

GENETIC PORTRAITS OF KHANTY AND MANSI BASED ON THE Y CHROMOSOME HAPLOGROUPS IN THE CONTEXT OF GENE POOLS OF RUSSIA

Ponomarev GYu ✉, Agdzhoyan AT, Potanina AYu, Adamov DS, Balanovska EV

Research Centre for Medical Genetics, Moscow, Russia

Khanty and Mansi are small indigenous peoples of Western Siberia with the unique cultural, anthropological, and linguistic characteristics. The study of their gene pool will make it possible to reconstruct the genetic structure of the Ugric-speaking population, of which in modern times only Khanty and Mansi remain, along with Hungarians, in whose gene pool there are traces of medieval migration of the Ugric-speaking Magyars. The detailed characterization of the gene pool of Khanty and Mansi is important for reconstruction of Ugric populations and genetic history of the region. The study was aimed to assess representative samples of Khants ($n = 83$) and Mansi ($n = 74$) based on the standard panel of 60 SNP markers and the extended panel of 74 Y-chromosomal SNP markers by statistical and cartographic methods in the context of indigenous population of Urals and Western Siberia. The differences between the gene pools of Khanty and Mansi have been revealed based on both standard panel of Y chromosome haplogroups and branches of haplogroups N2 and N3a4. Most of the Khanty gene pool is evenly distributed between N2-Y3195 (26%), N2-VL67 (23%), and N3a4-Z1936 (23%). The "Western" branch N2-Y3195 predominates in the Mansi gene pool (69%). Mansi gravitate towards populations of the Urals-Volga region in the multidimensional genetic space. Based on the standard panel of Y haplogroups, Khanty are close to the populations of Western and South Siberia. However, the analysis of branches N3a4 has shown that Khanty are intermediate between the "Uralic" and "Siberian" clusters: when the ancestors of Khanty moved from the Ural region to the northeast, these acquired both genetic components. The gene geographic maps of 10 haplogroup N3a4 branches in the populations of Urals and Western Siberia reflect the dynamic changes of the gene pool that took place 4–2 kyr.

Keywords: Khanty, Mansi, gene pool, Y-chromosome, SNP, haplogroup, N3a4-Z1936, N2-Y3195, N2-VL67

Funding: State Assignment of the Ministry of Science and Higher Education of the Russian Federation for the Research Centre for Medical Genetics.

Acknowledgements: the authors would like to thank all participants of the expedition survey (donors of samples) and Biobank of North Eurasia for access to DNA collections.

Author contribution: Balanovska EV — management; Agdzhoyan AT — contribution to the expedition survey of Ob Ugrians; Ponomarev GYu — Y-SNP marker genotyping; Ponomarev GYu, Potanina AYu, Adamov DS — statistical and cartographic analysis, manuscript formatting; Balanovska EV, Ponomarev GYu — study design and manuscript writing.

Compliance with ethical standards: the study was approved by the Ethics Committee of the Research Centre for Medical Genetics (protocol No. 1 dated 29 June 2020).

✉ **Correspondence should be addressed:** Georgy Yu. Ponomarev
Moskvorechie, 1, 115522, Moscow, Russia; st26i900@gmail.com

Received: 09.09.2024 **Accepted:** 01.10.2024 **Published online:** 28.10.2024

DOI: 10.24075/brsmu.2024.044

ГЕНЕТИЧЕСКИЕ ПОРТРЕТЫ ХАНТОВ И МАНСИ ПО ГАПЛОГРУППАМ Y-ХРОМОСОМЫ В КОНТЕКСТЕ ГЕНОФОНДОВ РОССИИ

Г. Ю. Пономарёв ✉, А. Т. Агджоян, А. Ю. Потанина, Д. С. Адамов, Е. В. Балановская

Медико-генетический научный центр имени академика Н. П. Бочкова, Москва, Россия

Ханты и манси — коренные малочисленные народы Западной Сибири с уникальными культурными, антропологическими и лингвистическими особенностями. Изучение их генофонда позволяет реконструировать генетическую структуру населения, говорившего на угорских языках, от которого в современности сохранились только ханты, манси, а также венгры, в чьем генофонде присутствуют следы средневековой миграции угроязычных мадьяр. Детальная характеристика генофонда хантов и манси важна для реконструкции угорских популяций и генетической истории региона. Целью исследования было изучить репрезентативные выборки хантов ($n = 83$) и манси ($n = 74$) на основе базовой панели из 60 SNP-маркеров и углубленной панели 74 SNP-маркеров Y-хромосомы статистическими и картографическими методами в контексте коренного народонаселения Урала и Западной Сибири. Выявлены различия генофондов хантов и манси и по базовой панели гаплогрупп Y-хромосомы, и по ветвям гаплогрупп N2 и N3a4. Основная часть генофонда хантов равномерно распределена между N2-Y3195 (26%), N2-VL67 (23%) и N3a4-Z1936 (23%). В генофонде манси резко преобладает «западная» ветвь N2-Y3195 (69%). В многомерном генетическом пространстве манси тяготеют к популяциям Урало-Поволжья. Ханты по базовой панели Y-гаплогрупп сближаются с популяциями Западной и Южной Сибири. Анализ ветвей N3a4 показал, что ханты занимают промежуточное положение между «уральским» и «сибирским» кластерами: при движении предков хантов из Приуралья на северо-восток они включили оба генетических компонента. Геногеографические карты 10 ветвей гаплогруппы N3a4 в популяциях Урала и Западной Сибири отражают динамику генофонда в период от 4 до 2 тыс. лет назад.

Ключевые слова: ханты, манси, генофонд, Y-хромосома, SNP, гаплогруппа, N3a4-Z1936, N2-Y3195, N2-VL67

Финансирование: государственное задание Министерства науки и высшего образования РФ для Медико-генетического научного центра имени академика Н. П. Бочкова

Благодарности: авторы благодарят всех участников экспедиционного обследования (доноров образцов) и АНО «Биобанк Северной Евразии» (за предоставление коллекций ДНК).

Вклад авторов: Е. В. Балановская — руководство; А. Т. Агджоян — участник экспедиционного обследования обских угров; Г. Ю. Пономарёв — генотипирование Y-SNP маркеров; Г. Ю. Пономарёв, А. Ю. Потанина, Д. С. Адамов — статистический и картографический анализ, оформление статьи; Е. В. Балановская, Г. Ю. Пономарёв — дизайн исследования и написание статьи.

Соблюдение этических стандартов: исследование одобрено этическим комитетом Медико-генетического научного центра имени Н. П. Бочкова (протокол № 1 от 29 июня 2020 г.).

✉ **Для корреспонденции:** Георгий Юрьевич Пономарёв
ул. Москворечье, д. 1, 115522, г. Москва, Россия; st26i900@gmail.com

Статья получена: 09.09.2024 **Статья принята к печати:** 01.10.2024 **Опубликована онлайн:** 28.10.2024

DOI: 10.24075/vrgmu.2024.044

Most representatives of the Ob Ugrians, Khanty and Mansi, live in the Khanty-Mansi Autonomous Okrug – Yugra of the Tyumen Region (hereinafter, KhMAO). According to the 2010 census, the population of Mansi is 11,614 people, and the population of Khanty is 29,277 people. They speak languages of the Ugric branch of the Uralic language family. Among contemporary peoples, only Hungarians, who have inherited a negligible genetic trace from medieval Magyars, speak these languages [1]. The Ob Ugrians have a unique combination of anthropological traits distinguished as a distinct “Uralic race” [2], and their culture comprises two components: the ancient northern, hunter, taiga component (inherited from the indigenous population) and the southern pastoralist one (associated with migration of the Indo-Iranian-speaking population) [3]. The population structure of the Ob Ugrians took shape in the 1st millennium A. D. on both sides of the Ural Mountains: between the Kama region and the Ob River basin. Their range has been declining gradually and shifting to the northeast since 14th–15th century under the pressure of the Komi people, who migrated due to expansion of the Russian population.

Four ethnographic groups characterized by high endogamy levels were distinguished in Mansi: in the 18th–19th centuries, the average share of endogamous marriages in these groups exceeded 80%. Later, part of Southern Mansi was assimilated by Tatars and Russians, and the Western Mansi were absorbed into the northern and eastern groups of Mansi. The largest northern group of Mansi included the populations of Western and Southern Mansi, as well as Khanty; marriages with the Nenets people added some Mongoloid traits to their genetic portrait. The Eastern Mansi maintain the anthropological type close to that of the Finnish-speaking peoples of the Urals-Volga region [2].

In Khanty, the dialects of three ethnographic groups differed so much that these were considered as independent languages, and the average endogamy level reached 90% in the 18th–19th century [2]. The culture of Northern Khanty is most similar to that of Northern Mansi and the Nenets people, while the culture of Eastern Khanty is most close to that of the Selkup people. To date, the southern group has been lost, however, it is assumed that the “Ugric” substrate of the culture [4, 5] and gene pool [6] of the Zabolotny Siberian Tatars is inherited from the southern groups of Mansi and Khanty.

Both genetic components, the Western Eurasian and Eastern Eurasian, are found in the Y gene pool of the Ob Ugrians [1, 7–14]; the eastern component prevails (77% in Khanty, 89% in Mansi), the origin of which is considered to be associated with the Upper Paleolithic migration from South Siberia and Central Asia [10]. An opposite trend is reported based on the mtDNA data: the maximum share of the Eastern Eurasian component (59% in Khanty, 69% in Mansi) [10]. A genetic relationship between Khanty and the Komi-Zyrians, Samoyedic peoples (Enets, Tundra and Forest Nenets, Nganasans) against the background of genetic differences between the Ob Ugrians and the majority of peoples of Siberia (Altaians, Buryats, Kets, Selkup people, Evenki, Dolgans, Yakuts) and Central Asia was demonstrated [15].

In the study of the haplogroup N phylogeography and the model of expansion of the haplogroup N carriers, the data on the Ob Ugrians provided are based on the sum of haplogroups N2 and N3 [9] (according to the haplogroup N nomenclature introduced in the studies [12]). Later it was shown that the haplogroup N2 frequency decreased from north to south in the Western Siberia, while the haplogroup N3 frequency increased [12]. The analysis of the N3a4 sub-variants in the gene pool of Hungarians [1, 12] showed that the ancestors of the Ob Ugrians and Magyars split up about 2.7–2.9 kya.

In Khanty, significant genetic differences between the ethnographic groups are likely to persist: if in the northern group the frequencies of N2 and N3 are equal (38% each) [7], then in the southern group the frequency of N3 is almost twice higher (64%) [10], although the differences can be due to small samples: $n = 47$ in the northern and $n = 28$ in the southern group.

The recent study of the Khanty and Mansi gene pool using the standard panel of 60 Y-chromosomal markers suggests that there are three components in the gene pool of the Ob Ugrians: Siberian, Uralic, and Northern European [16]. The Siberian component predominates (80%), in which haplogroup N2 is dominant, and the second place is shared by N3a4 (Northern European component) and N3a1 (Uralic). The identified differences between the gene pools of Mansi and Khanty [16] can reflect various aspects of the genetic history of Ugrians and require a more thorough analysis.

In this study we considerably expanded the range of the studied populations and improved the genotyping level, and maximum attention was given to the gene geography of different haplogroup N2 and N3a4 variants. N2 dominating in the Ob Ugrians is spread from the Volga region to Far East, and assessment of the genetic landscape of various haplogroup branches can provide new information about the genetic history of Ugrians. Haplogroup N3a4 that is more frequent in the Ob Ugrians, than in the majority of other populations of Urals and Siberia, is no less important. Despite the fact that the today's N3a4 area is centered round the north of the East European Plain, it is considered to originate from Western Siberia [1, 11, 12]. Therefore, construction of the detailed genetic portraits of the Ob Ugrians in the context of the broad range of surrounding populations is essential for understanding the genetic history of a large number of indigenous peoples of Russia.

METHODS

The population-based samples of Khanty ($n = 83$) and Mansi ($n = 74$) were assessed within the framework of the expedition survey (RFBR project No. 16-06-00303a) with the help of the Biobank of North Eurasia; the samples covered a broad spectrum of local populations (Fig. 1). Despite the existing genetic differences between local groups of Khanty (Fig. 1), we combined these groups into a single sample based on the shared ethnicity and cultural identity for the purposes of this study. Such a combination made it possible to perform statistical analysis showing the genetic characteristics of Khanty as a single ethnic population. Only unrelated (at least to the third generation) males, whose fathers and grandfathers were born in the studied populations of this ethnic group, were included in the study. The data on the comparison groups were provided by the Biobank of North Eurasia.

The Qiagen QIASymphony SP automated nucleic acid isolation and purification system was used to extract DNA from the venous blood samples; genotyping was performed by the OpenArray method using the QuantStudio 12 Flex PCR system (Thermo Fisher Scientific; USA). The basic analysis involved 60 Y-chromosomal SNP markers most typical for the population of North Eurasia: D-M174, E-M35, E-M78, C-M217, C-F3791, C-F5481, C-F3918, C-M48, C-SK1066, C-M407, G-M201, G1-M285, G2-P15, G2-FGC595, G2-M406, G2-P303, H-M69, I-M170, I-M253, I-P37.2, I-M223, J1-M267, J1-P58, J2-M172, J2-M12, J2-M67, J2-M9, L-M20, L-M317, T-M70, N-M231, N-M128, N2-Y3205, N-M178, N3a1-B211, N3a2-M2118, N3a3-CTS10760, N3a4-Z1936, N3a5a-F4205, N-B202, N-B479, O-P186, O-M119, O-P31, O-M122, O-P201, O-M134, Q-M242, R1a-M198, R1a-PF6202, R1a-Y2395,

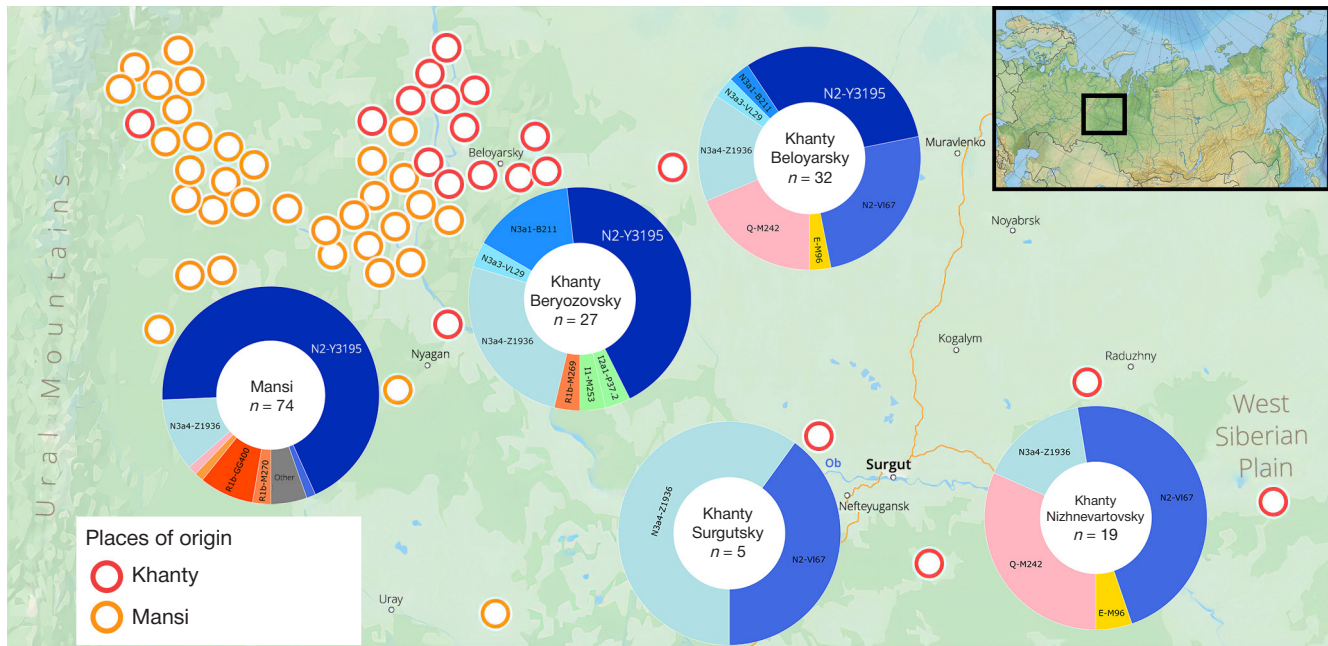


Fig. 1. Gene pool composition and the sites, where the samples of Mansi and various groups of Khanty were originate. Northern Mansi were surveyed in the Beryozovsky district of KhMAO and around, along the Sosva and Lyapin rivers (ancestors in the direct male line were born in villages of Beryozovo, Sosva, Vanzetur, Karym, Kimkyasui, Lombovozh, Verkhniye, Novinskaya, Nyaksimvol, Saranpaul, Khulimsunt, Patrasui, Rakhtynia, Sartyria, Khanglytur, Shhekurya, Yasunt, Unyugan, as well as the sites of Mansi settlements). Khanty were surveyed in the Beloyarsky ($n = 32$), Beryozovsky ($n = 27$), Surgutsky ($n = 5$), and Nizhnevartovsky ($n = 19$) districts of KhMAO

R1a-CTS1211, R1a-Z92, R1a-Z93, R1b-M343, R1b-Y13887, R1b-M269, R1b-L51, R1b-Z2105, R2-M124.

The samples were further genotyped by another 14 branches specifically for this study: haplogroups N2 (Y3195, VL67) and N3a4 (CTS1223, Y13850, Y24370, Y24360, L1034, L1442, Y28540, Y28544, Z1924, YP5259, Z35275, Z1928).

The genetic relationships and age to the origin of the haplogroups were obtained from the YFull Y-chromosome tree [17], unless otherwise specified. The evolutionary age of branches was based on the analysis of TMRCA (time to the most recent common ancestor) with the 95% confidence intervals (CI). Statistical analysis was performed using the Statistica 7.0 software (StatSoft; USA); the Nei's genetic distances were calculated in DJGenetic [18]. The gene geographic maps were created using the original GeneGeo software tool [19] by the Shepard's inverse distance weighting method with the weight function $K = 2$ and the radius of influence of 1100 km. Haplogroup N2 was mapped using the color chart with the frequency maximum of 60%, while N3a4 was mapped using the chart with the frequency maximum of 30%. Sampling for comparative analysis was performed based on the objectives of studying the genetic structure of Khanty and Mansi in Western Siberia and the Ural region on a large scale. In this regard, determination of grading for each population was driven by the combination of such factors, as the large population area and population heterogeneity, since it was necessary to ensure the widest possible geographical coverage to construct the detailed and reliable haplogroup distribution maps. The data of

the Biobank of North Eurasia were complemented by the data of the earlier reported studies [1, 11–14, 20–23].

RESULTS

Genetic portraits of the Ob Ugrians in the context of surrounding peoples

A total of 13 Y chromosome haplogroup variants have been identified in the Ob Ugrian gene pools (Table 1, Fig. 1, 2) showing that the gene pools of Khanty and Mansi are dramatically different, despite their geographical and linguistic proximity. Despite the fact that haplogroup N2 predominates in both ethnic groups, its frequency is significantly higher in Mansi (70%), than in Khanty (49%); N3a4 is twice more common in Khanty (23%) compared to Mansi (11%). There are different proportions of rare haplogroups in Mansi and Khanty. Particularly noteworthy are the increased Q frequency in Khanty (14%) and the increased R1b frequency in Mansi (11%).

Haplogroups N2 and N3 that predominate in the genetic portraits of the Ob Ugrians are common in the indigenous populations of the Ural region and Western Siberia (Fig. 2): these predominate in many populations of the region suggesting similar historical processes and migration routes.

The pattern showing the haplogroup N2 predominance and a significant share of N3a4, that is typical for the Ob Ugrians and many indigenous peoples of the region suggests the importance of the Eastern Eurasian genetic component present

Table 1. Frequencies (%) of Y chromosome haplogroups in the gene pools of Khanty and Mansi

CORE	Размер выборки	C-M217 (xM48, M407)	E-M96	I1-M253	I2a1-P37.2	N2		N3a1-B211	N3a3-VL29	N3a4-Z1936	Q-M242	R1a-M198 (xM458, Z93)	R1b	
						Y3195	VL67						GG400	M269(xGG400)
Mansi	74	1.35	0	1.35	1.35	68.92	1.35	1.35	0	10.81	1.35	1.35	8.11	2.7
Khanty	83	0	2.38	1.19	1.19	26.19	22.62	5.95	2.38	22.62	14.29	0	1.19	0

Note: the gradient from minimum to maximum frequencies is highlighted in shades of red and pink: the more intense the cell color, the higher the frequency

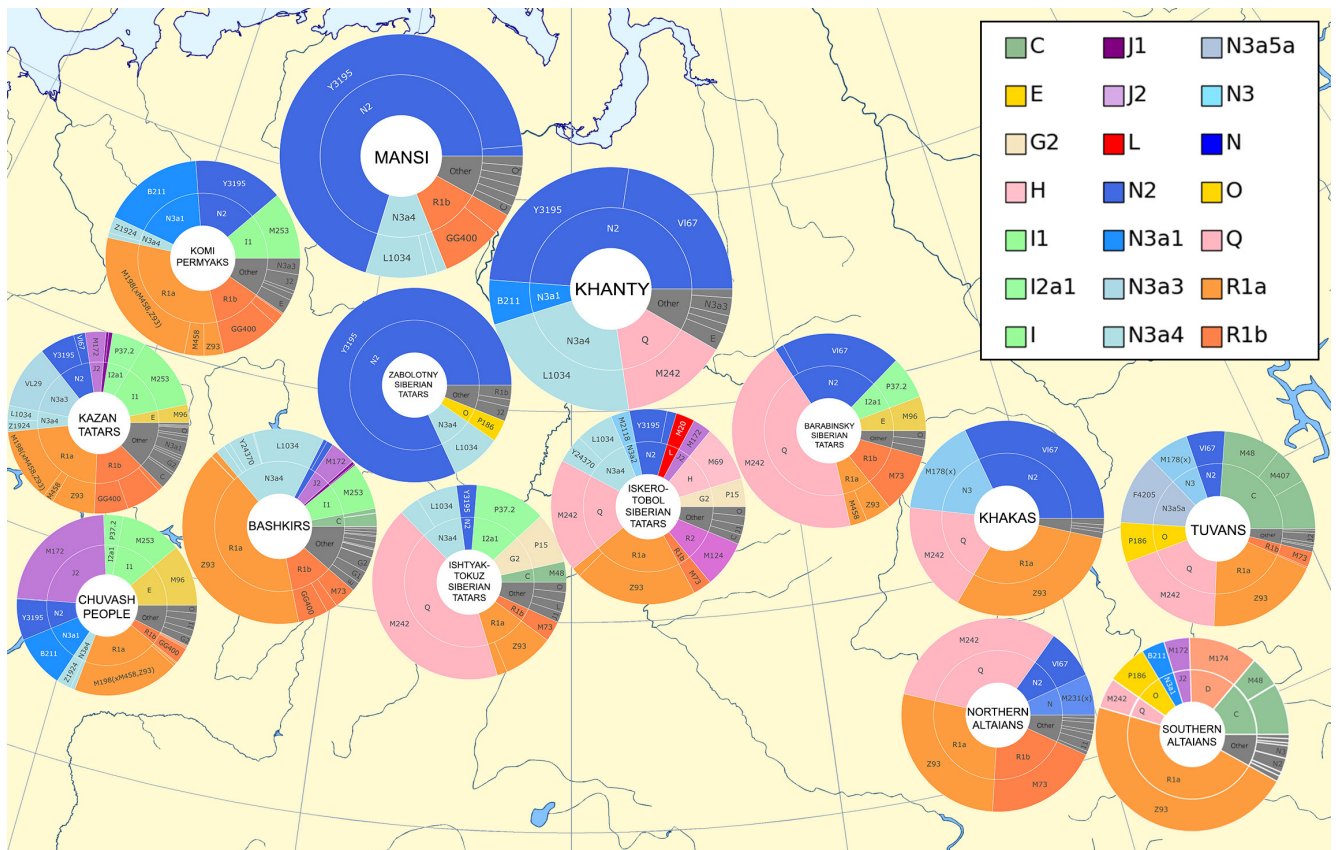


Fig. 2. Genetic portraits of the Ob Ugrians and surrounding indigenous peoples of the Ural region and Western Siberia. The haplogroup name is provided in the inner ring, while the outer ring provides markers determining the haplogroup; the outer ring division into sectors shows the sub-haplogroup frequency in the population. The color of each Y chromosome haplogroup is specified in the map key.

in their gene pools and possibly reflects ancient migration waves from Siberia and Central Asia.

The Mansi gene pool originality manifests itself in high frequency of haplogroup R1b (11%), specifically of its GG400 branch (8%). The Zabolotny Siberian Tatars, who are different from Mansi only by the fact of having the East Asian haplogroup O, are most close to the genetic portrait of Mansi. Among the populations of the Ural region, Komi Permyaks are most close to Mansi, however, in Komi Permyaks, the greater Western Eurasian influence is reflected in the increased frequency of R1a, N3a1, and I1.

Along with the large share of haplogroups N2 and N3a4, high frequency of the haplogroup Q reflecting great Eastern Eurasian

influence is the feature of the Khanty gene pool. Haplogroup Q found in the Altaians, Khakas, Tuvans, most groups of Siberian Tatars can reflect the genetic relationships between the Ob Ugrians and the population, which has migrated from the Central and East Asia.

Gene geography of haplogroups N2 and N3a4

Cartographic analysis was performed for the branches of haplogroups N2 and N3a4 most typical for the Ob Ugrians (Fig. 3, 4). The analysis makes it possible to reveal spatial patterns even when the data on the Y chromosome frequencies in the surrounding populations are incomplete, which makes it

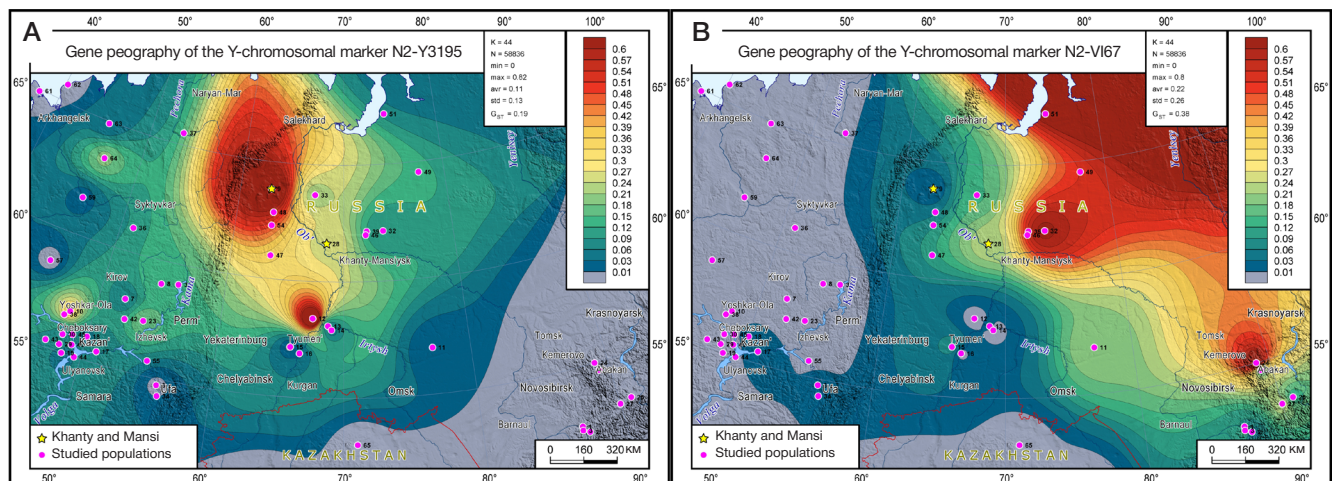


Fig. 3. Gene geography of the Y chromosome haplogroup N2 major branches: western branch Y3195 of the haplogroup N2 (A) and eastern branch VL67 of the haplogroup N2 (B)

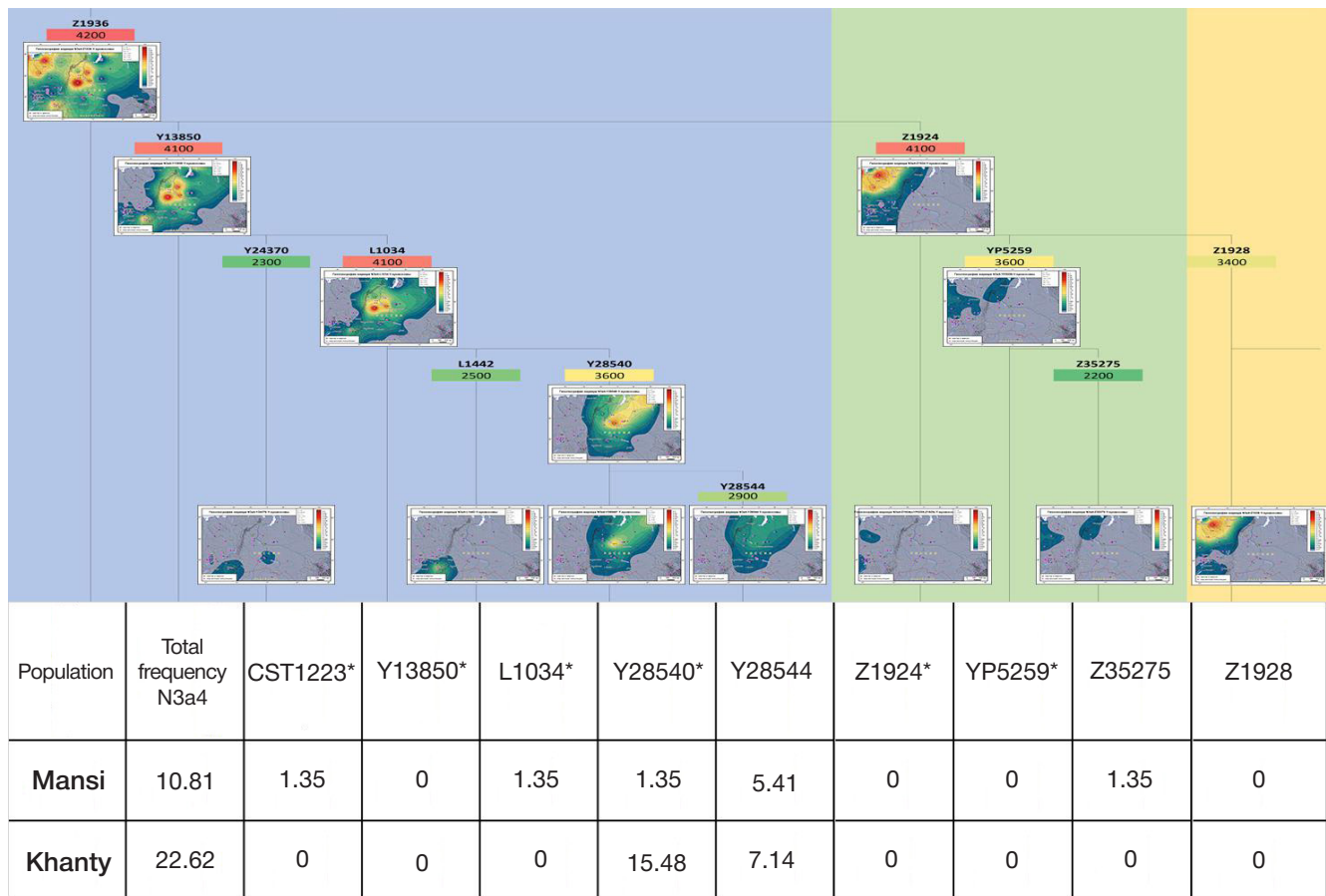


Fig. 4. Phylogenetic structure of the Ural-Siberian cluster N3a4 and frequency of distinct lineages in the populations of Khanty and Mansi. The numbers provided under the names of the branches represent their age in thousands of years ago (kya); the 95% confidence intervals are provided in the text

possible to include the published data about haplogroups N2 and N3a4 in the analysis. Despite different provenance, these haplogroups are contemporaries: the haplogroups emerged about 4.5 kya (5.1–3.8 kya) and began growing phylogenetically.

Haplogroup N2 is divided into the western branch Y3195 (Fig. 3A) and the eastern branch VL67 (Fig. 3B). The eastern branch reaches high frequencies in the Eastern Khanty (60%) and the Gydan Nenets (72%) of the Tazovsky District (Yamalo-Nenets Autonomous Okrug), who maintain their traditional social framework [14]. Our findings suggest a different frequency range: the maximum frequency of VL67 (42%) is typical for Tofalars, the second place is shared by Khanty and Khakas (22%). However, when the Eastern Khanty (Surgutsky and Nizhnevartovsky districts) are specially highlighted, the VL67 frequency increases to 43%, reaching 61% in our small sample of the Forest Nenets.

In Mansi, on the contrary, the eastern branch N2 is the least frequent (1%), and the diversity of N2 is represented by the western branch Y3195 (69%). It reaches its maximum frequency in the Zabolotny Siberian Tatars (81%), while medium frequencies are found in Mari (32%), Khanty (26%), Udmurts (17%), and Komi Permyaks (14%).

The distribution of both branches shows a clear geographical correlation and a common area of overlap at the Ob and Irtysh river basins.

Haplogroup N3a4 is represented by two major branches, Y13850 and Z1924, in the area of Western Siberia and the Ural region most important for genetic history of Ugrians: in the tree of the haplogroup N3a4 (Fig. 4), the Y13850 branch is marked with the blue background, and the Z1924 branch is marked with

the green background. Branch Y13850 emerged about 4.1 kya (4.8–3.4 kya) [1, 14, 23] and gave rise to the Ural-Siberian cluster of the Y-chromosomal lineages. The second branch, Z1924, emerged around that time (4.8–3.4 kya) and spread across the western regions of North Eurasia, showing low frequency in the Ural region, where the areas of both branches (Y13850 and Z1924) overlap. In the studied Ob Ugrian populations almost all samples belong to the large Ural-Siberian cluster Y13850. Exception is only one Mansi sample with the marker Z35275 belonging to the western cluster Z1924 (Fig. 4).

The main branch Y13850 (Fig. 5A) is important as a bridge between the Ugric-speaking peoples: the Ob Ugrians and medieval Magyars (who were among the ancestors of Hungarians) [11, 13]. In the gene pool of the Ob Ugrians, the frequency of the Y13850 lineage reaches 23% in Khanty and 8% in Mansi. It is divided into two sub-branches: Y24370 (not found in the Ob Ugrians) and L1034, to which almost all the studied samples of Khanty and Mansi with the haplogroup N3a4 belong (Fig. 5B). The sub-branch L1034 also gives rise to two descended lineages: L1442 (not found in the Ob Ugrians) and Y28540 that emerged 3.6 kya (4.5–2.7 kya) and included all carriers of N3a4 among Khanty (23%) and the majority of Mansi (7%). The branch Y28540, in turn, is divided into two variants (Fig. 4): the descended variant Y28544 (Fig. 5D) emerged 2.9 kya (3.9–2.1 kya), it is almost equally common in Khanty and Mansi; the root variant Y28540(xY28544) (Fig. 5C) predominates in Khanty (15%) and is rare in Mansi (1%).

Several descended lineages, which can be identified by whole genome sequencing, are likely to hide under high frequency of the root variant in Khanty. Maybe Khanty and Mansi have inherited both branches (Y28544 and basal Y28540(xY28544))

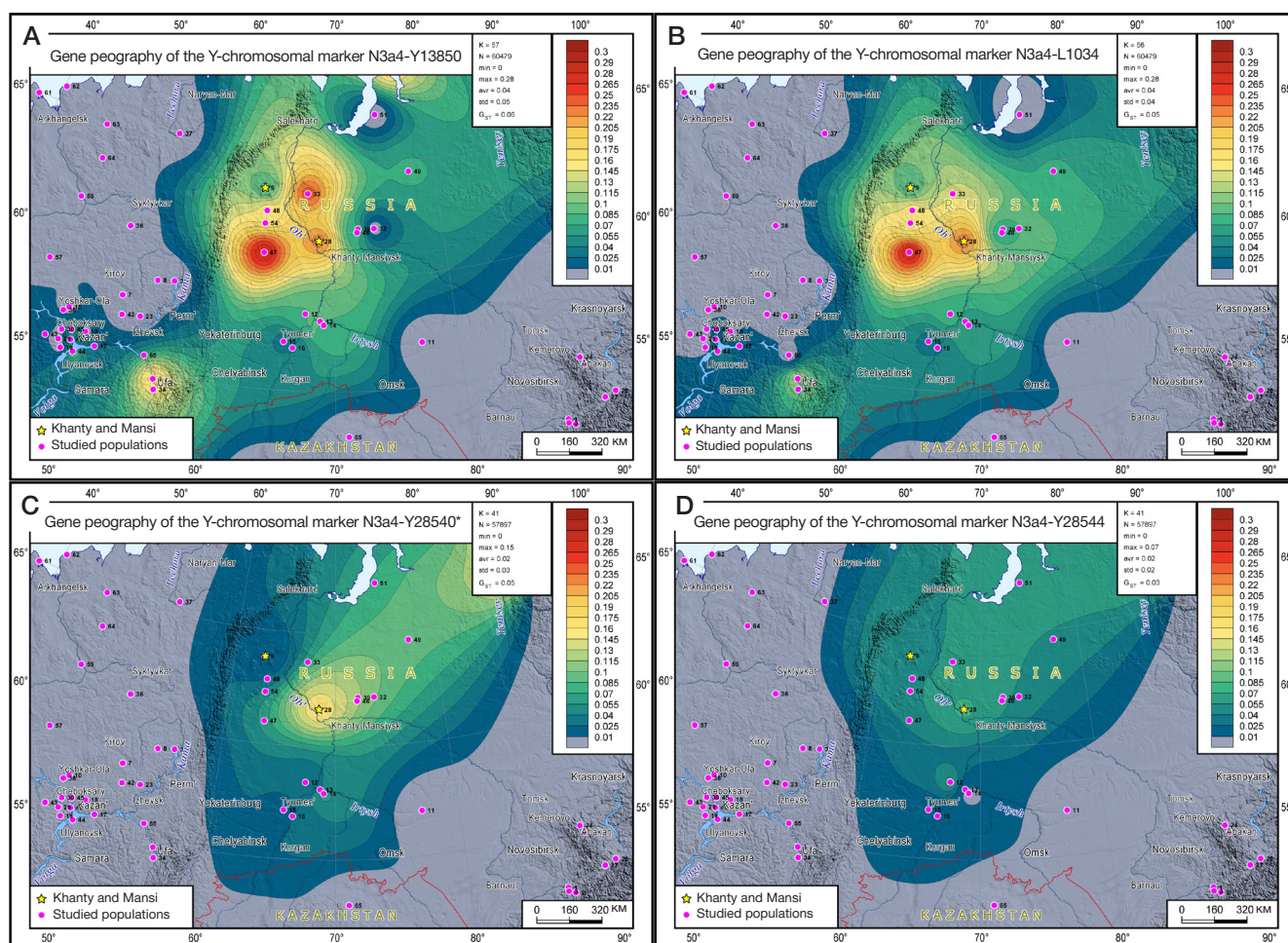


Fig. 5. Geographical distribution and frequency of the haplogroup N3a4 sub-branches. **A.** Gene geography of the haplogroup N3a4 branch Y13850. **B.** Gene geography of the haplogroup N3a4 branch L1034. **C.** Gene geography of the haplogroup N3a4 branch Y28540*. **D.** Gene geography of the haplogroup N3a4 branch Y28544

from their common ancestors, but Khanty maintained high frequency of the root variant due to larger population size and larger area of various endogamous ethnographic groups.

Mapping of the branch L1034 demonstrates that it's common in the Western Siberia and the Urals-Volga region, as well as that there are two frequency surges. The first L1034 maximum is observed in the Khanty included in this study (~23% of frequency), while the second was reported in the Southern Mansi (~27% of frequency), first described in the paper [13] and mentioned in other papers [1, 24]. Such high values of Southern Mansi are dramatically different from both our sample of Northern Mansi (8%) presented in this study and the sample [11] (15%). The differences in the branch L1034 frequency between the geographically distant subpopulations within Khanty and Mansi can be explained by the impact of genetic drift and small sample size, which have eventually shifted the haplogroup frequency ratios.

Position of the Ob Ugrians in the genetic space of indigenous peoples of Russia

The analysis of the position of Ob Ugrians in the genetic space by multidimensional scaling was performed twice: 1) based on the standard panel of 38 Y-SNP markers allowing us to include the literature data on the Nenets and Eastern Khanty [14, 23] (Table 2, Fig. 6A); 2) analysis of the populations of our team based on the extended panel of 48 Y-SNP markers (Table 2, Fig. 6B). The latter included 10 additional haplogroup N3a4 sub-branches,

playing a key role in segregation of the genetic landscapes of Siberia and Europe. The analysis conducted in such a detailed manner for the first time makes it possible to trace the dynamic changes of the gene pool over time. The analysis based on the standard panel of 38 Y-SNP markers considers the pattern of genetic relationships that started to develop more than 4 kya (before the haplogroup N3a4 division into sub-branches). The analysis based on the panel of 48 Y-SNP markers extended through the haplogroup N3a4 subtyping reflects the features of the population of the studied territory in more recent times.

Analysis based on the standard panel of 38 Y-SNP markers. The multidimensional scaling plot based on the standard panel revealed two clusters of populations (Fig. 6A). The Uralic cluster included Finno-Ugric and Turkic populations of the Ural region; Mansi and Zabolotny Siberian Tatars are close to this cluster. The Siberian cluster included all populations of Khanty and Nenets, along with the Turkic-speaking peoples of Siberia. The populations of South Siberia (Todzhints, Tofalars, Tuvans, Khakas) are genetically close to Khanty, but dramatically different from Mansi (Fig. 6A, Table 2). The "geographical principle" is violated by the populations of the Yalutorovsky Siberian Tatars gravitating towards the Uralic cluster and the populations of Bashkirs that are most close to the Siberian cluster.

Analysis based on the extended panel of 48 Y-SNP markers. Position of the Ob Ugrians in the genetic space is clarified when comparing the multidimensional scaling plots (Fig. 6A and Fig. 6B) due to division of the taxonomically important haplogroup N3a4 into sub-branches. It is clear that the Siberian and Uralic clusters

Table 2. Pattern of genetic distances between the Ob Ugrians and surrounding peoples of Russia

Genotyping panel		Standard panel			Expanded panel		
Ethnic and sub-ethnic groups	Sample size (people)	38 SNP			48 SNP		
		Mansi	Khanty	Ob Ugrians	Mansi	Khanty	Ob Ugrians
Mansi	74	0.00	0.41	0.05	0.00	0.41	0.05
Zabolotny Siberian Tatars	83	0.01	0.45	0.07	0.01	0.43	0.06
Ob Ugrians	157	0.05	0.14	0.00	0.05	0.15	0.00
Mari	316	0.26	0.61	0.29	0.25	0.54	0.27
Khanty	84	0.41	0.00	0.14	0.41	0.00	0.15
Komi Permyaks	367	0.79	1.15	0.83	0.79	1.09	0.81
Kazan Tatars	432	1.20	1.24	1.13	1.25	1.35	1.20
Udmurts	260	1.22	1.23	1.14	1.22	1.17	1.12
Iskero-Tobol Siberian Tatars	64	1.47	0.74	1.04	1.58	0.95	1.21
Chuvash people of Tatarstan	80	1.47	1.57	1.42	1.49	1.50	1.41
Chuvash people of the Chuvash Republic	419	1.52	1.90	1.57	1.52	1.84	1.54
Mishar Tatars	158	1.84	2.03	1.82	1.85	1.97	1.81
Kryashen Tatars	66	1.91	2.06	1.88	1.91	2.00	1.85
Besermyan	49	2.00	1.53	1.70	1.99	1.47	1.68
Ishtyak-Tokuz Siberian Tatars	59	2.18	0.88	1.38	2.32	0.88	1.44
Bashkirs	943	2.57	1.54	1.95	3.31	3.20	3.18
Yalutorovsky Siberian Tatars	77	2.57	2.12	2.28	2.66	2.24	2.40
Barabinsky Siberian Tatars	70	2.90	0.68	1.38	2.90	0.62	1.36
Todzhints	69	3.88	1.16	1.92	3.87	1.10	1.90
Tyumen Siberian Tatars	59	3.90	1.97	2.62	4.21	2.04	2.76
Tofalars	42	3.91	0.89	1.68	3.91	0.83	1.65
Khakas	178	3.92	0.82	1.61	3.93	0.76	1.59
Tuvans	390	3.95	1.31	2.06	3.95	1.25	2.04
Northern Altaians	161	4.02	1.18	1.95	4.03	1.12	1.93
Southern Altaians	175	4.63	2.67	3.33	4.63	2.61	3.30
Kazym Khanty [14]	54	0.72	0.15	0.37	—	—	—
Ruskinskaya Khanty [14]	64	1.85	0.56	1.06	—	—	—
Gydan Nenets [23]	322	2.23	0.43	1.06	—	—	—

become more distinct and drift further apart in the genetic space. Furthermore, Khanty come out of the Siberian cluster and get closer to the Uralic cluster. Bashkirs, on the contrary, move further away from both Uralic cluster and Khanty: after turning to the extended panel their genetic distances from Khanty increase 2-fold: from $d = 1.5$ to $d = 3.2$ (Table 2) due to significant contribution of haplogroup N3a4 to their gene pool. Mansi together with the Zabolotny Siberian Tatars, Khanty and Finnish-speaking populations of the Ural region (Mari, $d = 0.2$; Komi Permyaks, $d = 0.8$; Udmurts, $d = 1.2$) form the new Proto-Uralic cluster (Fig. 6B).

DISCUSSION

Continuing research on the Ugric peoples, we have significantly refined and expanded the conclusions drawn in the previous paper [16].

Further analysis of the sub-branches N2 and N3a4 and the position of the Ob Ugrians in the genetic space based on the extended panel of the Y-chromosomal markers confirmed the dominant influence of the Eastern Eurasian component on their gene pools. However, the N2 branch gene geography has revealed significant differences: in Khanty, the eastern VI67 and the western Y3195 branches are almost equally frequent (about a quarter of the gene pool each); in Mansi, the western branch constitutes two thirds of the entire Mansi Y gene pool, while the eastern branch is almost absent. Considering migrations of the Mansi population, this can indicate that the Proto-Mansi populations lived far to the west of the current area during the period of the N2 sub-branch development, so the effect of the Eastern Eurasian component is small. This is also confirmed by the position of Mansi in the genetic space: the Finno-Ugric peoples of the Volga region are genetically most close to Mansi.

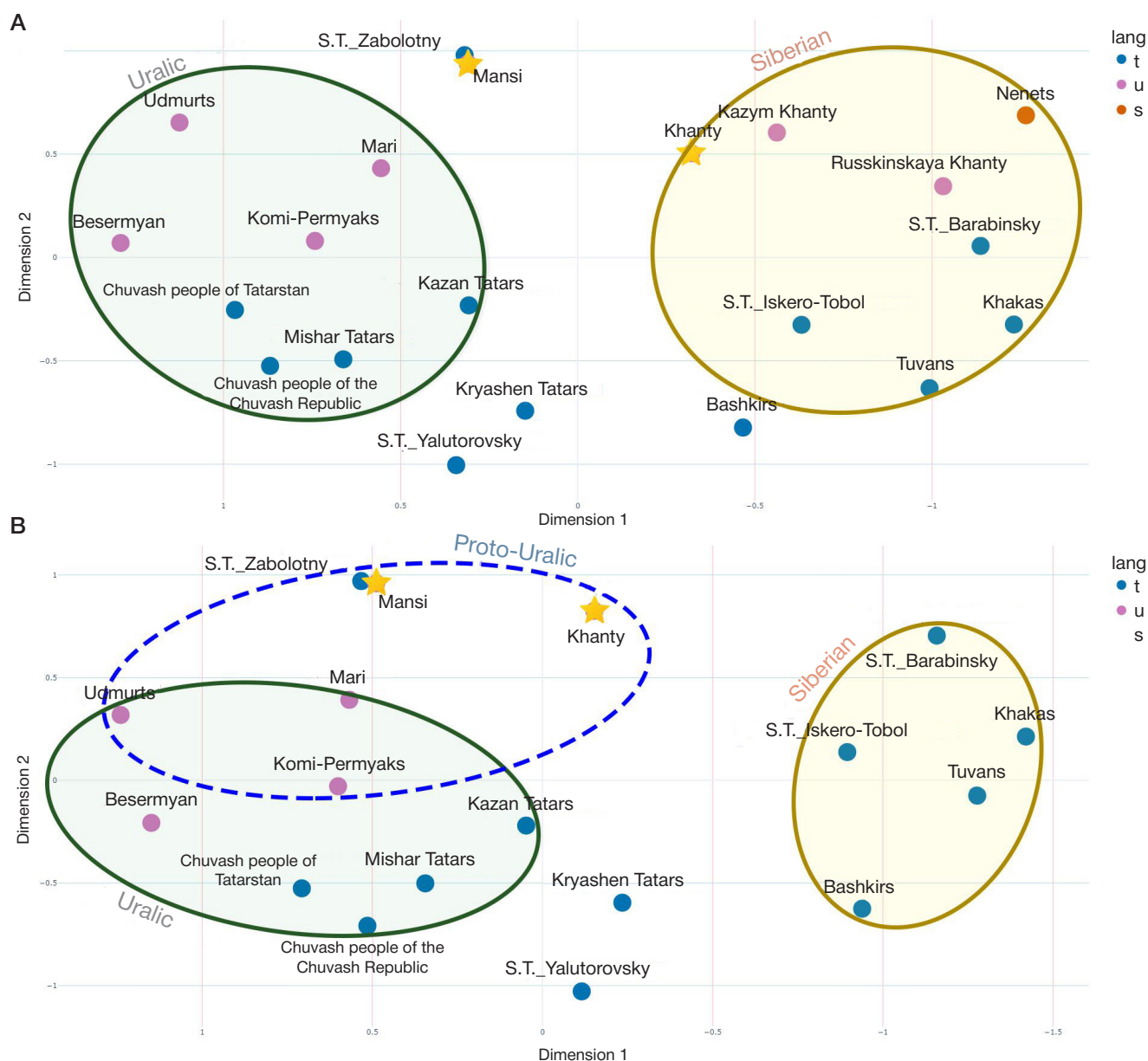


Fig. 6. Position of the Ob Ugrians in the genetic space of surrounding peoples of Russia. **A.** Multidimensional scaling plot based on the standard panel of 38 Y-chromosomal markers (stress coefficient = 0.14; alienation coefficient = 0.16). **B.** Multidimensional scaling plot based on the extended panel of 48 Y-chromosomal markers (stress coefficient = 0.16; alienation coefficient = 0.17). The peoples speaking languages of the Uralic linguistic family are highlighted in pink; those belonging to the Turkic branch of the Altaic language family are highlighted in blue

The Khanty genetic history reconstruction seems to be more challenging. Considering the large expanse of territory and equal contributions of the Eastern Eurasian and Uralic components, several hypotheses about the development of their gene pool can be proposed. Maybe the ancestors of Khanty originate from the territory of the Ural region; they assimilated local populations, in the gene pool of which the haplogroup Q was highly significant, when moving northeast. This hypothesis is based on the equal contribution of both N2 branches and high frequency of the eastern N2 branch in the Northern and Eastern Khanty, as well as on the increased frequency of N3a4. The multidimensional scaling plot confirms this hypothesis, showing that Khanty get closer to the peoples of the Finno-Ugric Volga region due to the haplogroup N3a4 contribution in Khanty. According to another hypothesis, the Proto-Khanty populations initially had both N2 branches, the Uralic cluster N3a4, and the Siberian haplogroup Q. This hypothesis is based on the fact that various groups have a similar spectrum of dominant haplogroups, despite broad ethnographic area of Khanty.

We can say unequivocally that the genetic relationships of the Ob Ugrians associated with the haplogroup N3a4 are ancient; these are important for the Ugrian gene pool reconstruction, along with the haplogroup N2 branches. Thus, our study is the first to demonstrate that the phylogenetic structure of haplogroups N3a4 and N2 clearly divides the gene pools of Siberia and Europe, which is an important step on the way to understanding the population history of the region.

CONCLUSIONS

We conducted a comprehensive analysis of the Khanty and Mansi gene pools based on the broad spectrum of the Y chromosome haplogroups in the context of indigenous populations of the Western Siberia and Urals. Thorough investigation of the haplogroup N3a4 eastern cluster phylogenetic structure has made it possible to clarify the time frame and the directions of migration in the region. Position of the Ob Ugrians among indigenous peoples of Siberia and Urals has been determined:

Mansi are genetically close to the populations of the Urals-Volga region; Khanty are intermediate between the Uralic and Siberian clusters, which reflects the complex historical interactions and mixing of genetic components. Thus, the aim of the study has

been achieved; the findings extend understanding of the genetic history of the Ob Ugrians and emphasize the importance of the in-depth analysis of the Y chromosome haplogroup branches for reconstruction of historical processes in the region.

References

1. Post H, Németh E, Klima L, Flores R, Fehér T, Türk A, et al. Y-chromosomal connection between Hungarians and geographically distant populations of the Ural Mountain region and West Siberia. *Scientific Reports*. 2019; 9 (1): 7786.
2. Gemuev IN, i dr., redaktory. *Narody Zapadnoj Sibiri: Hanty. Mansi. Sel'kupy. Nency. Jency. Nganasany. Kety. M.*: Nauka, 2005; 804 s. Russian.
3. Sokolova ZP. *Hanty i mansi. Vzgljad iz XXI veka. M.*: Nauka, 2009; 755 s. Russian.
4. Tomilov NA. *Tjurkojazychnoe naselenie Zapadno-Sibirskoj ravniny v konce XVI pervoj chetverti XIX vv. Izd-vo TGU. Tomsk, 1981; 276 s. Russian.*
5. Bakieva GT, Kvashnin YuN. *Tjurki, samodijcy i ugray v Tobol'skom Priirtysh'e (k voprosu ob jetnogeneze zabolotnyh tatar). Bol'she, chem jetnograf. Sbornik nauchnyh statej, posvjashhennyh pamjati professora V. V. Pimenova. M.*: Izd-vo MGU, 2015. Russian.
6. Agdzhoyan AT, Balanovska EV, Padyukova AD, Dolinina DO, Kuznetsova MA, Zaporozhchenko VV, et al. Gene pool of Siberian Tatars: Five ways of origin for five subethnic groups. *Molecular biology*. 2016; 50: 860–73.
7. Karafet TM, Osipova LP, Gubina MA, Posukh OL, Zegura SL, Hammer MF. High Levels of Y-Chromosome Differentiation among Native Siberian Populations and the Genetic Signature of a Boreal Hunter-Gatherer Way of Life. *Human Biology*. 2002; 74 (6): 761–89.
8. Karafet TM, Osipova LP, Savina OV, Hallmark B, Hammer MF. Siberian genetic diversity reveals complex origins of the Samoyedic-speaking populations. *American Journal of Human Biology*. 2018; 30 (6): e23194.
9. Rootsi S, Zhivotovsky LA, Baldovitch M, Kayser M, Kutuev IA, Khusainova R, et al. A counter-clockwise northern route of the Y-chromosome haplogroup N from Southeast Asia towards Europe. *European Journal of Human Genetics*. 2007; 15 (2): 204–11.
10. Pimenoff VN, Comas D, Palo JU, Vershubsky G, Kozlov A, Sajantila A. Northwest Siberian Khanty and Mansi in the junction of West and East Eurasian gene pools as revealed by uniparental markers. *European Journal of Human Genetics*. 2008; 16 (10): 1254–64.
11. Fehér T, Németh E, Vándor A, Kornienko IV, Csáji LK, Pamjav H. Y-SNP L1034: limited genetic link between Mansi and Hungarian-speaking populations. *Molecular Genetics and Genomics*. 2015; 290 (1): 377–86.
12. Ilumäe AM, Reidla M, Chukhryaeva M, Järve M, Post H, Karmin M, et al. Human Y Chromosome Haplogroup N: A Non-trivial Time-Resolved Phylogeography that Cuts across Language Families. *The American Journal of Human Genetics*. 2016; 99 (1): 163–73.
13. Dudás E, Vágó-Zalán A, Vándor A, Saypasheva A, Pomozi P, Pamjav H. Genetic history of Bashkirian Mari and Southern Mansi ethnic groups in the Ural region. *Molecular Genetics and Genomics*. 2019; 294 (4): 919–30.
14. Kharkov VN, Kolesnikov NA, Valikhova LV, Zarubin AA, Svarovskaya MG, Marusin AV, Khitrinskaya IY, Stepanov VA, i dr. Svjaz' genofonda hantov s narodami Zapadnoj Sibiri, Predural'ja i Altaja-Sajana po dannym o polimorfizme avtosomnyh lokusov i Y-hromosomy. *Vavilovskij zhurnal genetiki i selekcii*. 2023; 27 (1): 46–54. Russian.
15. Karafet TM, Mendez FL, Meilerman MB, Underhill PA, Zegura SL, Hammer MF. New binary polymorphisms reshape and increase resolution of the human Y chromosome haplogroup tree. *Genome Research*. 2008; 18 (5): 830–38.
16. Agdzhoyan AT, Ponomarev GYu, Salnikova AD, Lavryashina MB, Pylev VYu, Balanovskaya EV. Genofond obshchih ugrov v kontekste narodonaseleniya severnoj Evrazii: genogeografija vetvej gaplogruppy N Y-hromosomy. Kul'turnoe nasledie narodov Zapadnoj Sibiri: ugray. V sbornike: Materialy IX Vserossijskogo simpoziuma «Kul'turnoe nasledie narodov Zapadnoj Sibiri: ugray»; 8–9 dekabrya 2022 g.; Kirov: Izd-vo MCITO, 2023; s. 31–36. Russian.
17. Tree YTree v12.03.00 (cited 31 august 2024) Available from: <https://www.yfull.com/sc/tree/>
18. Balanovsky O, Rootsi S, Pshenichnov A, Kivisild T, Churnosov M, Evseeva I, et al. Two Sources of the Russian Patrilineal Heritage in Their Eurasian Context. *The American Journal of Human Genetics*. 2008; 82 (1): 236–50.
19. Koshel SM. *Geoinformacionnye tehnologii v genogeografii. V knige: I. K. Lur'e, V. I. Kravcova, redaktory. Sovremennaja geograficheskaja kartografija. M.*: Data+, 2012; s. 158–66. Russian.
20. Tambets K, Yunusbayev B, Hudjashov G, Ilumäe AM, Rootsi S, Honkola T, et al. Genes reveal traces of common recent demographic history for most of the Uralic-speaking populations. *Genome Biology*. 2018; 19 (1): 139.
21. Triska P, Chekanov N, Stepanov V, Khusnutdinova EK, Kumar GPA, Akhmetova V, et al. Between Lake Baikal and the Baltic Sea: genomic history of the gateway to Europe. *BMC Genetics*. 2017; 18 (S1): 110.
22. Wong EHM, Khrunin A, Nichols L, Pushkarev D, Khokhrin D, Verbenko D, et al. Reconstructing genetic history of Siberian and Northeastern European populations. *Genome Research*. 2017; 27 (1): 1–14.
23. Kharkov VN, Valihova LV, Yakovleva EL, Serebrova VN, Kolesnikov NA, Petelina TI, i dr. Rekonstrukcija proishozhdeniya gydanskikh nencev na osnove geneticheskogo analiza ih rodovoj struktury s pomoshh'ju novogo nabora YSTR-markerov. *Genetika*. 2021; 57 (12): 1403–14. Russian.
24. Pamjav H, Dudás E, Krizsán K, Galambos A. A Y-chromosomal study of mansi population from konda River Basin in Ural. *Forensic Science International: Genetics Supplement Series*. 2019; 7 (1): 602–03.

Литература

1. Post H, Németh E, Klima L, Flores R, Fehér T, Türk A, et al. Y-chromosomal connection between Hungarians and geographically distant populations of the Ural Mountain region and West Siberia. *Scientific Reports*. 2019; 9 (1): 7786.
2. Гемуев И. Н. и др., редакторы. *Народы Западной Сибири: Ханты. Манси. Селькупы. Ненцы. Энцы. Нганасаны. Кеты. М.*: Наука, 2005; 804 с.
3. Соколова З. П. *Ханты и манси. Взгляд из XXI века. М.*: Наука, 2009; 755 с.
4. Томилов Н. А. *Тюркоязычное население Западно-Сибирской равнины в конце XVI первой четверти XIX вв. Изд-во ТГУ. Томск, 1981; 276 с.*
5. Бакиева Г. Т., Квашнин Ю. Н. *Тюрки, самодийцы и угры в Тобольском Прииртышье (к вопросу об этногенезе заболотных татар). Больше, чем этнограф. Сборник научных статей, посвященных памяти профессора В. В. Пименова. М.*: Изд-во МГУ, 2015.
6. Агджоян А. Т., Балановская Е. В., Падюкова А. Д., Долинина Д. О., Кузнецова М. А., Запороженко В. В. и др. *Генофонд сибирских татар: пять субэтносов — пять путей этногенеза.*

- Молекулярная биология. 2016; 50 (6): 978–91.
7. Karafet TM, Osipova LP, Gubina MA, Posukh OL, Zegura SL, Hammer MF. High Levels of Y-Chromosome Differentiation among Native Siberian Populations and the Genetic Signature of a Boreal Hunter-Gatherer Way of Life. *Human Biology*. 2002; 74 (6): 761–89.
 8. Karafet TM, Osipova LP, Savina OV, Hallmark B, Hammer MF. Siberian genetic diversity reveals complex origins of the Samoyedic-speaking populations. *American Journal of Human Biology*. 2018; 30 (6): e23194.
 9. Rootsi S, Zhivotovsky LA, Baldovici M, Kayser M, Kutuev IA, Khusainova R, et al. A counter-clockwise northern route of the Y-chromosome haplogroup N from Southeast Asia towards Europe. *European Journal of Human Genetics*. 2007; 15 (2): 204–11.
 10. Pimenoff VN, Comas D, Palo JU, Vershubsky G, Kozlov A, Sajantila A. Northwest Siberian Khanty and Mansi in the junction of West and East Eurasian gene pools as revealed by uniparental markers. *European Journal of Human Genetics*. 2008; 16 (10): 1254–64.
 11. Fehér T, Németh E, Vándor A, Kornienko IV, Csáji LK, Pamjav H. Y-SNP L1034: limited genetic link between Mansi and Hungarian-speaking populations. *Molecular Genetics and Genomics*. 2015; 290 (1): 377–86.
 12. Ilumäe AM, Reidla M, Chukhryaeva M, Järve M, Post H, Karmin M, et al. Human Y Chromosome Haplogroup N: A Non-trivial Time-Resolved Phylogeography that Cuts across Language Families. *The American Journal of Human Genetics*. 2016; 99 (1): 163–73.
 13. Dudás E, Vágó-Zalán A, Vándor A, Saypasheva A, Pomozi P, Pamjav H. Genetic history of Bashkirian Mari and Southern Mansi ethnic groups in the Ural region. *Molecular Genetics and Genomics*. 2019; 294 (4): 919–30.
 14. Харьков В. Н., Колесников Н. А., Валихова Л. В., Зарубин А. А., Сваровская М. Г., Марусин А. В. и др. Связь генофонда хантов с народами Западной Сибири, Предуралья и Алтая-Саян по данным о полиморфизме аутосомных локусов и Y-хромосомы. *Вавиловский журнал генетики и селекции*. 2023; 27 (1): 46–54.
 15. Karafet TM, Mendez FL, Meilerman MB, Underhill PA, Zegura SL, Hammer MF. New binary polymorphisms reshape and increase resolution of the human Y chromosomal haplogroup tree. *Genome Research*. 2008; 18 (5): 830–38.
 16. Агджоян А. Т., Пономарёв Г. Ю., Сальникова А. Д., Лавряшина М. Б., Пылёв В. Ю., Балановская Е. В. Генофонд обских угров в контексте народонаселения северной Евразии: геногеография ветвей гаплогруппы N Y-хромосомы. Культурное наследие народов Западной Сибири: угры. В сборнике: Материалы IX Всероссийского симпозиума «Культурное наследие народов Западной Сибири: угры»; 8–9 декабря 2022 г.; Киров: Изд-во МЦИТО, 2023; с. 31–36.
 17. Tree YTree v12.03.00 (cited 31 august 2024) Available from: <https://www.yfull.com/sc/tree/>
 18. Balanovsky O, Rootsi S, Pshenichnov A, Kivisild T, Churnosov M, Evseeva I, et al. Two Sources of the Russian Patrilineal Heritage in Their Eurasian Context. *The American Journal of Human Genetics*. 2008; 82 (1): 236–50.
 19. Кошель С. М. Геоинформационные технологии в геногеографии. В книге: И. К. Лурье, В. И. Кравцова, редакторы. Современная географическая картография. М.: Дата+, 2012; с. 158–66.
 20. Tambets K, Yunusbayev B, Hudjashov G, Ilumäe AM, Rootsi S, Honkola T, et al. Genes reveal traces of common recent demographic history for most of the Uralic-speaking populations. *Genome Biology*. 2018; 19 (1): 139.
 21. Triska P, Chekanov N, Stepanov V, Khusnutdinova EK, Kumar GPA, Akhmetova V, et al. Between Lake Baikal and the Baltic Sea: genomic history of the gateway to Europe. *BMC Genetics*. 2017; 18 (S1): 110.
 22. Wong EHM, Khrunin A, Nichols L, Pushkarev D, Khokhrin D, Verbenko D, et al. Reconstructing genetic history of Siberian and Northeastern European populations. *Genome Research*. 2017; 27 (1): 1–14.
 23. Харьков В. Н., Валихова Л. В., Яковлева Е. Л., Сереброва В. Н., Колесников Н. А., Петелина Т. И. и др. Реконструкция происхождения гыданских ненцев на основе генетического анализа их родовой структуры с помощью нового набора YSTR-маркеров. *Генетика*. 2021; 57 (12): 1403–14.
 24. Pamjav H, Dudás E, Krizsán K, Galambos A. A Y-chromosomal study of mansi population from konda River Basin in Ural. *Forensic Science International: Genetics Supplement Series*. 2019; 7 (1): 602–03.