенный уровень гомоцистеина — фактор риска повреждения сосудов [19]. Данный полиморфизм также был исследован в метаанализе высокого уровня, где

была показана его ассоциация с риском развития ишемического инсульта [17].

Ангиотензинпревращающий фермент играет важную роль в регуляции давления крови путем превращения ангиотензина I в ангиотензин II. Как было показано ранее, полиморфизм *rs1799752* приводит к изменению активности фермента [20], что ведет к повышению тонуса сосудов и развитию атеросклероза. Согласно результатам метаанализа [17] данный полиморфизм ассоциирован с риском развития ишемического инсульта.

Полиморфизм генов *F5*, *F2*, *MTHFR* и *ACE* необходимо учитывать при разработке тест-систем для определения индивидуального риска развития ишемического инсульта, поскольку в метаанализе высокого уровня была показана их ассоциация с данным заболеванием.

Ассоциацию полиморфизма генов *F7*, *F13B*, *APOE*, *GP1IIa*, *eNOS*, *PAI*, *GP1BA*, *ITGA2B* и *LPL* также исследовали в ходе метаанализа высокого уровня [17], однако не было обнаружено их значимой ассоциации с риском развития ишемического инсульта. Полиморфизм этих генов не следует учитывать при разработке тест-систем для определения индивидуального риска развития ишемического инсульта.

Несмотря на отсутствие ассоциации полиморфизма гена аполипопротеина E (*APOE*) с ишемическим инсультом в общей популяции, была показана его значимая ассоциация с данным заболеванием у людей в возрасте до 45 лет [21]. Полиморфизм этого гена является функционально значимым и играет важную роль в неврологических патологиях и нарушениях липидного спектра. Аллель *e4* гена *APOE* ассоциирован с повышенным уровнем общего холестерина в крови, а также толщиной комплекса интима-медиа сонной артерии. Кроме того, аллель *e4* значимо ассоциирован с риском развития некоторых неврологических заболеваний и нарушений (болезнь Альцгеймера, сотрясение мозга, более длительное восстановление после травмы головы) [22]. В нервной системе продукт гена *APOE* осуществляет выведение через гематоэнцефалический барьер β-амилоида, образующегося в результате повреждений отдельных нейронов (травмы, гематомы и пр.) и оказывающего

токсическое действие на здоровые клетки. Аллель *e4* снижает сродство APOE к β -амилоиду и, как следствие, приводит к его накоплению и гибели нейронов. Данный полиморфизм может быть рассмотрен в качестве функционально значимого, но может учитываться только при разработке тест-систем, специфичных к ранней форме ишемического инсульта.

Ген *ITGA2* кодирует $\alpha 2$ -субъединицу интегринов — белков, осуществляющих связь тромбоцитов с тканевыми белками, обнажаемыми при повреждении стенки сосудов. Образование монослоя из тромбоцитов в месте повреждения ткани сосуда является необходимым условием запуска каскада реакций свертывания крови. Полиморфизм *rs1126643 (c.759C>T)* приводит к увеличению скорости адгезии тромбоцитов и ассоциирован с риском развития тромбофилии [23]. Данный полиморфизм непосредственно влияет на скорость протекания патофизиологических процессов, являющихся факторами риска ишемического инсульта, и может быть рассмотрен в качестве функционально значимого.

Ген *IL6* кодирует интерлейкин 6 и активно экспрессируется в атеросклеротических бляшках. *IL6* совместно с другими медиаторами воспаления оказывает значительное влияние на жесткость сосудов, даже находящихся на значительном расстоянии от очага инсульта [24]. Несмотря на влияние *IL6* на развитие и прогрессирование инсульта, функциональная значимость полиморфизма *rs1800795* в этом гене не очевидна. Данный полиморфизм находится в промоторной области гена и влияет на концентрации *IL6* и С-реактивного белка. В метаанализе также не была подтверждена ассоциация этого полиморфизма с риском развития инсульта [25], поэтому его не следует рассматривать в качестве генетического маркера риска ишемического инсульта.

Параоксоназа (ген *PON1*) — фермент, играющий важную роль в предотвращении атеросклероза, препятствуя окислению липидов в ЛПНП путем их гидролиза, дифференцировке моноцитов в макрофаги, захвату макрофагами окисленных ЛПНП и превращению макрофагов в пенные клетки [26]. Полиморфизм *rs854560* приводит к снижению концентрации параоксоназы, что может быть рассмотрено как фактор риска атеросклероза и, как следствие, инсульта. Однако результаты метаанализа не подтвердили ассоциацию данного полиморфизма с риском развития инсульта [27], поэтому данный генетический маркер не должен учитываться при разработке тест-системы для определения индивидуального риска развития ишемического инсульта.

ВЫВОДЫ

В соответствии с настоящими критериями при разработке прогностических тест-систем, основанных на анализе генетического полиморфизма, отбор ДНК-маркеров необходимо осуществлять, руководствуясь их статистически значимой ассоциацией с исследуемой МФО или функциональной значимостью для проявления этой особенности.

Существующие на сегодняшний день подходы к разработке прогностических тест-систем предполагают учет

либо генетических маркеров, показавших статистически достоверную ассоциацию с МФО, либо учет некоторого числа функционально значимых полиморфизмов. Оба подхода имеют существенные недостатки, которые сказываются на прогностических свойствах тест-системы. Отбор генетических маркеров на основании статистически значимой ассоциации [11] может приводить к исключению функционально значимых полиморфизмов в процессе отправки на множественность либо из-за их сравнительно невысокой частоты встречаемости.

В то же время механизмы формирования многих МФО сегодня изучены недостаточно хорошо, поэтому не для всех ассоциированных генетических маркеров известен механизм влияния на развитие МФО. Учет только полиморфизмов, для которых установлена функциональная значимость и ассоциация которых была подтверждена в некоторых исследованиях, посвященных изучению генов-кандидатов МФО, может приводить к снижению чувствительности и специфичности разрабатываемой тест-системы. Например, оценка индивидуального риска развития ишемического инсульта только по полиморфизму гена *PDE4D* [28] будет приводить к значительному количеству ложноотрицательных результатов, поскольку, как было показано выше, полиморфизм значительно большего количества генов ассоциирован с патофизиологическими процессами, приводящими к инсульту. Данный подход также будет приводить к ложноположительным результатам, поскольку в метаанализе [29] не была подтверждена ассоциация полиморфизма гена *PDE4D* с риском развития инсульта в европейских популяциях. Функциональная значимость генетических маркеров в данном гене также не установлена.

Указанные выше недостатки устраняются в данных критериях путем анализа целесообразности включения в прогностические тест-системы обоих типов маркеров. При этом если в настоящее время не установлена функциональная значимость генетического маркера, за статистически достоверную ассоциацию принимается полногеномная ассоциация с подтверждением на независимой выборке. Данный подход позволяет уменьшить включение генетических маркеров, для которых была получена ложноположительная ассоциация. Учет данных метаанализов позволяет решить вопрос целесообразности включения в разрабатываемые тест-системы генетических маркеров, ассоциация которых не достигла полногеномной значимости в ходе GWAS. В случае отсутствия GWAS и метаанализа генетический маркер учитывается при разработке тест-системы только в случае установленного влияния на развитие МФО.

Дальнейшая валидация представленного метода будет осуществлена путем составления перечня генетических маркеров, ассоциированных с МФО в соответствии с данными критериями, а также критериями других авторов. На основании упомянутых перечней маркеров станет возможным построение нескольких предсказательных моделей в зависимости от используемых критериев. Сравнение данных моделей на реальных данных генотипирования позволит установить целесообразность использования предлагаемых критериев.

Литература

1. Giardiello FM, Allen JL, Axilbund JE, Boland CR, Burke CA, Burt RW, et al. Guidelines on genetic evaluation and management

of Lynch syndrome: a consensus statement by the US Multi-society Task Force on colorectal cancer. *Am J Gastroenterol.*

- 2014 Aug; 109 (8): 1159–79.
2. Wang TJ, Zhang F, Richards JB, Kestenbaum B, van Meurs JB, Berry D, et al. Common genetic determinants of vitamin D insufficiency: a genome-wide association study. *Lancet*. 2010 Jul 17; 376 (9736): 180–8.
 3. Lin JS, Thompson M, Goddard KAB, Piper MA, Heneghan C, Whitlock EP. Evaluating genomic tests from bench to bedside: a practical framework. *BMC Med Inform Decis Mak*. 2012 Oct 19; 12: 117.
 4. Андреевко Е. Ю., Балацкий А. В., Бойцов С. А., Самоходская Л. М., Ткачук В. А., авторы; Факультет фундаментальной медицины Московского государственного университета имени М. В. Ломоносова, Российский кардиологический научно-производственный комплекс, патентообладатели. Способ определения наследственной предрасположенности к развитию инфаркта миокарда у лиц без клинических проявлений ишемической болезни сердца. Патент РФ № 2469096. 10.12.2012. Доступно по: http://www1.fips.ru/fips_servl/fips_servlet?DB=RUPAT&DocNumber=2469096&TypeFile=html
 5. Dowaidar M, Settin A. Risk of myocardial infarction related to factor V Leiden mutation: a meta-analysis. *Genet Test Mol Biomarkers*. 2010 Aug; 14 (4): 493–8.
 6. Bito LZ, Matheny A, Cruickshanks KJ, Nondahl DM, Carino OB. Eye color changes past early childhood. *The Louisville Twin Study. Arch Ophthalmol*. 1997 May; 115 (5): 659–63.
 7. Nuzzo R. Scientific method: statistical errors. *Nature*. 2014 Feb 13; 506 (7487): 150–2.
 8. Pearson TA, Manolio TA. How to interpret a genome-wide association study. *JAMA*. 2008 Mar 19; 299 (11): 1335–44.
 9. Moll S, Varga EA. Homocysteine and MTHFR Mutations. *Circulation*. 2015 Jul 7; 132 (1): e6–9.
 10. Kingah PL, Luu HN, Volcik KA, Morrison AC, Nettleton JA, Boerwinkle E. Association of NOS3 Glu298Asp SNP with hypertension and possible effect modification of dietary fat intake in the ARIC study. *Hypertens Res*. 2010 Feb; 33 (2): 165–9.
 11. Hsu A, Naughton B, Wu Sh. 23andMe. White Paper 23-03. Guidelines on Vetting Genetic Associations [файл из Интернета]. 2007- [изменено 10.06.2010; процитировано 20.12.2016]. 115 KB. Доступно по: https://23andme.https.internapcdn.net/res/pdf/trmm3vmf11BU5d3Qw_qlGg_23-03_Vetting_Genetic_Associations_2010_06.pdf
 12. Wang WYS, Barratt BJ, Clayton DG, Todd JA. Genome-wide association studies: theoretical and practical concerns. *Nat Rev Genet*. 2005 Feb; 6 (2): 109–18.
 13. Russo MW. How to Review a Meta-analysis. *Gastroenterol Hepatol (N Y)*. 2007 Aug; 3 (8): 637–42.
 14. Ikram MA, Seshadri S, Bis JC, Fornage M, DeStefano AL, Aulchenko YS, et al. Genomewide association studies of stroke. *N Engl J Med*. 2009 Apr 23; 360 (17): 1718–28.
 15. Dichgans M, Malik R, König IR, Rosand J, Clarke R, Gretarsdottir S, et al. Shared genetic susceptibility to ischemic stroke and coronary artery disease: a genome-wide analysis of common variants. *Stroke*. 2014 Jan; 45 (1): 24–36.
 16. Gretarsdottir S, Thorleifsson G, Manolescu A, Styrkarsdottir U, Helgadottir A, Gschwendtner A, et al. Risk variants for atrial fibrillation on chromosome 4q25 associate with ischemic stroke. *Ann Neurol*. 2008 Oct; 64 (4): 402–9.
 17. Casas JP, Hingorani AD, Bautista LE, Sharma P. Meta-analysis of genetic studies in ischemic stroke: thirty-two genes involving approximately 18,000 cases and 58,000 controls. *Arch Neurol*. 2004 Nov; 61 (11): 1652–61.
 18. Franco RF, Trip MD, ten Cate H, van den Ende A, Prins MH, Kastelein JJ, et al. The 20210G → A mutation in the 3'-untranslated region of the prothrombin gene and the risk for arterial thrombotic disease. *Br J Haematol*. 1999 Jan; 104 (1): 50–4.
 19. Wald DS, Law M, Morris JK. Homocysteine and cardiovascular disease: evidence on causality from a meta-analysis. *BMJ*. 2002 Nov 23; 325 (7374): 1202.
 20. Agerholm-Larsen B, Nordestgaard BG, Tybjaerg-Hansen A. ACE gene polymorphism in cardiovascular disease: meta-analyses of small and large studies in whites. *Arterioscler Thromb Vasc Biol*. 2000 Feb; 20 (2): 484–92.
 21. Xin X-Y, Song Y-Y, Ma J-F, Fan C-N, Ding J-Q, Yang G-Y, et al. Gene polymorphisms and risk of adult early-onset ischemic stroke: A meta-analysis. *Thromb Res*. 2009 Nov; 124 (5): 619–24.
 22. Farrer LA, Cupples LA, Haines JL, Hyman B, Kukull WA, Mayeux R, et al. Effects of age, sex, and ethnicity on the association between apolipoprotein E genotype and Alzheimer disease. A meta-analysis. APOE and Alzheimer Disease Meta Analysis Consortium. *JAMA*. 1997 Oct 22–29; 278 (16): 1349–56.
 23. Santoro SA, Zutter MM. The alpha 2 beta 1 integrin: a collagen receptor on platelets and other cells. *Thromb Haemost*. 1995 Sep; 74 (3): 813–21.
 24. Tuttolomondo A, Di Raimondo D, di Sciacca R, Pinto A, Licata G. Inflammatory cytokines in acute ischemic stroke. *Curr Pharm Des*. 2008; 14 (33): 3574–89.
 25. Jin XF, Wang DL, Zhou Y, Xiong H. Association between the interleukin-6-174 G/C polymorphism and risk of ischemic stroke: a meta-analysis. *Genet Mol Res*. 2015 Oct 26; 14 (4): 13076–83.
 26. Watson AD, Berliner JA, Hama SY, La Du BN, Faull KF, Fogelman AM, et al. Protective effect of high density lipoprotein associated paraoxonase. Inhibition of the biological activity of minimally oxidized low density lipoprotein. *J Clin Invest*. 1995 Dec; 96 (6): 2882–91.
 27. Shao P, Qu D-J, Song R-Y, Chen M-L, Wang L-H. Association between PON1 L55M polymorphism and ischemic stroke: a systematic review and meta-analysis. *Int J Clin Exp Med*. 2015 Mar 15; 8 (3): 3429–37.
 28. Бондаренко Е. А., Тупицына Т. В., Шамалов Н. А., Шлетова И. М., Скворцова В. И., Наседкина Т. В. и др., авторы; Российский государственный медицинский университет, Российская Федерация, от имени которой выступает Министерство образования и науки Российской Федерации, патентообладатели. Аллель SNP41 гена PDE4D, его применение для прогнозирования индивидуальной предрасположенности к инсульту в русской популяции, применение молекулярно-генетического маркера индивидуальной предрасположенности к инсульту и способ прогнозирования индивидуальной предрасположенности к инсульту. Патент РФ № 2422523. 27.06.2011. Доступно по: http://www1.fips.ru/fips_servl/fips_servlet?DB=RUPAT&DocNumber=2422523&TypeFile=html
 29. Domingues-Montanari S, Fernández-Cadenas I, del Rio-Espinola A, Corbeto N, Krug T, Manso H, et al. Association of a genetic variant in the ALOX5AP with higher risk of ischemic stroke: a case-control, meta-analysis and functional study. *Cerebrovasc Dis*. 2010; 29 (6): 528–37.

References

1. Giardiello FM, Allen JI, Axilbund JE, Boland CR, Burke CA, Burt RW, et al. Guidelines on genetic evaluation and management of Lynch syndrome: a consensus statement by the US Multi-society Task Force on colorectal cancer. *Am J Gastroenterol*. 2014 Aug; 109 (8): 1159–79.
2. Wang TJ, Zhang F, Richards JB, Kestenbaum B, van Meurs JB, Berry D, et al. Common genetic determinants of vitamin D insufficiency: a genome-wide association study. *Lancet*. 2010 Jul 17; 376 (9736): 180–8.
3. Lin JS, Thompson M, Goddard KAB, Piper MA, Heneghan C, Whitlock EP. Evaluating genomic tests from bench to bedside: a practical framework. *BMC Med Inform Decis Mak*. 2012 Oct 19; 12: 117.
4. Андреевко Е. Ю., Балатский А. В., Бойцов С. А., Самоходская Л. М., Ткачук В. А., авторы; Факультет фундаментальной медицины Московского государственного университета имени М. В. Ломоносова, Российский кардиологический научно-производственный комплекс, патентообладатели. Способ определения наследственной предрасположенности к развитию инфаркта миокарда у лиц без клинических проявлений ишемической болезни сердца. Патент РФ № 2469096. 10.12.2012. Доступно по: http://www1.fips.ru/fips_servl/fips_servlet?DB=RUPAT&DocNumber=2469096&TypeFile=html

- инфаркта миокарда у лиц без клинических проявлений ишемической болезни сердца. Russian Federation patent RU 2469096. 2012 Dec 10. Available from: http://www1.fips.ru/fips_servl/fips_servlet?DB=RUPAT&DocNumber=2469096&TypeFile=html. Russian.
5. Dowaidar M, Settin A. Risk of myocardial infarction related to factor V Leiden mutation: a meta-analysis. *Genet Test Mol Biomarkers*. 2010 Aug; 14 (4): 493–8.
 6. Bito LZ, Matheny A, Cruickshanks KJ, Nondahl DM, Carino OB. Eye color changes past early childhood. *The Louisville Twin Study. Arch Ophthalmol*. 1997 May; 115 (5): 659–63.
 7. Nuzzo R. Scientific method: statistical errors. *Nature*. 2014 Feb 13; 506 (7487): 150–2.
 8. Pearson TA, Manolio TA. How to interpret a genome-wide association study. *JAMA*. 2008 Mar 19; 299 (11): 1335–44.
 9. Moll S, Varga EA. Homocysteine and MTHFR Mutations. *Circulation*. 2015 Jul 7; 132 (1): e6–9.
 10. Kingah PL, Luu HN, Volcik KA, Morrison AC, Nettleton JA, Boerwinkle E. Association of NOS3 Glu298Asp SNP with hypertension and possible effect modification of dietary fat intake in the ARIC study. *Hypertens Res*. 2010 Feb; 33 (2): 165–9.
 11. Hsu A, Naughton B, Wu Sh. 23andMe. White Paper 23-03. Guidelines on Vetting Genetic Associations [file on the Internet]. 2007- [revised 2010 Jun 10; cited 2016 Dec 20]. 115 KB. Available from: https://23andme.https.internapcdn.net/res/pdf/trmm3vmf11BU5d3Qw_qIGg_23-03_Vetting_Genetic_Associations_2010_06.pdf
 12. Wang WYS, Barratt BJ, Clayton DG, Todd JA. Genome-wide association studies: theoretical and practical concerns. *Nat Rev Genet*. 2005 Feb; 6 (2): 109–18.
 13. Russo MW. How to Review a Meta-analysis. *Gastroenterol Hepatol (N Y)*. 2007 Aug; 3 (8): 637–42.
 14. Ikram MA, Seshadri S, Bis JC, Fornage M, DeStefano AL, Aulchenko YS, et al. Genomewide association studies of stroke. *N Engl J Med*. 2009 Apr 23; 360 (17): 1718–28.
 15. Dichgans M, Malik R, König IR, Rosand J, Clarke R, Gretarsdottir S, et al. Shared genetic susceptibility to ischemic stroke and coronary artery disease: a genome-wide analysis of common variants. *Stroke*. 2014 Jan; 45 (1): 24–36.
 16. Gretarsdottir S, Thorleifsson G, Manolescu A, Styrkarsdottir U, Helgadóttir A, Gschwendtner A, et al. Risk variants for atrial fibrillation on chromosome 4q25 associate with ischemic stroke. *Ann Neurol*. 2008 Oct; 64 (4): 402–9.
 17. Casas JP, Hingorani AD, Bautista LE, Sharma P. Meta-analysis of genetic studies in ischemic stroke: thirty-two genes involving approximately 18,000 cases and 58,000 controls. *Arch Neurol*. 2004 Nov; 61 (11): 1652–61.
 18. Franco RF, Trip MD, ten Cate H, van den Ende A, Prins MH, Kastelein JJ, et al. The 20210 G→A mutation in the 3'-untranslated region of the prothrombin gene and the risk for arterial thrombotic disease. *Br J Haematol*. 1999 Jan; 104 (1): 50–4.
 19. Wald DS, Law M, Morris JK. Homocysteine and cardiovascular disease: evidence on causality from a meta-analysis. *BMJ*. 2002 Nov 23; 325 (7374): 1202.
 20. Agerholm-Larsen B, Nordestgaard BG, Tybjaerg-Hansen A. ACE gene polymorphism in cardiovascular disease: meta-analyses of small and large studies in whites. *Arterioscler Thromb Vasc Biol*. 2000 Feb; 20 (2): 484–92.
 21. Xin X-Y, Song Y-Y, Ma J-F, Fan C-N, Ding J-Q, Yang G-Y, et al. Gene polymorphisms and risk of adult early-onset ischemic stroke: A meta-analysis. *Thromb Res*. 2009 Nov; 124 (5): 619–24.
 22. Farrer LA, Cupples LA, Haines JL, Hyman B, Kukull WA, Mayeux R, et al. Effects of age, sex, and ethnicity on the association between apolipoprotein E genotype and Alzheimer disease. A meta-analysis. APOE and Alzheimer Disease Meta Analysis Consortium. *JAMA*. 1997 Oct 22–29; 278 (16): 1349–56.
 23. Santoro SA, Zutter MM. The alpha 2 beta 1 integrin: a collagen receptor on platelets and other cells. *Thromb Haemost*. 1995 Sep; 74 (3): 813–21.
 24. Tuttolomondo A, Di Raimondo D, di Sciacca R, Pinto A, Licata G. Inflammatory cytokines in acute ischemic stroke. *Curr Pharm Des*. 2008; 14 (33): 3574–89.
 25. Jin XF, Wang DL, Zhou Y, Xiong H. Association between the interleukin-6-174 G/C polymorphism and risk of ischemic stroke: a meta-analysis. *Genet Mol Res*. 2015 Oct 26; 14 (4): 13076–83.
 26. Watson AD, Berliner JA, Hama SY, La Du BN, Faull KF, Fogelman AM, et al. Protective effect of high density lipoprotein associated paraoxonase. Inhibition of the biological activity of minimally oxidized low density lipoprotein. *J Clin Invest*. 1995 Dec; 96 (6): 2882–91.
 27. Shao P, Qu D-J, Song R-Y, Chen M-L, Wang L-H. Association between PON1 L55M polymorphism and ischemic stroke: a systematic review and meta-analysis. *Int J Clin Exp Med*. 2015 Mar 15; 8 (3): 3429–37.
 28. Bondarenko EA, Tupitsyna TV, Shamalov NA, Shetova IM, Skvortsova VI, Nasedkina TV, et al, inventors; Russian State Medical University, the Ministry of Education and Science of the Russian Federation on Russian Federation behalf, assignees. Allel' SNP41 gena PDE4D, ego primeneniye dlya prognozirovaniya individual'noi predraspolozhennosti k insul'tu v russkoi populyatsii, primeneniye molekulyarno-geneticheskogo markera individual'noi predraspolozhennosti k insul'tu i sposob prognozirovaniya individual'noi predraspolozhennosti k insul'tu. Russian Federation patent RU 2422523. 2011 Jun 27. Available from: http://www1.fips.ru/fips_servl/fips_servlet?DB=RUPAT&DocNumber=2422523&TypeFile=html. Russian.
 29. Domingues-Montanari S, Fernández-Cadenas I, del Rio-Espinola A, Corbeto N, Krug T, Manso H, et al. Association of a genetic variant in the ALOX5AP with higher risk of ischemic stroke: a case-control, meta-analysis and functional study. *Cerebrovasc Dis*. 2010; 29 (6): 528–37.