

## GENE POOL OF THE URAL-VOLGA REGION: GENETIC HISTORY OF MORDOVIA'S POPULATION BASED ON THE Y-CHROMOSOMAL HAPLOGROUP N3A1-Y23475 PHYLOGEOGRAPHY

Agdzhoyan AT<sup>1</sup>✉, Adamov DS<sup>1</sup>, Potanina AY<sup>1</sup>, Voronina MM<sup>1</sup>, Gorin IO<sup>1</sup>, Shtrunov-Shlykov AG<sup>1</sup>, Koshel SM<sup>2</sup>, Balanovska EV<sup>1</sup>, Ponomarev GYu<sup>1</sup>

<sup>1</sup> Research Centre for Medical Genetics, Moscow, Russia

<sup>2</sup> Lomonosov Moscow State University, Moscow, Russia

It has been shown that Y-haplogroup N3a1-B211 is common in the Finnish-speaking peoples of the Ural-Volga region. The study aimed to investigate gene geography and phylogeography of the westernmost variant of this haplogroup: the N3a1-Y23475 branch. Comprehensive genotyping of 395 haplogroup N3a1-B211 carriers from 29 populations of Eastern Europe, Ural-Volga region, and Siberia revealed 78 carriers of its western branch reaching its maximum frequency in Mordovia's populations (8% in Moksha, 9% in Erzya, 25% in Tengushevsky Erzya-Shoksha). Low N3a1-Y23475 frequencies in the Turkic-speaking and Slavic populations suggest the role of the Finnish-speaking substrate in their gene pools. According to the phylogenetic analysis data, the N3a1-Y23475 branch emerged 2.3–2.7 thousand years ago, but active accumulation of its current diversity took place mainly in the populations of Mordovia during the last millennium. We performed DNA genotyping in 74 haplogroup N3a1-Y23475 carriers using the 37 Y-STR panel. The Y-STR haplotype phylogenetic network created suggests two periods of population growth in ancestors of Mordovia's indigenous population: about 1000 years ago in the populations of proto-Erzya and proto-Shoksha, about 500 years ago in the populations of Moksha and Shoksha. The fact of finding haplogroup N3a1-Y23475 in the Northern and Southern Altaians requires further research. Position of Northern Altaians-Kumandins in the phylogenetic network presumably reflects migration of Mordovia's population to Altai in the 19<sup>th</sup>–20<sup>th</sup> centuries. The age estimates reported for Southern Altaians-Telengits can indicate the association with the haplogroup N3a1 ancestral homeland in South Siberia before resettlement of its ancient carriers in the Ural-Volga region about 1.7 thousand years ago.

**Keywords:** gene pool, gene geography, phylogeography, Ural region, Mordovia, Y chromosome, haplogroup N3a1-Y23475, TMRCA estimates

**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 (sample donors) and Biobank of North Eurasia (for access to DNA collections).

**Author contribution:** Balanovska EV — management; Shtrunov-Shlykov AG — expedition survey of the populations of Mordovia; Ponomarev GYu, Voronina MM, Adamov DS — Y-SNP and Y-STR marker genotyping; Agdzhoyan AT, Ponomarev GYu, Adamov DS, Gorin IO, Potanina AY, Koshel SM — statistical, phylogenetic, cartographic analysis; Balanovska EV, Agdzhoyan AT — 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:** Anastasia T. Agdzhoyan  
Moskvorechye, d. 1, 115522, Moscow, Russia; aagdzhoyan@gmail.com

**Received:** 10.12.2024 **Accepted:** 20.12.2024 **Published online:** 29.12.2024

**DOI:** 10.24075/brsmu.2024.066

## ГЕНОФОНД УРАЛО-ПОВОЛЖЬЯ: ГЕНЕТИЧЕСКАЯ ИСТОРИЯ НАСЕЛЕНИЯ МОРДОВИИ ПО ДАННЫМ ФИЛОГЕОГРАФИИ ГАПЛОГРУППЫ N3A1-Y23475 Y-ХРОМОСОМЫ

А. Т. Агджоян<sup>1</sup>✉, Д. С. Адамов<sup>1</sup>, А. Ю. Потанина<sup>1</sup>, М. М. Воронина<sup>1</sup>, И. О. Горин<sup>1</sup>, А. Г. Штрунов-Шлыков<sup>1</sup>, С. М. Кошель<sup>2</sup>, Е. В. Балановская<sup>1</sup>, Г. Ю. Пономарёв<sup>1</sup>

<sup>1</sup> Медико-генетический научный центр имени Н. П. Бочкова, Москва, Россия

<sup>2</sup> Московский государственный университет имени М. В. Ломоносова, Москва, Россия

Показано, что Y-гаплогруппа N3a1-B211 распространена в популяциях финноязычных народов Урало-Поволжья. Целью исследования было изучение геногеографии и филогеографии самого западного варианта этой гаплогруппы: ветви N3a1-Y23475. При детальном генотипировании 395 носителей гаплогруппы N3a1-B211 из 29 популяций Восточной Европы, Урало-Поволжья и Сибири выявлено 78 носителей ее западной ветви, которая достигает максимальных частот в популяциях Мордовии (8% — у мокши, 9% — у эрзи, 25% — у теньгушевских эрзи-шокши). Низкие частоты N3a1-Y23475 в тюркоязычных и славянских популяциях могут указывать на роль финноязычного субстрата в их генофондах. По данным филогенетического анализа ветвь N3a1-Y23475 возникла 2,3–2,7 тыс. лет назад, но активное накопление ее современного разнообразия происходило преимущественно в популяциях Мордовии в течение последнего тысячелетия. Проведено генотипирование ДНК 74 носителей гаплогруппы N3a1-Y23475 по панели 37 Y-STR. Созданная филогенетическая сеть Y-STR гаплотипов указывает на два периода демографического роста у предков коренного населения Мордовии: около 1000 лет назад в популяциях праэрзи и прашокши, около 500 лет назад — в популяциях мокши и шокши. Выявленное наличие гаплогруппы N3a1-Y23475 у северных и южных алтайцев требует продолжения исследования. Положение северных алтайцев-кумандинцев на филогенетической сети предположительно отражает миграцию мордовского населения на Алтай в XIX–XX вв. Генетическая датировка для южных алтайцев-теленгитов может указывать на связь с прародиной гаплогруппы N3a1 в Южной Сибири до расселения ее древних носителей в Урало-Поволжье около 1,7 тыс. лет назад.

**Ключевые слова:** генофонд, геногеография, филогеография, Уральский регион, Мордовия, Y-хромосома, гаплогруппа N3a1-Y23475, генетические датировки TMRCA

**Финансирование:** Государственное задание Министерства науки и высшего образования РФ для Медико-генетического научного центра им. 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. Г. Штрунов-Шлыков — экспедиционное обследование популяций Мордовии; Г. Ю. Пономарёв, М. М. Воронина, Д. С. Адамов — генотипирование Y-SNP и Y-STR маркеров; А. Т. Агджоян, Г. Ю. Пономарёв, Д. С. Адамов, И. О. Горин, А. Ю. Потанина, С. М. Кошель — статистический, филогенетический, картографический анализ; Е. В. Балановская, А. Т. Агджоян — 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:** Анастасия Торосовна Агджоян  
ул. Москворечье, д. 1, 115522, г. Москва, Россия; aagdzhoyan@gmail.com

**Статья получена:** 10.12.2024 **Статья принята к печати:** 20.12.2024 **Опубликована онлайн:** 29.12.2024

**DOI:** 10.24075/vrgmu.2024.066

Y-haplogroup N3 represents one of the basic components of the North Eurasian gene pool and is considered to be a marker of the ancient population expansion, during which the Uralic languages spread [1–2]. Most of the haplogroup N3 range is the territory of Russia, where there are currently more than 1.6 million speakers of languages of the Uralic linguistic group belonging to 20 ethnic groups [3]. More than 90% of the population of the Uralic language-speakers in the Russian Federation is represented by the Finnish-speaking peoples of the Ural-Volga region of three language groups: Mari (Mari), Mordovian (Moksha, Shoksha, Erzya), and Permian (Besermyan, Komi-Zyryans, Komi-Permyaks, Udmurts). Among branches of haplogroup N3 most common in Europe (N3a1, N3a3, N3a4), high significant correlation with the Uralic ancestral component of the autosomal genome was reported for haplogroup N3a1 only [4].

Haplogroup N3a1 is common in the gene pools of peoples of the Ural region and is rare outside the region: in Udmurts (67%), Komi-Zyryans (18–43%), Chuvash (20%), Khanty and Mansi (19%), Komi-Permyaks (12%), Mari (14%), and Mordvins (5–10%) [1]. The spread of N3a1 is characterized by considerable frequency variation within its range: from 1% on the outskirts of the range (in Bashkirs, Belarusians, Karelians, Russians, Khakas) to the world's maximum in Udmurts (67%). This haplogroup, like other N3 lineages, could be brought to Europe by the population related to the Seima-Turbino transcultural phenomenon in the Bronze Age, despite the fact that this hypothesis has not yet been confirmed by direct paleoDNA analysis [5–6].

The structure of haplogroup N3a1-B211, common in the Finnish-speaking peoples of the Ural-Volga region, and the time of origin of its branches are poorly understood, but basic understanding is ensured by the YFull open source data [7]. Among the commercial testing participants, a common one in the populations of the Ural-Volga region is the N3a1-Y23475 branch spread across the western part of the region: in Mordvins-Erzya, Volga Tatars, and Russians (Bryansk, Nizhny Novgorod, Penza, Kirov, and Sverdlovsk regions). According to the YFull data, active accumulation of today's diversity within N3a1-Y23475 took place on average in the last 2.4–2.7 thousand years.

The study aimed to investigate gene geography and phylogenetic structure of haplogroup N3a1-Y23475 based on the extensive data on the populations of North Eurasia.

## METHODS

Biological samples were collected after obtaining the informed consent from the donors during the expedition surveys managed by Professor E.V. Balanovska and RAS Professor O.P. Balanovsky in accordance with the same program that had been described earlier [8]. The donors were unrelated adult males, whose ancestors for at least three generations considered themselves to belong to this ethnic group and were born into this population. In Mordovia, the Moksha and Erzya ethnic groups were assessed, including the distinct group of Erzya living in the Tengushevsky District, hereinafter referred to as Shoksha.

DNA was isolated from the venous blood samples using the QIAAsymphony SP nucleic acid purification system or by phenol-chloroform extraction using proteinase K; the sample preparation phases had been described earlier [9]. Among 4051 samples (Table 1) there were 395 samples of the haplogroup N3a1-B211 carriers from 29 populations of Eastern Europe, Ural-Volga region, and Siberia. For these genotyping

of the Y23475 SNP marker was performed by real-time PCR using the TaqMan probes and the OpenArray technique in the QuantStudio 12 Flex thermocycler (Thermo Fisher Scientific, USA). A total of 78 carriers of the N3a1-Y23475 branch were identified. Of those the results of fragment analysis (37 YSTR markers) were obtained for 74 samples using the commercially available Yfiler Plus (Thermo Fisher Scientific, USA), Powerplex Y23 (Promega, USA) kits and the Nanophore 05 genetic analyzer (Syntol, Russia).

Cartographic analysis was performed using the original GeneGeo cartographic software package [10] developed under the leadership of E. V. Balanovska and O. P. Balanovsky. The gene geographic map of the haplogroup N3a1-Y23475 spread was created based on the genotyping data by the weighted average interpolation (radius 400 km, and weight function degree 3).

The N3a1-Y23475 phylogenetic network was constructed based on the median-joining principle [11] using the Network v.10.2.0.0 software tool (Fluxus Technology Ltd, UK). The network image was visualized in Network Publisher v.2.1.2.5 (Fluxus Technology Ltd, UK). The weight of each of 37 STR markers was considered to be 10 with  $\epsilon = 0$ . The time to the most recent common ancestor (TMRCA) for the entire phylogenetic network and the clusters distinguished (Table 2) was calculated by the ASD method [12]. We excluded the DYF387S1b locus due to partial AZFc deletions in the P1 palindromic sequence in the haplogroup N3-M178 samples. The mutation rate constant for the 36-marker haplotype was selected based on the world's data [13–14]: 0.0038 per locus per generation. The average generation interval for males was considered to be 31.5 years [15].

## RESULTS

### Gene geography of haplogroup N3a1-Y23475

The range of haplogroup N3a1-Y23475 extends from northwest to southeast (from the Vologda Russians to Mongols), but shows intermittent and irregular frequency distribution (Table 1). Most of the haplogroup carriers (84%) are in the Finnish-speaking populations of the Ural-Volga region, where the N3a1-Y23475 frequency varies between 0.4% and 25% (Fig. 1). The haplogroup frequency reaches its maximum in the populations of Mordovia, where there more than 70% of its carriers with the following frequency values: 8% in Moksha, 9% in Erzya, and the maximum value (25%) in Tengushevsky Erzya-Shoksha (hereinafter, Shoksha). N3a1-Y23475 is almost an order of magnitude more rare in the neighbouring Turkic-speaking peoples: Mishar Tatars (2%), Kazan Tatars (1%), and extremely rare in Bashkirs (0.4%). The haplogroup is rare in the Russian populations: 3% in the Vologda Region and 1% in the Belgorod, Kaluga, Oryol regions (Table 1). In Siberia, haplogroup N3a1-Y23475 has been found in Altaians (6%) and one Mongol.

The reported haplogroup N3a1-Y23475 gene geography necessitates consideration of two issues: 1) of the factors that led to its accumulation in Mordovia's populations; 2) of its sources in Altaians. The search for answers to these questions was performed using phylogenetic analysis.

### Phylogenetic structure and chronology of clusters/branches

The samples of the haplogroup N3a1-Y23475 carriers were assessed using the 37 Y-STR marker panel in all populations,

**Table 1.** Spread of haplogroup N3a1-Y23475 across the populations of the Ural-Volga region and adjacent regions

Linguistic characteristics	Populations	Total sample	Number of the N3a1 test samples	Number of the N3a1-Y23475 branch carriers	N3a1-Y23475 branch frequency, %
Eastern Europe					
Slavic group of Indo-European language family	Russians of Vologda Region	80	2	2	3
	Russians of Kaluga Region	96	1	1	1
	Russians of Belgorod Region	341	2	2	1
	Russians of Oryol Region	173	1	1	1
Ural-Volga region					
Mari group of Uralic language family	Mountain Mari	65	14	0	0
	Meadow Mari	161	33	0	0
Mordovian group of Uralic language family	Moksha Mordvins**	194	16	16	8
	Erzya Mordvins**	280	26	26	9
	Shoksha Mordvins**	52	13	13	25
Permic group of Uralic language family	Besermyan	45	20	0	0
	Eastern Udmurts	70	50	0	0
	Northern Udmurts	42	13	0	0
	Southern Udmurts	113	83	0	0
	Zyuzdinski Komi (Kirov Region)	32	8	0	0
	Northern Komi-Permyaks	74	7	0	0
	Southeastern Komi-Permyaks	74	23	0	0
	Southwestern Komi-Permyaks	175	19	0	0
Turkic branch of Altaic language family	Kazan Tatars	382	11	5	1
	Mishar Tatars	154	12	3	2
	Chuvash of Tatarstan	66	12	0	0
	Bashkirs of Bashkortostan	719	7	3	0.4
Siberia and Cenral Asia					
Ugric group of Uralic language family	Khanty	83	5	0	0
	Mansi	76	1	0	0
Turkic branch of Altaic language family	Siberian Tatars	544	6	0	0
	Northern Altaians-Kumandins	44	1	1	2
	Northern Altaians-Chelkans	56	1	0	0
	Southern Altaians-Telengits	54	4	4	7
	Tofalars people	51	1	0	0
Mongolic branch of Altaic language family	Buryats of Trans-Baikal Territory, Duldurginsky District	96	2	0	0
	Mongols	636	1	1	0.2
Total:		5028	395	78	

**Note:** \*\* — significant differences in haplogroup frequency between the Mordvin populations ( $p < 0.01$ )

where the haplogroup was found. A total of 74 haplotypes were obtained (Table in Appendix), based on which a phylogenetic network was constructed (Fig. 2).

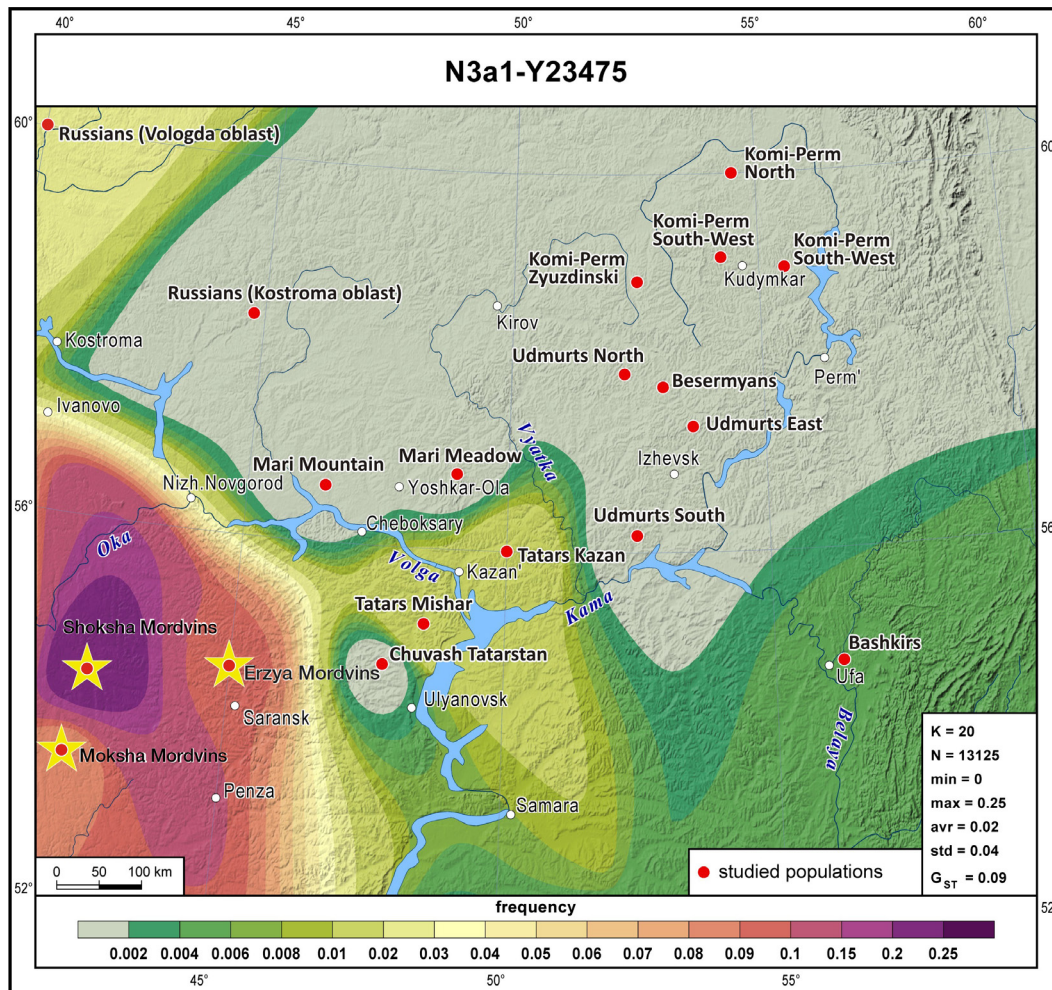
The use of 37 Y-STR markers allowed us to achieve high phylogenetic resolution: most haplotypes were distributed across six clusters (Fig. 2). Four clusters (B, C, D, E) are characterized by absolute specificity: each cluster is formed by the samples belonging to only one Mordovia's ethnic group, i.e. Erzya, Moksha or Shoksha. Clusters A and F include three specific subclusters, two of which (A1 and F1) are formed by the haplotypes from Mordovia's populations. Chronology of their emergence was calculated for all clusters and subclusters (Table 2 — time to the most recent common ancestor, TMRCA).

Subcluster A1 and cluster B (Fig. 2) of Moksha were formed within the same period: about 500 years ago (Table 2). Subcluster A1 includes for different Moksha haplotypes from

the Insarsky District of Mordovia (hereinafter, birthplaces of the assessed individuals' paternal grandfathers are specified). Cluster B is represented by the Moksha samples from three neighboring districts of Mordovia: four samples from the Insarsky District, three samples from the Atyuryevsky District, and one sample from the Kovytkinsky District.

Clusters C and E (Fig. 2) of Erzya were formed about 900–1000 years ago (Table 2). Cluster C includes representatives of the Ichalkovsky District of Mordovia, and cluster E includes mostly those of the adjacent Chamzinsky District.

Cluster D and subcluster F1 of Shoksha are far from each other in the phylogenetic network, despite the fact that these originate from the populations of the Tengushevsky District in northwestern Mordovia. The dates differ almost two times: subcluster F1 was formed about 500 years ago, cluster D about 900 years ago (Table 2).



**Fig. 1.** Spread of haplogroup N3a1-Y23475 across indigenous populations of the Ural-Volga region. The studied populations are marked with red dots; yellow dots represent Mordovia's populations. The haplogroup frequency is shown according to the color chart in the map legend

All haplotypes of the Kazan Tatars and Mishar Tatars are located outside the clusters. A casual resemblance to these is reported for two Moksha samples and two Bashkir samples (Fig. 2), as well as for subcluster D of Shoksha.

The samples of Russians also do not form any separate subcluster; these are included in the most heterogeneous cluster F, along with the haplotypes of Moksha, Bashkir, and Mongol and subclusters of Shoksha and Telengt Altaians.

Haplotypes of Altaians are distributed across clusters A and F in the phylogenetic network. The age of cluster A including only one sample of the Kumandin Altaian is about 1150 years. This value twice exceeds the value of the Moksha subcluster A1, the Altaian haplotype originates from. Samples of southern Altaians-Telengits are merged into the specific subcluster F2, which suggests their descent from a common ancestor. The time to the most recent common ancestor for subcluster F2 is about 800 years (Table 2).

The age of the entire haplogroup N3a1-Y23475 calculated based on the Y-STR haplotypes ( $2340 \pm 330$  years) within the margins of error is consistent with the estimates of the YFull team [7] obtained based on SNP markers ( $2700 \pm 300$  years). Matches in the YFull phylogenetic tree can be found for a half of the Y-STR clusters identified in the population-based study (Table 2).

## DISCUSSION

The detected accumulation of haplogroup N3a1-Y23475 in the gene pools of Mordovia's population and two population

growth momenta reflect their demographic history. The population growth about 500 years ago could result from liberation of the population from the system of dependence on the Mongol Empire and the Golden Horde (that lasted from mid-13<sup>th</sup> century to late 15<sup>th</sup> century). The population growth momentum about 1000 years ago is considered to be associated with the formation of ethnic foci of Moksha (on the Tsna River and in Prisurie) and Erzya (in Poteshie) in the 10<sup>th</sup> century. The late Ryazan-Oka traditions followed by the ceramic complex of the Shokshinsky burial ground developed within the Shoksha range in an isolated manner, inheriting and developing the features of ceramics of the preceding period [16].

Peculiarity of the Shoksha gene pool is traced through high frequency of N3a1-Y23475 (25% vs. 8–9% in Moksha and Erzya) and the analysis of autosomal gene pool. The ADMIXTURE method revealed two ancestral components [17]. The first ancestral component (Moksha-Erzya) merges the populations of Moksha and Erzya only, while the second (Shoksha) is typical only for Shoksha populations. Both ancestral components are found in the genomes of most Russian populations, suggesting the contribution of pre-Slavic population to the Russian gene pool (Table 3).

Accumulation of haplogroup N3a1-Y23475 in the population of Mordovia can be explained by consecutive effects of two factors: migration and the founder effect. The presence of this lineage in all Mordovia's populations suggests that it was inherited by the Mordvin proto-population from the same source (probably, from the alien carriers of haplogroup N3a1).

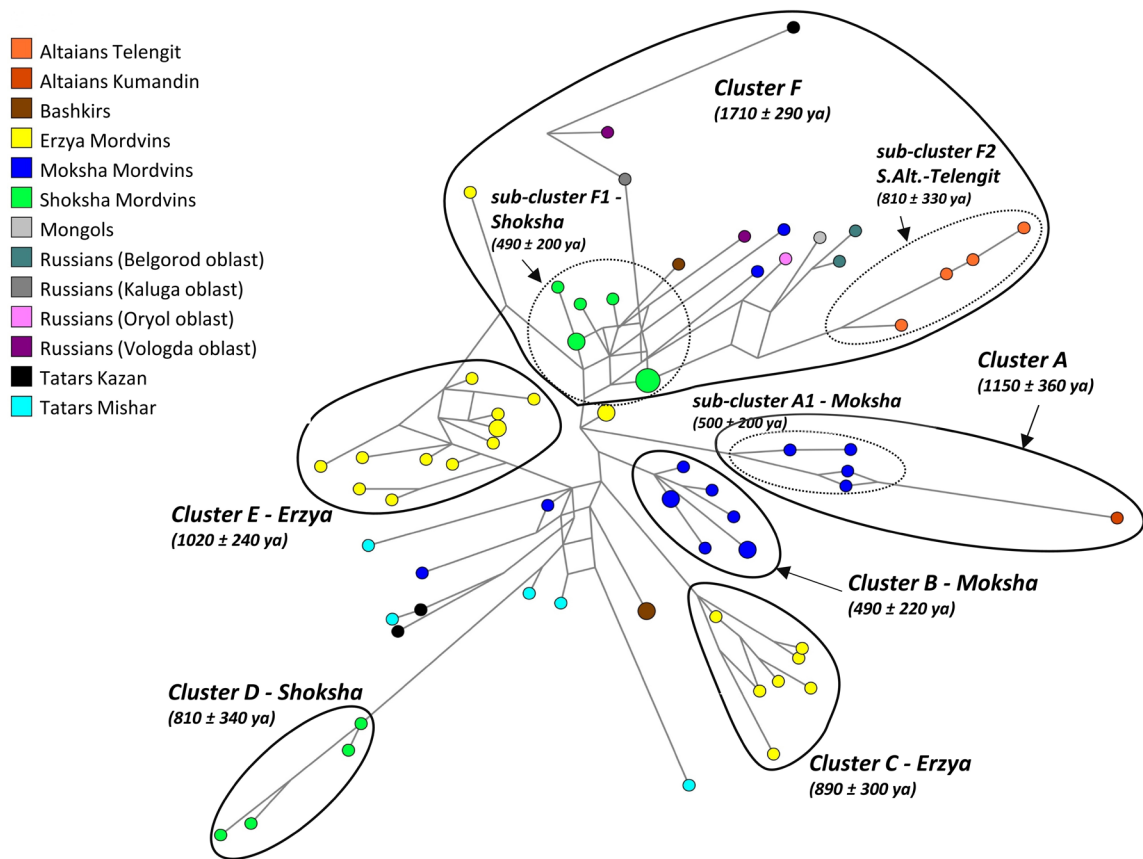


Fig. 2. Phylogenetic network of haplogroup N3a1-Y23475

The founder effect presumably manifested itself after isolation of Mordovia’s ethnic groups. This explains high specificity of clusters and subclusters (accumulation of haplotype diversity within each population) and their structure in the phylogenetic network (Fig. 2).

The N3a1-Y23475 phylogenetic analysis results make it possible to give an interim answer to the question about the sources of origin of haplogroup N3a1-Y23475 in Altaians. Haplotype of the Northern Altaian-Kumandin is most close to the Moksha haplotypes (Fig. 2, cluster A). Since the population of Kumandins is small (2400 people), a single Moksha lineage could emerge in Kumandins due to mass resettlement of Mordvins to Altai in the 19<sup>th</sup>–20<sup>th</sup> centuries [18]: in the early 20<sup>th</sup> century, Mordvins ranked third (following Russian and

Ukrainians) among ethnic groups based on the population size in the Altai Region.

The issue of the population being the source of branch N3a1-Y23475 in Southern Altaians is related to the issue of its ancestral homeland that could hypothetically be located in the Ural region (“Uralic ancestral homeland”) or South Siberia (“Siberian ancestral homeland”).

Southern Altaians-Telengits form their own subcluster F1 in the phylogenetic network, which, along with two Belgorod samples and the Mongol sample can be traced back to the most common Shoksha haplotype (Fig. 2, cluster F). The subcluster F2 chronology suggests that the ancestor of today’s N3a1-Y23475 branch carriers could emerge within the range of Southern Altaians-Telengits about 800 years ago (500–1100 years ago).

Table 2. Genetic chronology (TMRCA) of the haplogroup N3a1-Y23475 clusters

Cluster	<i>n</i>	TMRCA, years ago	YFull estimate	Match between the STR cluster and the YFull tree branch
A (Moksha and the Kumandin Altaian)	5	1150 ± 360		not found
A1 (Moksha)	4	500 ± 200		not found
B (Moksha)	8	490 ± 220	800 ± 210	N-BY9737
C (Erzya)	7	890 ± 300	1400 ± 260	N-Y164241
D (Shoksha)	4	810 ± 340		not found
E (Erzya)	12	1020 ± 240		not found
E	25	1710 ± 290	1250 ± 260	N-Y39435
F1 (Shoksha)	9	490 ± 200	425 ± 110	N-Y39282
F2 (Altaians-Telengits)	4	810 ± 330		not found
N3a1-Y23475	74	2340 ± 330	2700 ± 300	N-Y23472
in general				

Note: *n* — number of samples the cluster consists of, TMRCA — time to the most recent common ancestor.

**Таблица 3.** Share of individual autosomal genomes (%) in Russian populations, where the ADMIXTURE ancestral components typical for Moksha and Erzya (Moksha-Erzya) and Shoksha (Shoksha) are found

Regions with the Russian population	Share of Mordovia's ancestral components in Russian populations	
	Moksha-Erzya (%)	Shoksha (%)
Tver	90	50
Kaluga	80	60
Smolensk	60	50
Oryol	65	25
Tambov	50	40
Kostroma	50	80
Ryazan	50	80
Nizhny Novgorod	45	80

Reasons for the “Uralic ancestral homeland” hypothesis. The greatest diversity of both clusters and single N3a1-Y23475 haplotypes is observed in today's populations of the Ural-Volga region, which indirectly indicates the haplogroup origin can be in this area. In this case the emergence of haplogroup N3a1-Y23475 in Southern Altaians and Mongols is associated with the eastward migration from the Ural region. Furthermore, migration to Altai had to occur no later as 500 years ago, since subcluster F2 is significantly isolated from the main pool of haplotypes from the populations of the Ural-Volga region.

Reasons for the “Siberian ancestral homeland” hypothesis. An alternative hypothesis places the haplogroup N3a1-Y23475 ancestral homeland to South Siberia. This hypothesis is supported by the previously reported general vector of the haplogroup N spread from east to west [1]. In this case, subcluster F2 of Southern Altaians-Telengits preserves genetic memories about the ancient haplogroup N3a1-Y23475 carriers, who migrated to the Ural region from South Siberia. Cluster F differs from other clusters of the phylogenetic network (Fig. 2) by the highest population heterogeneity and the most ancient date ( $1710 \pm 290$  years ago; Table 2). There is a large number of reticulations in the structure of cluster F, along with rather distant relationships between haplotypes from different regions and both subclusters. The chronology calculated suggests that cluster F was formed on average only 600 years later than the entire haplogroup N3a1-Y23475. Both observations suggest that in the past there were many clusters within cluster F, from which only distinct haplotypes survived by today.

The hypothesis about the South Siberian ancestral homeland of haplogroup N3a1-Y23475 is not contradicted by the pattern of its spread across today's populations, since high frequency and haplotype diversity in Mordovia's populations formed as late as in the last millennium. Chronology of cluster F suggests that the South Siberian, some Volga region, and

all Russian populations had the same root about 1700 years ago. During this period N3a1-Y23475 could emerge in South Siberia in the population of haplogroup N3a1-B211 carriers, who migrated to the west. The subcluster F2 founder can be a descendant of this population, which left sporadic marks in the gene pool of today's Southern Altaians-Telengits.

The available data do not yet provide unambiguous confirmation of validity of one or another hypothesis. That is why it is planned to analyze the entire complex of the Y chromosome haplogroups and haplotypes in order to trace genetic links between the gene pools of the Ural region and South Siberia.

## CONCLUSIONS

Haplogroup N3a1-Y23475 was formed 2.3–2.7 thousand years ago, but intense accumulation of the haplogroup took place mostly in Mordovia's populations since the 10th century. High share of the haplogroup in all Mordovia's populations suggests that it originates from the common source. Two population growth momenta in the ancestors of Mordvins can be traced: about 1000 years ago in Erzya, about 500 years ago in Moksha; both momenta have been reported for the Shoksha population. Low N3a1-Y23475 frequency in the populations of Tukkic-speaking and Slavic-speaking peoples is associated with the presence of the substrate layer of Finnish-speaking peoples in their gene pools. The presence of haplogroup N3a1-Y23475 in Northern Altaians-Kumandins is likely to be associated with migration of Mordovia's population to Altai in the 19<sup>th</sup>–20<sup>th</sup> centuries. The source of N3a1-Y23475 in Southern Altaians-Telengits requires verification of two hypotheses: of the “Uralic ancestral homeland” with negligible migration to the east and of the “Siberian ancestral homeland” with migration from South Siberia to the Ural-Vplga region about 1700 years ago.

## References

1. 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.
2. Lamnidis TC, Majander K, Jeong C, Salmela E, Wessman A, Moiseyev V, et al. Ancient Fennoscandian genomes reveal origin and spread of Siberian ancestry in Europe. *Nature Communications*. 2018; 9 (1): 5018.
3. All-Russian Population Census 2020. Volume 5 National composition and language proficiency. Table 1. National composition of the population (cited 10 November 2024). Available from: [https://rosstat.gov.ru/vpn/2020/Tom5\\_Nacionalnyj\\_sostav\\_i\\_vladenie\\_yazykami](https://rosstat.gov.ru/vpn/2020/Tom5_Nacionalnyj_sostav_i_vladenie_yazykami).
4. 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 Biol*. 2018; 19 (1): 139.
5. Zeng TC, Vyazov LA, Kim A, Flegontov P, Sirak K, Maier R, et al. Postglacial genomes from foragers across Northern Eurasia reveal prehistoric mobility associated with the spread of the Uralic and Yeniseian languages. *bioRxiv*. 2023.
6. Childebayeva A, Fricke F, Rohrlach AB, Huang L, Schiffels S, Vesakoski O, et al. Bronze age Northern Eurasian genetics in the context of development of metallurgy and Siberian ancestry. *Communications Biology*. 2024; 7 (1): 723.

7. Tree YTree v12.03.00 (cited 20 November 2024). Available from: <https://www.yfull.com/sc/tree/>.
8. Balanovska EV, Agdzhoyan AT, Chukhryaeva MI, Skhalyakho RA, Zaporozhchenko VV, Romanov AG, et al. Population biobanks: organizational models and prospects of application in gene geography and personalized medicine. *Russian Journal of Genetics*. 2016; 52 (12): 1371–87.
9. Agdzhoyan A, Iskandarov N, Ponomarev G, Pylev V, Koshel S, Salaev V, et al. Origins of East Caucasus Gene Pool: Contributions of Autochthonous Bronze Age Populations and Migrations from West Asia Estimated from Y-Chromosome Data *Genes*. 2023; 14 (9): 1780.
10. Koshel SM. Geoinformacionnye tehnologii v genogeografii. V knige: *Sovremennaja geograficheskaja kartografiya*. M.: Data+, 2012; s. 158–166. Russian.
11. Bandelt HJ, Forster P, Röhl A. Median-joining networks for inferring intraspecific phylogenies. *Molecular Biology and Evolution*. 1999; 16 (1): 37–48.
12. Zhivotovsky LA, Underhill PA, Cinnioglu C, Kayser M, Morar B, Kivisild T, et al. The effective mutation rate at Y chromosome short tandem repeats, with application to human population-divergence time. *American Journal of Human Genetics*. 2004; 1: 50–61.
13. Willuweit S, Roewer L. The new Y Chromosome Haplotype Reference Database. *Forensic Science International: Genetics*. 2015; 15: 43–48.
14. Ballantyne KN, Goedbloed M, Fang R, Schaap O, Lao O, Wollstein A, et al. Mutability of Y-chromosomal microsatellites: rates, characteristics, molecular bases, and forensic implications. *American Journal of Human Genetics*. 2010; 87 (3): 341–53.
15. Fenner J. Cross-cultural estimation of the human generation interval for use in genetics-based population divergence studies. *American Journal of Physical Anthropology*. 2005; 128 (2): 415–23.
16. Grishakov V. V. *Keramika pogrebal'nykh pamyatnikov Oksko-Surskogo Mezhdurech'ya v I tys. n.e.* [dissertation]. M., 1993. Russian.
17. Balanovska EV, Gorin IO, Ponomarev GY, Pylev VY, Petrushenko VS, Markina NV, et al. Footprints of interaction among Finnic-speaking, Slavic, and Turkic-speaking populations in modern gene pool and their reflection in pharmacogenetics. *Bulletin of RSMU*. 2022; 2: 18–27.
18. Ovcharova MA. *Mordva Altaya: istoriya i etnokul'turnye protsessy (XIX - nachalo XXI v.)* [dissertation]. Barnaul, 2009. Russian.

## Литература

1. 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.
2. Lamnidis TC, Majander K, Jeong C, Salmela E, Wessman A, Moiseyev V, et al. Ancient Fennoscandian genomes reveal origin and spread of Siberian ancestry in Europe. *Nature Communications*. 2018; 9 (1): 5018.
3. Итоги Всероссийской переписи населения 2020 года. Том 5 Национальный состав и владение языками. Таблица 1. Национальный состав населения (дата обращения: 10 November 2024). Доступно по ссылке: [https://rosstat.gov.ru/vpn/2020/Tom5\\_Nacionalnyj\\_sostav\\_i\\_vladenie\\_yazykami](https://rosstat.gov.ru/vpn/2020/Tom5_Nacionalnyj_sostav_i_vladenie_yazykami).
4. 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 Biol*. 2018; 19 (1): 139.
5. Zeng TC, Vyazov LA, Kim A, Flegontov P, Sirak K, Maier R, et al. Postglacial genomes from foragers across Northern Eurasia reveal prehistoric mobility associated with the spread of the Uralic and Yeniseian languages. *bioRxiv*. 2023.
6. Childebayeva A, Fricke F, Rohrlach AB, Huang L, Schiffels S, Vesakoski O, et al. Bronze age Northern Eurasian genetics in the context of development of metallurgy and Siberian ancestry. *Communications Biology*. 2024; 7 (1): 723.
7. Tree YTree v12.03.00 (дата обращения: 20 November 2024). Доступно по ссылке: <https://www.yfull.com/sc/tree/>.
8. Балановская Е. В., Жабалин М. К., Агджоян А. Т., Чухряева М. И., Маркина Н. В., Балаганская О. А. и др. Популяционные биобанки: принципы организации и перспективы применения в геногеографии и персонализированной медицине. *Генетика*. 2016; 52 (12): 1371–87.
9. Agdzhoyan A, Iskandarov N, Ponomarev G, Pylev V, Koshel S, Salaev V, et al. Origins of East Caucasus Gene Pool: Contributions of Autochthonous Bronze Age Populations and Migrations from West Asia Estimated from Y-Chromosome Data *Genes*. 2023; 14 (9): 1780.
10. 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.
11. Bandelt HJ, Forster P, Röhl A. Median-joining networks for inferring intraspecific phylogenies. *Molecular Biology and Evolution*. 1999; 16 (1): 37–48.
12. Zhivotovsky LA, Underhill PA, Cinnioglu C, Kayser M, Morar B, Kivisild T, et al. The effective mutation rate at Y chromosome short tandem repeats, with application to human population-divergence time. *American Journal of Human Genetics*. 2004; 1: 50–61.
13. Willuweit S, Roewer L. The new Y Chromosome Haplotype Reference Database. *Forensic Science International: Genetics*. 2015; 15: 43–48.
14. Ballantyne KN, Goedbloed M, Fang R, Schaap O, Lao O, Wollstein A, et al. Mutability of Y-chromosomal microsatellites: rates, characteristics, molecular bases, and forensic implications. *American Journal of Human Genetics*. 2010; 87 (3): 341–53.
15. Fenner J. Cross-cultural estimation of the human generation interval for use in genetics-based population divergence studies. *American Journal of Physical Anthropology*. 2005; 128 (2): 415–23.
16. Гришаков В. В. *Керамика погребальных памятников Окско-Сурского Междуречья в I тыс. н.э.* [диссертация]. М., 1993.
17. Балановская Е. В., Горин И. О., Пономарев Г. Ю., Пылёв В. Ю., Петрушенко В. С., Маркина Н. В. и др. Следы взаимодействия финноязычного, славянского и тюркоязычного населения в современном генофонде и их отражение в фармакогенетике. *Вестник РГМУ*. 2022; 2: 20–9.
18. Овчарова М. А. *Мордва Алтая: история и этнокультурные процессы (XIX — начало XXI в.)* [диссертация]. Барнаул, 2009.