

УДК 575.8:612.6.05

ГЕНЕТИКА И ГЕНОМИКА ЧЕЛОВЕКА. ПОПУЛЯЦИИ И ЭТНОСЫ В ПРОСТРАНСТВЕ И ВРЕМЕНИ: ЭВОЛЮЦИОННЫЕ И МЕДИЦИНСКИЕ АСПЕКТЫ

© 2013 г. С.А. Боринская, Н.К. Янковский

Институт общей генетики имени Н.И. Вавилова РАН, Москва, Россия,
e-mail: nick.yankovsky@vigg.ru

Поступила в редакцию 9 августа 2013 г. Принята к публикации 1 ноября 2013 г.

Генетическая адаптация популяций человека к локальным условиям среды может быть представлена как возникновение новых аллелей в результате мутаций и последующее изменение частот аллелей в поколениях вследствие естественного отбора признаков, ассоциированных с этими аллелями и важных для выживания и успешной репродукции человека. Помимо процессов адаптации, на изменение частот аллелей влияют генетический дрейф и миграции. Исследование генетической адаптации человека занимает одно из центральных мест в биологической антропологии, генетике человека и эволюционной биологии и уже дало значительный вклад в понимание взаимодействия средовых и генетических факторов, влияющих на здоровье человека.

Ниже будут описаны подходы к поиску аллелей, определяющих приспособленность человека к действию факторов окружающей среды. Эти подходы потенциально приложимы не только к человеку, но и к другим животным и растениям. Исходной посылкой для применения этих подходов является перекрывание ареала действия конкретного фактора среды и ареала повышенной популяционной частоты аллеля, что позволяет предполагать их связь, хотя и не доказывает ее. Такое совпадение ареалов, после вычитания вклада случайных причин (генетический дрейф, генетическая подразделенность популяций), является основой для выдвижения и дальнейшей проверки гипотезы о роли данного фактора отбора в повышении частоты данного аллеля в популяциях человека. Прямое доказа-

тельство таких гипотез может быть получено, если удастся показать связь генотипа и фенотипа с приспособленностью к данному фактору отбора (выживаемость или репродуктивный успех) (Hancock *et al.*, 2008, 2010a).

Мы привыкли сравнивать генетические особенности индивидов друг с другом. Индивид как единица учета был и остается необходимым, но позволяет установить лишь наличие аллеля, но не его частоту. Для придания «смысла» гену и конкретному аллелю нужно применить эволюционный подход. Для этого единицей учета должна стать популяция, а учитываемым признаком станет частота аллеля. При этом сравнивать друг с другом нужно признаки популяций (условия среды и частоты аллелей), а не признаки индивидов.

ИСТОРИЯ ВОПРОСА

Исследования генетических причин адаптации к локальной среде ведутся уже более 60 лет. Для этого сравнивали распределение частот встречаемости фенотипов по отношению к величине значений средовых переменных, предположительно являющихся факторами отбора. Такой подход был использован в классических исследованиях географического распределения аллелей бета-талассемии (Haldane, 1949), серповидноклеточной анемии (Allison, 1954) и дефицита глюкозо-6-фосфатдегидрогеназы человека (более детально описанного ниже), придающих устойчивость к малярии. Путь от гипотезы о протективном действии аллеля,

выдвинутой на основе совпадения ареалов распространения аллеля с ареалом эндемичной инфекции, до четкого доказательства, включающего понимание молекулярных механизмов, занимает порой десятки лет. Рассмотрим этот путь на примере гена глюкозо-6-фосфат дегидрогеназы (*G6PD*).

В начале 1950-х годов во время корейской войны американским солдатам в обязательном порядке выдавался примахин. Это лекарство должно было защитить солдат от малярии, но у части из них прием препарата вызывал тяжелое побочное действие – острую гемолитическую анемию. Внезапно развивавшееся заболевание крови приводило даже к гибели. Исследование показало, что смертность от «примахиновой» анемии различалась у представителей разных этнических групп. Особенно высокой (до 10–15 %) она была среди солдат, предки которых имели средиземноморское происхождение (выходцы из северной Африки, Италии, Испании, Ближнего Востока). Вскоре были открыты причины «примахиновой» анемии. Она возникала у людей с наследственным дефицитом фермента глюкозо-6-фосфатдегидрогеназы (*G6PD*) (Carson *et al.*, 1956). Фермент *G6PD* участвует в гексоз-монофосфатном пути, единственном процессе, генерирующем НАДФ в зрелых эритроцитах, в которых отсутствует цикл Кребса. Этот фермент катализирует окисление глюкозо-6-фосфата в 6-фосфоглюконат. В этой реакции образуется восстановленный никотинамидадениндинуклеотидфосфат (НАДФН), который в дальнейшем используется для восстановления глутатиона (при участии глутатионредуктазы), а также частично метгемоглобина в гемоглобин. Восстановленный глутатион защищает гемоглобин и тиоловые ферменты, поддерживающие нормальную проницаемость мембран эритроцитов, от окислительного действия различных веществ, в том числе и лекарственных препаратов. При недостаточности *G6PD* прием некоторых лекарственных средств (примахин, сульфаниламиды и др.) или продуктов питания (бобы) ведет к массивному разрушению эритроцитов (гемолитические кризы) вследствие падения содержания в них восстановленного глутатиона и дестабилизации мембран. Размножение малярийного плазмодия в эритроцитах также дестабилизирует их мембраны. Сочетание на-

следственного дефекта и инфекции приводит к гибели зараженных эритроцитов, и паразит не успевает развиваться. В результате у людей с дефицитом фермента малярия протекает в более легкой форме. Показано, что в культуре клеток эритроциты, дефицитные по *G6PD*, поддерживают рост малярийного плазмодия в 3 раза хуже, чем нормальные клетки (Roth *et al.*, 1983). Эпидемиологические данные показали, что распространенная в Африке форма дефицита *G6PD* ассоциирована со снижением риска тяжелой малярии в 2 раза, причем как у гемизиготных мужчин, имеющих дефектный ген в своей единственной X-хромосоме, так и у гетерозиготных женщин, носительниц одного нормального и одного дефектного гена (Ruwando *et al.*, 1995). Мутации в гене *G6PD*, ведущие к дефициту фермента, возникали неоднократно и широко распространились в малярийных зонах под действием отбора на устойчивость к этой инфекции.

ГЕНОМ ЧЕЛОВЕКА И БАЗЫ ДАННЫХ ЧАСТОТ АЛЛЕЛЕЙ

Для проверки гипотез о вовлечении определенного фенотипа в процесс адаптации и выявлении факторов среды, вызывающих эту адаптацию, необходим сбор фенотипических данных, что дорого и требует значительного времени, а также экспериментальный анализ больших выборок для достижения статистически значимых показателей (Hancock *et al.*, 2008, 2010a). В то же время в последнее десятилетие интенсивно развиваются подходы к выявлению особенностей геномов людей, указывающих на действие отбора. Участки генома, подверженные действию отбора, отличаются по ряду характеристик от участков с нейтральными генетическими вариациями (Gillespie *et al.*, 1991). Недавнее развитие технологий и ресурсов (проекты HarMap, «1000 геномов» и др.) для исследования генетических вариаций человека в беспрецедентном масштабе позволяют исследователям сканировать полные геномы человека в поиске сигналов позитивного отбора (см обзор: Hancock *et al.*, 2008, 2010b).

Созданы базы данных, описывающие разнообразие генома человека. Состав и частоты аллелей отражают специфику популяции, ее

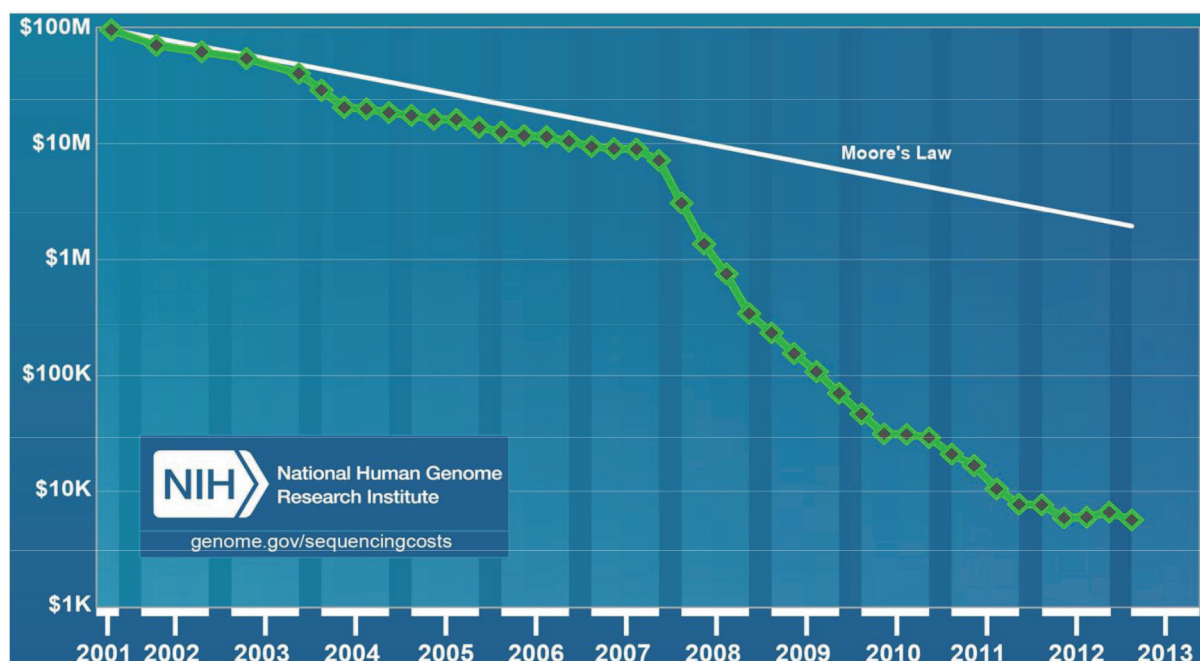


Рис. 1. Изменение стоимости секвенирования генома человека (genome.gov/sequencingcosts).

происхождение, историю, миграции и родство с другими популяциями. Выявление новых аллелей в геноме человека происходит сейчас в основном в результате секвенирования. Это секвенирование отдельных участков генома, представляющих интерес для исследователя, и секвенирование всего генома индивида целиком. Стоимость секвенирования генома человека составляет около 5000 дол. на июль 2013 г. и продолжает быстро падать (рис. 1). Для человека частоты аллелей и гаплотипов всех популяций мира будут, вероятно, установлены в представимом будущем полногеномным секвенированием.

Для многих локусов, выявленных методом геномного сканирования как предполагаемых мишеней отбора, связь с фенотипическими признаками пока не известна (Coop *et al.*, 2009). Неизвестны и факторы отбора аллелей в этих локусах. Наиболее очевидными факторами является адаптация к геоклиматическим условиям, особенностям традиционного питания, эндемичным инфекциям и особенностям образа жизни.

Для выявления таких факторов необходимо сопоставление вариаций частот аллелей в популяциях и присутствия или наличия потенциального фактора отбора в среде, либо интенсивности его действия. Различия между

популяциями по составу и частоте аллелей формируются либо в результате локус-специфичных процессов (различные виды отбора, направленные на один или несколько локусов, определяющих адаптивный фенотип), либо в результате событий популяционного уровня (колебания численности и экспансия популяций, метисация и др.), которые затрагивают множество локусов одновременно (рис. 2).

Средовые факторы, отбор, популяционные факторы (миграции, дрейф) действуют одновременно на многие гены в популяции. Чтобы приписать аллелю адаптивную и медицинскую роль, нужно уметь: 1) отличить действие отбора от действия стохастических факторов (дрейф, серфинг аллеля – серия «бутылочных горлышек») и 2) выявить аллель, причинный для заболевания, среди всех аллелей причинного гаплотипа, подхваченного отбором (Gene lifting).

К указаниям на возможное действие отбора на аллель и ген, к которому он принадлежит, относят высокую популяционную частоту либо гомозиготность по эволюционно «молодым» аллелям, присутствие в популяции протяженных гаплотипов с частотой, превышающей ожидаемую для случайного распределения размеров гаплотипов, экстремально высокий уровень

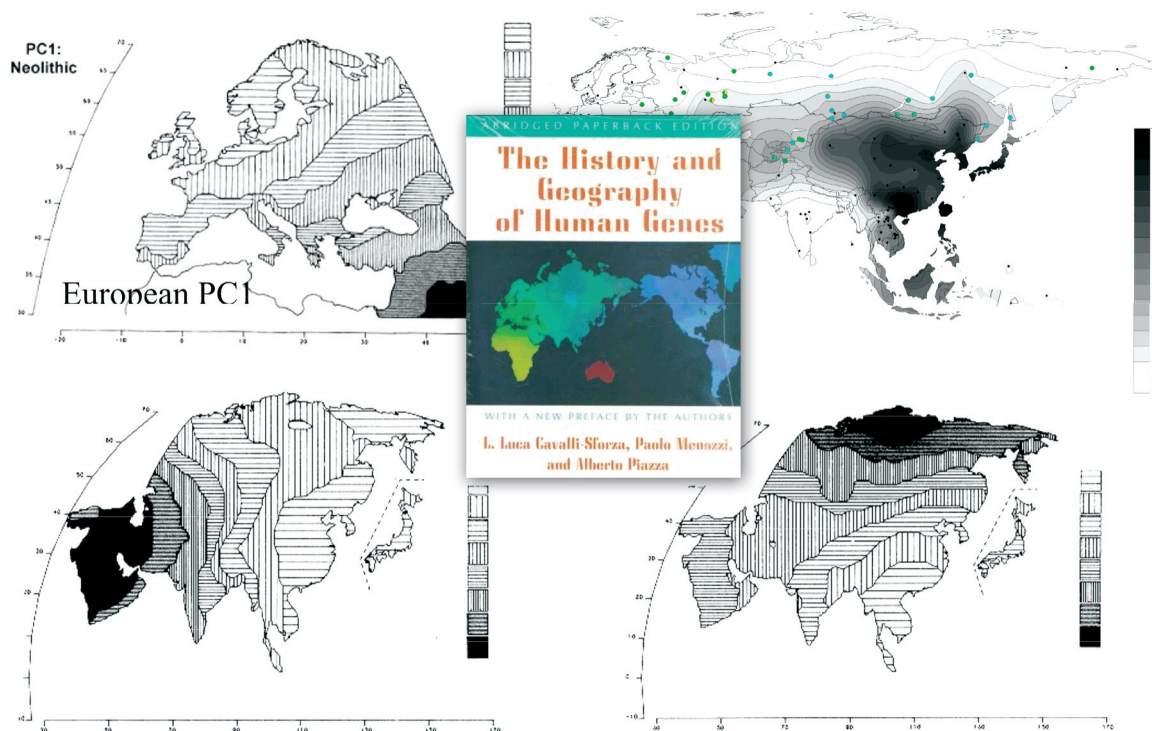


Рис. 2. Различные главные компоненты в распределении частот аллелей в Евразии, формирующиеся под действием природных факторов и миграций (Cavalli-Sforza *et al.*, 1994; Excoffier, Ray, 2008; Slatkin, Excoffier, 2012).

Карта частот аллеля гена *ADH1B* приведена по: (Borinskaya *et al.*, 2009. P. 89–92).

Профили географического распределения популяционных частот аллелей различны для разных аллелей и гаплотипов (набор сцепленных аллелей). Исходным источником генетического разнообразия являются мутации, например SNP. Новый аллель может накопиться в поколениях и сформировать уникальную географическую карту распределения частоты встречаемости, которая определяется как случайными причинами (генетический дрейф) так и отбором (адаптация).

межпопуляционных различий по частоте аллеля (высокий уровень *F_{st}*) (Sabeti *et al.*, 2006). Совпадение ареалов действия конкретных факторов среды и высокой частоты тех или иных аллелей позволяет предполагать их связь, но не доказывает. Однако именно такое совпадение после контроля возможных причин корреляции на популяционном уровне (контроля генетической структурированности популяций) является основой для выдвижения и дальнейшей проверки гипотез о факторах отбора, влияющих на частоты аллелей в популяциях человека. Следует подчеркнуть, что геномный анализ может указать на гаплотип, ген или даже его аллель, который подвержен действию отбора, но этот анализ не способен указать на фактор среды, который определил адаптивность данного аллеля и явился причиной его накопления в данной популяции.

КАКОВ ВКЛАД МУТАЦИЙ РАЗНОГО ЭВОЛЮЦИОННОГО ВОЗРАСТА В БОЛЕЗНИ ЧЕЛОВЕКА И В ЧАСТОТУ СООТВЕТСТВУЮЩЕГО АЛЛЕЛЯ В ПОПУЛЯЦИИ?

Считается, что чем чаще встречается данный аллель в популяции, тем меньше величина его вклада в развитие заболевания (рис. 3).

Подавляющее большинство одиночных мутаций не имеют морфологического проявления и не приводят к менделирующим заболеваниям человека. Мутации, принадлежащие к генам менделирующих наследственных заболеваний человека, немногочисленны и составляют порядка одной сотысячной от всех десятков миллионов мутаций человека, зарегистрированных в базах данных.

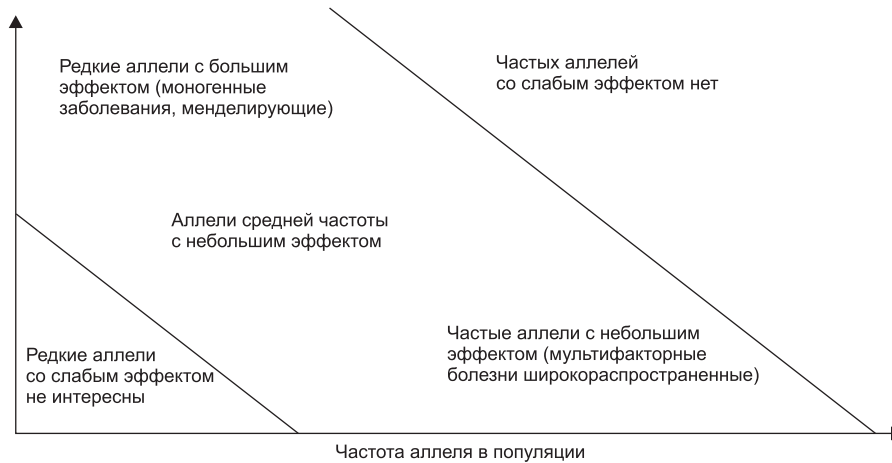


Рис. 3. Вклад аллеля в формирование признака и частота его встречаемости в популяции.

Ребенка от родителей отличают 63 новых генеративных мутации (по данным секвенса, 78 триад ребенок + оба родителя) (Kong *et al.*, 2012). То есть вероятность возникновения мутации в данной позиции генома в данном поколении составляет $1,2 \times 10^{-8}$. Другими словами, чтобы хотя бы один индивид стал носителем новой генеративной мутации в данной позиции генома, численность популяции должна составлять 10^8 человек (а в расчете на белок-кодирующую часть данного гена – около 10^5 человек). Таким образом, мутации *de novo* не дают большого вклада в частоту встречаемости наследственных болезней человека.

Из 3 млрд позиций в нашем геноме примерно каждая тысячная полиморфна (Venter *et al.*, 2001; Venter, 2010, 2011). Каждый из нас несет 2–3 млн мутаций (сумма различий папиных и маминых хромосом индивида), т. е. 4–6 млн аллелей. Напомним, что в данном поколении в геноме возникает менее 100 новых генеративных мутаций и никакие даже самые ужасные условия внешней среды не способны увеличить эту частоту хотя бы в 10 раз. Таким образом, более чем 99,99 % генеративных мутаций (аллелей) в геноме каждого из нас уже существовали в геномах наших предков и возникли давно – иногда миллионы лет назад (например, группы крови человека и приматов). Именно эти «старые» мутации и их сочетания, генотипы, определяют практически все генетически контролируемые особенности людей.

Какую долю от разнообразия всех закрепившихся в популяциях генеративных мута-

ций несет каждый из нас, один из миллиардов людей, живущих сейчас на Земле? Эта доля совсем не одна миллиардная, а на несколько порядков больше. Число всех описанных на сегодня однонуклеотидных полиморфных участков генома (SNP) составляет несколько десятков миллионов мутаций, а геном каждого из нас несет несколько миллионов мутаций. То есть каждый из нас несет заметную долю всех мутаций, описанных для всех обследованных на данный момент людей.

Основная часть мутаций в генофонде человечества присутствует уже давно, возможно, до времени выхода наших предков из Африки, и большинство аллелей присутствуют в большинстве популяций мира, хотя и с очень разной частотой. «Старые» мутации – это основной потенциал и нашей устойчивости, и чувствительности к болезням, в том числе к мультифакторным.

ФАКТОРЫ ОТБОРА, ДЕЙСТВУЮЩИЕ В ПОПУЛЯЦИЯХ ЧЕЛОВЕКА, И ПОДХОДЫ К ИХ ВЫЯВЛЕНИЮ

Как вид *Homo sapiens* уникален в том, что имеет широчайший ареал расселения и за счет биологических и культурных адаптаций приспособлен к самым различным экорегионам, использованию различных стратегий ведения хозяйства и источников пищи. На протяжении большей части своего существования как вида человек вел образ жизни охотника-собирателя. С появлением производящего хозяйства возникли

новые векторы давления отбора – переход на новый тип питания, изменение демографических характеристик обществ, появление новых источников инфекций в результате перехода к оседлому образу жизни, развитию технологий хранения продуктов, одомашнивания животных.

Возможно, что у охотников-собирателей отбор действовал в пользу сохранения аллелей генов, обеспечивших наиболее эффективное использование дефицитных нутриентов (сахаров, жиров и соли). Таков постулат гипотезы об «экономичных генах», точнее – экономичных аллелях («thrifty» гипотеза (Neel, 1999)), которые, после развития производства доступных продуктов в индустриализованных обществах, стали аллелями риска соответствующих заболеваний (диабета, атеросклероза, гипертонии).

Выявление аллелей генов, являющихся факторами риска развития широко распространенных заболеваний в одних условиях среды и протективных или нейтральных в других условиях среды, актуально как с точки зрения генетики человека и эволюционной биологии, так и для фундаментальных аспектов антропологии и практического здравоохранения.

Для исследования влияния природных и антропогенных факторов на частоты аллелей в популяциях человека и их связь с важными для здоровья признаками актуальными являются междисциплинарные исследования, дающие новые подходы к пониманию путей и механизмов адаптации популяций человека к изменяющимся условиям среды. Сравнение частот аллелей в популяциях и условий их проживания начаты более полувека тому назад профессором Л. Кавалли-Сфорца (сводка дана: Cavalli-Sforza *et al.*, 1994). Наиболее интенсивно велись исследования адаптации к климатическим условиям (см. краткий обзор: Соор *et al.*, 2010; Hancock *et al.*, 2011).

Ранее для таких сравнений нужно было, исходя из функций гена, предположить, по каким параметрам среды следует сравнивать популяции, различающиеся по частотам аллелей этого гена. В настоящее время можно сравнивать базы данных частот аллелей (даже если функциональная значимость аллеля неизвестна) с базами данных, описывающих различные характеристики среды, которые могут служить факторами отбора.

Новые возможности открываются при использовании массивов данных о факторах среды, влияющих на организм человека (помимо климата – особенности диеты, распространение инфекций и многие другие), которые сформированы исследователями в различных областях наук (эпидемиологии, антропологии, этнографии и др.).

Этнографы описывают не популяции, а этнические группы, выделяемые на основе не биологических, а культурных признаков. Популяция не эквивалентна этнической группе. Но популяции, которые представляют один этнос, имеют ряд общих признаков (регион проживания, традиции питания, другие хозяйственно-культурные особенности), и эти признаки формировались на протяжении многих веков адаптации к локальным условиям – адаптации как культурной, так и биологической. Поэтому некоторые особенности, зафиксированные этнографами, отражают влияние факторов среды, к которым проходила генетическая адаптация популяций.

Для проведения анализа корреляций необходимо, чтобы словесное описание было формализовано, т. е. представлено в бинарном виде (фактор присутствует или отсутствует) или в виде численной характеристики, отражающей вклад или интенсивность действия фактора. Такие формализованные этнографические описания были созданы Дж.П. Мердоком на основе картотеки этнографических описаний «Human Relations Area Files» – основного центра аккумуляции этнографической информации в мире. Созданный Дж.П. Мердоком «Этнографический атлас» был опубликован в виде серии статей в журнале «Ethnology» в период с 1960 г. по 1973 г. и дополнен описаниями народов России с участием одного из авторов (Korotayev *et al.*, 2004). «Атлас» не имеет аналогов по полноте описания народов мира и широко используется не только для кросс-культурных и этнографических исследований, но и для исследований в других научных областях, включая медико-антропологические. «Атлас» ранее использовался в популяционно-генетических исследованиях, но лишь как компендиум сведений о языковой принадлежности или экологической классификации среды (Guglielmino *et al.*, 1995) либо наличия отдельного признака (например, молочного животноводства) (Bersaglieri *et al.*, 2004).

В последние годы введены в обращение формализованные описания эпидемиологической нагрузки (Murray, Schaller, 2010), созданные на основе эпидемиологических атласов (Rodenwaldt, Bader, 1952–1961; Simmons *et al.*, 1944).

Это новый подход к поиску факторов отбора, которые могли привести к увеличению популяционной частоты аллеля, обеспечивающего повышенную приспособленность к описанным условиям среды.

Экспериментально доступно определение аллельного состояния для сотен тысяч SNPs (панель Illumina). Такое число аллелей достаточно, чтобы маркировать основные группы неравновесия по сцеплению у европейцев.

Но столь детальные данные в открыто доступных базах данных относятся лишь к нескольким десяткам популяций (HGDP-CEPH populations, 1000 Genomes Project) из тысяч существующих популяций. По единичным генам частоты аллелей известны для сотен популяций (например, > 300 популяций для генов *APOE*, *ADH1B*, для сотен популяций известны частоты аллелей групп крови и других «классических» генетических маркеров).

Сопоставление массивов этих данных имеет ряд ограничений, в частности при наличии сотен популяций, для которых доступны генетические данные, и тысяч этнических групп, для которых доступны базы данных с этнографическими и эпидемиологическими характеристиками, эти массивы перекрываются по немногим десяткам популяций.

Исходным источником генетического разнообразия являются мутации, например SNP. Новый аллель может накопиться в поколениях и сформировать уникальную географическую карту распределения частоты встречаемости, которая определяется как случайными причинами (генетический дрейф в сочетании с миграциями), так и отбором (адаптация).

Средовые факторы, отбор, популяционные факторы (миграции, дрейф) действуют одновременно на многие гены в популяции. Чтобы приписать аллелю адаптивную роль и медицинскую роль, нужно уметь отличить действие отбора от действия стохастических факторов (дрейф, аллель-серфинг – серия «бутылочных горлышек») (Excoffier, Ray, 2008; Slatkin, Excoffier, 2012), а также уметь выявить аллель,

который является причиной заболевания среди всех аллелей гаплотипа, связанного с заболеванием и подхваченного отбором (такое явление называется Gene lifting).

КАКОВЫ КЛЮЧЕВЫЕ ФАКТОРЫ СРЕДЫ, К КОТОРЫМ ПРОХОДИЛА ГЕНЕТИЧЕСКАЯ АДАПТАЦИЯ ПОПУЛЯЦИЙ ЧЕЛОВЕКА?

Это прежде всего климат, инфекции, тип хозяйствования, особенности питания (Cavalli-Sforza *et al.*, 1994; Гончарова и др., 2006; Hancock *et al.*, 2008, 2010 a, b; Coop *et al.*, 2009; Fumagalli *et al.*, 2011). Наиболее обширная сводка сведений об этих факторах содержит база данных Мердока «Этнографический атлас» (рис. 4) (Murdock, 1967). Такие базы пока очень мало использовались в популяционно-генетических исследованиях человека. Они содержат информацию о более чем 1000 этнических групп, которые в случае географической компактности обитания могут рассматриваться как популяции. Использование данных «Этнографического атласа» для выявления корреляций с частотами аллелей популяций, обитающих в тех же регионах, может дать информацию о генах, аллели которых важны для адаптации к факторам среды. Рассмотрим пример такого анализа в следующем разделе.

ГЕНЫ, КОНТРОЛИРУЮЩИЕ МЕТАБОЛИЗМ АЛКОГОЛЯ

Основной способ окисления (до 80–90 % экзогенного алкоголя) связан с действием печеночных ферментов. На первом этапе этанол превращается в ацетальдегид под действием алкогольдегидрогеназы (АДГ), кодируемой геном *ADH1B*. Затем и ацетальдегид окисляется до ацетата под действием ацетальдегиддегидрогеназы (АльдГ), кодируемой геном *ALDH2*.

Оба гена полиморфны. Замена аминокислоты аргинин в позиции 48 полипептидной цепи АДГ на гистидин (аллель *48His*) приводит к 100-кратному увеличению в скорости работы фермента (Jornvall *et al.*, 1984; Matsuo *et al.*, 1989). В результате при потреблении алкоголя происходит более быстрое накопление токсичного альдегида, что связано с неприятными

Факторы среды (природной и культурной) описаны в базах данных

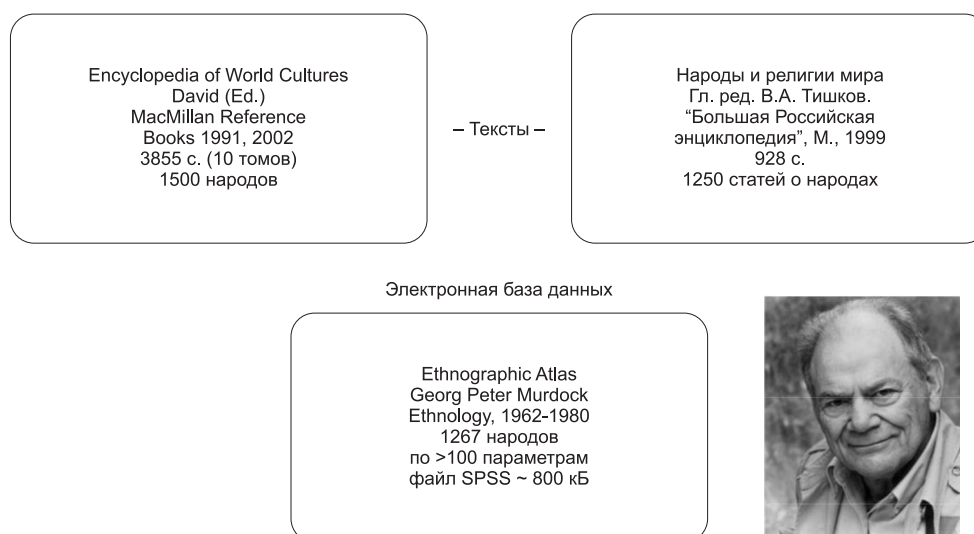


Рис. 4. Сопоставление объема сводок по этнографии народов.

ощущениями. Носители аллеля *48His* в среднем потребляют меньше алкоголя и реже страдают алкоголизмом (Osier *et al.*, 1999; Tolstrup *et al.*, 2008). Еще более сильный эффект дает аллель гена *ADLH2*, у носителей которого имеется замена глутамина на лизин *Glu504Lys*. Аллель *504Lys* определяет синтез неактивного фермента. У гомозигот по данному аллелю фермент нефункционален, а у гетерозигот активность составляет лишь около 6 % от уровня активности данного фермента у гомозигот по эволюционно более древнему аллелю *504Glu* (Crabb *et al.*, 1989). При потреблении даже небольших доз алкоголя у носителей аллеля *ALDH2*504Lys* часто развивается комплекс симптомов (головкружение, учащение сердцебиений, пототделение, тошнота, покраснение кожи лица), препятствующий продолжению приема алкоголя. Частота аллеля *ALDH2*504Lys* снижена у алкоголиков, а гомозиготы по данному аллелю среди них крайне редки. Гетерозиготные носители аллеля *ALDH2*504Lys* потребляют меньше алкоголя, чем те индивиды, у которых аллель отсутствует (Wall *et al.*, 2000; Kim *et al.*, 2008). Популяционные различия по частотам этих аллелей довольно велики (карты распределения частот представлены на рис. 5). Частоты аллеля *48His* варьируют от 80 % в популяциях Восточной Азии до практически полного отсутствия в европейских популяциях. Имеется также локальный максимум (30–40 % частота) в популя-

циях Ближнего Востока (Borinskaya *et al.*, 2009). Аллель *ALDH2*504Lys* представлен с частотой до 40 % в популяциях Китая и Японии и менее 1 % – в Европе, Западной Азии и Африке (Li *et al.*, 2009) (рис. 5). У коренного населения Нового Света оба аллеля отсутствуют.

В ряде работ показано, что в популяциях Юго-Восточной Азии оба аллеля – *ADH1B*48His* и *ALDH2*504Lys* – находятся под действием отбора (Oota *et al.*, 2004, 2007). Предположения о том, что частоты этих аллелей в популяциях Юго-Восточной Азии возросли под действием отбора, были выдвинуты вскоре после того, как было установлено, что различия между «типичными» *ADH1B*47Arg* и *ALDH2*504Glu* и «атипичными» (*ADH1B*48His* и *ALDH2*504Lys*) вариантами обусловлены единичными аминокислотными заменами, и что «атипичные» варианты являются эволюционно более поздними, чем распространенные у европеоидов «типичные» изоформы ферментов. Доводом в пользу отбора было то, что повышена частота производных аллелей, расположенных на разных хромосомах генов, которые контролируют одно звено метаболизма, и оба атипичных варианта ферментов повышают концентрацию ацетальдегида (Ikuta *et al.*, 1986). В качестве фактора отбора предполагались особенности диеты, в частности потребление содержащих алкоголь напитков и пищи. При этом подразумевалось, что адаптивную ценность этим аллелям обеспечивал именно их протектив-

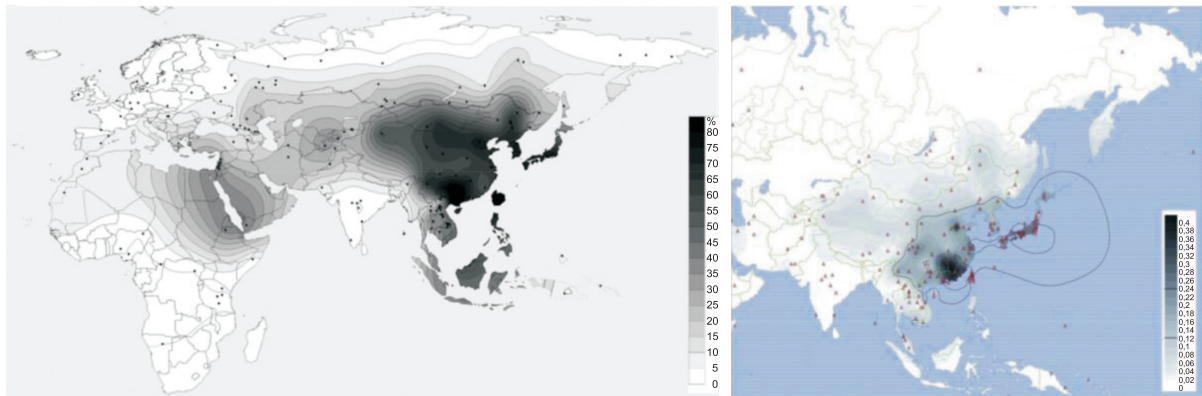


Рис. 5. Географическое распределение частот аллелей генов метаболизма алкоголя.

Аллель *ADH1B*48His* отвечает за быстрое накопление токсичного альдегида в популяциях на Ближнем Востоке и особенно в Юго-Восточной Азии (> 80 %). Аллель *ALDH2*504Lys* – за медленное разрушение этого яда в популяциях Юго-Восточной Азии (> 40 %).

ный эффект в отношении развития алкоголизма (Ikuta *et al.*, 1986). Гипотеза о том, что фактором отбора аллелей *ADH1B*47His* и *ALDH2*504Lys* могла быть какая-либо тропическая инфекция по типу малярии и серповидноклеточной анемии, была выдвинута Дж. Лонгом (Liberles, 2001). Сходное предположение было выдвинуто группой К. Кидда (Oota *et al.*, 2004).

Если какой-либо из перечисленных факторов среды влияет на частоту аллеля, то можно ожидать, что в популяциях, в среде обитания которых фактор присутствует, частоты аллеля будут в среднем выше, чем в тех, в которых фактор отсутствует. Чтобы проверить эту гипотезу, мы провели анализ корреляций частот аллеля *ADH1B*48His* с наличием эндемичных инфекций, скотоводства (одомашнивание скота послужило распространению инфекций) и земледелия. В качестве источника описаний факторов среды использовали «Этнографический атлас» (Murdock, 1967). Для двух переменных (скотоводство и филяриоз) получены значимые корреляции, но лишь для филяриоза корреляция остается значимой после поправки на множественное тестирование. Корреляционный анализ показывает, что частота аллеля *ADH1B*48His* повышена в популяциях, в которых филяриозы эндемичны (рис. 6).

В тех популяциях, для которых филяриозы эндемичны, частота аллеля *ADH1B*47His* составляет от 20 до 80 %. В тех популяциях, где эта инфекция не представлена, частота аллеля не превышает 25–30 % (рис. 6).

Полученные результаты позволяют выдвинуть гипотезу о том, что аллель *ADH1B*48His* может быть протективным в отношении развития филяриозной инфекции у человека. Для того, чтобы проверить, что именно аллель *ADH1B*48His* обеспечивает корреляцию с наличием филяриозов в среде проживания, мы определили корреляции частот аллелей, сцепленных с данным SNP (рис. 7). Соответствующие коэффициенты корреляции убывают по обе стороны от *ADH1B*Arg48His* (рис. 7). Это позволяет заключить, что если отбор на устойчивость к филяриозам действовал на locus *ADH*, то именно функциональный полиморфизм *ADH1B*Arg48His* был его мишенью. Для проверки этой гипотезы необходимо сравнить частоты данного аллеля в группе больных филяриозами и группе индивидов, проживающих в том же регионе, но не инфицированных на протяжении длительного времени. Если аллель обладает протективным эффектом, то частота его у здоровых людей будет выше, чем у больных. Возможна проверка предложенной гипотезы на животных моделях филяриозов (например на крысах) с применением ингибиторов АльДГ (например дисульфурама), так как блокирование работы АльДГ приведет к тому же фенотипическому эффекту, что и ускорение работы АДГ. В мире инфицировано филяриями более 120 млн человек (Taylor *et al.*, 2010), и выявление молекулярных основ природной устойчивости к инфекциям может открыть новые пути лечения и профилактики.

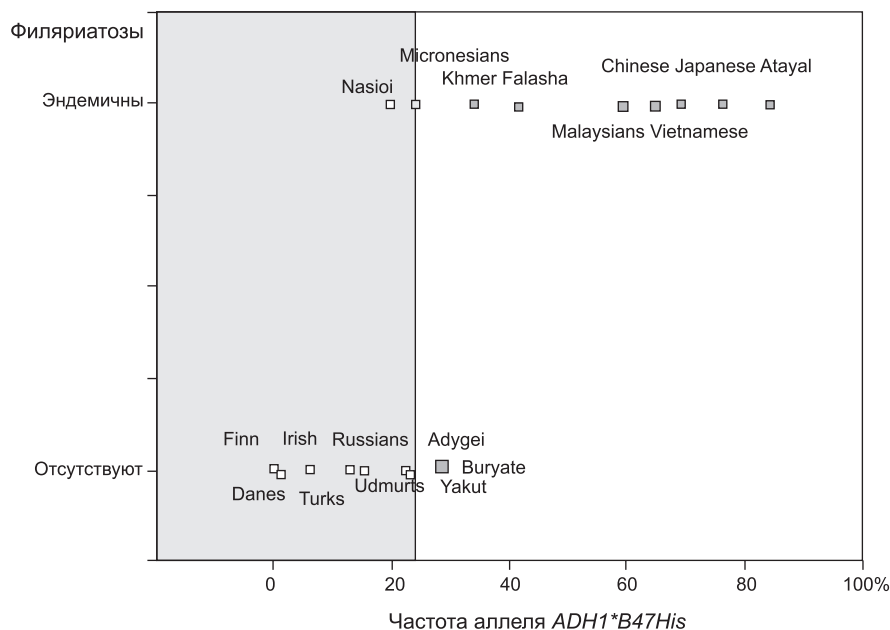


Рис. 6. Распределение частот аллеля *ADH1B*48His* в популяциях в зависимости от наличия филяриозов в регионе их проживания.

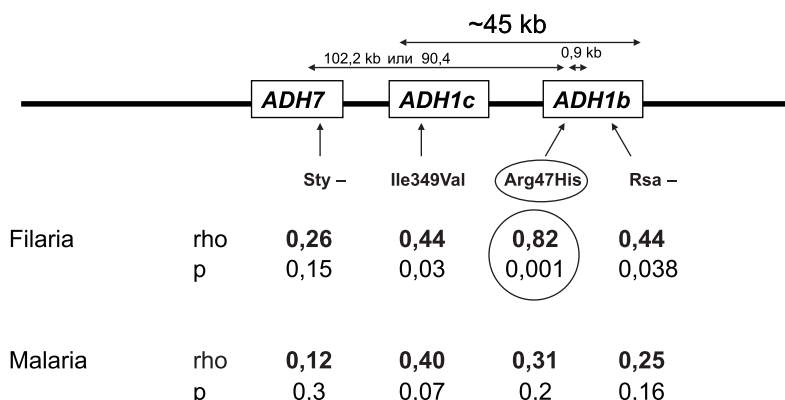


Рис. 7. Коэффициенты корреляции частот аллелей локуса ADH с наличием филяриозов в регионе проживания.

Для сравнения приведены аналогичные коэффициенты для малярии. Источник данных для частот аллелей – база ALFRED.

ИССЛЕДОВАНИЯ КОРРЕЛЯЦИЙ ЧАСТОТ АЛЛЕЛЕЙ В ПОПУЛЯЦИЯХ С УСЛОВИЯМИ СРЕДЫ

Сопоставление частот аллелей из генетических баз данных с кодированными описаниями среды из «Этнографического атласа» Мердока было предложено нами в 2002 г. (Borinskaya, 2002, 2003; Borinskaya, Korotaev, 2002; Yankovsky *et al.*, 2002; Боринская, Коротаев, 2003). Реализация такого подхода применена

в масштабах полногеномного анализа в исследованиях группы Анны ди Риенцо (Hancock *et al.*, 2010a, b). Эти исследования были основаны на предсказаниях модели отбора количественных признаков в том отношении, что такой отбор будет приводить к небольшим сдвигам в частоте аллелей по многим локусам до тех пор, пока популяция не достигнет своего оптимума частот аллелей по отношению к давлению данного фактора отбора. Примененный подход позволил выявить SNP, которые демонстрируют согласо-

ванные сдвиги частот аллелей в популяциях в соответствии с условиями среды их обитания. Внутригенные (особенно несинонимичные) SNP, а также **SNP в регуляторных участках** генов оказались избыточно представлены в подвыборке аллелей, проявляющих корреляцию с определенными условиями среды обитания популяции.

Наиболее четкие сигналы выявлены при адаптации к особенностям питания и климатическим условиям. Так, в популяциях, где основными продуктами питания являются корни и клубни, выявлены сигналы адаптации в генах, вовлеченных в метаболизм крахмала и сахарозы. Еще одна группа аллелей, адаптированных к диете, выявлена в популяциях, основой питания которых являются зерновые. Среди этих аллелей выявлены и те, которые имеют отношение к диабету второго типа. Выявлены аллели адаптации к холоду, влияющие на терморегуляцию организма человека (Hancock *et al.*, 2010a, b).

Аналогичный подход использован в широкомасштабном исследовании генов, вовлеченных в адаптацию к эндемичным инфекциям (Fumagalli *et al.*, 2011). Авторы использовали интегрированный индекс нагрузки патогенами, но не анализировали отдельные инфекции. При сопоставлении локальных адаптаций к особенностям питания, климату и нагрузке патогенами выяснилось, что именно патогены являются наиболее сильными факторами адаптации. Сигналы адаптации, ассоциированные с нагрузкой патогенами, выявлены в более чем 100 генах. Эта фракция генов обогащена генами, ассоциированными с аутоиммунными заболеваниями, такими как целиакия, диабет типа 1, рассеянный склероз. На основе этого авторы предполагают, что распространенность в популяциях человека аллелей предрасположенности к аутоиммунным заболеваниям может быть следствием отбора на устойчивость к инфекциям в прошлом (Fumagalli *et al.*, 2011).

Вклад конкретных генетических факторов в адаптацию к условиям среды зависит от числа генов, которые влияют на формирование селективируемого признака, степени вклада конкретного аллеля в адаптивность, а также от интенсивности и длительности действия фактора отбора. В популяциях человека частоты аллелей многих генов, важных для здоровья, на-

ходятся в движении под воздействием факторов среды, значительная часть которых возникла в результате развития производящего хозяйства и преобразования природной среды.

Пока в поиске причин заболеваний человека доминирует молекулярно-генетический подход. Популяционно-генетический подход в сочетании с биоинформатическим анализом, использованием медицинских и этнографических электронных баз данных, разработкой новых статистических подходов к выявлению сигналов адаптации станет очень важным дополнением для понимания генетических и средовых факторов, влияющих на здоровье человека. Эти подходы позволяют выявить функциональные **SNP, вовлеченные в адаптацию** к условиям среды, даже для тех генов и иных участков генома, функции которых не известны. Эти подходы могут быть усилены анализом генных сетей, различные звенья которых могут влиять на один и тот же фенотипический признак, методы анализа которых разработаны при исследовании заболеваний (см. например, Подколотная и др., 2010).

Генетика помогает подтвердить или опровергнуть гипотезы, сформулированные в пределах других наук, или выдвинуть новые гипотезы для их проверки этими науками. Поэтому генетика начинает выступать как «мост» между науками.

БЛАГОДАРНОСТИ

Работа выполнена в рамках Программы фундаментальных исследований Президиума РАН «Живая природа: современное состояние и проблемы развития», подпрограммы «Динамика и сохранение генофондов» и проекта РФФИ 13-04-02095.

ЛИТЕРАТУРА

- Боринская С.А., Коротаев А.В. Количественный подход к изучению ген-культурных взаимодействий // Антропология на пороге III тысячелетия. М.: Старый Сад, 2003. Т. 1. С. 503–517.
- Гончарова И.А., Фрейдин М.Б., Рудко А.А. и др. Геномные основы подверженности к инфекционным заболеваниям // Информ. вестн. ВОГиС. 2006. Т. 10. № 3. С. 540–552.
- Подколотная О.А., Яркова Е.Э., Деменков П.С. и др. Использование компьютерной системы ANDCell

- для реконструкции и анализа ассоциативных сетей потенциальных механизмов взаимосвязи миопии и глаукомы // *Информ. вестн. ВОГиС*. 2010. Т. 14. № 1. С. 106–115.
- Allison A.C. The distribution of the sickle-cell trait in East Africa and elsewhere, and its apparent relationship to the incidence of subtertian malaria // *Trans. Roy. Soc. Trop. Med. Hyg.* 1954. V. 48. No. 4. P. 312–318.
- Bersaglieri T., Sabeti P.C., Patterson N. *et al.* Genetic signatures of strong recent positive selection at the lactase gene // *Am. J. Hum. Genet.* 2004. V. 74. No. 6. P. 1111–1120.
- Borinskaya S., Korotaev A. Correlations between Genetic and Cultural Traits in Populations of Humans *Bioinformatics of Genome Regulation and Structure / Ed. N.A. Kolchanov et al.* Novosibirsk: BGRS, 2002. V. 4.
- Borinskaya S. Dopamine receptor gene allele and sociocultural characteristics: cross-cultural test // *Abstr. of 31st Annual Meeting Society for Cross-Cultural Research*. Santa Fe, 2002.
- Borinskaya S. Gene-culture relationships: cross-cultural tests of population genetic traits // *32nd Annual Meeting of Soc. for Cross-Cultural Research*. Charleston, SC, USA. Feb 19–24 2003. P. 13.
- Borinskaya S., Kal'ina N., Marusin A. *et al.* Distribution of alcohol dehydrogenase ADH1B*48His allele in Eurasia // *Am. J. Hum. Genet.* 2009. V. 84. No. 1. P. 89–92.
- Carson P.E., Flanagan C.L., Ickes C.E., Alving A.S. Enzymatic deficiency in primaquine-sensitive erythrocytes // *Science*. 1956. V. 124. P. 484–485.
- Cavalli-Sforza L.L., Menozzi P., Piazza A. *History and Geography of Human Genes*, Princeton. N.J.: Princeton Univ. Press, 1994.
- Coop G., Pickrell J.K., Novembre J. *et al.* The role of geography in human adaptation // *PLoS Genet.* 2009. V. 5. No. 6. e1000500.
- Coop G., Witonsky D., Di Rienzo A., Pritchard J.K. Using environmental correlations to identify loci underlying local adaptation // *Genetics*. 2010. V. 185. P. 1411–1423.
- Crabb D.W., Edenberg H.J., Bosron W.F., Li T.K. Genotypes for aldehyde dehydrogenase deficiency and alcohol sensitivity. The inactive ALDH2(2) allele is dominant // *J. Clin. Invest.* 1989. V. 83. P. 314–316.
- Excoffier L., Ray N. Surfing during population expansions promotes genetic revolutions and structuration // *Trends Ecol. Evol.* 2008. V. 23. No. 7. P. 347–351.
- Fumagalli M., Sironi M., Pozzoli U. *et al.* Signatures of environmental genetic adaptation pinpoint pathogens as the main selective pressure through human evolution // *PLoS Genet.* 2011. V. 7. No. 11. e1002355.
- Gillespie J. *The causes of molecular evolution*. N.Y., 1991.
- Guglielmino C.R., Viganotti C., Hewlett B., Cavalli-Sforza L.L. Cultural variation in Africa: role of mechanisms of transmission and adaptation // *Proc. Natl Acad. Sci. USA*. 1995. V. 92. No. 16. P. 7585–7589.
- Haldane J.B.S. Disease and evolution // *Ricercha. Sci.* 1949. V. 19. (Suppl.). P. 68–76.
- Hancock A.M., Alkorta-Aranburu G., Witonsky D.B., Di Rienzo A. Adaptations to new environments in humans: the role of subtle allele frequency shifts // *Philos. Trans. Roy. Soc. Lond. B. Biol. Sci.* 2010a.
- Hancock A.M., Witonsky D.B., Alkorta-Aranburu G. *et al.* Adaptations to climate-mediated selective pressures in humans // *PLoS Genet.* 2011. V. 7. No. 4. e1001375.
- Hancock A.M., Witonsky D.B., Ehler E. *et al.* Colloquium paper: human adaptations to diet, subsistence, and ecoregion are due to subtle shifts in allele frequency // *Proc. Natl Acad. Sci. USA*. 2010b.
- Hancock A.M., Witonsky D.B., Gordon A.S. *et al.* Adaptations to climate in candidate genes for common metabolic disorders // *PLoS Genetics*. 2008. V. 4. No. 2. P. 32.
- Ikuta T., Szeto S., Yoshida A. Three human alcohol dehydrogenase subunits: cDNA structure and molecular and evolutionary divergence // *Proc. Natl Acad. Sci. USA*. 1986. V. 83. P. 634–638.
- Jornvall H., Hempel J., Vallee B.L. *et al.* Human liver alcohol dehydrogenase: amino acid substitution in the beta-2 beta-2 Oriental isozyme explains functional properties, establishes an active site structure, and parallels mutational exchanges in the yeast enzyme // *Proc. Natl Acad. Sci. USA*. 1984. V. 81. P. 3024–3028.
- Kim D.J., Choi I.G., Park B.L. *et al.* Major genetic components underlying alcoholism in Korean population // *Hum. Mol. Genet.* 2008. V. 17. P. 854–858.
- Kong A., Frigge M.L., Masson G. *et al.* Rate of de novo mutations and the importance of father's age to disease risk // *Nature*. 2012. V. 488. No. 7412. P. 471–475.
- Korotayev A., Kazankov A., Borinskaya S. *et al.* *Ethnographic Atlas XXX: peoples of Siberia // Ethnology*. 2004. V. 43. No. 1. P. 83–92.
- Li H., Borinskaya S., Yoshimura K. *et al.* Refined geographic distribution of the oriental ALDH2*504Lys (nee 487Lys) variant // *Ann. Hum. Genet.* 2009. V. 73. No. 3. P. 335–345.
- Liberles D. SNPing variation from genomes (Meeting report) // *Genome Biol.* 2001. V. 3. No. 1. 4001. 1–3.
- Matsuo Y., Yokoyama R., Yokoyama S. The genes for human alcohol dehydrogenases beta-1 and beta-2 differ by only one nucleotide // *Europ. J. Biochem.* 1989. V. 183. P. 317–320.
- Murdock G.P. *Ethnographic Atlas: A Summary*. Pittsburgh: The Univ. of Pittsburgh Press, 1967.
- Neel J.V. The thrifty genotype in 1998 // *Nutr. Rev.* 1999. V. 57. P. S2–S7.
- Oota H., Dunn C.W., Speed W.C. *et al.* Conservative evolution in duplicated genes of the primate Class I ADH cluster // *Gene*. 2007. V. 1. No. 5. P. 64–76.
- Oota H., Pakstis A.J., Bonne-Tamir B. *et al.* The evolution and population genetics of the ALDH2 locus: random genetic drift, selection, and low levels of recombination // *Ann. Hum. Genet.* 2004. V. 68. P. 93–109.
- Osier M., Pakstis A.J., Kidd J.R. *et al.* Linkage disequilibrium at the ADH2 and ADH3 loci and risk of alcoholism // *Am. J. Hum. Genet.* 1999. V. 64. P. 1147–1157.
- Rodenwaldt E., Bader R.E. *World-atlas of epidemic diseases*. Hamburg, Germany: Falk-Verlag, 1952–1961.
- Roth E.F., Raventos-Suarez C., Rinaldi A., Nagel R. L. Glucose-6-phosphate dehydrogenase deficiency inhibits in vitro growth of *Plasmodium falciparum* // *Proc. Natl Acad. Sci. USA*. 1983. V. 80. P. 298–299.
- Ruwando C., Khea S.C., Snow R.W. *et al.* Natural selection

- of hemi- and heterozygotes for G6PD deficiency in Africa by resistance to severe malaria // *Nature*. 1995. V. 376. P. 246–249.
- Sabeti P., Schaffner S., Fry B. *et al.* Positive natural selection in the human lineage // *Science*. 2006. V. 312. P. 1614–1620.
- Simmons J.S., Whayne T.F., Anderson G.W. Horack H.M. *Global epidemiology*. Philadelphia: J.B. Lippincott, 1944.
- Slatkin M., Excoffier L. Serial founder effects during range expansion: a spatial analog of genetic drift // *Genetics*. 2012. V. 191. No. 1. P. 171–181.
- Taylor M.J., Hoerauf A., Bockarie M. Lymphatic filariasis and onchocerciasis // *Lancet*. 2010. V. 376. No. 9747. P. 1175–1185.
- Thornhill R., Fincher C.L., Murray D.R., Schaller M. Zoonotic and non-zoonotic diseases in relation to human personality and societal values: support for the parasite-stress model // *Evol. Psychol.* 2010. V. 8. No. 2. P. 151–169.
- Tolstrup J.S., Nordestgaard B.G., Rasmussen S. *et al.* Alcoholism and alcohol drinking habits predicted from alcohol dehydrogenase genes // *Pharmacogenomics J.* 2008. V. 8. P. 220–227.
- Venter J.C., Adams M.D., Myers E.W. *et al.* The sequence of the human genome // *Science*. 2001. V. 291. No. 5507. P. 1304–1351.
- Venter J.C. Genome-sequencing anniversary. The human genome at 10: successes and challenges // *Science*. 2011. V. 331. No. 6017. P. 546–547.
- Venter J.C. Multiple personal genomes await // *Nature*. 2010. V. 464. No. 7289. P. 676–677.
- Wall T.L., Horn S.M., Johnson M.L. *et al.* Hangover symptoms in Asian Americans with variations in the aldehyde dehydrogenase (ALDH2) gene // *J. Stud. Alcohol*. 2000. V. 61. P. 13–17.
- Yankovsky N., Borinskaya S., Rogaev E., Korotaev A. *et al.* Change of population genetic traits under different natural and cultural environment: new dimension in understanding of human diseases // *Abstr. of Human Origin and Disease Meeting*. Cold Spring Harbor, 2002.