

Перевод на английский язык <https://vavilov.elpub.ru/jour>

## Связь генофонда хантов с народами Западной Сибири, Предуралья и Алтая-Саян по данным о полиморфизме аутомомных локусов и Y-хромосомы

В.Н. Харьков , Н.А. Колесников, Л.В. Валихова, А.А. Зарубин, М.Г. Сваровская, А.В. Марусин, И.Ю. Хитринская, В.А. Степанов

Научно-исследовательский институт медицинской генетики, Томский национальный исследовательский медицинский центр Российской академии наук, Томск, Россия  
 [vladimir-kharkov@medgenetics.ru](mailto:vladimir-kharkov@medgenetics.ru)

**Аннотация.** Ханты – коренной сибирский народ, проживающий на территории Западной Сибири, в основном на территории Ханты-Мансийского и Ямало-Ненецкого автономных округов. Настоящее исследование направлено на комплексный анализ структуры генофонда хантов и их сравнение с другими популяциями коренного населения Южной и Западной Сибири. Для решения вопросов генетической близости хантов с другими коренными народами выполнено генотипирование широкого геномного набора аутомомных маркеров с помощью высокоплотных биочипов, а также расширенного набора SNP- и STR-маркеров Y-хромосомы у различных этнических групп: хакасов, тувинцев, южных алтайцев, сибирских татар, чулымцев (тюркская языковая семья) и кетов (енисейская языковая семья). Результаты анализа частот аутомомных SNP различными методами, сходства по составу гаплогрупп Y-хромосомы и YSTR-гаплотипов свидетельствуют, что генофонд хантов достаточно специфичен. При анализе аутомомных SNP в обеих выборках полностью доминирует угорский генетический компонент (до 99–100 %). Выборки хантов показали максимальное совпадение по IBD-блокам между собой, с выборкой кетов, чулымцев, тувинцев, томских татар, хакасов-качинцев и южных алтайцев. Степень совпадения IBD-блоков между хантами, кетами и томскими татарами согласуется с результатами распределения в этих популяциях частот аллелей и общих генетических компонентов. По составу гаплогрупп Y-хромосомы две выборки хантов значительно различаются между собой. Детальный филогенетический анализ различных гаплогрупп Y-хромосомы позволил описать и уточнить различия в филогении и структуре отдельных этноспецифичных сублиний, определить их родство, следы экспансии численности в генофонде хантов. Варианты разных гаплогрупп Y-хромосомы у хантов, хакасов и тувинцев восходят к общим для них предковым линиям. Результаты сравнительного анализа образцов мужчин также свидетельствуют о близком генетическом родстве между хантами и ненцами, коми, удмуртами и кетами. Специфичность гаплотипов, обнаружение различных терминальных SNP подтверждают, что ханты достаточно долго не имели контактов с другими этносами, кроме ненцев, в состав которых вошло много хантыйских родов. Ключевые слова: генофонд; популяции человека; генетическое разнообразие; генетические компоненты; Y-хромосома; ханты.

**Для цитирования:** Харьков В.Н., Колесников Н.А., Валихова Л.В., Зарубин А.А., Сваровская М.Г., Марусин А.В., Хитринская И.Ю., Степанов В.А. Связь генофонда хантов с народами Западной Сибири, Предуралья и Алтая-Саян по данным о полиморфизме аутомомных локусов и Y-хромосомы. *Вавиловский журнал генетики и селекции*. 2023;27(1):46-54. DOI 10.18699/VJGB-23-07

## Relationship of the gene pool of the Khants with the peoples of Western Siberia, Cis-Urals and the Altai-Sayan Region according to the data on the polymorphism of autosomic locus and the Y-chromosome

V.N. Kharkov , N.A. Kolesnikov, L.V. Valikhova, A.A. Zarubin, M.G. Svarovskaya, A.V. Marusin, I.Yu. Khitrinskaya, V.A. Stepanov

Research Institute of Medical Genetics, Tomsk National Research Medical Center of the Russian Academy of Sciences, Tomsk, Russia  
 [vladimir-kharkov@medgenetics.ru](mailto:vladimir-kharkov@medgenetics.ru)

**Abstract.** Khanty are indigenous Siberian people living on the territory of Western Siberia, mainly on the territory of the Khanty-Mansiysk and Yamalo-Nenets Autonomous Okrugs. The present study is aimed at a comprehensive analysis of the structure of the Khanty gene pool and their comparison with other populations of the indigenous population of Southern and Western Siberia. To address the issues of genetic proximity of the Khanty with other indigenous peoples, we performed genotyping of a wide genomic set of autosomal markers using high-density biochips, as well

as an expanded set of SNP and STR markers of the Y-chromosome in various ethnic groups: Khakas, Tuvans, Southern Altaians, Siberian Tatars, Chulymys (Turkic language family) and Kets (Yeniseian language family). The structure of the gene pool of the Khanty and other West Siberian and South Siberian populations was studied using a genome-wide panel of autosomal single nucleotide polymorphic markers and Y-chromosome markers. The results of the analysis of autosomal SNPs frequencies by various methods, the similarities in the composition of the Y-chromosome haplogroups and YSTR haplotypes indicate that the Khanty gene pool is quite specific. When analyzing autosomal SNPs, the Ugrian genetic component completely dominates in both samples (up to 99–100 %). The samples of the Khanty showed the maximum match in IBD blocks with each other, with a sample of the Kets, Chulymys, Tuvans, Tomsk Tatars, Khakas, Kachins, and Southern Altaians. The degree of coincidence of IBD blocks between the Khanty, Kets, and Tomsk Tatars is consistent with the results of the distribution of allele frequencies and common genetic components in these populations. According to the composition of the Y-chromosome haplogroups, the two samples of the Khanty differ significantly from each other. A detailed phylogenetic analysis of various Y-chromosome haplogroups made it possible to describe and clarify the differences in the phylogeny and structure of individual ethnospecific sublines, to determine their relationship, traces of population expansion in the Khanty gene pool. Variants of different haplogroups of the Y-chromosome in the Khanty, Khakas and Tuvans go back to their common ancestral lines. The results of a comparative analysis of male samples indicate a close genetic relationship between the Khanty and Nenets, Komi, Udmurts and Kets. The specificity of haplotypes, the discovery of various terminal SNPs confirms that the Khanty did not come into contact with other ethnic groups for a long time, except for the Nenets, which included many Khanty clans. Key words: gene pool; human population; genetic diversity; genetic components; Y-chromosome; Khanty.

**For citation:** Kharkov V.N., Kolesnikov N.A., Valikhova L.V., Zarubin A.A., Svarovskaya M.G., Marusin A.V., Khitrinskaya I.Yu., Stepanov V.A. Relationship of the gene pool of the Khants with the peoples of Western Siberia, Cis-Urals and the Altai-Sayan Region according to the data on the polymorphism of autosomal locus and the Y-chromosome. *Vavilovskii Zhurnal Genetiki i Seleksii = Vavilov Journal of Genetics and Breeding*. 2023;27(1):46-54. DOI 10.18699/VJGB-23-07

## Введение

Изучение структуры генофондов популяций различных сибирских регионов – одно из приоритетных направлений современной генетики человека, оно помогает раскрыть детально часть вопросов, связанных с их этногенезом.

Ханты – коренной народ, проживающий на территории Западной Сибири, в основном на территории Ханты-Мансийского (ХМАО) и Ямало-Ненецкого (ЯНАО) автономных округов, а также Тюменской области. Небольшие группы хантов проживают на севере Томской области и в Республике Коми. По данным Всероссийской переписи 2010 г., численность хантов составила 30943 человек, из которых 61.6 % проживают в ХМАО и 30.7 % – в ЯНАО. У хантов выделяются три большие этнографические группы, совпадающие с группами их языковых диалектов – северные, южные и восточные. Причем южные (прииртышские) ханты были тюркизированы и вошли в состав сибирских татар, смешавшись с ними, а также ассимилированы русскими переселенцами (Народы Западной Сибири, 2005).

Популяции хантов представляют значительный интерес для популяционно-генетических исследований как в силу относительно слабой их изученности с привлечением современных геномных технологий, так и по причине специфичности генофондов отдельных групп, развивавшихся в условиях длительной генетической изоляции.

Расселение хантов в древности было очень широким – от низовий Оби на севере до Барабинских степей на юге и от Енисея на Востоке до Зауралья, включая реки Северная Сосьва и Ляпин, а также часть рек Пелым и Конда на западе. С XIX в. за Урал из Прикамья и Приуралья стали переселяться манси, которых теснили коми-зыряне и русские. С более раннего времени к северу уходила и часть южных манси в связи с созданием в XIV–XV вв. Тюменского и Сибирского ханств – государств сибирских татар, а позднее (XVI–XVII вв.) и с освоением Сибири русскими. В XVII–XVIII вв. на Пелыме и Конде уже жили манси.

Часть хантов также переселялись из западных областей на восток и север (на Обь с ее левых притоков), это фиксируется статистическими данными архивов. Их места занимали манси. Так, к концу XIX в. на реках Северная Сосьва и Ляпин не осталось остяцкого населения, которое либо переселилось на Обь, либо слилось с новоприсельцами (Народы России, 1994).

На севере ханты вступали в контакт с ненцами, часть их была ассимилирована ими, что подтверждается этнографическими данными, а также нашим исследованием родовой структуры гыданских ненцев по маркерам Y-хромосомы (Харьков и др., 2021). Миграция хантов на север и восток продолжалась и в XX в. К XX в. южные ханты были почти полностью ассимилированы татарами и русскими.

Исторически сложилось так, что хантыйское население не было однородным ни по языку, ни по культуре. Одни ученые разделяют хантыйский язык на две крупные группы – западную и восточную, а другие еще подразделяют западные диалекты на южные и северные. В антропологическом отношении ханты – наиболее характерные представители уральского типа, к которому относятся также манси, селькупы, ненцы, барабинские татары, шорцы, северные алтайцы и хакасы. Самые близкие родственники хантов по происхождению, языку и культуре – манси (Брук, 1986).

Целью настоящего исследования стали комплексный анализ структуры генофонда хантов и реконструкция их происхождения в сравнении с другими популяциями коренного населения Южной и Западной Сибири. Для решения вопросов генетической близости хантов с другими коренными народами было выполнено генотипирование широкого геномного набора аутосомных маркеров с помощью высокоплотных биочипов, а также расширенного набора SNP- и STR-маркеров Y-хромосомы у различных этнических групп: хакасов, тувинцев, южных алтайцев, сибирских татар, чулымцев (тюркская языковая семья) и кетов (енисейская языковая семья).

## Материалы и методы

Материалом исследования были образцы ДНК мужчин и женщин из двух популяций хантов деревни Русскинская, Сургутского района и села Казым, Белоярского района ХМАО. Забор первичного биологического материала (венозной крови) у доноров производили с соблюдением процедуры письменного информированного согласия на проведение исследования. На каждого донора составляли анкету с краткой родословной, указанием этнической принадлежности и места рождения предков. Индивиды относили к данной этнической группе на основании этнической идентификации – его собственной, его родителей, и места рождения.

Для анализа Y-хромосомных гаплогрупп и гаплотипов хантов использованы 120 образцов ДНК мужчин д. Русскинская ( $N = 64$ ) и с. Казым ( $N = 54$ ) ХМАО. Для генотипирования на высокоплотных микрочипах были выбраны неродственные образцы с. Казым ( $N = 30$ ) и д. Русскинская ( $N = 26$ ). Другие популяции коренного населения Сибири представлены: чулымцами ( $N = 22$ ), хакасами (сагайцами Таштыпского района,  $N = 29$ , и качинцами Ширинского района,  $N = 26$ ), южными алтайцами (с. Бешпельтир Чемальского района,  $N = 24$ , и с. Кулада Онгудайский района,  $N = 25$ ), кетами (пос. Келлог Туруханского района Красноярского края,  $N = 15$ ), томскими татарами (пос. Черная Речка, пос. Эушта и пос. Тахтамышево Томского района,  $N = 20$ ), тувинцами (с. Тээли Бай-Тайгинского кожууна,  $N = 28$ ).

Данные широкогеномного генотипирования получены с использованием микрочипов Infinium Multi-Ethnic Global-8 (Illumina) для SNP-генотипирования, включающего свыше 1.7 млн маркеров. Материал депонирован в биоресурсной коллекции «Биобанк населения Северной Евразии».

Кластеризацию массива генотипов аутосомных SNP (single nucleotide polymorphism) и контроль качества выполняли с помощью протокола, разработанного Y. Guo с коллегами (2014), с использованием GenomeStudio (Illumina. GenomeStudio, модуль генотипирования v2.0.3), программного пакета, который компания Illumina создала для различных геномных анализов. Для фильтрации, нормализации и расчета стандартных геномных статистик и показателей оптимальным был стандартный набор программ, включающий vcftools, bcftools, и plink.

Для анализа идентичных по происхождению блоков сцепления использовали алгоритм Refined IBD (Browning B.L., Browning S.R., 2013), показывающий более точные результаты по сравнению с встроенными в plink алгоритмами. Предварительно генотипы были фазированы с помощью программного обеспечения Beagle 5.1 (Browning S.R., Browning B.L., 2007). Для сравнения популяций были получены суммы средних длин идентичных по происхождению блоков (сегментов IBD – identical by descent) между парами индивидов.

Для анализа генетических взаимоотношений между популяциями применяли метод tSNE. Для определения компонентного состава и количества примесей у отдельных индивидов и популяций использованы методика NGSadmix (Skotte et al., 2013) и программа ADMIXTURE (Alexander et al., 2009; Alexander, Lange, 2011).

Для изучения состава и структуры гаплогрупп Y-хромосомы в исследование были включены две системы генетических маркеров: диаллельных локусов, представленных SNP, и полиаллельных высоковариабельных микросателлитов (YSTR). С помощью 138 SNP-маркеров определяли принадлежность образцов к различным гаплогруппам. Классификация гаплогрупп соответствует данным Международного общества генетической генеалогии (сайт [www.isogg.org](http://www.isogg.org)).

Анализ STR-гаплотипов внутри гаплогрупп проводили с использованием 44 STR-маркеров нерекombинирующей части Y-хромосомы (DYS19, 385a, 385b, 388, 389I, 389II, 390, 391, 392, 393, 426, 434, 435, 436, 437, 438, 439, 442, 444, 445, 448, 449, 456, 458, 460, 461, 481, 504, 505, 518, 525, 531, 533, 537, 552, 570, 576, 635, 643, YCAIIa, YCAIIb, GATA H4.1, Y-GATA-A10, GGAAT1B07). STR-маркеры генотипировали с помощью капиллярного электрофореза на генетическом анализаторе ABI Prism 3730. Генотипирование SNP-маркеров проводили с применением полимеразной цепной реакции и последующего анализа фрагментов ДНК с помощью ПДРФ-анализа (полиморфизм длин рестрикционных фрагментов).

Экспериментальные исследования осуществлены на базе Центра коллективного пользования научно-исследовательским оборудованием «Медицинская геномика» (НИИ медицинской генетики Томского НИМЦ). Построение медианных сетей гаплотипов Y-хромосомы выполняли с использованием программы Network v.10.2.0.0 (Fluxus Technology Ltd; [www.fluxus-engineering.com](http://www.fluxus-engineering.com)) по методу медианных сетей Бандельта (Bandelt et al., 1999). Оценку возраста генерации наблюдаемого разнообразия гаплотипов в гаплогруппах проводили методом ASD (Zhivotovskiy et al., 2004) на основании средних квадратичных отличий в числе повторов между всеми маркерами.

## Результаты и обсуждение

Полученный в результате генотипирования высокоплотных микрочипов большой массив данных по аутосомным SNP в выборках хантов и других коренных сибирских народов позволяет максимально детально охарактеризовать генофонд исследованных выборок с применением различных методик. Генотипирование расширенного набора специфичных SNP Y-хромосомы из различных гаплогрупп дает возможность гораздо точнее описать молекулярно-филогенетическую и филогеографическую структуры отдельных гаплогрупп Y-хромосомы.

После обработки данных по результатам микрочипового исследования для фильтрации прогенотипированных образцов и выполнения дальнейших расчетов был проведен поиск среди хантов метисов с использованием программы NGSadmix. Алгоритм этой программы позволяет определять соотношение предковых компонентов по данным NGS с относительно малой глубиной покрытия. Принцип расчетов аналогичен другим программам, таким как FRAPPE и ADMIXTURE, но NGSadmix, в отличие от них, эффективно работает, когда имеется статистическая неточность в отдельных генотипах. Метод NGSadmix при запуске на сформированном нами массиве данных показал, что практически все образцы хантов не имеют метисации, что полностью соответствует данным анкетирования.

рования ДНК-доноров. Метисация с русскими (до 30 %) обнаружена только для одного мужчины из д. Русскинская. Его принадлежность к европейской Y-хромосомной линии R1b1a1b-L407 подтверждает метисацию с отцовской стороны. Этот образец был исключен из дальнейших расчетов.

Полученные данные о частотах SNP в исследованных образцах были использованы для выяснения генетических взаимоотношений между включенными в работу популяционными выборками. Для понижения размерности, пространственного анализа и выявления генетических компонентов мы остановились на двух алгоритмах: tSNE и ADMIXTURE. Метод tSNE позволяет более четко разделить массив данных на отдельные этноспецифичные группы образцов по сравнению с методом главных компонент (PCA).

### Генетические взаимоотношения хантов с другими популяциями Западной и Южной Сибири

При анализе массива данных по частотам аутомсомных SNP с помощью метода tSNE на уровне отдельных образцов показано, что две выборки хантов очень близки, при этом образцы казымских и русскинских хантов не пересекаются на графике и отделены друг от друга (рис. 1).

Ханты характеризуются специфическими особенностями генофонда и не кластеризуются с другими популяциями. По сравнению с субэтническими группами хакасов и южных алтайцев из разных поселков более удаленные географически выборки хантов демонстрируют максимальную генетическую близость. Ближе всего к хантам расположены выборки кетов и томских татар. Генетические расстояния между хантами и популяциями Южной Сибири значительно больше. Этнически и территориально близкие друг другу выборки располагаются на рис. 1 довольно близко, но каждый образец входит в отдельный этноспецифичный кластер. Исключение составляют лишь несколько единичных образцов хакасов.

**Компонентный состав генофонда популяций.** Современные методы, применяемые в геномных исследованиях, и новые биоинформационные подходы дают возможность достоверно выявлять предковые генетические компоненты разного происхождения в составе генофонда различных популяций и отдельных индивидов. Для обнаружения индивидуальных генетических компонентов в генофонде изучаемых популяций использована программа ADMIXTURE, позволяющая выявлять смешанность состава набора индивидов на основе данных о генотипах и тем самым строить предположения о происхождении популяции.

Моделирование с помощью ADMIXTURE в последнее время считается одним из основных методов анализа при исследовании генофондов современных и древних популяций человека, позволяет анализировать одни и те же данные на разных иерархических уровнях. При задании числа предковых компонентов больше двух в большинстве исследованных популяций обнаруживается специфичный для хантов генетический компонент, наиболее отчетливо проявляющийся на анализируемом массиве популяционных выборок при  $K = 8$ , который можно интерпретировать как «угорский» генетический пласт в генофонде современных популяций. Ханты характеризуются

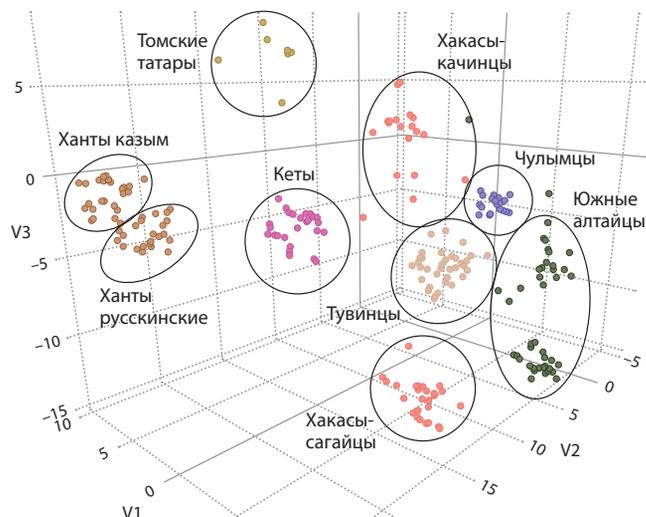


Рис. 1. Дифференциация геномов населения Южной и Западной Сибири по трем компонентам tSNE.

доминированием этого компонента, который является их генетической основой (до 99–100 % на уровне большинства индивидов). Значительна доля данного компонента также у кетов (до 45–50 % у отдельных индивидов) и томских татар (до 5–9 %). Ранее было показано, что существенную долю этот компонент занимает также в генофонде популяций Волго-Уральского региона – башкир (до 25 %), марийцев (до 20 %), коми, удмуртов и чувашей (до 15 %). С меньшей частотой он представлен практически во всех южносибирских выборках – у тувинцев, чулымцев, алтайцев и хакасов-сагайцев (от 5 до 10 %) (Харьков и др., 2020).

Доминирование угорского компонента у всех образцов хантов, начиная с  $K = 3$ , и практически полное отсутствие других генетических компонентов в их геномах на индивидуальном и популяционном уровне свидетельствуют, что их предковые популяции очень долгое время находились в генетической изоляции. Это говорит о том, что древнее угорское население современной территории расселения хантов не смешивалось с другими этносами, и подтверждает отсутствие других групп переселенцев с территории Южной Сибири и степной зоны.

Полученный результат показывает, что общая картина распределения компонентов хорошо согласуется с географическим местоположением исследуемых популяций, привязкой к конкретному региону, антропологическим и языковым различиям. Эта информация позволяет более точно судить о сходстве и различии сравниваемых популяций по составу предковых компонентов, а также о процессе формирования их генофонда.

**Идентичные по происхождению блоки сцепления.** В результате биоинформационной обработки данных генотипирования высокоплотных биочипов различных сибирских популяций проведен анализ совпадения общих по происхождению фрагментов ДНК между популяциями и отдельными индивидами. Сегментом, имеющим идентичные нуклеотидные последовательности, является IBD у двух или более лиц, если они унаследовали его от общего

предка без рекомбинации, т. е. у этих людей сегмент имеет общее происхождение. Ожидаемая длина сегмента IBD зависит от количества поколений с момента появления последнего общего предка. Одно из применений анализа общих по происхождению участков генома – количественная оценка степени родства между индивидами, что также может дополнить информацию о генетических связях популяций (Gusev et al., 2012).

Выборки хантов показали максимальное совпадение по IBD-блокам между собой (6 %), далее – с выборкой кетов (1.45 %), чулымцев (0.71 %), тувинцев (0.35 %), томских татар (0.33 %), хакасов-качинцев (0.32 %) и южных алтайцев (0.28 %). Среди хантов большее совпадение IBD-блоков наблюдается у русскиных хантов (23.5 %) по сравнению с казымскими (18.1 %).

Степень совпадения IBD-блоков между хантами, кетами и томскими татарами согласуется с результатами tSNE и ADMIXTURE по распределению в этих популяциях частот аллелей и общих генетических компонентов. При этом у русскиных хантов, имеющих наибольшую сумму средних длин сегментов IBD между парами индивидов, наибольший вклад вносят IBD длиной более 10 cM (42–46 %), что говорит о недавнем сильном инбридинге внутри популяции. Для подтверждения этого был рассчитан коэффициент инбридинга FROH для всех индивидов для трех классов блоков гомозиготности (ROH). Среди западно-сибирских популяций максимальные значения имеет популяция чулымцев (0.0292), близкие по значению – казымские (0.0280) и русскиные ханты (0.0266) и кеты (0.0259). Среди южносибирских популяций, включающих алтайцев, томских татар, тувинцев и хакасов, максимальное значение установлено и для выборки хакасов-сагайцев из предгорного Таштыпского района (0.0318), вдвое превосходящее у хакасов-качинцев равнинного Ширинского района. Минимальное значение характерно для томских татар (0.0071).

Наблюдается высокая корреляция для FROH > 1.5 с суммой средних длин сегментов IBD (IBD > 1.5 cM) между парами индивидов внутри сибирских популяций ( $r = 0.9246$ ,  $p < 5.612e-09$ ). Для расчета коэффициента

корреляции Спирмена использовали cor.test в программе R. Отношение суммы средних длин сегментов IBD (IBD > 1.5 cM) между парами индивидов к коэффициенту геномного инбридинга (FROH > 1.5) у русскиных хантов выше, чем у казымских. Эти показатели геномного инбридинга и распределения длин IBD внутри популяций хантов хорошо согласуются с их территориальной изолированностью и подтверждают отсутствие недавних потоков генов между популяциями на протяжении сотен лет.

**Гаплогруппы Y-хромосомы.** В результате анализа частоты встречаемости использованных SNP-маркеров в исследуемых выборках хантов было выявлено восемь гаплогрупп Y-хромосомы. По составу и частотам гаплогрупп выборки русскиных и казымских мужчин хантов существенно отличаются друг от друга. Только две гаплогруппы присутствуют в обеих выборках (см. таблицу).

К сублинии N1a2b1b1 у русскиных хантов принадлежат 39 образцов, а у казымских – всего 3. Терминальными для этой линии у хантов являются SNP Y68212, Y70717, Y70315, Y70327. Эта хантыйская сублиния близка к вариантам N1a2b1b1 у чулымцев (VL65, Z35095, Z35099, Z35102) и хакасов-качинцев (Z35093, Z35097, Z35103) (Валихова и др., 2022).

Гаплогруппа N1a2b1b1 у хантов – этноспецифичная, не совпадает по терминальным SNP и гаплотипам с доминирующей у ненцев N1a2b1b1a~ (B171, B170, Z35091, Z35092) (Харьков и др., 2021).

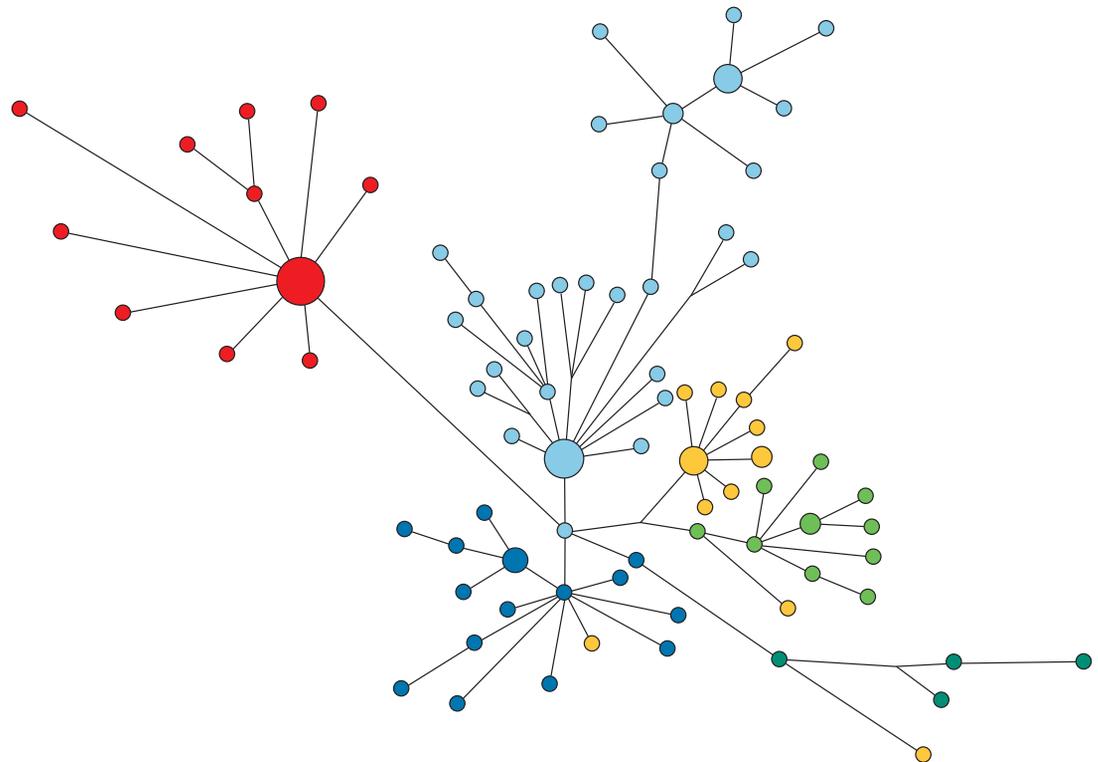
Особенность этнического состава большинства южно-сибирских народов – наличие родов (сеоков), где счет родства ведется по мужской линии. Такая родовая структура характерна для шорцев, хакасов, северных и южных алтайцев, телеутов. Все остальные образцы мужчин из различных западно- и южносибирских популяций (энцев, хакасов-сагайцев, шорцев, челканцев и тувинцев, а также хакаских сеоков, ранее входивших в состав бельтиров и бирюсинцев, ассимилированных в конце 19–начале 20 в.), относятся к другим сублиниям гаплогруппы N1a2b. Медианная сеть гаплотипов (рис. 2) демонстрирует у хантов звездообразную филогению с недавним эффектом основателя и преобладанием по частоте предкового гаплотипа.

Специфичный кластер гаплотипов хантов равноудален от всех сеоков хакасов-качинцев. Возраст этого кластера у хантов составил 858 лет (SD = 338 лет), что примерно в полтора-два раза превышает возраст кластеров качинских сеоков хасха – 487 лет (SD = 153 года), ызыр – 501 год (SD = 203 года), соххы – 585 лет (SD = 215 лет) (Харьков и др., 2020) и чулымских тюрок – 667 лет (SD = 194 года). Таким образом, ханты по этой гаплогруппе имеют прямую генетическую связь с качинцами, чулымцами и ненцами, предковые линии которых разошлись достаточно давно и отражают их связь с народами самодийской языковой группы.

Вторая общая для двух выборок хантов гаплогруппа – N1a2b2a1 (VL97, L1419, Y3185, Y3188, Y3189, Y3190, Y111190), ранее обозначавшаяся как европейская линия N1b-E. Эта сублиния обнаружена у башкир, казанских татар, коми, марийцев, карел, вепсов, финнов и русских (<https://www.yfull.com/>). Филогенетически ближе всего к хантам по этой линии расположены образцы коми. Этноспецифичные ветви хантов и коми объединяют SNP

Частота встречаемости гаплогрупп Y-хромосомы у хантов

Гаплогруппа	д. Русскиная % (N)	с. Казым % (N)
N1a2b1b1 (Y68212, Y70717, Y70315, Y70327)	60.9 (39)	5.5 (3)
N1a2b2a1 (VL97, L1419, Y3185, Y3188, Y3189, Y3190, Y111190)	9.4 (6)	16.6 (9)
N1a2b2b1~ (Z35076)	–	5.5 (3)
N1a1a1a2a1c1~ (Y13850, Y13852)	–	24.1 (13)
N1a2b1b1b1~ (B172, Z35108)	–	9.2 (5)
Q1b1a3b1a2~ (Z35974 xB32, B33, Z35993)	–	38.9 (21)
Q1a2b~ (M25, L716, YP1674, YP1676)	4.7 (3)	–
R1a1a1b2 (Y43850, S7280, FGC687, FGC38304)	25.0 (16)	–



**Рис. 2.** Медианная сеть YSTR-гаплотипов гаплогруппы N1a2b1b1 у хантов, чулымцев и хакасов-качинцев.

Голубым цветом обозначены ханты, красным – чулымцы, синим – хакасы сеока соххы, зеленым – хакасы сеока ызыр, желтым – хакасы сеока хасха, темно-зеленым – сеока пурут.

Y65017 и Y89655, не обнаруженные у других популяций. Ханты и коми имеют наименее древнего общего предка по этой гаплогруппе по сравнению с остальными европейскими популяциями.

По данным сайта YFull, эта ветвь отделилась от предковой линии примерно 2800 лет назад. Теоретически возможны два варианта появления этой гаплогруппы у современных хантов и коми: 1) наследование от общей древней предковой группы угорских племен; 2) недавнее смешение хантов с этническими коми, переселенцами в Сибирь. Но результаты анализа геномных данных методами NGSadmix, ADMIXTURE, IBD-блоков и различия по терминальным SNP Y-хромосом не подтверждают второй вариант. YSTR-гаплотипы этой линии у хантов и коми также отличаются на несколько мутаций. Ранее V.N. Pimenoff с коллегами (2008) выдвинули предположение, что когда обско-угорские ханты и манси ушли на западные склоны Уральских гор и на северо-запад Сибири, и было сформировано уникальное объединение N1b-A и N1b-E. Это сочетание сублиний N1b у хантов и манси предполагает недавнее слияние западной и восточной линий в Северо-Западной Сибири. Наши новые данные не противоречат этой версии.

Все остальные гаплогруппы представлены только в отдельных выборках хантов. К гаплогруппе N1a2b2b1~ (Z35076) относятся три образца казымских хантов. Наиболее близкая ей линия N1a2b2b1~ (B528, Y24382, Z35076, Z35077) также распространена у коми. У удмуртов, татар, чувашей и башкир присутствует ее более современная линия (B226). YSTR-гаплотипы этой гаплогруппы у коми

и удмуртов ближе друг другу, чем к образцам хантов. Присутствие у хантов и коми двух гаплогрупп, N1a2b2a1 и N1a2b2b1~, с этноспецифичными терминальными SNP и отличающимися гаплотипами говорит об их наследовании от достаточно древних общих предков, скорее всего, входивших в состав раннего угорского населения этих территорий.

Тринадцать образцов казымских хантов относятся к гаплогруппе N1a1a1a1a2a1c1~ (Y13850, Y13852), из них семь имеют фамилию Пяк, ненецкую по происхождению, относящейся к лесным ненцам. Все семь образцов имеют очень близкие гаплотипы и являются потомками относительно недавнего общего предка ненца. В анкетах этих мужчин, считающих себя хантами по отцовской линии, с разной глубиной были указаны ненецкие предки. Остальные шесть мужчин этой гаплогруппы отличаются гаплотипами от рода Пяк.

В нашем исследовании тазовских ненцев (Харьков и др., 2021) обнаружено, что к данной гаплогруппе полностью принадлежат все мужчины, представляющие хантыйские по происхождению рода Салиндер, Лар и Тибичи. Представители этих родов сформировались в XVIII–XIX вв. в низовьях Оби в результате развития ненецкого крупнотабунного оленеводства и вовлечения в него части северных хантов (Квашнин, 2003). Все гаплотипы казымских хантов этой гаплогруппы значительно отличаются от гаплотипов тазовских ненцев.

Другие пять образцов казымских хантов относятся к гаплогруппе N1a2b1b1b1~ (B172, Z35108). К ней принадлежат все ранее обследованные мужчины ненцы из

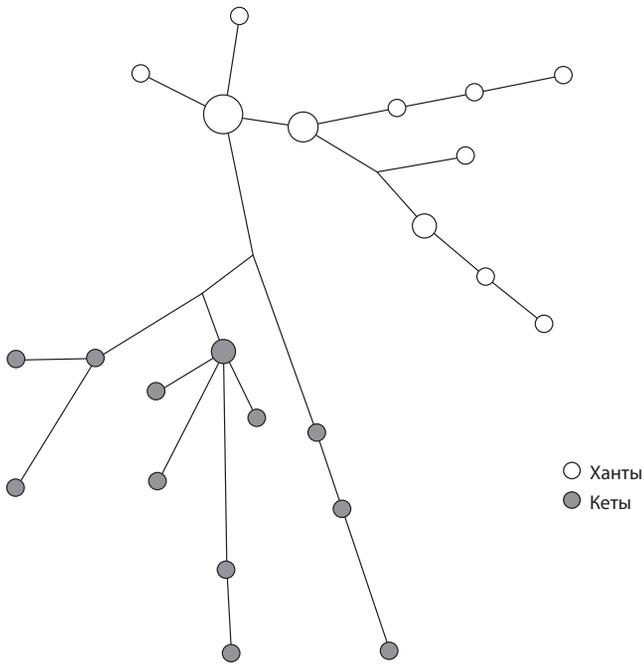


Рис. 3. Медианная сеть YSTR-гаплотипов гаплогруппы Q1b1a3b1a2~ у хантов и кетов.

фратрии Вануйто, относящиеся к родам Вануйто, Пуйко и Яунгат, и хантыйского по происхождению рода Пурунгуи. Четыре образца хантов отличаются по гаплотипам от ненцев, но один практически полностью совпадает с ними. Такое разделение на специфичные для хантов и близкие к ненцам гаплотипы совпадает с данными по гаплогруппе N1a1a1a1a2a1c1~. Очевидно, что в состав генофонда казымских хантов входят именно хантыйские по происхождению варианты этих гаплогрупп, но относительно недавно заключались браки и с лесными ненцами. Отсутствие этих гаплогрупп у русскиных хантов совпадает с данными распределения блоков IBD и коэффициента геномного инбридинга.

Распределение в исследованных популяциях различных гаплогрупп клады N Y-хромосомы согласуется с частотой угорского генетического компонента. Филогенетический анализ Y-хромосомных сублиний и гаплотипов различных гаплогрупп клады N показывает, что центр происхождения и расселения носителей угорского компонента в Южной, Западной Сибири и Восточной Европе – территория современного Алтая и Саян. Полученные результаты сопоставимы с данными этнологии, антропологии и лингвистики о вкладе уральского компонента в формирование различных народов Алтая и Саян и историческими ареалами угорских и других языков уральской языковой семьи.

Почти 40 % мужчин с. Казым принадлежит к гаплогруппе Q1b1a3b1a2~ (Z35974 xB32, B33, Z35993). Ближе всего к ней относится линия Q1b1a3b1a2~ (B33, Z35991), специфичная для популяции кетов. Этот вариант, кроме кетов, преобладает и у селькупов из Томской области и Красноярского края. Более удаленная линия Q1b1a3b1a~ (B30, YP1693 xZ35991) распространена в популяциях тувинцев с максимальной частотой в восточных горных районах Тывы (до 25 %). Образцы хантов демонстрируют

специфичный спектр гаплотипов с недавним эффектом основателя, который не наблюдается у кетов (рис. 3).

Распределение этих сублиний в популяциях хантов, кетов и тувинцев хорошо согласуется с долями совпадений по IBD-блокам между ними, графиком tSNE и по наличию угорского генетического компонента в этих популяциях по аутосомной части их генофонда. Присутствие этой линии у хантов связано не с недавним заимствованием у других аборигенных популяций (кетов и селькупов), а с тем, что она уже входила в состав расселявшихся предковых групп.

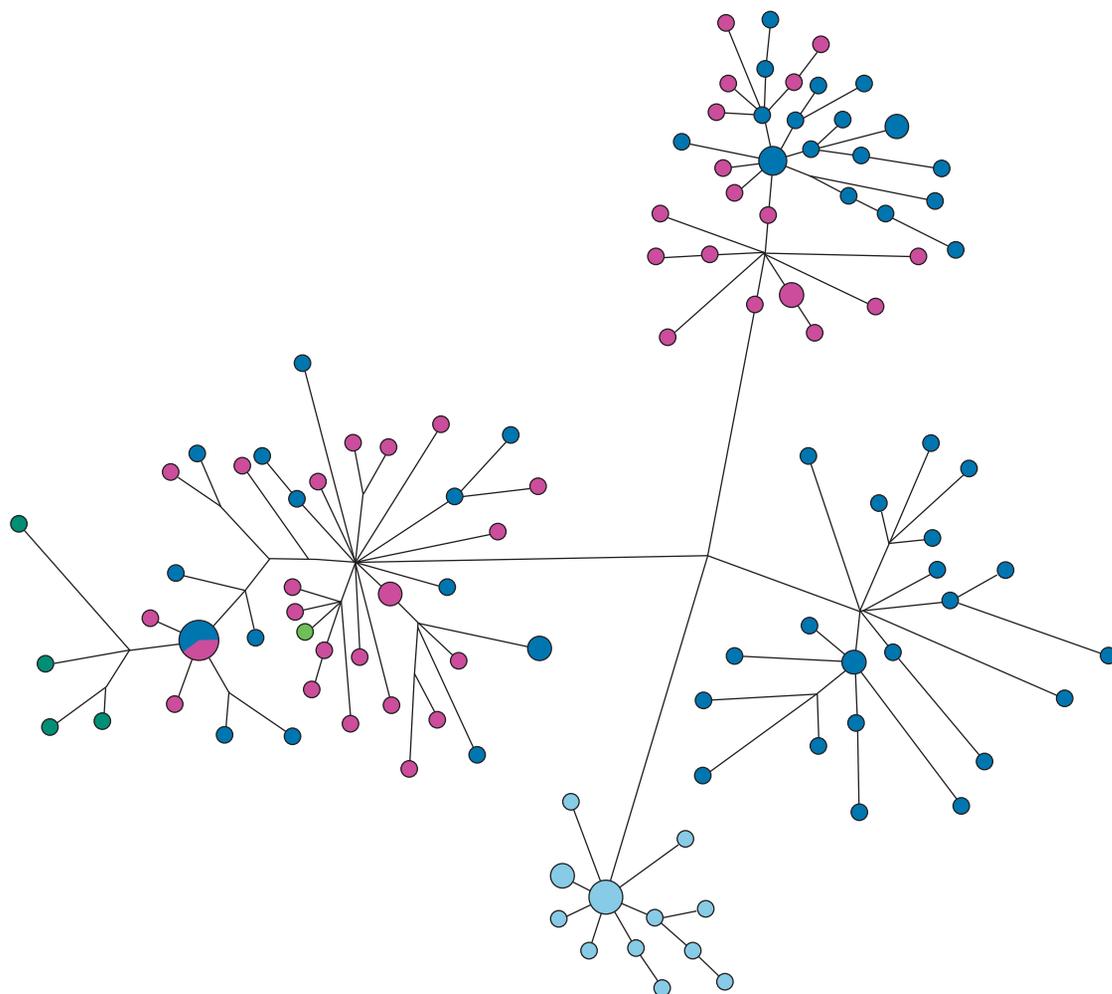
У трех мужчин д. Русскинская представлена совершенно другая гаплогруппа клады Q – Q1a2b~ (M25, L716, YP1674, YP1676). Это очень редкая гаплогруппа, не обнаруженная в других сибирских популяциях. С максимальной частотой она имеется у этнических туркмен Каракалпакстана, Ирана и Афганистана (Grugni et al., 2012; Схаляхо и др., 2016). У большинства других этносов ее частота очень низка. Гаплотипы хантов довольно сильно отличаются от других популяций. Скорее всего, присутствие у них этой линии – не следствие недавней метисации, а наследие угорских групп, переселившихся из Южной Сибири и Приуралья на север.

Последняя гаплогруппа, к которой относятся 16 мужчин хантов д. Русскинская, – R1a1a1b2-Y43850. Гаплотипы всех образцов достаточно близки, что говорит о недавнем эффекте основателя (рис. 4).

Специфичными терминальными SNP для хантов являются S7280, FGC687 и FGC38304. Наиболее близкие к этой линии варианты R1a1a1b2-Y43850 с большей частотой представлены у хакасов, шорцев и с меньшей – у тувинцев и северных алтайцев. По данным YFull, возраст этой гаплогруппы составляет примерно 3800 лет. Все эти образцы относятся к четырем различным линиям, которые разделились очень давно. Возраст кластера гаплотипов у хантов составил 933 года (SD = 336 лет), что примерно в полтора раза меньше возраста южносибирских линий. У хакасов сеока пилтир – 1469 лет (SD = 342 года) (Y39884 xY43109). Специфичная для бирюсинских хакасских сеоков туран, хызыл хая и шорских сеоков тарткын, шор-кызай и кара-шор линия этой гаплогруппы (Y62155.2) имеет примерно такой же возраст – 1315 лет (SD = 227 лет). Ветвь с более широким распространением в саяно-алтайских популяциях (Y43109) еще старше – 1566 лет (SD = 350 лет). Разница по SNP и STR между ее вариантами у хакасов, шорцев, тувинцев и северных алтайцев больше, чем с хантами.

Показана значительная неоднородность исследованных выборок хантов по составу и частотам гаплогрупп. Филогения разных линий двух гаплогрупп, N1a2b1b1 и R1a1a1b2-Y43850, указывает на их южносибирское происхождение в генофонде хантов. Территория Саян и Алтай была первичным очагом генерации разнообразия и экспансии численности предковых групп, носителей этих гаплогрупп на территории Сибири. Наиболее вероятно, что распространение большинства Y-хромосомных гаплогрупп у хантов происходило при начальном расселении угорских племен на север и запад.

Нужно учитывать тот факт, что ареал современных хантов находится севернее территории обитания их пред-



**Рис. 4.** Медианная сеть YSTR-гаплотипов гаплогруппы R1a1a1b2-Y43850 у хантов, хакасов, шорцев, тувинцев и алтайцев. Голубым цветом обозначены ханты, синим – хакасы, малиновым – шорцы, темно-зеленым – тувинцы, зеленым – алтайцы.

ков. Западно-Сибирский и Волго-Уральский регионы были местом вторичной генерации разнообразия, но не образования самой гаплогруппы N1a2. Окончательного мнения относительно места формирования этносов уральской языковой семьи на данный момент не существует, но многочисленные данные, в том числе и результаты исследований филогении и филогеографии гаплогрупп клады N, указывают на Южную Сибирь. Лингвистическая палеонтология описывает прауральский экологический ареал как территорию, ограниченную на западе Уральским хребтом, на севере – примерно Полярным кругом, на востоке – районом нижнего течения Ангары и Подкаменной Тунгуски и среднего течения Енисея, на юге – примерно современной южной границей западносибирской тайги от северных предгорий Саян и Алтая до нижнего течения Тобола и Среднего Урала включительно (Напольских, 2018).

### Заключение

Таким образом, генофонд двух популяций хантов представляет собой разнородный набор по составу гаплогрупп Y-хромосомы, но очень близкий по аутосомным маркерам. Расширенный состав терминальных SNP для выявленных гаплогрупп позволил детально описать и уточнить разли-

чия в филогении и структуре отдельных этноспецифичных сублиний, определить их родство, следы экспансии численности в генофонде хантов. Результаты сравнительного анализа образцов мужчин свидетельствуют о близком генетическом родстве хантов с алтае-саянскими хакасами и тувинцами, а также с ненцами, коми, удмуртами и кетами. Специфичность гаплотипов, обнаружение различных терминальных SNP говорят о том, что ханты достаточно долго не входили в контакты с другими этносами. Исключение составляют лишь ненцы, в составе которых много хантыйских родов. Для северной популяции хантов с. Казым по Y-хромосомным линиям показан небольшой вклад лесных ненцев.

Полученные результаты не противоречат общепринятым версиям этногенеза хантов, но позволяют по-новому взглянуть на этот процесс. Основным фактором формирования генофонда хантов были их территориальная генетическая изоляция и более позднее смешение с пришлым самодийским населением, что при переходе на тундровое оленеводство привело к сильному демографическому росту их родов в составе ненцев. Относительно низкое генетическое разнообразие по аутосомным SNP и достаточно высокий уровень инбридинга у хантов это под-

тверждают. Новая информация о структуре генофонда хантов является важным дополнением к существующим антропологическим, археологическим, этнологическим и лингвистическим данным о формировании и родственных связях хантов с другими народами.

## Список литературы / References

Брук С.И. Население мира. Этнодемографический справочник. М.: Наука, 1986.

[Brook S.I. The World Population. Ethnodemographic Reference Book. Moscow: Nauka Publ., 1986. (in Russian)]

Валихова Л.В., Харьков В.Н., Зарубин А.А., Колесников Н.А., Сваровская М.Г., Хитринская И.Ю., Штыгашева О.В., Волков В.Г., Степанов В.А. Генетическая взаимосвязь чулымских тюрков с хакасами и кетами по данным аутомомных SNP и гаплогруппам Y-хромосомы. *Генетика*. 2022;58(10):1177-1184. DOI 10.31857/S0016675822100113.

[Valikhova L.V., Kharkov V.N., Zarubin A.A., Kolesnikov N.A., Svarovskaya M.G., Khitrinskaya I.Yu., Shtygasheva O.V., Volkov V.G., Stepanov V.A. Genetic interrelation of the Chulym Turks with Khakass and Kets according to autosomal SNP data and Y-chromosome haplogroups. *Russ. J. Genet.* 2022;58(10):1228-1234. DOI 10.1134/S1022795422100118.]

Квашин Ю.Н. Гыданские ненцы: история формирования современной родовой структуры (XVIII–XX вв.). М.: ИЭА РАН, 2003.

[Kvashin Yu.N. Gydan Nenets: the History of the Formation of the Modern Genetic Structure (18-20th Centuries). Moscow: IEA RAS Publ., 2003. (in Russian)]

Напольских В.В. Очерки по этнической истории. Казань: Институт археологии им. А.Х. Халикова, 2018.

[Napolskikh V.V. Essays on Ethnic History. Kazan: Khalikov Institute of Archeology, 2018. (in Russian)]

Народы Западной Сибири: Ханты. Манси. Селькупы. Ненцы. Эңцы. Нганасаны. Кеты. М.: Наука, 2005.

[Peoples of West Siberia: Khanty, Mansi, Selkups, Nenets, Enets, Nganasans, Kets. Moscow: Nauka Publ., 2005. (in Russian)]

Народы России: Энциклопедия. М.: Большая Российская энциклопедия, 1994.

[The Peoples of Russia: Encyclopedia. Moscow: Bol'shaya Rossiyskaya Entsiklopediya Publ., 1994. (in Russian)]

Схалихо Р.А., Жабалин М.К., Юсупов Ю.М., Агджоян А.Т., Сабитов Ж.М., Гурьянов В.М., Балаганская О.А., Далимова Д.А., Давлетчурин Д.Х., Турдикулова Ш.У., Чухряева М.И., Асылгузин Р.Р., Акильжанова А.Р., Балановский О.П., Балановская Е.В. Генофонд туркмен Каракалпакстана в контексте популяций Центральной Азии (полиморфизм Y-хромосомы). *Вестн. Моск. ун-та. Серия 23. Антропология*. 2016;3:86-96.

[Skhalyakho R.A., Zhabagin M.K., Yusupov Yu.M., Agdzhojan A.T., Sabitov Zh.M., Gurianov V.M., Balaganskaya O.A., Dalimova D.A., Davletchurin D.Kh., Turdikulova Sh.U., Chukhryaeva M.I., Asilgujin R.R., Akilzhanova A.R., Balanovsky O.P., Balanovska E.V. Gene pool of Turkmens from Karakalpakstan in their Central Asian context (Y-chromosome polymorphism). *Vestnik Moskovskogo Universiteta. Seriya 23. Antropologiya = Moscow University Anthropology Bulletin*. 2016;3:86-96. (in Russian)]

Харьков В.Н., Валихова Л.В., Яковлева Е.Л., Сереброва В.Н., Колесников Н.А., Петелина Т.И., Хитринская И.Ю., Степанов В.А. Реконструкция происхождения гыданских ненцев на основе генетического анализа их родовой структуры с помощью нового

набора YSTR-маркеров. *Генетика*. 2021;57(12):1403-1414. DOI 10.31857/S0016675821120067.

[Kharkov V.N., Valikhova L.V., Yakovleva E.L., Serebrova V.N., Kolesnikov N.A., Petelina T.I., Khitrinskaya I.Yu., Stepanov V.A. Reconstruction of the origin of the Gydan Nenets based on genetic analysis of their tribal structure using a new set of YSTR markers. *Russ. J. Genet.* 2021;57(12):1414-1423. DOI 10.1134/S1022795421120061.]

Харьков В.Н., Новикова Л.М., Штыгашева О.В., Лузина Ф.А., Хитринская И.Ю., Волков В.Г., Степанов В.А. Генофонд хакасов и шорцев по маркерам Y-хромосомы: общие компоненты и генетическая структура родов. *Генетика*. 2020;56(7):826-833. DOI 10.31857/S0016675820070073.

[Kharkov V.N., Novikova L.M., Shtygasheva O.V., Luzina F.A., Khitrinskaya I.Y., Volkov V.G., Stepanov V.A. Gene pool of Khakass and Shors for Y chromosome markers: common components and tribal genetic structure. *Russ. J. Genet.* 2020;56(7):849-855. DOI 10.1134/S1022795420070073.]

Alexander D.H., Lange K. Enhancements to the ADMIXTURE algorithm for individual ancestry estimation. *BMC Bioinformatics*. 2011; 12:246. DOI 10.1186/1471-2105-12-246.

Alexander D.H., Novembre J., Lange K. Fast model-based estimation of ancestry in unrelated individuals. *Genome Res*. 2009;19(9):1655-1664. DOI 10.1101/gr.094052.109.

Bandelt H.J., Forster P., Röhl A. Median-joining networks for inferring intraspecific phylogenies. *Mol. Biol. Evol.* 1999;16(1):37-48. DOI 10.1093/oxfordjournals.molbev.a026036.

Browning B.L., Browning S.R. Improving the accuracy and efficiency of identity-by-descent detection in population data. *Genetics*. 2013; 194(2):459-471. DOI 10.1534/genetics.113.150029.

Browning S.R., Browning B.L. Rapid and accurate haplotype phasing and missing-data inference for whole-genome association studies by use of localized haplotype clustering. *Am. J. Hum. Genet.* 2007;81(5):1084-1097. DOI 10.1086/521987.

Grugin V., Battaglia V., Hooshar Kashani B., Parolo S., Al-Zahery N., Achilli A., Olivieri A., Gandini F., Houshmand M., Sanati M.H., Torroni A., Semino O. Ancient migratory events in the Middle East: new clues from the Y-chromosome variation of modern Iranians. *PLoS One*. 2012;7(7):e41252. DOI 10.1371/journal.pone.0041252.

Guo Y., He J., Zhao S., Wu H., Zhong X., Sheng Q., Samuels D.C., Shyr Y., Long J. Illumina human exome genotyping array clustering and quality control. *Nat. Protoc.* 2014;9(11):2643-2662. DOI 10.1038/nprot.2014.174.

Gusev A., Palamara P.F., Aponte G., Zhuang Z., Darvasi A., Gregersen P., Pe'er I. The architecture of long-range haplotypes shared within and across populations. *Mol. Biol. Evol.* 2012;29(2):473-486. DOI 10.1093/molbev/msr133.

Pimenoff V.N., Comas D., Palo J.U., Vershubsky G., Kozlov A., Sajantila A. Northwestern Siberian Khanty and Mansi in the junction of West and East Eurasian gene pools as revealed by uniparental markers. *Eur. J. Hum. Genet.* 2008;16(10):1254-1264. DOI 10.1038/ejhg.2008.101.

Skotte L., Korneliusen T., Albrechtsen A. Estimating individual admixture proportions from next generation sequencing data. *Genetics*. 2013;195(3):693-702. DOI 10.1534/genetics.113.154138.

Zhivotovskiy L.A., Underhill P.A., Cinnioglu C., Kayser M., Morar B., Kivisild T., Scozzari R., Cruciani F., Destro-Bisol G., Spedini G., Chambers G.K., Herrera R.J., Yong K.K., Gresham D., Tournev I., Feldman M.W., Kalaydjieva L. The effective mutation rate at Y-chromosome STRs with application to human population divergence time. *Am. J. Hum. Genet.* 2004;74(1):50-61. DOI 10.1086/380911.

## ORCID ID

V.N. Kharkov orcid.org/0000-0002-1679-2212  
N.A. Kolesnikov orcid.org/0000-0001-8855-577X  
V.A. Stepanov orcid.org/0000-0002-5166-331X

**Благодарности.** Исследование выполнено за счет гранта Российского научного фонда № 22-64-00060 (<https://rscf.ru/project/22-64-00060/>).

**Конфликт интересов.** Авторы заявляют об отсутствии конфликта интересов.

Поступила в редакцию 14.10.2022. После доработки 23.11.2022. Принята к публикации 27.11.2022.