
**ОБЗОРНЫЕ
И ТЕОРЕТИЧЕСКИЕ СТАТЬИ**

УДК 575.174:599.9

**МАРКЕРЫ Y-ХРОМОСОМЫ В ПОПУЛЯЦИОННОЙ ГЕНЕТИКЕ:
ФУНДАМЕНТАЛЬНЫЕ И ПРИКЛАДНЫЕ РЕЗУЛЬТАТЫ
ЭТНОГЕНОМНЫХ ИССЛЕДОВАНИЙ**

© 2021 г. В. Н. Харьков*

*Научно-исследовательский институт медицинской генетики, Томский национальный исследовательский
медицинский центр Российской академии наук, Томск, 634050 Россия*

*e-mail: vladimir.kharkov@medgenetics.ru

Поступила в редакцию 31.08.2020 г.

После доработки 23.02.2021 г.

Принята к публикации 19.03.2021 г.

Y-хромосома является уникальным инструментом для изучения структуры генофондов и эволюционной истории популяций человека. Патрилинейность обеспечивает последовательное накопление мутаций, в результате чего Y-хромосома представляет собой достаточно простую маркерную систему для реконструкции эволюционных изменений, миграционных и других демографических процессов в популяциях человека. Анализ гаплогрупп Y-хромосомы может быть эффективно применен как для описания генетико-демографической истории различных популяций, так и для масштабных филогенетических и филогеографических исследований отдельных монофилетических линий разного иерархического уровня. Обнаружение все новых популяционно-специфичных ДНК-маркеров этой части генома человека позволяет очень подробно анализировать популяционную, субэтническую и родовую структуру различных этносов на протяжении всех периодов ее формирования. Определение аллельного профиля образцов мужчин в работах по палеогенетике особенно важно для реконструкции миграций, компонентного состава, преемственности генетического наследия древних популяций и уточнения филогении различных сублиний. Накопление большого массива данных по специфичному для конкретных популяций распределению различных гаплогрупп и особенностей их профилей по SNP и STR-маркерам кроме разрешения фундаментальных проблем позволяет успешно решать практические задачи по генетической генеалогии каждого мужчины, а также этнической идентификации неизвестного индивида по образцу его ДНК и разработке тест-систем для ДНК-идентификации.

Ключевые слова: Y-хромосома, генофонд, популяция, генетическое разнообразие, этногеномика, палеогенетика, ДНК-идентификация.

DOI: 10.31857/S0016675821090046

Развитие новых технологий масштабного генотипирования и биоинформационного анализа за последние несколько лет позволили перейти на новый уровень изучения генетической структуры популяций человека. Активно накапливающиеся данные по индивидуальным геномам из различных научных проектов появляются в свободном доступе и позволяют интегрировать эти результаты при статистической обработке различных по масштабу массивов данных. Несмотря на то что большую часть этих “больших данных” составляют различные аутосомные маркеры, результаты генотипирования мтДНК и Y-хромосомы занимают особое место среди маркерных генетических систем в силу их однородительского наследования. Принципиальная важность ДНК-маркеров этих специфических участков генома человека состо-

ит в том, что они позволяют исследовать раздельно материнскую и отцовскую линии, поскольку мтДНК передается потомкам только от матери, а Y-хромосома наследуется только от отца к сыну.

Благодаря отсутствию рекомбинации и линейному наследованию этих гаплоидных участков анализ их генетического полиморфизма является важным методом изучения геномного разнообразия и генетической истории популяций человека. Анализ линий мтДНК и Y-хромосомы применяется для выявления особенностей генетической структуры локальных популяций, истории и закономерностей формирования региональных генофондов, этапов заселения человеком различных континентов и, наконец, этапов эволюции человека как вида [1, 2].

ОСОБЕННОСТИ Y-ХРОМОСОМЫ КАК ИНСТРУМЕНТА ДЛЯ ПОПУЛЯЦИОННЫХ ИССЛЕДОВАНИЙ

Преимущество Y-хромосомы заключается в том, что она, имея все те же удобные для эволюционных и популяционных исследований качества, что и мтДНК, несет гораздо более широкий спектр маркеров, имеющих разные темпы мутирования. Это позволяет проводить анализ мужских линий на разных уровнях разрешения — от определения крупных клад до персональной идентификации каждой конкретной хромосомы в популяции. К настоящему времени установлено, что распределение маркеров Y-хромосомы специфично для большинства популяций мира [3].

Наибольшей информативностью обладают гаплотипы Y-хромосомы, построенные на основе сочетания максимального числа высокополиморфных микросателлитных локусов и диаллельных маркеров (в основном SNP). При этом крупные кластеры гаплотипов строятся на основании композиции диаллельных локусов, частота мутирования которых гораздо ниже, а их более детальное дробление проводится с помощью набора STR-маркеров. Сходство или различие в спектре и частотах гаплотипов между двумя или более популяциями позволяет оценить степень их родства и уточнять их происхождение [4]. Сочетание в гаплотипе SNP и STR-маркеров дает возможность с помощью различных статистических методов получить количественные оценки возраста линий Y-хромосомы, а также долю генофонда, унаследованную этносом от различных предковых групп [5].

Спецификой Y-хромосомных линий является их высокая разрешающая способность при анализе генетической дифференциации популяций. Вследствие низкой эффективной численности пула Y-хромосом (как и мтДНК) в популяции (в 4 раза меньше, чем для аутосом) эта хромосома в значительной степени подвержена эффектам генетического дрейфа [6]. К уменьшению эффективного размера приводит и значительная вариабельность числа потомков у мужчин. Этот фактор влиял на Y-хромосомные линии гораздо интенсивнее, чем на мтДНК, поскольку на протяжении большей части своей истории люди придерживались полигении и лишь относительно недавно перешли к моногамным отношениям, что было подтверждено анализом различных Y-хромосомных гаплогрупп [7]. При этом оценка статистических различий в числе потомков между мужчинами и женщинами показывает значительные диспропорции в разные периоды человеческой истории, которые связаны с развитием древних цивилизаций и изменениями социальной организации обществ. Было показано, что увеличение эффективной численности мужской части популяции про-

изошло значительно позже, чем женской. Первоначально время максимального роста популяции на основе данных, полученных при исследовании мтДНК, было оценено в 80–60 тыс. л.н., а данные анализа Y-хромосомы дали гораздо меньшие значения — 18 тыс. л.н. [8]. Позднее был подтвержден рост разнообразия 50 тыс. л.н., и по мужским и по женским линиям, что свидетельствует об увеличении эффективного размера популяции в период расселения по Евразии. Но был обнаружен и сильный спад разнообразия Y-хромосомных линий 10–8 тыс. л.н. и уточнен последующий рост по разным регионам в период 8–4 тыс. л.н., с формированием множества регионально-специфических кластеров [3].

В популяциях человека особенности заключения браков, когда мужчины, как правило, чаще остаются в местах своего рождения, чем женщины, обуславливают феномен патрилокальности — явление, которое отмечено практически во всех этносах. Генетическая дифференциация большинства популяций характеризуется высоким разнообразием по мтДНК и низким по маркерам Y-хромосомы внутри групп популяций и, наоборот, большими различиями по маркерам Y-хромосомы и относительно малыми по мтДНК между группами. Это дает в результате более высокий уровень географической дифференциации вариантов Y-хромосомы по сравнению с мтДНК и аутосомами, который может быть более эффективно использован для исследования миграций [1]. Распределение отдельных гаплогрупп может служить маркером различных компонент популяционного генофонда, позволяя исследователю проводить количественный анализ генетической вариабельности непосредственно по их частоте и составу.

К настоящему времени обнаружено более ста тысяч полиморфных SNP [9, 10] и несколько сотен STR-маркеров [11, 12], которые позволяют решать популяционные и эволюционные задачи различного масштаба — от характеристики мирового генофонда до высокоразрешающего анализа генетически близких популяций.

КЛАССИФИКАЦИЯ ГАПЛОГРУПП

За три последних десятилетия накоплено огромное количество данных, полученных при анализе нерекombинирующей части Y-хромосомы. Первые работы в этой области относятся к середине 1980-х гг. [13, 14]. В последующие годы маркеры Y-хромосомы стали использоваться для эволюционных исследований [15], в судебно-медицинской экспертизе, медицинской генетике [6, 16] и реконструкции родословных [17]. Обнаружение диаллельных ДНК-маркеров на различных участках Y-хромосомы позволило открыть и классифицировать монофилетические гаплогруппы и

начать подробное изучение генофонда различных популяций.

За первые десять лет работы прогресс был не слишком велик, и количество Y-хромосомных ДНК-маркеров оставалось относительно небольшим. К концу 1996 г. было известно лишь около 60 полиморфных маркеров [18, 19]. Затем были опубликованы сведения о 19 новых SNP, которые были обнаружены при помощи денатурирующей жидкостной хроматографии при высоком давлении (DHPLC) [20]. С тех пор с помощью этого метода было обнаружено несколько сотен SNP на Y-хромосоме. Далее на первый план при поиске новых маркеров Y-хромосомы выходит методика секвенирования ДНК [3, 21–23].

С увеличением числа известных SNP возрастало как число исследований, выполненных с их использованием, так и число номенклатурных систем. Со временем появилось семь различных вариантов названий гаплогрупп Y-хромосомы. Главная проблема носила технический характер, поскольку разные исследовательские коллективы использовали разные наборы маркеров и различную номенклатуру гаплогрупп, что затрудняло сравнение результатов разных работ. Это побудило исследователей, работающих в данной области, к созданию единой номенклатуры, систематизирующей весь ранее накопленный материал и при этом достаточно гибкой для включения каких-либо новых линий [24]. Было выделено 18 основных мутационных кластеров в хромосоме, так называемых клад, обозначаемых латинскими буквами от A до R. Порядок букв отражал последовательность возникновения определяющих их мутаций. Эти клады в свою очередь разветвлялись на гаплогруппы, пронумерованные цифрами и буквами. Дальнейшее увеличение количества открытых SNP способствовало уточнению топологии древа и все большей его детализации [2, 25–28 и др.].

Современная номенклатура Y-хромосомных линий представляет собой многоуровневую буквенно-цифровую систему обозначений, классифицирующихся в соответствии с пошаговой мутационной моделью. Иерархическое местоположение конкретной линии определяется сцеплением мутантного и исходного вариантов ДНК-маркера с лежащими выше и ниже SNP. Для обозначения субклад используются две системы: либо на основе буквенно-цифровой номенклатуры (например, E1b1, N1a1a1a2, R1a1a1b2a2, R1a1a1b2a2a3c, T1a2b и др.), либо по номеру терминальной мутации, определяющей данную гаплогруппу (например, Q-BZ99 или N-B172) [9, 10]. Первая система может быть упрощена до 3–5 буквенно-цифровых уровней, при этом промежуточные узлы древа обозначаются сцеплением именованных узлов, например наличие Y1'3 и Y4 указывает, что Y1, Y2, Y3

имеют общего предка, а уже этот предок — общего предка с Y4 [3].

В настоящее время эта классификация является основной для всех коллективов, занимающихся исследованиями с привлечением маркеров Y-хромосомы. При этом постоянно выходят в свет работы, уточняющие топологию отдельных монофилиетических кластеров [21, 23, 25], вплоть до полной ревизии филогении базовых клад A и B [22].

Не менее значимым для классификации гаплогрупп является ресурс YFull.com, который был основан в 2013 г. и является сервисом по биоинформатическому анализу исходных данных секвенирования. Древо гаплогрупп на этом сайте является сейчас наиболее подробным и включает максимальное количество SNP, с привязкой к конкретным образцам. Обозначение линий при этом отличается от буквенно-цифровой кодировки и содержит только привязку к конкретным маркерам, начиная от самых базовых до терминальных для конкретного образца [10].

ФИЛОГЕНИЯ И ФИЛОГЕОГРАФИЯ ГАПЛОГРУПП

Кроме характеристики генетического разнообразия, дифференциации и компонентного состава популяционного генофонда, исследование маркеров Y-хромосомы очень важно для филогенетических реконструкций. Использование множества различных полиморфных сайтов позволяет прослеживать комбинации аллелей, представляющие собой последовательную запись мутаций в ряду поколений. При этом под линией понимают группу гаплотипов, связанных общим происхождением, где каждый вариант отличается от соседнего на один мутационный шаг [2]. Анализ филогении и филогеографии всего Y-хромосомного древа человечества, его отдельных клад и гаплогрупп является сутью этой области исследований.

Частота возникновения SNP на Y-хромосоме примерно в 2 раза выше, чем в среднем по геному, и уступает по скорости мутирования только мтДНК [5, 29]. Поскольку нерекombинирующая часть Y-хромосомы значительно превосходит мтДНК по размеру, а число повторных и обратных мутаций для отдельных позиций намного ниже, то филогения мужских линий может быть гораздо более детально структурирована по сравнению с женскими. Накопление все большего количества просеквенированных образцов Y-хромосом из различных популяционных выборок для всех гаплогрупп позволяет увеличивать информативность и популяционную специфичность их филогенетической реконструкции от самых древних линий до недавно возникших в локальных популяционных группах.

В момент возникновения гаплогруппы она представлена только у одного индивида, поскольку определяющая ее мутация происходит на конкретной хромосоме, а затем частота гаплогруппы начинает стохастически меняться. Новая возникшая сублиния с большой вероятностью может исчезнуть из популяции, если мужской род, ведущий свое начало от родоначальника этой мутации, прервется в каком-либо поколении. Большинство таких новых линий элиминируются стохастическим процессом генетического дрейфа. Но некоторые из гаплогрупп остаются в популяции и увеличивают свою частоту. На фоне стохастического роста численности гаплогруппы в ней с течением времени начинает накапливаться изменчивость по другим SNP и микросателлитным локусам [30]. Изначально гаплогруппа представлена одним определенным гаплотипом-основателем, характеризующимся определенным значением аллелей в различных STR-локусах. Затем, с течением времени, у потомков появляются и другие аллели за счет мутаций и возникают кластеры эволюционно родственных гаплотипов.

Наибольшей информативностью обладают гаплотипы, построенные на основе сочетания большого числа высокополиморфных микросателлитных локусов (STR) и SNP. В дальнейшем более детальное дробление и анализ уровня разнообразия хромосом, несущих определенные варианты диаллельного локуса, проводятся с помощью набора микросателлитов и дополнительных SNP, позволяющих детализировать структуру субклад исследуемой гаплогруппы.

Аналогично, вполне реальным является проведение оценки времени коалесценции к общему предку микросателлитных гаплотипов внутри гаплогрупп Y-хромосомы. Таким образом, представляется возможным оценить время происхождения гаплотипического разнообразия внутри линий, т.е. их возраст. Такие исследования проводились для отдельных клад, в частности для гаплогрупп E, J, G [25, 26], I [27], R1a [21], R1b [28], N [23].

Филогеографический подход к картированию всего разнообразия сублиний конкретной клады и уровня гаплотипического разнообразия в пределах гаплогрупп позволяет решать специфические задачи: выявлять места происхождения и пути распространения линий, проводить реконструкцию древних миграций [31]. Анализ распределения частот и оценка разнообразия гаплогрупп позволяют подробно описывать генофонды популяций и определять причины, которые привели к тем или иным структурным изменениям.

Чем эволюционно ближе друг другу популяции, тем шире должен быть спектр перекрывающихся сублиний гаплогрупп и STR-гаплотипов. И наоборот, при сходных частотах основных гап-

логрупп спектр STR-гаплотипов в пределах отдельных линий может значительно различаться, что является следствием длительной изоляции популяций и формирования на основе различных генетических составляющих, маркируемых специфичными кластерами гаплотипов, относящихся к одним и тем же диаллельным гаплогруппам.

Не только SNP, но и STR-маркеры позволяют детально реконструировать филогенетические взаимоотношения между отдельными Y-хромосомами, принадлежащими к одной гаплогруппе, и дать оценку возраста наименее древнего общего предка. Анализ STR-гаплотипов Y-хромосомы успешно применяют для сравнительного анализа популяций, относящихся к близким этносам, что позволяет выявить у них определенную степень гетерогенности, которую не всегда удастся обнаружить с помощью аутосомных маркеров. Степень детализации анализа зависит практически лишь от количества YSTR, используемых для построения гаплотипа: если для высокоразрешающего анализа индивидов и генеалогических реконструкций требуется от 30 и более маркеров (в качестве основного примера выступает компания Family Tree DNA), то для общего описания популяционного генофонда в большинстве работ используются 15–20 YSTR [3, 23].

Наблюдаемое в гаплогруппах разнообразие микросателлитных гаплотипов формировалось на протяжении определенного промежутка времени, под воздействием мутационного процесса и фиксации или элиминации из генофонда популяций тех или иных вариантов хромосом под воздействием других факторов популяционной динамики. Исследуя их современное разнообразие в популяциях и зная особенности и скорость мутирования изучаемых ДНК-маркеров, можно с определенной степенью точности оценить время его генерации.

При расчетах возраста генерации гаплотипического разнообразия по STR-маркерам в пределах конкретных гаплогрупп очень важным является определение гаплотипа-основателя. На протяжении первых поколений нарастания изменчивости аллель-основатель еще остается наиболее представительным в выборке: частоты других аллелей не успевают увеличиться до значительных величин. Поэтому в качестве гаплотипа-основателя предлагается принимать модальный гаплотип, частота которого максимальна в исследуемой выборке [32, 33]. Но распределение частот аллелей YSTR-локусов, даже в отдельной сублинии в пределах гаплогруппы, не всегда является унимодальным. Мутации, миграции и генетический дрейф могут привести к тому, что гаплотип-основатель перестает быть модальным и может вообще исчезнуть из состава генного пула конкретных популяций. Поэтому часто более подходящей является медианная оценка числа повторов в гаплоти-

пе-основателе: для каждого локуса она вычисляется как медиана числа повторов по всем аллелям, которые обнаружены в данной выборке [30]. Но это может приводить к заниженной оценке возраста на эволюционно более старых гаплогруппах, поскольку за большие промежутки времени распределение числа повторов в дочерних гаплотипах может значительно отклоняться от аллельной композиции гаплотипа-основателя.

Гаплотипы-основатели, определенные медицинским методом, и основанные на них оценки возраста гаплогрупп, сделанные по данным одной выборки или небольшого числа выборок из географически ограниченной области, могут относиться не к истинному моменту возникновения этой гаплогруппы, а к моменту появления ее на данной территории. В случае же если группа мигрантов на момент появления в генофонде локальной популяции имела по данной гаплогруппе сформировавшееся разнообразие гаплотипов, оценка возраста гаплогруппы будет относиться не к изучаемой, а к предковой популяции [30]. Таким образом, достижение наиболее адекватных оценок возраста гаплогрупп возможно при формировании как можно более многочисленных выборок, обязательно включающих популяции из региона предполагаемого происхождения этих гаплогрупп, и должно опираться на информацию по максимально возможному числу генетических маркеров [34–36].

При определении прародины гаплогрупп необходимо использовать несколько критериев. Во-первых, частота гаплогруппы должна достигать высоких значений в месте ее происхождения. Но максимальная частота не может служить однозначным показателем центра зарождения гаплогруппы, поскольку и на периферии ареала гаплогруппы популяции могут испытывать сильное влияние дрейфа, что может привести к эффекту основателя по данной линии у конкретного этноса или группы популяций.

Примеры очень высокой частоты гаплогруппы в популяциях на сильном удалении от места ее происхождения широко известны. Такая ситуация характерна дляклады N у финнов [37] и якутов [38],клады D у японцев [39], гаплогруппы R1b1b1 у башкир [40]. Но в большинстве случаев миграции ее носителей разносят гаплогруппу по соседним популяциям, не подвергающимся резким колебаниям эффективной численности, и частоты там будут ниже, чем в популяциях, близких к месту происхождения SNP, определяющего эту линию.

Во-вторых, наличие предковой по отношению к исследуемой линии и близкородственных сестринских гаплогрупп также указывает на прародину. В большинстве случаев местом возникнове-

ния гаплогруппы может быть регион с высокой частотой базальных ветвей этой клады.

Третьим критерием является разнообразие YSTR-гаплотипов в пределах изучаемой гаплогруппы. Территорию с наибольшим разнообразием какой-то родственной группы гаплотипов принято рассматривать как вероятную зону их общего происхождения. Там, где гаплогруппа впервые достигла значимой частоты и откуда она затем расширяла свой ареал, должны были накопиться многочисленные мутации и сформироваться множество дочерних гаплотипов. В месте происхождения показатели разнообразия должны быть максимальными. В отдаленных популяциях гаплогруппы гаплогруппы разнообразие должно быть мало, поскольку с миграционными потоками они получают не репрезентативную выборку гаплотипов предковой популяции, а лишь случайную часть спектра [41].

Анализ микросателлитной изменчивости в пределах гаплогруппы Y-хромосомы можно проводить и для определения времени, прошедшего с момента отделения ныне существующих популяций от общей предковой популяции. Поскольку в начале процесса накопления STR-изменчивости ее дисперсия равна нулю, можно получить верхнюю границу для времени дивергенции популяций (в пределах изучаемой гаплогруппы). Эта оценка наиболее корректна для популяций, у которых момент разделения отстоит от момента возникновения STR-изменчивости в данной гаплогруппе на наименьший отрезок времени. Для недавно разошедшихся популяций эта верхняя граница может значительно отстоять от реального времени дивергенции. В этом случае для оценки времени дивергенции популяции применима TD-статистика [42, 43]. Очень важное ее отличие перед другими мерами генетического расстояния в том, что ее оценка не зависит от изменения численности популяций во времени и устойчива к слабым миграциям.

Обнаружение все большего числа информативных маркеров позволяет максимально подробно исследовать филогенетические взаимосвязи отдельных линий внутри разных клад Y-хромосомы. В то же время это открывает и широкие возможности для подробной характеристики популяционных генофондов. При этом уже накоплен значительный массив данных по варибельности десятков микросателлитных маркеров в тысячах образцов ДНК, тестируемых коммерческими ДНК-генеалогическими компаниями. Это позволяет научно-исследовательским группам гораздо эффективнее планировать работу по анализу Y-хромосомных гаплогрупп. Если 10–15 лет тому назад генотипирование новых SNP и YSTR далеко не всегда приводило к значительному увеличению информативности выбранного набора маркеров, то сейчас можно отбирать ДНК-маркеры, широкий полиморфизм

которых уже подтвержден. Более того, можно выбирать как быстро мутирующие STR, для подробного анализа относительно молодых сублиний, так и маркеры с гораздо более медленным темпом мутирования, выделяющие в пределах отдельных гаплогрупп региональные и этноспецифичные кластеры гаплотипов.

Установлено, что для большинства гаплогрупп Y-хромосомы наблюдается разделение не только на этноспецифические линии, но и на более узкие сублинии и кластеры гаплотипов. Анализ микросателлитных гаплотипов Y-хромосомы успешно применяют для сравнительного описания популяций, относящихся к близким этносам, что позволяет выявить у них определенную степень гетерогенности, которую не удастся обнаружить с помощью классических маркеров, а также аутомомных маркеров. Большинство этнических генофондов характеризуются преобладанием или даже полным доминированием специфичных SNP по всем крупным кладам, исследованным в настоящее время [3, 23 и др.]. Степень детализации зависит практически лишь от количества SNP и STR-маркеров, используемых для построения гаплотипа.

Аллели разной длины в одном и том же локусе могут различаться по темпам мутирования. Более длинные аллели могут чаще мутировать, предпочтительно к аллелям с меньшим числом повторов [30, 44–46]. Возможность того, что аллели большей длины могут предпочтительно мутировать в короткие аллели и, наоборот, короткие аллели — в длинные, может являться эволюционным механизмом, удерживающим изменчивость по микросателлитным локусам в определенных пределах [44, 47, 48]. Очевидно, что в разных гаплогруппах темп мутирования разных локусов неодинаков и может также отличаться на разную глубину поколений. Для более точных оценок темпа по отдельным линиям необходимы калибровки по точно датированным событиям в жизни различных популяционных групп.

Новые SNP Y-хромосомы в различных гаплогруппах все активнее выявляются в ходе работы проектов по полногеномному секвенированию и специальному поиску таких маркеров научными группами и различными компаниями. Однако точных данных по их филогенетическим взаимоотношениям в пределах различных гаплогрупп на большом объеме популяционных выборок зачастую нет, многие SNP подтверждены лишь на ограниченном наборе образцов и данные о частотах определяемых ими сублиний в реальных этнических группах отсутствуют или являются весьма приблизительными в силу нерепрезентативности исследованных выборок. Пока отсутствуют результаты генотипирования многих обнаруженных маркеров на материале различных территориальных и этнических групп, значительно затруднен их филогеографический ана-

лиз, достоверные реконструкции возникновения и распространения вновь открытых линий.

ЭТНОГЕНОМИКА НАСЕЛЕНИЯ СЕВЕРНОЙ ЕВРАЗИИ

В настоящее время научные коллективы по всему миру проводят анализ структуры генофондов различных популяций. Одним из наиболее актуальных направлений этногеномных исследований является изучение региональной и внутри-этнической структурированности генетического разнообразия коренного населения разных регионов. Генофонд населения Северной Евразии представляет собой уникальную систему с точки зрения исследования популяционно- и эволюционно-генетических процессов, анализа генетического разнообразия и реконструкции генетической истории популяций. Огромные пространства этого региона и малочисленность аборигенного населения способствовали формированию его значительной территориальной и генетической подразделенности. Относительная изоляция населения отдельных регионов, таких как Сибирь и Дальний Восток, от основных миграционных потоков на территории Евразии обеспечила сохранение в их генофонде древних генетических пластов, хранящих информацию о важнейших этапах заселения человеком современного типа не только Сибири, но и Европы, Кавказа, Центральной Азии и Америки. Генофонд коренного населения хранит в себе следы миграций, экспансий численности, метисационных событий, генетико-демографических процессов и изменения генетического состава популяций. Разнообразие генофонда Северной Евразии по сравнению с другими регионами мира достаточно велико и включает множество генетических компонент. Результаты его анализа очень важны для изучения генофонда всего человечества.

Одними из наименее исследованных территорий в настоящее время остаются Сибирь и Дальний Восток. Современный этнический состав их коренного населения постепенно складывался в результате длительных этногенетических процессов. Археологические, этнографические, антропологические данные указывают на то, что население Сибири формировалось на протяжении тысячелетий при объединении самых различных компонент. Ее заселение было комплексным и продолжительным процессом вследствие миграций с территории Восточной Европы, Средней, Центральной и Восточной Азии. Пересечение различных миграционных путей привело к тому, что народы, населяющие этот обширный регион, весьма разнообразны по своему антропологическому и языковому составу. Огромное влияние на процесс заселения Сибири оказывало изменение климатических условий, вызывавшее смену биоценозов, что отражалось на

возможности использования их человеком и заселении им новых территорий.

Популяции коренных этносов Сибири представляют значительный интерес для такого рода исследований как в силу относительно слабой их изученности с привлечением современных ДНК-маркеров, так и по причине специфичности их генофондов, развивавшихся, зачастую, в условиях длительной генетической изоляции. Уникальный генофонд коренного населения Сибири представляет собой ценный источник генетической информации, важный как для выяснения особенностей генофондов популяций Северной Евразии и заселения Америки, так и для анализа молекулярной эволюции многих гаплогрупп Y-хромосомы и мтДНК.

Первичные данные по сибирским популяциям были получены в ранних работах, описывающих распределение новых диаллельных молекулярно-генетических маркеров Y-хромосомы [19, 49]. После этого большинство исследований, где приводились результаты по сибирским популяциям, касались проблемы заселения Америки. Некоторые научные работы включали материал сибирских популяций, но не рассматривали их как самостоятельные объекты. Затем изучение истории популяций, этнических и региональных генофондов населения Сибири с использованием ДНК-маркеров Y-хромосомы стало самостоятельной задачей в целом ряде исследований [50–53]. Различными группами исследователей была охарактеризована общая структура генофонда коренных этносов Сибири, проведены анализ филогенетических взаимоотношений сибирских популяций и изучение филогеографии отдельных гаплогрупп. К настоящему времени уже накоплен значительный фактический материал о составе гаплогрупп и YSTR-гаплотипов.

Однако полученная картина далеко не полно отражает реальную структуру сибирского генофонда. Это связано, в первую очередь, с относительной малочисленностью исследованных выборок и проблемой репрезентативности выборок в отношении всего генофонда изучаемых этносов. По-прежнему остаются неисследованными генофонды некоторых малочисленных коренных этнических групп и субэтносов. Кроме того, многие сибирские этносы антропологически и генетически довольно гетерогенны (например, алтайцы, тувинцы, якуты, эвенки, буряты), а исследованные до настоящего времени выборки не отражают всей многокомпонентности генофонда, характеризуя лишь часть имеющегося разнообразия. Для коренных этносов Сибири актуальность проведения исследований их генофондов с помощью маркеров Y-хромосомы обусловлена также недостаточной степенью детализации гаплогрупп по сравнению с европейскими, ближневосточными и африкан-

скими популяциями и отсутствием данных по широкому набору YSTR в большинстве работ.

Для многих гаплогрупп Y-хромосомы гаплотипическое разнообразие в сибирских популяциях ниже величин, характеризующих современное население других регионов (Восточная Европа, Средняя, Центральная и Юго-Восточная Азия). Принимая во внимание все возможные погрешности при подобного рода оценках, можно, тем не менее, заключить, что накопление гаплотипического разнообразия в гаплогруппах, составляющих большую часть суммарного генофонда сибирских этносов, сопровождалось снижением эффективной численности популяций. При этом накопление гаплотипического разнообразия в гаплогруппах Y-хромосомы, составляющих большую часть суммарного генофонда сибирских этносов, сопровождалось значительным дрейфом генов и многочисленными событиями популяционного “горлышка бутылки” в генетической истории коренных популяций.

Были выявлены локальные эффекты основателя по различным гаплогруппам Y-хромосомы как для отдельных этносов, так и для различных территориальных групп в пределах конкретных сублиний. С привлечением оценки темпов мутирования YSTR было рассчитано время накопления разнообразия YSTR-гаплотипов внутри наиболее частых гаплогрупп. По всей видимости, это является следствием значительного роста численности популяций, связанного с изменением климатических условий и типа ведения хозяйства.

Изучение вариативности Y-хромосомы в современных популяциях, эволюции гаплогрупп и их географического распределения позволило прояснить проблемы происхождения и расселения человека, реконструировать пути древних миграций, описать структуру и происхождение генетического разнообразия в общемировом масштабе. На новом этапе на первый план выходят работы по анализу внутри этнической генетической дифференциации и особенностей локальных генофондов субэтнических и этнотерриториальных групп населения. Детальное исследование этнических генофондов требует увеличения численности популяционных выборок и подробности описания исследуемого материала.

Таким образом, в настоящее время одним из наиболее актуальных направлений этногенетических исследований является изучение региональной и внутриэтнической структурированности генетического разнообразия коренного населения Северной Евразии, с привлечением материала по крайней мере нескольких географически дистанцированных выборок для каждого этноса. Изучение множества локальных популяций позволит оценить степень внутриэтнической дифференциации генофонда. Новая информация о структуре Y-хромосомной составляющей генофонда коренных

популяций этого региона будет являться важным дополнением к существующим антропологическим, археологическим, лингвистическим данным, а также результатам, полученным при исследовании других маркерных генетических систем.

АНАЛИЗ РОДОВОЙ СТРУКТУРЫ НАРОДОВ

Одним из актуальных направлений популяционной генетики человека является изучение внутриэтнической структурированности генетического разнообразия различных этносов, с привлечением материала входящих в их состав субэтнических и родоплеменных групп. Комплексный междисциплинарный подход приносит значительные результаты для такого рода работ. Методы популяционной генетики в последние два десятилетия активно используются на стыке таких ранее не связанных между собой наук, как молекулярная генетика, с одной стороны, и антропология, археология, этнография и лингвистика, с другой.

Особенностью этнического состава многих народов является наличие родов, где счет родства ведется по мужской линии. Такая родовая структура характерна для кланов Шотландии и Ирландии, различных народов Центральной и Средней Азии, Волго-Уральского региона, Сибири и Дальнего Востока. Проблема родовой организации разных народов достаточно хорошо изучена с этнографической и социокультурной точек зрения. Многочисленные работы этнографов подробно освещают особенности родоплеменной организации. В то же время их генетическая структура оставалась практически неизученной.

Применение Y-хромосомных маркеров в этой связи дает богатейшие возможности для исследования проблем, связанных с родоплеменной структурой, поскольку накопление большого массива генотипических данных по множеству новых SNP и STR позволяет вывести этногенетические исследования на совершенно новый уровень. Если два человека принадлежат к разным гаплогруппам Y-хромосомы, между ними не может быть близкого родства по мужской линии. Если же два человека принадлежат к одной гаплогруппе, то возможность родства между ними существует. Степень этого родства устанавливается путем анализа STR-гаплотипов и специфичных SNP Y-хромосомы у исследуемых индивидов. Определение специфичных гаплотипов по большому числу STR и значительная детализация сублиний внутри гаплогрупп позволяют проверить, является ли род кровнородственным объединением, имеющим одного родоначальника по мужской линии, или осознание такой общности связано лишь с социокультурными особенностями этой группы, когда род является общностью людей, проживающих на одной территории, но не связанных генетическим родством по отцовским генеалогическим линиям.

Современные молекулярно-филогенетические методы позволяют безошибочно выполнить такого рода сравнительный анализ, а обширная база мировых генетических данных дает возможность провести подробное сравнение обнаруженных генетических вариантов на разных иерархических уровнях. Генотипирование широкого набора SNP и YSTR-маркеров, подобранных с учетом имеющейся информации о их информативности для изучаемых популяционных групп, позволяет выявлять генетическую структуру патрилинейных обществ, оценивать родственные связи отдельных индивидов, родов, групп родов и субэтносов, а также отслеживать общие генетические компоненты между разными народами и реконструировать источник, время и направление миграций носителей отдельных сублиний внутри исследуемых гаплогрупп.

Одними из первых заметных публикаций по анализу популяций в контексте родоплеменных отношений были работы по генетическому наследию династии монголов [54], центральноазиатских популяций [55] и маньчжурской династии Цин [56]. Результаты этих исследований послужили толчком для дальнейшего применения Y-хромосомных маркеров для более подробной характеристики родов, за счет увеличения количества маркеров и популяционных выборок.

В настоящее время достаточно подробно охарактеризованы родовые группы башкир [57], казахов [58, 59], туркмен [59], хакасов и шорцев [60], тувинцев [61]. Для большинства исследованных родов в составе разных народов показаны значительные генетические различия по составу гаплогрупп и спектру STR-гаплотипов. Для многих родов обнаружены доминирование одной гаплогруппы и недавние эффекты основателя, оценка возраста которых, как правило, хорошо согласуется с данными этнографии. Это свидетельствует об однозначной родоплеменной подразделенности большинства исследованных популяций по мужской линии. Анализ STR-гаплотипов показывает, что большинство представителей одного рода являются родственниками по мужской линии и имеют родоначальника, жившего в относительно недалеком прошлом. Именно родовой уровень организации генофонда является наиболее точно характеризующим его популяционную структуру.

Перспективы дальнейших работ по этому направлению достаточно обнадеживающи, поскольку родовую принадлежность у представителей разных этносов во многих случаях возможно установить не только по родовой самоидентификации, но и по их фамилиям, так как большинство современных фамилий коренного населения Сибири повторяют название рода в исходной форме, либо с небольшими изменениями, а также ранее привязаны этнографами к различным родам и фратриям. Такая

ситуация характерна для башкир, хантов, ненцев, шорцев, хакасов, северных и южных алтайцев, телеутов, нанайцев, удэгейцев и других народов. Практическое выполнение таких работ требует значительного пополнения популяционных библиотек и расширения этнических выборок с привязкой к конкретным патрилинейным группам.

ПАЛЕОГЕНЕТИКА

Палеогенетика основана на анализе геномной информации, содержащейся в биологических останках ископаемых организмов. Данное направление возникло на стыке антропологии, археологии и молекулярной генетики и позволяет не только описывать генетические портреты древних образцов, но и более детально реконструировать генетическую историю заселения различных регионов. Анализ древних образцов также включает в себя филогенетический и филогеографический подходы, в том числе с помощью анализа линий мтДНК и Y-хромосомы.

Изучение генетического разнообразия древнего населения предоставляет новую информацию о динамике развития генофондов того или иного региона. Вероятную прародину гаплогруппы должны указывать и данные о гаплотипах образцов из захоронений [62]. Популяционно-генетические исследования с использованием древней ДНК ориентированы в основном на решение проблем заселения территорий, изучение древних миграций и установление генетической преемственности древнего и современного населения.

В большинстве ранних работ внимание было сосредоточено на анализе гаплотипов мтДНК, а данные по маркерам Y-хромосомы отсутствовали из-за гораздо более низкой ее копийности. Но за последнее десятилетие опубликовано довольно много работ, где определена принадлежность древних образцов к различным гаплогруппам Y-хромосомы как по YSTR-гаплотипам, так и по широкому набору SNP с достаточной плотностью покрытия информативных участков.

Определение принадлежности образцов древних популяций к конкретным сублиниям Y-хромосомы позволяет более точно, по сравнению с анализом аутосомных генотипов, описать филогенетическое положение конкретного образца, внутри- и межпопуляционные взаимосвязи, подробно реконструировать древние миграции и этногенез популяций.

Прародиной гаплогруппы может быть регион с высокой частотой предковой для нее линии. Если где-то наблюдается высокая концентрация исследуемой гаплогруппы и та же самая гаплогруппа с близкими YSTR-гаплотипами выявлена в образцах ДНК из захоронений многовековой давности, расположенных в том же регионе, это указывает на

генетическую преемственность населения и длительное присутствие этого варианта Y-хромосомы на данной территории. Исследования древних образцов во многих случаях позволяют обнаружить базальные варианты различных линий, что полностью соответствует филогенетическим схемам их происхождения и позволяет уточнять расселение предков их современных носителей. Так, было установлено, что образец из сибирской стоянки Мальта принадлежит к предковой линии гаплогруппы R, которая сейчас распространена в пределах Западной Евразии, Южной Азии и в Южной Сибири [63, 64].

В Западном Иране обнаружен самый древний из известных на сегодня носитель гаплогруппы G1a (образец I1674), относящийся к эпохе энеолита (4500–3500 до н.э.) [65], что полностью согласуется со снижением гаплотипического разнообразия от Западного Ирана к востоку Юго-Западной Азии и далее на север к Евразийским степям в современных популяциях. Это подтверждает гипотезу о локализации прародины гаплогруппы G1 в западной части Ирано-Армянского нагорья [66]. Носители различных субкластеров гаплогруппы Q1a Y-хромосомы достоверно выявлены среди представителей древних кочевых групп гунно-сарматского времени на территории Горного Алтая [67].

Также важно установить и проанализировать совпадения и отличия в филогеографии различных сублиний у древних и современных носителей. Показано, что мужское население окуневской культуры принадлежит к различным вариантам клады Q1a. Гаплогруппы клады Q представлены в составе ранних и поздних археологических комплексов. Один образец принадлежит к Q1a1b1 (xQ1a1b1a), другой – к Q1a2b-L940 (xQ1a2b1, Q1a2b2), которая сейчас имеет преимущественно среднеазиатское распространение. Еще два образца относятся к Q1a2-M346. Также один образец окуневской культуры и один из стоянки Усть-Ида принадлежат Q1a2a1, родственной линиям коренных американских популяций. Несколько образцов относятся к Q1a2alc-L330 (xQ1a2alc1), которая в настоящее время выявлена у многих коренных народов Южной Сибири – кетов, алтайцев, хакасов, тувинцев, хантов [68]. Три образца афанасьевской культуры и один из окуневской принадлежат к гаплогруппе R1b1a1a2a-L23, а три других окуневских мужчины относятся к очень редкой в настоящее время линии NO, предковой для основной сибирской клады N Y-хромосомы [69].

У древних образцов из Европы была найдена очень редкая в современных популяциях гаплогруппа C1a2, которая маркирует раннее заселение этого региона группой, близкородственной мигрантам, заселявшим по аустерическому пути Австралию, Южную и Восточную Азию. Анализ Y-хромосомного разнообразия позволил уточ-

нить формирование генофонда населения Европы, после начала освоения земледелия. В европейских неолитических образцах доминировали гаплогруппы G и I, а пришлые анатолийские земледельцы принесли ближневосточные гаплогруппы, которые затем клинально распространились на запад. Это свидетельствует о смешении земледельцев с местными охотниками-собираателями, в соответствии с их географической локализацией [70].

Изучение гаплогрупп Y-хромосомы на материале древних популяций позволяет точнее проследить и особенности генетико-демографических событий в истории различных регионов. Установлено, что пришлые неолитические земледельцы в Центральной Европе и на Иберийском полуострове смешивались с местными охотниками-собираателями, и это смешение в разных регионах различалось не только по интенсивности, но и по различному соотношению полов. Дисбаланс в смешении сильно выражен в пользу мужского вклада охотников-собираателей, по сравнению с женским в среднем неолите, что подтверждается маркерами X-хромосомы и высокой частотой Y-хромосомных гаплогрупп охотников-собираателей. В среднем европейском неолите насчитывалось 35–50% предкового вклада охотников-собираателей мужчин и только до 5% приходилось на долю женщин, т.е. местные мужчины-охотники брали в жены женщин из земледельческих общин, а не наоборот [71].

В случае более поздней степной миграции в Европу на примере древних образцов из Эстонии установлено, что у пяти индивидов мужского пола все образцы принадлежали к линии R1a. Гаплогруппа N, которая является наиболее частой в современной Эстонии, обнаружена не была. Это вместе с анализом других маркеров свидетельствует об обратном сценарии: из степей мигрировали преимущественно мужчины, и по мере продвижения по Европе они брали в жены местных женщин из общин европейских земледельцев [72].

В последние годы выходят работы, которые включают гораздо больше образцов, чем более ранние, что дает более достоверные статистические результаты и однозначные выводы. При анализе более 100 геномов бронзового века из множества захоронений степной зоны Евразии установлено распределение различных сублиний гаплогрупп R1a, R1b и Q1a. Выявлено, что генофонд населения Европы и Центральной Азии именно ближе к концу бронзового века отражает современную евразийскую генетическую структуру в такой степени, которая не наблюдалась в предыдущие периоды. Большая часть основы современного евразийского генетического ландшафта была сформирована именно в бронзовом веке, в ходе миграций, смешения и вытеснения различных групп населения [67]. Изученные образцы срубной культуры несут

только гаплогруппу R1a, и большая часть из них и образцов полтавкинской культуры принадлежит к азиатской ветви R1a-Z93, которая сейчас распространена в Южной Сибири, Центральной и Южной Азии, Волго-Уральском регионе и на Кавказе, но отсутствует во всех изученных к настоящему времени древних образцах Центральной Европы.

Развитие этого направления науки очень актуально для раскрытия и уточнения многих вопросов эволюции человека и истории формирования современных популяций. Значительный прогресс в генотипировании древних образцов с помощью полногеномного секвенирования, накопление большого массива палеогенетических данных и новые биоинформатические инструменты позволят в ближайшем будущем гораздо достовернее и точнее раскрыть многие аспекты при совместном анализе современных и палеогеномов, в том числе благодаря генотипированию Y-хромосомных маркеров.

ДНК-ИДЕНТИФИКАЦИЯ

Помимо фундаментальных вопросов популяционной генетики применение Y-хромосомных маркеров имеет важное значение и для практического использования в области генетической генеалогии и разработки методов ДНК-идентификации биологических образцов человека. Каждый индивидуальный организм с его уникальным геномом отличается неповторяющимся сочетанием множества полиморфных ДНК-маркеров. Каждая популяция человека отличается от других популяций не только частотами аутосомных SNP, специфичностью идентичных по происхождению блоков сцепления, но и особенностями представленности различных филогенетических линий Y-хромосомы и специфичностью кластеров их STR-гаплотипов [73].

Накопление данных индивидуальных генотипов по тысячам образцов ДНК из различных мировых популяций в общедоступных интернет-ресурсах позволяет использовать эту информацию и интегрировать экспериментальные данные научных коллективов и доступные генотипы образцов из открытых баз генеалогических компаний, расширяя возможности по определению происхождения человека и его привязке к конкретной этно-территориальной группе. Генотипирование маркеров Y-хромосомы может быть эффективно в криминалистике и судебной медицине и служить альтернативой ДНК-идентификации по аутосомным локусам. Дополнительный анализ гаплотипов Y-хромосомы наиболее значим в случаях биологических образцов со смесью мужской и женской ДНК. Значительная специфичность Y-хромосомных сублиний играет важную роль при

определении популяционной принадлежности неизвестного индивида по образцу его ДНК.

Обнаружение огромного количества SNP-маркеров, которые определяют филогенетически недавние ветви для различных гаплогрупп Y-хромосомы, и данные об их распространении в современных популяциях, позволяют провести отбор специфичных маркеров не только для большинства популяций и этносов, но зачастую и для отдельных патрилинейных родовых групп. Дополнительное использование широкого набора YSTR совместно с генотипированием таких терминальных SNP позволяет в большинстве случаев очень точно идентифицировать популяционную принадлежность образца. Такая идентификация касается только привязки Y-хромосом и не затрагивает весь геном исследуемого индивида.

В данном случае для популяционной идентификации применимы только информативные высоко-специфичные терминальные SNP, поскольку большинство базовых SNP для всех основных сублиний большинства гаплогрупп, которые возникли тысячи лет назад, присутствуют у представителей самых различных народов. Использование таких высокоинформативных этноспецифичных SNP-маркеров Y-хромосомы является одним из наиболее перспективных инструментов для анализа региональных и этнических популяционных генофондов на уровне субэтносов, родов и локальных территориальных групп, а также для разработки и применения идентификационных тест-систем для криминалистики и исследования этнического происхождения мужчин, заинтересованных в определении своей генетической генеалогии [74].

Использование маркеров Y-хромосомы позволяет повысить эффективность идентификации личности в сочетании с другими известными в криминалистике методами. Получение информации для проведения оперативно-розыскных мероприятий на основе развития генетических технологий становится возможным практически по любым биологическим следам, остающимся на местах преступлений, а при идентификации индивида генетические технологии обеспечивают максимальную точность. За счет наибольшей межпопуляционной дифференциации населения именно по Y-хромосомным маркерам разработка тест-систем на их основе очень важна и эффективна, поскольку в этом случае наборы для генотипирования будут основаны на значительно большем числе маркеров, по сравнению с аутосомными.

Очевидно, что при разработке новых наборов Y-хромосомных SNP и STR для определения популяционной принадлежности помимо дискриминирующего потенциала выбранных маркеров необходимо учитывать и степень представленности информации о генотипах по этим маркерам в доступных генетических базах данных. Поиск

совпадающих гаплотипов в базах данных, этническая и географическая привязка таких образцов могут являться важной составляющей работы по криминалистическому анализу биологических образцов.

Работа поддержана грантом РФФИ № 19-14-50421 “Маркеры Y-хромосомы в популяционной генетике: фундаментальные и прикладные результаты этногеномных исследований”.

Настоящая статья не содержит каких-либо исследований с использованием в качестве объекта животных.

Настоящая статья не содержит каких-либо исследований с участием в качестве объекта людей.

СПИСОК ЛИТЕРАТУРЫ

1. *Underhill P.A., Shen P., Lin A.A. et al.* Y chromosome sequence variation and the history of human populations // *Nat. Genet.* 2000. V. 26. P. 358–361. <https://doi.org/10.1038/81685>
2. *Jobling M.A., Tyler-Smith C.* The human Y chromosome: an evolutionary marker comes of age // *Nat. Rev. Genet.* 2003. V. 4. P. 598–612. <https://doi.org/10.1038/nrg1124>
3. *Karmin M., Saag L., Vicente M. et al.* A recent bottleneck of Y chromosome diversity coincides with a global change in culture // *Genome Res.* 2015. P. 459–466. <https://doi.org/10.1101/gr.186684.114>
4. *Underhill P.A.* Inferring human history: clues from Y-chromosome haplotypes // *Cold Spring Harb. Symp. Quant. Biol.* 2003. V. 68. P. 487–493. <https://doi.org/10.1101/sqb.2003.68.487>
5. *Балановский О.П., Запороженко В.В.* Хромосома-летописец: датировки генетики, события истории, соблазн ДНК-генеалогии // *Генетика.* 2016. Т. 52. № 7. С. 810–830. <https://doi.org/10.1134/S1022795416070048>
6. *Jobling M.A., Tyler-Smith C.* New uses for new haplotypes the human Y chromosome, disease and selection // *Trends in Genetics.* 2000. V. 16. P. 356–362. [https://doi.org/10.1016/s0168-9525\(00\)02057-6](https://doi.org/10.1016/s0168-9525(00)02057-6)
7. *Dupanloup I., Pereira L., Bertorelle G. et al.* A recent shift from polygyny to monogamy in humans is suggested by the analysis of worldwide Y-chromosome diversity // *Mol. Evol.* 2003. V. 57. P. 85–97. <https://doi.org/10.1007/s00239-003-2458-x>
8. *Dupanloup I., Pereira L., Bertorelle G. et al.* A recent shift from polygyny to monogamy in humans is suggested by the analysis in worldwide Y-chromosome diversity // *J. Mol. Evol.* 2003. V. 57. P. 85–97. <https://doi.org/10.1007/s00239-003-2458-x>
9. <http://www.isogg.org/>.
10. <http://www.yfull.com/>.
11. *Ballantyne K.N., Keel V., Wollstein A. et al.* A new future of forensic Y-chromosome analysis: Rapidly mutating Y-STRs for differentiating male relatives and paternal lineages // *Forensic. Sci. Int. Genet.* 2012. V. 6. P. 208–218. <https://doi.org/10.1016/j.fsigen.2011.04.017>

12. Burgarella C., Navascues M. Mutation rate estimates for 110 Y-chromosome STRs combining population and father-son pair data // *Eur. J. Hum. Genet.* 2011. V. 19. P. 70–75.
<https://doi.org/10.1038/ejhg.2010.154>
13. Casanova M., Leroy P., Boucekkine C. et al. A human Y-linked DNA polymorphism and its potential for estimating genetic and evolutionary distance // *Science*. 1985. № 230. P. 1403–1406.
<https://doi.org/10.1126/science.2999986>
14. Ngo K.Y., Vergnaud G., Johnsson C. et al. A DNA probe detecting multiple haplotypes of the human Y chromosome // *Am. J. Hum. Genet.* 1986. V. 38. P. 407–418. PMID: PMC1684810
15. Rosser Z.H., Balaesque P., Jobling M.A. Gene conversion between the X chromosome and the male-specific region of the Y chromosome at a translocation hotspot // *Am. J. Hum. Genet.* 2009. V. 85. P. 130–134.
<https://doi.org/10.1016/j.ajhg.2009.06.009>
16. Krausz C., Quintana-Murci L., Rajpert-De Meyts E. et al. Identification of a Y chromosome haplogroup associated with reduced sperm counts // *Hum. Mol. Genet.* 2001. V. 10. P. 1873–1877.
<https://doi.org/10.1093/hmg/10.18.1873>
17. King T.E., Ballereau S.J., Schürer K.E., Jobling M.A. Genetic signatures of coancestry within surnames // *Curr. Biol.* 2006. V. 16. P. 384–388.
<https://doi.org/10.1016/j.cub.2005.12.048>
18. Hammer M.F. A recent insertion of an Alu element on the Y chromosome is a useful marker for human population genetics // *Mol. Biol. Evol.* 1994. V. 11. P. 749–761.
<https://doi.org/10.1093/oxfordjournals.molbev.a040155>
19. Underhill P.A., Li Jin, Zemans R. et al. A pre-Columbian Y chromosome-specific transition and its implications for human evolutionary history // *Proc. Natl Acad. Sci. USA*. 1996. V. 93. P. 196–200.
<https://doi.org/10.1073/pnas.93.1.19>
20. Underhill P.A., Li Jin, Lin A.A. et al. Detection of numerous Y chromosome biallelic polymorphisms by denaturing high-performance liquid chromatography // *Genome Res.* 1997. V. 7(10). P. 996–1005.
<https://doi.org/10.1101/gr.7.10.996>
21. Underhill P.A., Myres N.M., Rootsi S. et al. Separating the post-Glacial coancestry of European and Asian Y chromosomes within haplogroup R1a // *Eur. J. Hum. Genet.* 2010. V. 18. P. 479–484.
<https://doi.org/10.1038/ejhg.2009.194>
22. Cruciani F., Trombetta B., Massaia A. et al. A revised root for the human Y chromosome phylogenetic tree: The origin of patrilineal diversity in Africa // *Am. J. Hum. Genet.* 2011. V. 88. № 6. P. 814–818.
<https://doi.org/10.1016/j.ajhg.2011.05.002>
23. Illumäe A.-M., Reidla M., Chukhryaeva M. et al. Human Y-chromosome haplogroup N: A non-trivial time-resolved phylogeography that cuts across language families // *Am. J. Hum. Genet.* 2016. V. 99. P. 163–173.
<https://doi.org/10.1016/j.ajhg.2016.05.025>
24. The Y-chromosome Consortium. A nomenclature system for the tree of human Y-chromosomal binary haplogroups // *Genome Res.* 2002. V. 12. P. 339–348.
<https://doi.org/10.1101/gr.217602>
25. Cinnioglu C., King R., Kivisild T. et al. Excavating Y-chromosome haplotype strata in Anatolia // *Hum. Genet.* 2004. V. 114. P. 127–148.
<https://doi.org/10.1007/s00439-003-1031-4>
26. Semino O., Magri M., Benuzzi G. et al. Origin, diffusion, and differentiation of Y-chromosome haplogroups E and J: Inferences on the neolithization of Europe and Later Migratory Events in the Mediterranean Area // *Am. J. Hum. Genet.* 2004. V. 74. P. 1023–1034.
<https://doi.org/10.1086/386295>
27. Rootsi S., Magri C., Kivisild T. et al. Phylogeography of Y-chromosome haplogroup I reveals distinct domains of prehistoric gene flow in Europe // *Am. J. Hum. Genet.* 2004. V. 75. P. 128–137.
<https://doi.org/10.1086/422196>
28. Myres N.M., Rootsi S., Lin A. et al. A major Y-chromosome haplogroup R1b Holocene era founder effect in Central and Western Europe // *Eur. J. Hum. Genet.* 2011. V. 19. P. 95–101.
<https://doi.org/10.1038/ejhg.2010.146>
29. Kivisild T. Maternal ancestry and population history from whole mitochondrial genomes // *Investig Genet.* 2015. V. 6(3).
<https://doi.org/10.1186/s13323-015-0022-2>
30. Животовский Л.А. Микросателлитная изменчивость в популяциях человека и методы ее изучения // *Вестник ВОГиС*. 2006. Т. 10. № 1. С. 74–96.
31. Underhill P.A., Passarino G., Lin A.A. et al. The phylogeography of Y chromosome binary haplotypes and the origins of modern human populations // *Ann. Hum. Genet.* 2001. V. 65. P. 43–62.
<https://doi.org/10.1046/j.1469-1809.2001.6510043.x>
32. Thomas M.G., Skorecki K., Ben-Ami H. et al. Origins of old testament priests // *Nature*. 1998. V. 394. P. 138–140.
<https://doi.org/10.1038/28083>
33. Stumpf M.P.H., Goldstein D.B. Genealogical and evolutionary inference with human Y chromosome // *Science*. 2001. V. 291. P. 1738–1742.
<https://doi.org/10.1126/science.291.5509.1738>
34. Zhivotovsky L.A., Feldman M.W. Microsatellite variability and genetic distances // *Proc. Natl Acad. Sci. USA*. 1995. V. 92. P. 11549–11552.
<https://doi.org/10.1073/pnas.92.25.11549>
35. Goldstein D.B., Zhivotovsky L.A., Nayar K. et al. Statistical properties of the variation at linked microsatellite loci: Implications for the history of human Y chromosomes // *Mol. Biol. Evol.* 1996. V. 13. P. 1213–1218.
36. Jorde L.B., Rogers A.R., Bamshad M. et al. Microsatellite diversity and the demographic history of modern humans // *Proc. Natl Acad. Sci. USA*. 1997. V. 94. P. 3100–3103.
<https://doi.org/10.1073/pnas.94.7.3100>
37. Lahermo P., Savontaus M.L., Sistonen P. et al. Y-chromosomal polymorphisms reveal founding lineages in the Finns and the Saami // *Eur. J. Hum. Genet.* 1999. V. 7. P. 447–458.
<https://doi.org/10.1038/sj.ejhg.5200316>
38. Пузырев В.П., Степанов В.А., Голубенко М.В. и др. Линии мтДНК и Y-хромосомы в популяции якутов // *Генетика*. 2003. Т. 39. № 7. С. 975–981.

39. *Hammer M.F., Karafet T.M., Park H. et al.* Dual origins of the Japanese: Common ground for hunter-gatherer and farmer Y chromosomes // *J. Hum. Genet.* 2006. V. 51. P. 47–58.
<https://doi.org/10.1007/s10038-005-0322-0>
40. *Юсупов Ю.М., Балановская Е.В., Жабагин М.К. и др.* Генофонд юго-западных башкир по маркерам Y-хромосомы: опыт междисциплинарного анализа // *Генетика*. 2018. Т. 54. № 13. Приложение. С. S94–S97.
<https://doi.org/10.1134/S0016675818130222>
41. *Балановская Е.В., Балановский О.П.* Русский генофонд на русской равнине. М.: ООО “Луч”, 2007. 416 с.
42. *Zhivotovsky L.A.* Estimating divergence time with the use of microsatellite genetic distances: Impacts of population growth and gene flow // *Mol. Biol. Evol.* 2001. № 18. P. 700–709.
<https://doi.org/10.1093/oxfordjournals.molbev.a003852>
43. *Zhivotovsky L.A., Underhill P.A., Cinnioglu C. et al.* The effective mutation rate at Y chromosome short tandem repeats, with application to human population-divergence time // *Am. J. Hum. Genet.* 2004. V. 74. P. 50–61.
<https://doi.org/10.1086/380911>
44. *Xu X., Peng M., Fang Z., Xu X.* The direction of microsatellite mutations is dependent upon allele length // *Nat. Genet.* 2000. V. 24. P. 396–399.
<https://doi.org/10.1038/74238>
45. *Huang, Q.Y., Xu F.H., Shen H. et al.* Mutation patterns at dinucleotide microsatellite loci in humans // *Am. J. Hum. Genet.* 2002. V. 70. P. 625–634.
<https://doi.org/10.1086/338997>
46. *Dupuy B.M., Stenersen M., Egeland T., Olaisen B.* Y-chromosomal microsatellite mutation rates: Differences in mutation rate between and within loci // *Hum. Mutat.* 2004. V. 23. P. 117–124.
<https://doi.org/10.1002/humu.10294>
47. *Garza J.C., Freimer N.B.* Homoplasmy for size at microsatellite loci in humans and chimpanzees // *Genome Res.* 1996. V. 6. P. 211–217.
<https://doi.org/10.1101/gr.6.3.211>
48. *Feldman M.W., Kumm J., Pritchard J.K.* Mutation and migration in models of microsatellite evolution // *Microsatellites: Evolution and Applications* / Eds Goldstein D.G., Schlotterer C. Oxford Univ. Press, 1999. P. 98–115.
49. *Zerjal T., Dashnyam B., Pandya A. et al.* Genetic relationships of Asians and Northern Europeans, revealed by Y-chromosomal DNA analysis // *Am. J. Hum. Genet.* 1997. V. 60(5). P. 1174–1183. PMID: 9150165. PMCID: PMC1712423.
50. *Степанов В.А.* Этногеномика населения Северной Евразии. Томск: Печатная мануфактура, 2002. 244 с.
51. *Karafet T.M., Osipova L.P., Gubina M.A. et al.* High levels of Y-chromosome differentiation among native Siberian populations and the genetic signature of a boreal hunter-gatherer way of life // *Hum. Biol.* 2002. V. 74. № 6. P. 761–789.
<https://doi.org/10.1353/hub.2003.0006>
52. *Derenko M., Malyarchuk B., Denisova G. et al.* Contrasting patterns of Y-chromosome variation in South Siberian population from Baikal and Altai-Sayan regions // *Hum. Genet.* 2006. V. 118. P. 591–604.
<https://doi.org/10.1007/s00439-005-0076-y>
53. *Derenko M., Malyarchuk B., Denisova G. et al.* Y chromosome haplogroup N dispersals from south Siberia to Europe // *J. Hum. Genet.* 2007. V. 52(9). P. 763–770.
<https://doi.org/10.1016/j.ajhg.2016.05.025>
54. *Zerjal T., Xue Y., Bertorelle G. et al.* The genetic legacy of the Mongols // *Am. J. Hum. Genet.* 2003. V. 72. P. 717–721.
<https://doi.org/10.1086/367774>
55. *Chaix R., Austerlitz F., Khégay T. et al.* The genetic or mythical ancestry of descent groups: Lessons from the Y chromosome // *Am. J. Hum. Genet.* 2004. V. 75. P. 1113–1116.
<https://doi.org/10.1086/425938>
56. *Xue Y., Zerjal T., Bao W. et al.* Male demography in East Asia: A North-South contrast in human population expansion times // *Genetics*. 2006. V. 172. P. 2431–2439.
<https://doi.org/10.1534/genetics.105.054270>
57. *Балановская Е.В., Юсупов Ю.М., Схаляхо Р.А. и др.* Генетические портреты семи кланов северо-западных башкир: вклад финно-угорского компонента в генофонд башкир // *Вестник Моск. ун-та. Серия XXIII. Антропология*. 2017. № 3. С. 94–103.
58. *Abilev S., Malyarchuk B., Derenko M. et al.* The Y-chromosome C3*Star-Cluster attributed to Genghis Khan’s Descendants Is Present at High frequency in the Kerey Clan from Kazakhstan // *Hum. Biol.* 2012. V. 84(1). P. 79–89.
<https://doi.org/10.3378/027.084.0106>
59. *Zhabagin M.K., Dibirova H.D., Frolova S.A. et al.* The relation between the Y-chromosomal variation and the clan structure: The gene pool of the steppe aristocracy and the steppe clergy of the Kazakhs // *Moscow Univ. Anthropol. Bul.* 2014. V. 1. P. 96–101.
60. *Схаляхо Р.А., Жабагин М.К., Юсупов Ю.М. и др.* Генофонд туркмен Каракалпакстана в контексте популяций Центральной Азии (полиморфизм Y-хромосомы) // *Вестник Моск. ун-та. Серия XXIII. Антропология*. 2016. № 3. С. 83–93.
61. *Харьков В.Н., Новикова Л.М., Лузина Ф.А. и др.* Генофонд хакасов и шорцев по маркерам Y-хромосомы: общие компоненты и генетическая структура родов // *Генетика* 2020. Т. 56. № 7. С. 826–833.
<https://doi.org/10.31857/S0016675820070073>
62. *Дамба Л.Д., Балановская Е.В., Жабагин М.К. и др.* Оценка вклада монгольской экспансии в генофонд тувинцев // *Вавиловский журн. генетики и селекции*. 2018. Т. 22(5). С. 611–619.
<https://doi.org/10.18699/VJ18.402>
63. *Keyser C., Bouakaze C., Crubezy E. et al.* Ancient DNA provides new insights into the history of south Siberian Kurgan people // *Hum. Genet.* 2009. V. 126. P. 395–410.
<https://doi.org/10.1007/s00439-009-0683-0>
64. *Raghavan M., Skoglund P., Graf K.E. et al.* Upper palaeolithic siberian genome reveals dual ancestry of Native Americans // *Nature*. 2014. V. 505(7481). P. 87–91.
<https://doi.org/10.1038/nature12736>
65. *Lazaridis I., Patterson N., Mitnik A. et al.* Ancient human genomes suggest three ancestral populations for

- present-day Europeans // *Nature*. 2014. V. 513(7518). P. 409–413.
<https://doi.org/10.1038/nature13673>
66. *Balanovsky O., Zhabagin M., Agdzhoyan A. et al.* Deep phylogenetic analysis of haplogroup G1 provides estimates of SNP and STR mutation rates on the human Y-chromosome and reveals migrations of Iranic speakers // *PLoS One*. 2015. V. 10(4): e0122968.
<https://doi.org/10.1371/journal.pone.0122968>
 67. *Allentoft M.E., Sikora M., Sjögren K.G. et al.* Population genomics of Bronze Age Eurasia // *Nature*. 2015. V. 522(7555). P. 167–172.
<https://doi.org/10.1038/nature14507>
 68. *Damgaard P., Martiniano R., Kamm J. et al.* The first horse herders and the impact of early Bronze age steppe expansions into Asia // *Science*. 2018. V. 360(6396). P. 1422.
<https://doi.org/10.1126/science.aar7711>
 69. *Hollard C., Zvenigorosky V., Kovalev A. et al.* New genetic evidence of affinities and discontinuities between bronze age Siberian populations // *Am. J. Phys. Anthropol.* 2018. № 167. P. 97–107.
<https://doi.org/10.1002/ajpa.23607>
 70. *Lipson M., Szécsényi-Nagy A., Mallick S. et al.* Parallel palaeogenomic transects reveal complex genetic history of early European farmers // *Nature*. 2017. V. 551(7680). P. 368–372.
<https://doi.org/10.1038/nature24476>
 71. *Mathieson I., Alpaslan-Roodenberg S., Posth C. et al.* The genomic history of southeastern Europe // *Nature* 2018. V. 555(7695). P. 197–203.
<https://doi.org/10.1038/nature25778>
 72. *Saag L., Varul L., Scheib C.L. et al.* Extensive farming in Estonia started through a sex-biased migration from the steppe // *Curr. Biol.* 2017. V. 27(14). P. 2185–2193.
<https://doi.org/10.1016/j.cub.2017.06.022>
 73. *Amorim A., Pinto N.* Big data in forensic genetics // *Forensic Sci. Int. Genet.* 2018. № 37. P. 102–105.
<https://doi.org/10.1016/j.fsigen.2018.08.001>
 74. *Харьков В.Н., Зарубин А.А., Вагайцева К.В. и др.* Y-хромосома как инструмент для ДНК-идентификации и определения популяционной принадлежности // *Генетика*. 2020. Т. 56. № 9. С. 1056–1074.
<https://doi.org/10.31857/S0016675820090118>

Y-Chromosome Markers in Population Genetics: Fundamental and Applied Results of Ethnogenomic Research

V. N. Kharkov*

*Tomsk National Research Medical Center of the Russian Academy of Sciences,
Research Institute of Medical Genetics, Tomsk, 634050 Russia*

*e-mail: vladimir.kharkov@medgenetics.ru

The Y-chromosome is a unique tool for studying the structure of gene pools and the evolutionary history of human populations. Patrilinearity ensures the sequential accumulation of mutations, as a result of which the Y-chromosome is a fairly simple marker system for reconstructing evolutionary changes, migration, and other demographic processes in human populations. The analysis of Y-chromosome haplogroups can be effectively applied both for the analysis of the genetic and demographic history of various populations and for large-scale phylogenetic and phylogeographic studies of individual monophyletic lineages of different hierarchical levels. The discovery of more and more population-specific DNA markers of this part of the human genome makes it possible to analyze in great detail the population, subethnic and generic structure of various ethnic groups throughout all periods of its formation. Determination of the allelic profile of male samples in works on paleogenetics is especially important for the reconstruction of migrations, component composition, continuity of the genetic heritage of ancient populations, and clarification of the phylogeny of various sublines. The accumulation of a large array of data on the distribution of various haplogroups specific for different populations and the features of their profiles by SNP and STR markers, in addition to solving fundamental problems, allows us to successfully solve practical problems on the genetic genealogy of each man, as well as the ethnic identification of an unknown individual based on his DNA sample and the development systems for DNA identification.

Keywords: Y-chromosome, gene pool, population, genetic diversity, ethnogenomics, paleogenetics, DNA identification.