

УДК 575.174:599.9

ЭТНОГЕНОМИКА

© 2013 г. Э.К. Хуснутдинова

Федеральное государственное бюджетное учреждение науки Институт биохимии и генетики
Уфимского научного центра РАН, Уфа, Россия, e-mail: elzakh@mail.ru;
ФГБОУ ВПО Башкирский государственный университет, Уфа, Россия

Поступила в редакцию 26 апреля 2013 г. Принята к публикации 1 ноября 2013 г.

ВВЕДЕНИЕ

Этногеномика – это раздел популяционной генетики, изучающий особенности геномного полиморфизма и геномного разнообразия отдельных популяций, этносов и реконструкция на этой основе их генетической истории.

Эволюция популяций человека, их происхождение, родство, историческое развитие всегда были в центре внимания многих наук. Для решения этих проблем необходимо исследовать множество полиморфных признаков в большом числе популяций и этно-территориальных групп. В качестве таких признаков в распоряжении исследователей в течение долгого времени находились полиморфные белки, которые использовались как генетические маркеры в популяционных исследованиях. С помощью таких маркеров получено довольно много интересных сведений о популяциях различных регионов мира. Подлинный переворот в популяционных исследованиях произошел при появлении нового инструмента в виде полиморфных маркеров ДНК.

Огромное множество полиморфных ДНК-маркеров, выявленных при расшифровке генома человека, стало мощным инструментом для описания на новом уровне генетических особенностей народов; восстановления истории их формирования, а также становления человека как биологического вида в целом. На основе развития этих исследований в рамках геномики возник новый раздел науки – этногеномика. Геном человека, состоящий примерно из 3 млрд нуклеотидных пар, расшифрован почти полностью. Однако завершение гигантского по

замыслу и грандиозного по реализации международного научного проекта по расшифровке структуры генома человека отнюдь не означает, что процесс познания генома завершен. Скорее, это только начало. Уже сейчас очевидно, что не существует какого-то «усредненного» генома человека: каждый геном, как и каждый человек, сугубо индивидуален. Эта индивидуальность генома проявляется на уровне не только отдельной личности, но и этнических групп, отдельных сообществ и рас. Различия между двумя людьми на уровне ДНК составляют в среднем 1 нуклеотид на 1000. Именно эти различия обуславливают наследственные индивидуальные особенности каждого человека. Заметим, что различия между ДНК человека и шимпанзе – его ближайшего сородича в животном мире – на порядок больше.

ОСНОВНЫЕ СИСТЕМЫ ГЕНЕТИЧЕСКИХ МАРКЕРОВ ДЛЯ ЭВОЛЮЦИОННЫХ ИССЛЕДОВАНИЙ

В исследованиях по эволюции популяций человека используются различные системы генетических маркеров, отличающиеся локализацией в геноме, уровнем вариабельности и характером мутирования.

Преимущества изучения генетического полиморфизма на уровне ДНК по сравнению с белковым уровнем колоссальны. Наиболее важным преимуществом ДНК является наличие в ней полиморфизма разного типа, каждый из которых имеет свои особенности. Это однонуклеотидные замены (SNPs), инсерционно-

Этнос – исторически сложившаяся общность людей с общими культурой, языком и самосознанием.

Генетический дрейф – изменение частот гена в популяции в результате случайных колебаний в выборке гамет, образующих следующее поколение популяции.

Эффект основателя – один из видов генетического дрейфа, обусловленный возникновением новой популяции от относительно небольшого числа основателей, из-за чего генетическая изменчивость в новой популяции оказывается суженной.

Эффект «горлышка бутылки» – сокращение генетического разнообразия популяции вследствие прохождения периода, во время которого по различным причинам происходит критическое уменьшение ее численности, в дальнейшем восстановленное. Сокращение генетического разнообразия приводит к изменению относительных и абсолютных частот аллелей генов.

Гаплотип – совокупность аллелей на локусах одной хромосомы, обычно наследуемых вместе.

Гаплогруппа – группа схожих гаплотипов, имеющих общего предка, у которого в обоих гаплотипах имела место одна и та же мутация – однонуклеотидный полиморфизм.

Время коалесценции/дивергенции – время существования наиболее близкого общего предка для группы последовательностей.

Alu-повторы – семейство коротких повторяющихся элементов с числом копий порядка 500 тыс. на гаплоидный геном.

Микросателлиты – фрагменты ДНК с большим количеством tandemно повторяющихся «мотивов», или «повторов» – коротких последовательностей из нескольких пар нуклеотидов.

делеционный полиморфизм, мини- и микросателлиты (short tandem repeats, STR). Спектр селективных свойств ДНК-маркеров намного шире, чем классических маркеров белкового полиморфизма: он варьирует от истинно селективно нейтральных генов до генов, подверженных самому мощному отбору, как стабилизирующему, так и дифференцирующему. Кроме того, по ДНК-маркерам можно исследовать гаплотипы – сочетания аллелей тесно сцепленных полиморфных локусов. Такие хромосомные участ-

ки весьма невелики, поэтому обнаруживают очень малую степень рекомбинаций и ведут себя как единые блоки, мало меняющиеся во времени и поэтому имеющие довольно древнее происхождение. Таким образом, размер сохранившегося неизменным гаплотипа может служить мерой времени, которое прошло от какого-то момента в прошлом. В общем случае суть метода состоит в поиске неравновесного сцепления между собой локусов вследствие эффекта основателя. Анализ частоты и возраста появления в популяции хромосомы-основателя позволяет проследить ее историю, а вместе с ней и популяционные события, сопутствующие ее распространению. Очевидно, что подобные данные представляют несомненный интерес и огромную научную значимость для изучения истории современных народов; характеристики генофондов и оценки основных направлений эволюции всего человечества.

Для исследований геномов людей используют разные системы ДНК-маркеров: маркеры, расположенные на парных хромосомах (аутосомные), на митохондриальной ДНК и непарной Y-хромосоме.

Изучение **аутосомных маркеров**, которые наследуются по обеим (женской и мужской) линиям и в которых представлена подавляющая часть генома человека, чрезвычайно важно, ибо это дает исследователям маркеры для изучения сочетанной изменчивости, одновременно привносимой и с отцовской, и с материнской сторон. Данные маркеры ДНК характеризуют сообщества в целом, не выделяя генетического вклада каждого из полов. Использование определенных типов полиморфизма ДНК позволяет оценить те или иные временные события, происходившие в истории данной популяции. В настоящее время для изучения генофонда и генетической истории популяций человека наиболее широко применяют однонуклеотидные замены (SNPs), микросателлиты (STR) и Alu-повторы.

Особую роль играют маркеры митохондриальной ДНК (мтДНК) и ДНК Y-хромосомы, поскольку они позволяют проследить генетическую историю человечества отдельно по женской и мужской линиям. Это дает новые, несуществовавшие раньше возможности в этногенетических исследованиях: проследить

и сопоставить историю женской и мужской части популяции и оценить их вклад в популяционный генофонд. Передаваясь из поколения в поколение только по одной из родительских линий и не участвуя в рекомбинации, данные маркеры позволяют, по крайней мере теоретически, реконструировать генетические события от наиболее популярных предков современного человека – «Y-хромосомного Адама» и «митохондриальной Евы» до современных популяций. Полиморфизм этих маркеров определяется факторами микроэволюции (миграция, отбор, мутации), однако характер их вариабельности по-разному отражает действие и результат этих процессов.

Митохондриальная ДНК (мтДНК) представляет собой небольшую молекулу кольцевой формы размером 16569 п.н., содержащуюся в митохондриях эукариотических клеток (рис. 1). Число копий мтДНК в соматической клетке достигает 10 тыс. Митохондриальная ДНК построена по принципу максимальной экономии: она практически лишена интронов, а кодирующие последовательности почти соприкасаются или даже слабо перекрываются. МтДНК характеризуется рядом особенностей в сравнении с ядерной: материнским характером наследования, отсутствием рекомбинации (обмен участками гомологичных хромосом в процессе мейоза) и относительно высокой скоростью накопления мутаций. В митохондриальном геноме мутации возникают в десятки раз быстрее, чем в ядерном. Наиболее вариабельной частью митохондриального генома человека является главная не кодирующая область или контрольный регион, имеющий протяженность 1122 п.н., в котором расположены гипервариабельные сегменты 1 (ГВС1) и

2 (ГВС2) мтДНК. По сравнению с аутосомными локусами мтДНК имеет в 4 раза меньшую эффективную численность в популяции, что определяет большую подверженность случайным флуктуациям и воздействию генетического дрейфа. Соответственно, мтДНК позволяет уловить эффекты основателя или «бутылочного горлышка» в популяции, неразличимые на уровне аутосомных локусов. Передача мтДНК по материнской линии без рекомбинации (от матери ко всем ее потомкам и далее – только дочерями) определяет то, что мутации, возникшие в мтДНК однажды, сохраняются и передаются неизменными в ряду поколений как единый локус, представленный множеством аллелей – гаплотипов. Определенные группы гаплотипов соответствуют определенным группам сцепления между конкретными мутациями (т. е. гаплогруппам).

Изучение изменчивости митохондриального генома в настоящее время проводится с использованием комбинированного подхода – выявления однонуклеотидных замен мтДНК классическим методом анализа полиморфизма длины рестрикционных фрагментов ДНК (ПДРФ) в сочетании с изучением нуклеотидной последовательности ГВС1 и ГВС2 главной некодирующей области мтДНК. Такой подход выявил важную особенность мтДНК, имеющую значение для изучения молекулярной эволюции. В частности, установлено, что определенным группам рестрикционных типов мтДНК, ключевые мутации которых расположены в различных участках молекулы, соответствуют вполне определенные типы нуклеотидных последовательностей гипервариабельного участка мтДНК. Эта особенность молекул мтДНК имеет важное значение для изучения молекулярной

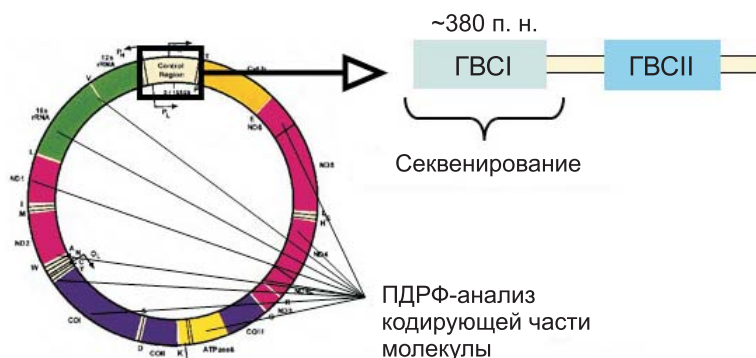


Рис. 1. Карта мтДНК человека (<http://www.mitomap.org>).

эволюции, поскольку разнообразие митохондриального генофонда хранит в себе множество их комбинаций, по которым можно проследить изменения молекул мтДНК во времени и классифицировать молекулярные изменения в приложении к эволюции популяций. Последовательное накопление мутаций позволяет отслеживать историю индивидов к так называемой точке коалесценции, т. е. некоему общему предку. Зная скорость мутирования, можно рассчитать время коалесценции – время, когда у индивидов был общий предок. По спектру гаплотипов мтДНК в популяции можно проследить эволюционные взаимосвязи между древними и вновь возникающими гаплотипами и более того, реконструировать генетическую историю женского генофонда популяции. В отличие от локусов ядерной ДНК, где эволюционные изменения прослеживаются главным образом по вариативности частот различных аллелей в популяциях, митохондриальная ДНК дает возможность восстановить действительную филогению, т. е. последовательность возникновения носителей различных мт-гаплотипов в эволюционном ряду. Филогенетические взаимосвязи между гаплотипами изображаются в виде дерева, имеющего корень, ствол, более древние большие ветви и мелкие веточки, возникшие в относительно недавнем прошлом. Корень гаплотипов мтДНК человека был выявлен исходя из сравнения с последовательностью мтДНК шимпанзе. Для большего удобства отсчет ведется не от предполагаемой предковой популяции, а от последовательности митохондриального генома человека, так называемой кембриджской референсной последовательности (rCRS – Cambridge reference sequence). Дальнейшее построение сети зависит от числа гаплотипов. Реконструкция эволюционного древа с учетом времени коалесценции линий и в совокупности с географическими, палеонтологическими и археологическими данными составляет основу филогеографического подхода, который в последнее время получил широкое распространение в исследованиях происхождения человека, рас и отдельных этносов.

Основы универсальной номенклатуры мтДНК были заложены в 1993 г. в работах А. Торрони с соавторами, посвященных изучению происхождения коренного населения

Америки. Тогда же было предложено использовать буквенную номенклатуру для построения филогенетических деревьев. Сегодня буквенная номенклатура, основанная на использовании как ПДРФ-анализа, так и прямого поиска мутаций в ГВС1 и ГВС2 путем секвенирования, является общепринятой и широко используется генетиками. Мутации, определяющие гаплогруппы, обозначаются буквенным и цифровым индексами. При этом основные ветви с обнаружением новых мутаций в митохондриальном геноме разделяются на более мелкие, что позволяет выявлять специфические особенности популяций человека в тех или иных регионах мира, отличающие их друг от друга по женской линии. Субклады, располагающиеся внутри мажорных гаплогрупп, обозначаются цифрами. Надежность классификации зависит от количества имеющейся информации об изменчивости мтДНК, и, естественно, в идеале необходима информация о нуклеотидных последовательностях целых митохондриальных геномов, относящихся к разным филогенетическим группам. Накопленные к настоящему времени данные в области популяционной митохондриальной геномики (секвенировано уже около 17 тыс. митохондриальных геномов) позволили существенно улучшить представления о топологии филогенетического дерева мтДНК человека и детализировать классификацию мтДНК с учетом данных о региональных особенностях эволюции митохондриального генома. К настоящему времени разработана детальная классификация гаплогрупп мтДНК (van Oven, Kayser, 2009) (рис. 2).

Согласно современным данным, филогенетическое дерево мтДНК человека представлено двумя крупными дочерними ветвями – L0 и L1'2'3'4'5'6 (L1'5). L1'5 распространена более широко, в нее входят почти все обнаруженные линии мтДНК. Все разнообразие неафриканских гаплогрупп, за исключением тех, что были привнесены миграциями в течение последних нескольких тысяч лет, в том числе с интенсивной работорговлей, сводится к вариациям внутри базальной гаплогруппы L3, разделившейся впоследствии на две ветви: гаплогруппы M и N. От последней в свою очередь почти сразу же отделилась гаплогруппа R. Эти события произошли, предположительно, около 60000–65000 лет

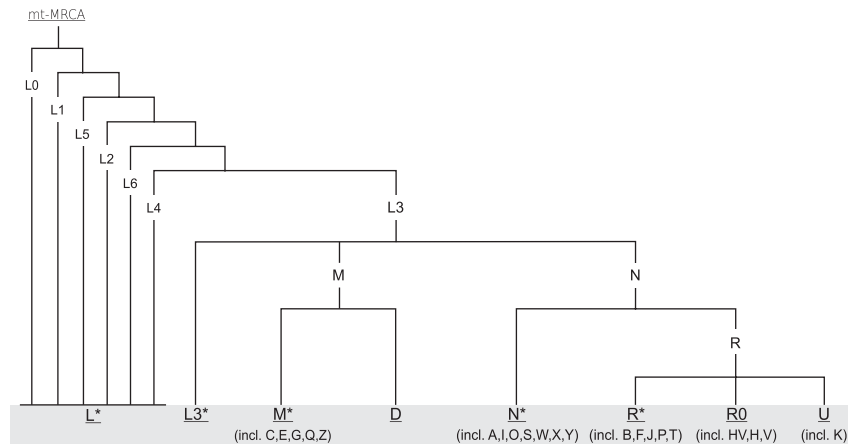


Рис. 2. Филогенетическое дерево мтДНК человека.

Показан порядок появления гаплогрупп мтДНК в процессе эволюции (приведено согласно (<http://www.phylotree.org>) с модификацией).

назад в районе между Восточной Африкой и Персидским заливом. Важно отметить, что в распределении гаплогрупп мтДНК наблюдается выраженная региональная специфичность (рис. 3). Ветви R0, U и JT, а также минорные ветви гаплогруппы N (N1, N2 и X) объясняют большую часть разнообразия мтДНК в Европе, в то время как в восточно-евразийский пул мтДНК-гаплогруппа N и специфичная для

Азии макрогаплогруппа M приносят примерно одинаковый вклад. Данная классификация постоянно обновляется, что позволяет рассматривать различия между популяциями с высокой степенью филогенетического разрешения.

Y-хромосома – самая маленькая в геноме человека – имеет размер около 60 млн п.н., половина приходится на эухроматиновые домены, а другая – на гетерохроматиновый блок,

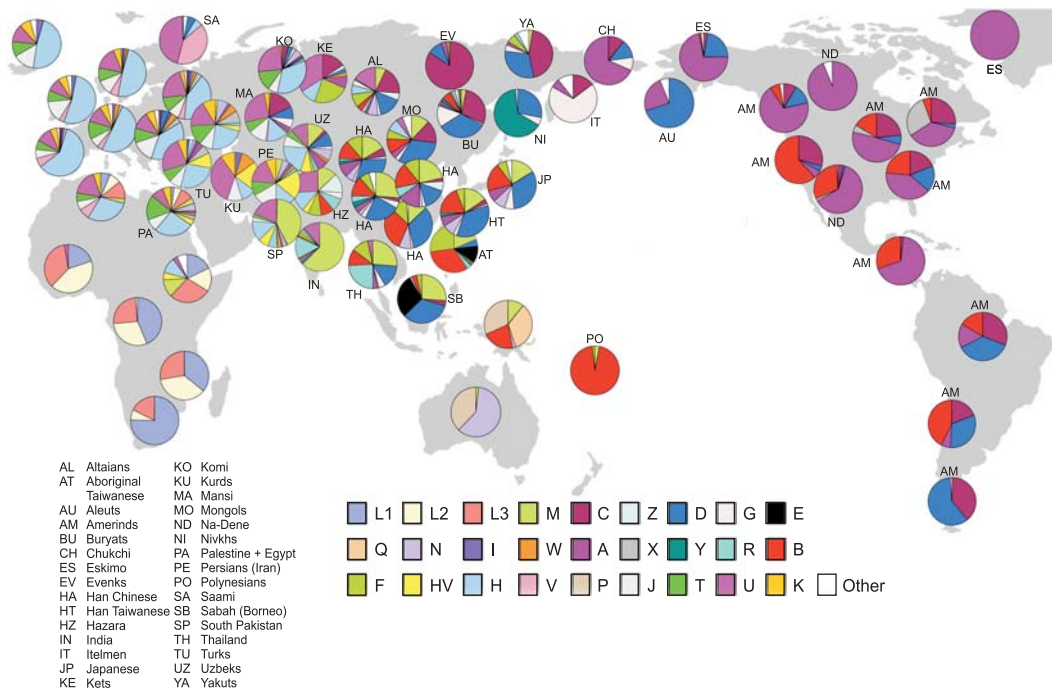


Рис. 3. Карта распространения гаплогрупп мтДНК в популяциях мира (<http://www.scs.illinois.edu/~mcdonald/WorldHaplogroupsMaps.pdf>).

который может сильно варьировать по размеру у разных индивидов. Уникальным отличием Y-хромосомы является то, что она определяет пол, специфична для мужчин и передается от отца к сыновьям без рекомбинации большей ее части. Нерекombинирующая часть Y-хромосомы (NRY) не подвергается обмену участками с X-хромосомой в процессе мейоза и составляет 95 % от общей длины. Мутации, возникшие в Y-хромосоме, сохраняются и передаются единым блоком в интактном виде от поколения к поколению. Нерекombинирующая часть и последовательность накопления в ней мутаций поддаются расшифровке, подобно записи исторических событий на древнем пергаменте, описывающей происхождение и эволюцию отцовских линий. В отличие от мтДНК Y-хромосома имеет гораздо большие размеры, потенциально является более полиморфной и, как следствие, более информативной системой. Генетические маркеры в нерекombинирующей части Y-хромосомы можно разделить на две основные категории – диаллельные и мультиаллельные, которые различаются по темпам эволюции и уровню генетического разнообразия и позволяют выявить дифференциацию популяций и филогеографию линий Y-хромосомы на различных уровнях иерархии. К первой категории относятся SNP (точечные мутации, замены оснований) и более редкие инсерции и делеции. Темп мутирования таких локусов низок: 5×10^{-7} на сайт на поколение. Диаллельные маркеры используются для выделения гаплогрупп. Вторая категория маркеров – мультиаллельные полиморфизмы – включает микросателлиты, или короткие tandemные повторы (STR), эволюционирующие на несколько порядков быстрее. Темп их мутирования гораздо выше: для Y-сцепленных STR он составляет примерно 2×10^{-3} на локус на поколение. Сейчас на Y-хромосоме описано несколько десятков микросателлитов. Мультиаллельные маркеры удобно использовать для анализа разнообразия гаплотипов внутри гаплогрупп и для более детальной реконструкции филогении и происхождения линий.

Помимо вышеперечисленных особенностей еще одной является то, что численность Y-хромосом в популяции по сравнению с аутосомами составляет 1 : 4 и по сравнению с X-хромосомами – 1 : 3. Поэтому Y-хромосомы так же, как

и мтДНК, более подвержены эффекту генетического дрейфа, сильно меняющего частоты различных гаплотипов в популяциях с малой эффективной численностью. Как следствие, степень генетической подразделенности популяций по Y-хромосоме намного выше, чем по аутосомным локусам.

На географическое распределение вариантов Y-хромосомы помимо генетических факторов (дрейф генов, эффект основателя в популяциях) сильное влияние оказывают демографические и социальные факторы. Примерно 70 % современных обществ характеризуются патрилокальностью. Это означает, что мужчины живут ближе к месту их рождения, чем женщины: при заключении брака, как правило, женщина переезжает на местожительство мужа, а не наоборот. Со временем этот фактор увеличивает различия в распределении вариантов Y-хромосом и может приводить к градиентному распределению линий в стабильных популяциях большого размера. Следствием патрилокальности объяснялись данные распределения типов Y-хромосомы в Европе и на островах Юго-Восточной Азии. Влияние социальных факторов может иметь прямо противоположный эффект. Например, поток генов при экспансии европейцев на территорию Америки или Океании за последние 500 лет происходил в основном за счет мужчин и сильно повлиял на спектр вариантов Y-хромосомы, но не мтДНК в популяциях Полинезии, Гренландии и Южной Америки.

В 2002 г. Консорциумом по Y-хромосоме была предложена единая классификация на основе диаллельных маркеров. Дерево, предложенное Консорциумом, включало 237 полиморфных сайтов. Всего, включая обратные мутации в различных ветвях дерева, для определения 153 гаплогрупп было использовано 245 мутационных событий. Определение корня дерева было произведено исходя из сравнения NRY человека с гомологичными последовательностями близкородственных видов: шимпанзе, гориллы, орангутанга. Для удобства номенклатура была представлена с помощью прописных латинских букв, обозначающих базальные ветви, начиная с буквы А, приданной ветви, наиболее близко находящейся к корню, и далее по алфавиту до буквы R. Буквой Y была обозначена наиболее крупная ветвь, включающая все гаплогруппы

от А до R. Порядок букв отражает последовательность возникновения мутаций. Эти клады в свою очередь разветвляются на гаплогруппы, которые нумеруются цифрами и буквами. Следует заметить, что использованные до сих пор термины – «клады», «ветви», «линии», «базальные линии», «субклады» и т. д. – применимы ко всем ветвям древа, зачастую относящимся к различным иерархическим уровням. Субклады, располагающиеся внутри мажорных гаплогрупп, обозначаются цифрами, а в более мелких ветвях к цифре прибавляется строчная латинская буква. Так, внутри гаплогруппы R можно выделить субклады R1 и R2, а внутри R1 – R1a. Для дальнейшего обозначения применяется чередование цифр и строчных букв, например: R1b1b1. Консорциумом по Y-хромосоме была принята терминология, по которой термин «гаплогруппа» означает линии NRY, определенные на основе диаллельных маркеров, а «гаплотип» подразумевает линии

гаплогрупп, выявленные различными вариациями микросателлитных локусов NRY. Мутации, определяющие диаллельные гаплогруппы, обозначаются буквенным и цифровым индексом. В 2003 г. принятая Консорциумом классификация была немного модифицирована, а в 2008 г. подверглась существенной переработке. При этом, однако, принципы классификации и основная номенклатура остались неизменными. Недавно опубликованное Т. Карафет (Karafet *et al.*, 2008) филогенетическое дерево Y-хромосомы человека содержит 20 основных макрокластеров, 311 гаплогрупп, определенных с использованием 600 диаллельных маркеров (рис. 4).

Полиморфизм Y-хромосомы изучен уже в нескольких сотнях популяций по всему миру, и результаты этих исследований свидетельствуют о том, что гаплогруппы Y-хромосомы имеют вполне определенное географическое распространение, причем точность географической характеристики той или иной гаплогруппы во

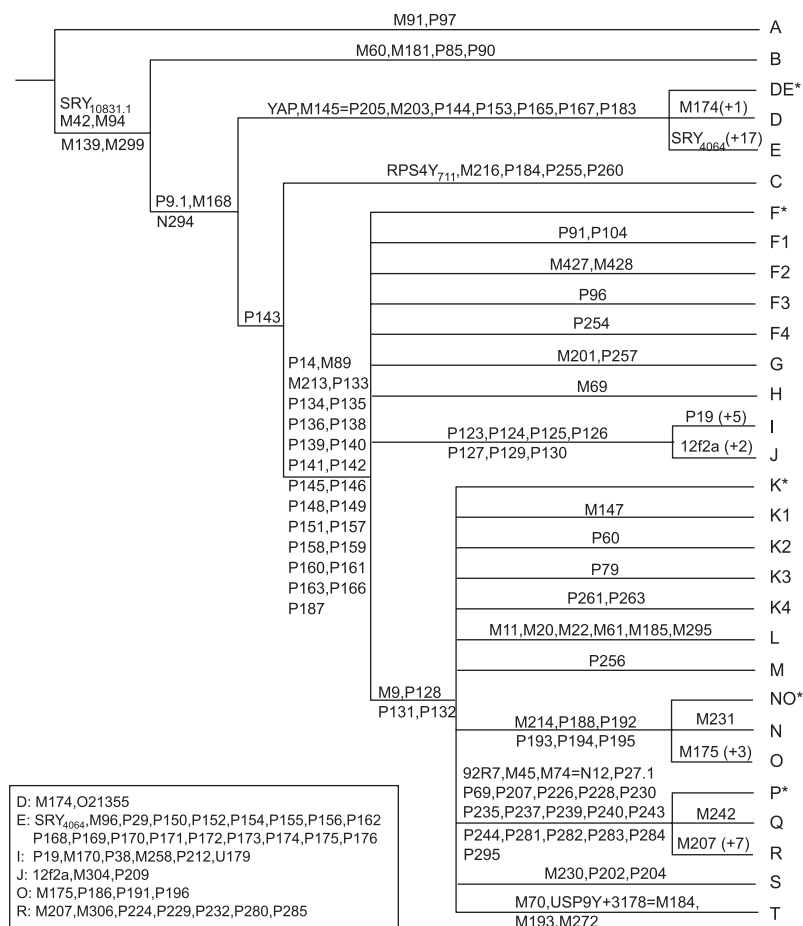


Рис. 4. Филогенетическое дерево основных гаплогрупп Y-хромосомы (Karafet *et al.*, 2008. P. 830–838).

многим зависит от использования маркеров с определенной степенью филогенетического разрешения. Однако анализ одних только частот гаплогрупп Y-хромосомы имеет существенный недостаток: высокая частота гаплогруппы может быть не следствием ее происхождения на данной территории, а результатом иных событий в истории популяции (дрейф, эффект основателя, эффект «бутылочного горлышка»). В результате оценка распределения гаплогрупп только на основе их частот может привести к некорректным выводам. Ввиду этого целесообразно использовать комплексный анализ как однонуклеотидных замен в нерекombинирующей области Y-хромосомы, так и микросателлитных локусов (STR). Они обладают достаточно высокой степенью изменчивости для выявления места и времени происхождения определенной гаплогруппы. Ожидаемо, что в том регионе или в популяции, где возникла данная гаплогруппа, ее частота и микросателлитная изменчивость будут максимальны по сравнению с другими популяциями.

Изучение полиморфизма ДНК за последние годы приобрело широкий размах и проводится во многих странах мира. Эти исследования позволили выявить значительные внутри- и межпопуляционные различия в частотах полиморфных маркеров ДНК во многих географических районах мира, что стало одной из важнейших характеристик генетической структуры человеческих сообществ. За последнее десятилетие генетиками собраны и проанализированы коллекции мтДНК и Y-хромосом представителей народов почти всего мира. По ним восстановлены последовательность и время появления мутаций в ДНК человека. В достаточно полной мере на сегодняшний день классифицированы митохондриальные и Y-хромосомные гаплотипы у населения Западной и Восточной Евразии, Африки, Австралии и Америки. Накоплен большой массив данных об изменчивости мтДНК и Y-хромосомы в различных популяциях и этнических группах в глобальном масштабе. Установлено, что в распределении гаплогрупп Y-хромосомы также наблюдается выраженная региональная специфичность, что позволяет определять соотношения генетических компонентов различного происхождения в смешанных популяциях (рис. 5).

В России работы по этногеномике – одно из самых продуктивных генетических направлений, которое активно развивается в ряде научных центров. Разработками в области этногеномики и молекулярной филогеографии в России занимается целый ряд известных научных коллективов в Москве, Новосибирске, Томске, Магадане, Уфе и Якутске. Ими получены фундаментальные данные по изменчивости и эволюции мтДНК, Y-хромосомы и аутосомных ДНК-локусов в популяциях Волго-Уральского региона, Кавказа, Центральной России, Сибири, Средней Азии. Кроме того, результаты исследований позволили охарактеризовать структуру генофонда популяций России, получить генетические портреты отдельных этносов, сопоставить генетические реконструкции с историческими данными о происхождении и миграции коренных народов России и получить однозначную генетическую оценку по некоторым спорным вопросам этногенеза (Хуснутдинова, 1999; Лимборская и др., 2002; Степанов, 2002; Балановская, Балановский, 2007; Федорова, 2008; Деренко, Малярчук, 2010; Хуснутдинова, Федорова, 2010; Кутуев, Хуснутдинова, 2011).

Таким образом, для глубокого понимания эволюции популяций человека, их происхождения и миграции, а также этногенеза различных народов необходимо использовать все три системы генетических маркеров: аутосомные локусы, маркеры мтДНК и Y-хромосомы. Различные системы маркеров значительно дополняют друг друга, особенно в тех случаях, когда вследствие стохастических процессов в популяциях и/или особенностей их формирования та или иная система маркеров не может в достаточно полной мере ответить на поставленные вопросы.

ГЕНЕТИЧЕСКАЯ ИСТОРИЯ ПОПУЛЯЦИЙ ЧЕЛОВЕКА

Первым значительным вкладом ДНК-маркеров в решение проблемы происхождения и расселения современного человека явились исследования Ребекки Канн, Марка Стоункинга и Алана Уилсона (Cann *et al.*, 1987) мтДНК представителей различных рас: африканцев, европейцев, азиатов, австралийцев и жителей Новой Гвинеи. По количеству замен нуклео-

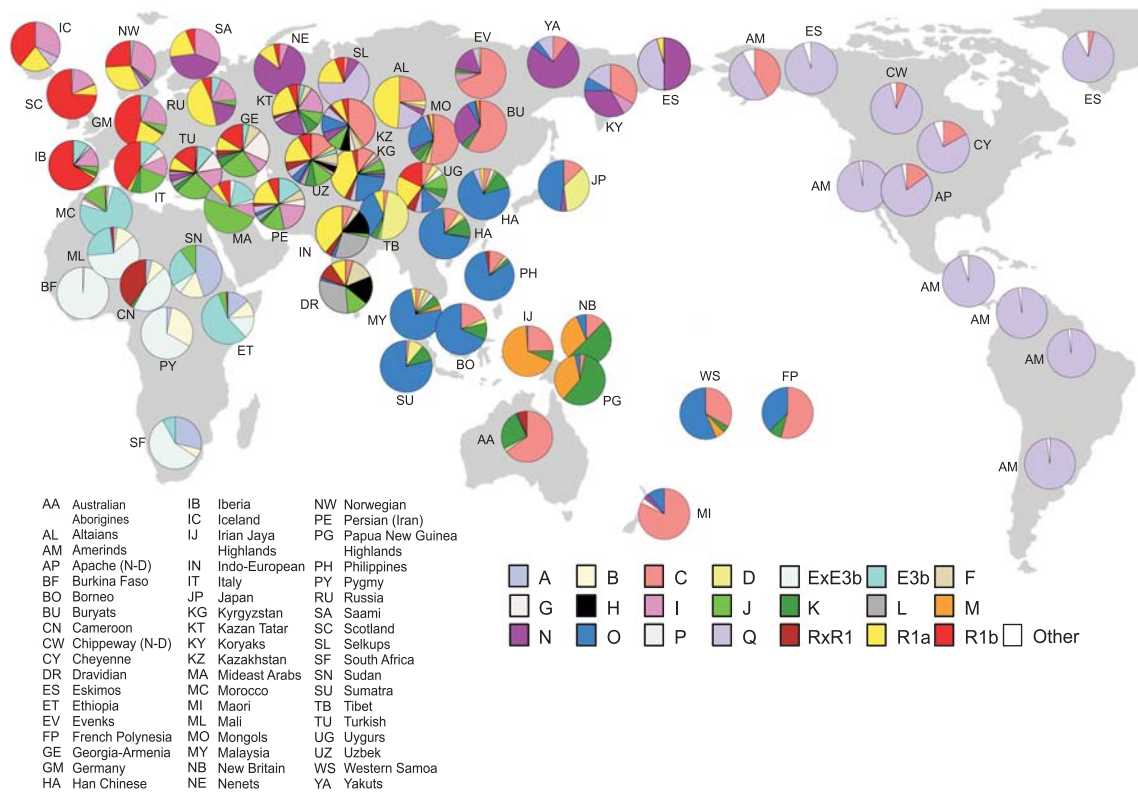


Рис. 5. Карта распространения гаплогрупп Y-хромосомы в популяциях мира (<http://www.scs.illinois.edu/~mcdonald/WorldHaplogroupsMaps.pdf>).

тидов в мтДНК определена степень родства различных групп людей и построено эволюционное дерево человечества (рис. 6). Самая ранняя точка ветвления на этом дереве отделяет от остальных людей группу африканцев, что указывает на африканское происхождение *Номо*

sapiens. Именно в Южной Африке у койсанов были найдены самые древние мутации и самое высокое разнообразие мтДНК. Митохондриальные ДНК у населения других континентов менее разнообразны, и сравнение их с мтДНК аборигенов Южной Африки показало, что они

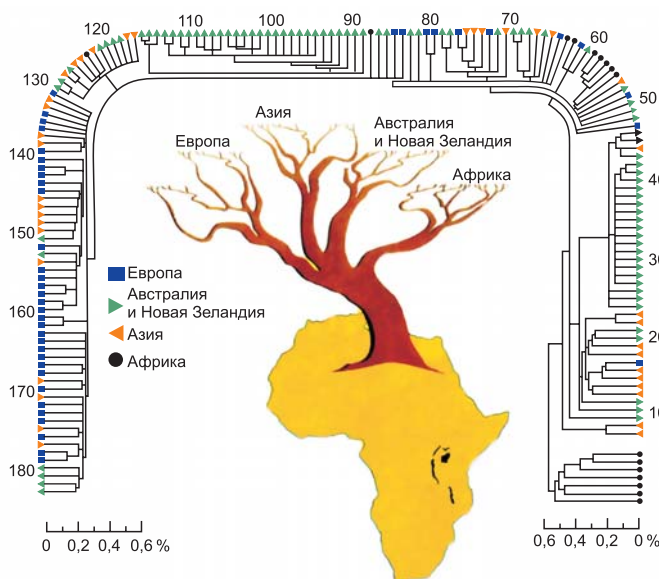


Рис. 6. Эволюционное дерево человечества, построенное по результатам исследований митохондриальной ДНК (Cann *et al.*, 1987. Р. 31–36).

возникли как мутационные изменения африканских типов после того, как человечество распространилось за пределы Африки.

Второй вывод Сапп с соавт. касался времени коалесценции (схождения к общему предку) мтДНК. Учитывая даты отделения ветви шимпанзе (5–7 млн лет назад) и приняв темп мутационной дивергенции равным 2–4 % за 1 млн лет, они вычислили продолжительность существования последней предковой мтДНК, общей для всех ныне живущих людей, примерно 185 тыс. лет. Последующие работы с использованием разных методов (секвенирование контрольного региона мтДНК и анализ полногеномной изменчивости мтДНК) подтвердили африканские корни дерева мтДНК современного человечества, хотя и остаются еще отдельные спорные моменты. По независимым оценкам нескольких групп исследователей, «митохондриальная Ева» жила в период резкого сокращения численности наших предков (до 10 тыс.), вызванного, по-видимому, изменениями климата. Именно этот период считают временем появления *Homo sapiens* как биологического вида. Сравнительное исследование мтДНК разных популяций современных людей позволило выдвинуть предположение, что еще до выхода из Африки, около 60–70 тыс. лет назад, предковая популяция разделилась по крайней мере на 3 группы, давшие начало трем расам: африканской, монголоидной и европеоидной.

Чуть позже, чем работы по мтДНК, появились данные по генеалогическим деревьям Y-хромосомы. Изучение небольшого участка Y-хромосомы свидетельствует о, возможно, гораздо более позднем происхождении «Y-хромосомного Адама» (140–175 тыс. лет). Все исследования указывают на его африканское происхождение. Различия между оценками, базирующимися на мтДНК и Y-хромосоме, могут быть объяснены как несходством демографической истории популяций по мужской и женской линиям, различным поведением женщин и мужчин при переселениях, завоеваниях и колонизациях, так и различиями этих геномов, например, в интенсивности отбора вариантов мтДНК и Y-хромосомы.

Гипотезу африканского происхождения современного человека подтверждают и наибольший уровень наследственного разнообразия в

Африке по сравнению с другими континентами по большей части типов полиморфизма ДНК, а также малые различия между популяциями (на долю межпопуляционного разнообразия приходится 10–15 % геномной вариабельности, а большая ее часть (85–90 %) сосредоточена внутри популяций), что отражает недавнее происхождение биологического вида. Результаты исследования мтДНК костных останков неандертальцев (уже более 20 индивидуумов) также свидетельствуют в пользу гипотезы африканского происхождения человека и о том, что неандертальцы не являлись предками анатомически современного человека. **Однако результаты секвенирования ядерного генома неандертальцев показали, что 2–4 % генома всех представителей современного населения за пределами Африки имеет неандертальское происхождение. Это свидетельствует о вкладе неандертальцев в генофонд современных людей вследствие их предполагаемой гибридизации.** Кроме того, полученные данные свидетельствуют о том, что дивергенция между людьми и неандертальцами произошла примерно 660 ± 140 тыс. лет назад.

В целом массив геномных данных наиболее соответствует гипотезе недавнего африканского происхождения современного человека и доказывает справедливость моноцентрической гипотезы. В то же время ни одна из групп генетических данных не является исчерпывающим и бесспорным доказательством этой гипотезы. Наряду с моноцентрической гипотезой существует и другая – полицентрическая, или гипотеза межрегиональной эволюции человека. Среди наиболее интересных результатов, полученных в последнее время в изучении формирования человека современного физического типа, выделяются материалы палеолитических стоянок российского Алтая. При изучении антропологических останков из Денисовой пещеры было сделано предположение о существовании ранее неизвестной группы древних людей (Krause *et al.*, 2010). Геном денисовского человека отклонился от генома человека на 11,7 %, а для неандертальца из пещеры Виндия (Хорватия) отклонение составило 12,2 %, т. е. среднее отклонение ядерного генома денисовца от современных людей такое же, как и неандертальцев. Анализ генома ископаемого человека из Дени-

совой пещеры показал его принадлежность к группе гоминидов, имеющей общего предка с неандертальцами, но разную историю развития популяции – их эволюционное расхождение произошло около 640 тыс. лет назад. Вместе с тем установлено, что отделение этой популяции от линии развития человека современного анатомического типа произошло в среднем около 800 тыс. лет назад. Новая популяция гоминидов – денисовцы, возможно, была широко распространена в восточной части Азии в период верхнего плейстоцена. Полученные результаты показывают, что на континенте Евразия в период верхнего плейстоцена вместе с человеком современного физического типа, возможно, существовали еще как минимум две формы гоминидов: форма Западной Евразии, которая обозначается как неандертальская, и восточная форма, к которой относятся денисовцы (Деревянко, 2011). Эти результаты требуют дальнейшего подтверждения.

Таким образом, результаты исследования полиморфизма аутосомных, митохондриальных и Y-хромосомных ДНК-маркеров в большей степени свидетельствуют в пользу гипотезы недавнего африканского происхождения современного человека и о том, что неандертальцы не являлись предками анатомически современного человека.

Заселение Евразии по данным мтДНК. Дальнейший путь расселения анатомически современного человека рассматривался в основном через призму двух гипотез. Согласно первой гипотезе, на основе исследования ископаемых останков предполагалось его проникновение из Северной Африки в Левант около 100 000 лет назад и дальнейшее расселение на Ближний Восток около 45 000 лет назад. Согласно второй, так называемой гипотезе «южного пути», первое расселение *Homo sapiens* из Африки шло через Эфиопию и Сомали на восток вдоль южного побережья Евразийского континента в Индию и далее – в Юго-Восточную Азию и Австралию. Предположение о двух, независимых друг от друга, путях противоречило наблюдаемому географическому распределению производных базальных гаплогрупп мтДНК М и N, а также наблюдаемому разнообразию гаплогрупп мтДНК в Центральной Азии, свидетельствовавшее о смешении

в данном регионе различных западно-евразийских, восточно-евразийских и юго-азиатских гаплогрупп, многообразие которых сводилось к гаплогруппам М, N и R. Изучение коренного населения Юго-Восточной Азии, в частности оранг-асли Малайзии, показало наличие в данных популяциях не встречающихся более нигде в мире «реликтовых» линий мтДНК, принадлежащих, тем не менее, к тем же трем основным гаплогруппам. Это привело исследователей к выводу о верности гипотезы «южного пути», причем скорость расселения анатомически современного человека по этому сценарию оказалась очень быстрой: расчеты, основанные на «традиционно используемой частоте мутаций», показали его прибытие в Индию около 66 тыс. лет назад и в Австралию около 63 тыс. лет назад. С учетом того что возраст линий мтДНК Западной Евразии близок к возрасту таковых в Индии, возникло предположение о том, что заселение Западной Евразии было результатом раннего ответвления от колонизации вдоль южного побережья на север и дальше – в Левант и Европу. Последнему событию, вероятно, после ответвления предшествовала затяжная пауза в расселении до улучшения климатических условий в данных регионах.

На основе распределения у разных народов частот различных мутаций в Y-хромосоме и мтДНК составлена карта расселения людей с африканской прародины (Cavalli-Sforza, 2000) (рис. 7). Первые волны расселения человека современного типа прошли из Африки через Азию в Австралию и Европу. Удивительно то, что эти данные в целом соответствуют археологическим находкам, которые сделаны раньше или делаются сейчас. Например, появление человека в Австралии и Новой Гвинее датируется 50–60 тыс. лет назад, согласно генетическим данным. Анализ изотопного состава химических элементов археологических находок показывает примерно то же самое время. В Центральной и Юго-Восточной Азии люди появились примерно 70 тыс. лет назад. Заселение Европы произошло позже, ~ 39–40 тыс. лет назад. Наиболее спорны оценки времени заселения Америки. Люди появились там гораздо позже, чем на других континентах, потому что нужно было пройти через Сибирь, добраться до Чукотки и воспользоваться тем моментом, когда

уровень моря в период оледенения позволял перейти нынешний Берингов пролив. Случилось это в промежуток времени от 12 до 35 тыс. лет назад. Позже под натиском ледника палеолитические европейцы несколько раз отступали на юг и юго-восток, возможно, даже возвращаясь обратно в Африку, о чем свидетельствуют результаты исследования гаплотипов Y-хромосомы в популяциях Африки. При сравнении спектра мутаций в ДНК современных европейцев и их азиатских соседей удалось установить, что 10–20 % генов было привнесено в Европу неолитическими переселенцами с Ближнего Востока около 10 тыс. лет назад. Вместе с ними в Европе появилось земледелие.

Разные расы и народы возникли после разделения предковых популяций. Эволюция вновь образовавшихся групп шла независимо. В каждой группе накапливались свои мутации, увеличивалась генетическая дистанция между группами. Сообщества приспосабливались к своим климато-географическим условиям и типу питания. В изолированных группах независимо шла также эволюция языка и культуры. На формирование современных народов влияли не только процессы разделения популяций, поскольку народы могут образовываться и при смешении нескольких исходных сообществ с разной расовой и языковой принадлежностью. При этом возникает генетически разнородная, но с единым типом культуры и общим языком этническая общность. В связи с этим все большую актуальность приобретают работы, связанные с изучением генетической истории популяций отдельных регионов, расово-этнических групп, генетической родословной современных этносов.

Таким образом, исследования полиморфизма митохондриальных и Y-хромосомных ДНК-маркеров внесли важный вклад в понимание путей происхождения человека и рас, расселения *Homo sapiens* по планете, а также в генетическую и демографическую историю отдельных этносов и популяций.

СОВРЕМЕННЫЕ ПОДХОДЫ В ЭТНОГЕНОМИКЕ

Несмотря на существенный прогресс в этногеномных исследованиях, остается ряд

Полногеномный анализ SNPs – генотипирование от сотен тысяч до нескольких миллионов однонуклеотидных полиморфных локусов в больших группах.

Секвенирование полных геномов – определение последовательности нуклеотидов ДНК всех хромосом.

нерешенных проблем: исследована лишь малая часть генома на небольших популяционных выборках, отсутствуют маркеры Y-хромосомы с достаточной разрешающей способностью и методы более точной калибровки молекулярных часов для разного набора маркеров мтДНК и Y-хромосомы. Решению некоторых из перечисленных проблем может помочь применение нового подхода в этногеномике – полногеномного генотипирования. В 2008 г. появились первые статьи, посвященные полногеномному анализу однонуклеотидных полиморфных локусов (SNPs), благодаря интенсивной разработке чиповых технологий разными компаниями. Всего известно около 47 млн SNPs. В настоящее время созданы микрочипы, которые позволяют одновременно анализировать от 96 тыс. до 5 млн SNPs в одном образце.

Изучение 650 000 SNPs у 1064 индивидов из 51 популяции мира – Африки, Европы, Ближнего Востока, Южной/Центральной Азии, Восточной Азии, Океании и Америки – позволило детально охарактеризовать генетическое разнообразие *Homo sapiens* в мире, дифференцировать все популяции по отдельным континентам и регионам и отнести индивидов к определенным популяциям (Li *et al.*, 2008) (рис. 8). Результаты анализа SNPs согласуются с моноцентрической гипотезой о последовательном эффекте основателя с единственным центром происхождения в Африке. Данные авторы первыми показали возможность оценки генетического предка каждого индивида без знания его популяционной принадлежности.

Несмотря на низкие средние уровни различий между европейскими популяциями было найдено полное соответствие между генетическими и географическими расстояниями при анализе генома 3192 человек по 500 000 SNPs. Индивиды из одного и того же географического региона образовывали один кластер, и основ-

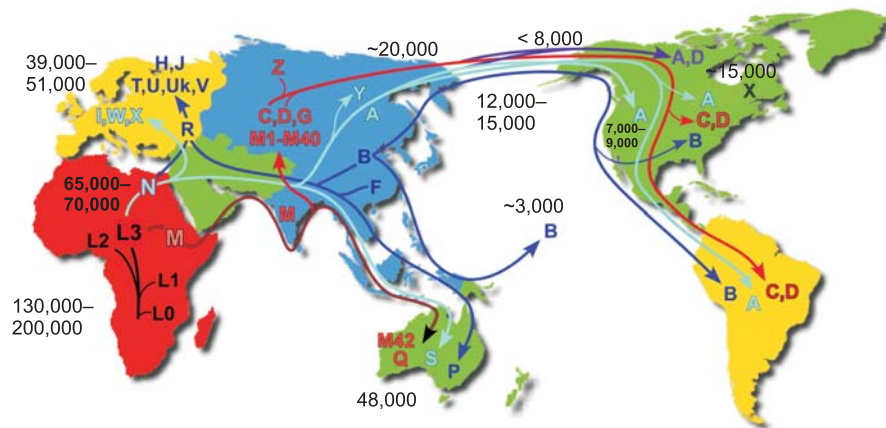


Рис. 7. Карта расселения современного человека с африканской прародины (<http://www.mitomap.org>).

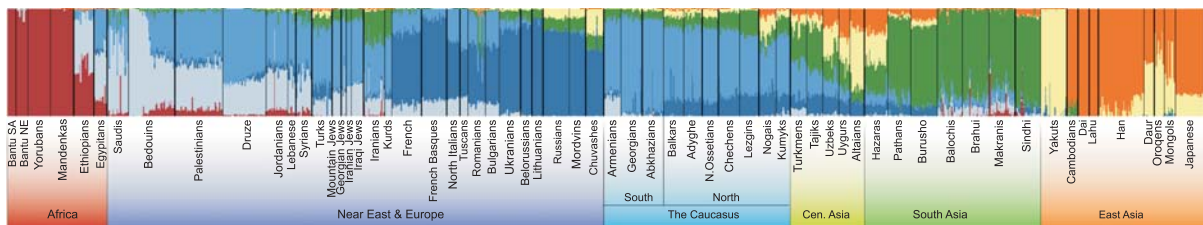


Рис. 8. Полногеномный анализ 610 000 SNPs в популяциях мира.

Различными цветами обозначены различные этнические компоненты (Yunusbayev *et al.*, 2012. Р. 359–365).

ные популяции четко разделялись. Различия по генетической структуре были обнаружены даже между группами людей, говорящими на французском, немецком и итальянском языках и проживающими в одном регионе – Швейцарии. Кроме того, было показано, что результаты в перспективе важны для индивидуального генетического тестирования предков. ДНК индивидов может быть отнесена к их географическому происхождению или положению с удивительной точностью – в пределах нескольких сотен километров.

Полногеномное генотипирование небольших популяционных выборок с использованием сотен тысяч SNPs становится важным для более глубокого изучения эволюционной и демографической истории отдельных регионов и этносов.

В результате полногеномного анализа 610 000 SNPs в 24 этнических группах Кавказа получено наиболее детальное и подробное представление об их генетической структуре и установлено, что их формирование происходило на основе гетерогенного генетического субстрата перед-

неазиатского происхождения с последующим существенным дрейфом генов в изолированных популяциях (рис. 8). Впервые показано существование отчетливого потока генов из популяций Восточной Европы в популяции Кавказа (Yunusbaev *et al.*, 2012).

Изучение генетических различий между популяциями является также чрезвычайно важным для проведения ассоциативных исследований (при поиске генов предрасположенности к определенным болезням), так как полученные ассоциации могут оказаться ложными в результате различий в частотах аллелей между популяциями.

Уже сегодня становится доступным **ресеквенирование полных геномов** человека. Это даст возможность изучать аллели с низкой частотой, и дальнейшее развитие статистических методов позволит нам использовать паттерны гаплотипического разнообразия. Прогресс в технологиях секвенирования, его скорости и стоимости колоссален. Если затраты на секвенирование генома человека в 2003 г. составили 300–400 млн долларов, то геномы, секвенируемые на платформах

второго поколения, стоили 200–500 тыс. долларов, а последние работы по ресеквенированию полных геномов довели стоимость анализа всего лишь до 1500 долларов США. В перспективе стоимость секвенирования индивидуального генома – 1000 долларов за один день.

В ближайшее время завершится проект «1000 геномов», ставящий своей целью получить полные геномы 2000 индивидов из различных популяций основных географических регионов мира – Африки, Европы, Азии и Америки. В первой фазе этого проекта уже секвенировано 1092 образца из разных популяций (www.1000genomes.org). Данные по ресеквенированию полных геномов подтверждают уровень индивидуальной вариабельности генома, оцененной в проекте «Геном человека»: 3 млн SNPs в среднем на геном из 3 млрд нуклеотидов дают уровень различий 1 нуклеотид на 1000 п.н. Полногеномное секвенирование индивидов из различных популяций мира позволит более глубоко изучить эволюционную и демографическую историю популяций и этносов отдельных регионов мира, а также поможет в разработке подходов в персонифицированной медицине.

ЗАКЛЮЧЕНИЕ

Таким образом, изучение геномного разнообразия в популяциях человека внесло значительный вклад в популяционную генетику в целом и позволило обнаружить неизвестные ранее факты и явления в области эволюционного развития вида. Изучение генетической структуры популяций человека необходимо для понимания эволюционной истории человека и для тщательного дизайна медико-генетических исследований. Дальнейшее развитие этногеномики в сочетании с палео- и археогеномикой, а также с усовершенствованием современного оборудования и биоинформатических подходов значительно расширит наши представления о генофонде человека, внесет весомый вклад в понимание вопросов исторического развития и эволюции человечества.

БЛАГОДАРНОСТИ

Исследование поддержано грантами РГНФ № 13-11-02014 и РФФИ № 11-04-00652_а.

ЛИТЕРАТУРА

- Балановская Е.В., Балановский О.П. Русский генофонд на Русской равнине. М.: ООО Луч, 2007. 415 с.
- Деренко М.В., Малайчук Б.А. Молекулярная филогеография населения Северной Евразии по данным об изменчивости митохондриальной ДНК. Магадан: СВНЦ ДВО РАН, 2010. 376 с.
- Деревянко А.П. Верхний палеолит в Африке и Евразии и формирование человека современного анатомического типа. Новосибирск: Изд-во ИАЭТ СО РАН, 2011. 560 с.
- Кутуев И.А., Хуснутдинова Э.К. Генетическая структура и молекулярная филогеография народов Евразии. Уфа: Гилем, 2011. 240 с.
- Лимборская С.А., Хуснутдинова Э.К., Балановская Е.В. Этногеномика и геногеография народов Восточной Европы. М.: Наука, 2002. 261 с.
- Степанов В.А. Этногеномика населения Северной Евразии. Томск: Изд-во «Печатная мануфактура», 2002. 243 с.
- Федорова С.А. Генетические портреты народов Республики Саха (Якутия): анализ линий митохондриальной ДНК и Y-хромосомы. Якутск: Изд-во ЯНЦ СО РАН, 2008. 235 с.
- Хуснутдинова Э.К. Молекулярная этногенетика народов Волго-Уральского региона. Уфа: Гилем, 1999. 238 с.
- Хуснутдинова Э.К., Федорова С.А. Этногеномика населения Евразии: состояние, проблемы и перспективы // Вестн. биотехнологии и физико-химической биологии им. Ю.А. Овчинникова. 2010. Т. 6. № 1. С. 40–49.
- Cann R.L., Stoneking M., Wilson A.C. Mitochondrial DNA and human evolution // *Nature*. 1987. V. 325. P. 31–36.
- Cavalli-Sforza L.L. *Genes, Peoples, and Languages*. N.Y.: North Point Press, 2000. 215 p.
- Karafet T.M., Mendez F.L., Meilerman M.B. *et al.* New binary polymorphisms reshape and increase resolution of the human Y chromosomal haplogroup tree // *Genome Res*. 2008. V. 18. P. 830–838.
- Krause J., Fu Q., Good J. *et al.* The complete mitochondrial DNA genome of unknown hominin from southern Siberia // *Nature*. 2010. V. 464. P. 894–897.
- Li J.Z., Absher D.M., Tang H. *et al.* Worldwide human relationships inferred from genome-wide patterns of variation // *Science*. 2008. V. 319. P. 1100–1104.
- Oven van M., Kayser M. Updated comprehensive phylogenetic tree of global human mitochondrial DNA variation // *Hum. Mutat*. 2009. V. 30 P. E386–E394.
- Yunusbayev B., Metspalu M., Järve M. *et al.* The Caucasus as an asymmetric semipermeable barrier to ancient human migrations // *Mol. Biol. Evol*. 2012. V. 29. No. 1. P. 359–65.