

## GENETIC PORTRAITS OF VOLGA–OKA REGION IN THE CONTEXT OF THE CENTRAL RUSSIA'S GENE POOL (Y-SNP POLYMORPHISM)

Ponomarev GYu<sup>1</sup>✉, Shlykov AG<sup>1</sup>, Gorin IO<sup>1</sup>, Voronina MM<sup>1</sup>, Potanina AYu<sup>1</sup>, Petrov VA<sup>1</sup>, Koshel SM<sup>2</sup>, Adamov DS<sup>1</sup>, Balanovska EV<sup>1</sup>

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

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

An urgent problem of the role of Slavic expansion in shaping gene pools of the population of Europe is being analyzed actively using various methods to study ancient and current populations. However, it is difficult to solve due to the lack of consolidated data on the Y-haplogroups in Slavic-, Finnish-, and Turkic-speaking populations of European Russia. The study aimed to look into genetic portraits of Mordovian populations and Russians of the Ryazan Region relative to the surrounding populations of indigenous peoples of Central Russia. For the first time the Y-gene pool of the Volga–Oka region (10 populations,  $n = 1136$  individuals) was assessed in a broad context of our own data on the gene pools of European Russia (based on a single panel of 35 Y-haplogroups, 80 populations,  $n = 9712$ ). The analysis was performed by multidimensional scaling (MDS) and computed cartography (GeneGeo). The produced series of 35 maps of the Y-gene pool of Central Russia and maps of genetic distances from peoples of the Volga–Oka region describe the gene-geographic landscape of the region in detail. It has been shown that all the assessed Russian populations belong to the common "Slavic" cluster that also includes representatives of Western Slavs. The cluster of populations of the Ural–Volga region including three Finnish-speaking populations of Mordovia (Erzya, Moksha, Shoksha) and 7 Turkic-speaking populations of the Chuvash and Mishar turned out to be the most genetically close to the Russian populations of Central Russia. It has been hypothesized that this group of populations can be traced back to the gene pool of the ancient indigenous Finnish-speaking population of the Volga–Oka region.

**Keywords:** gene pool, gene geography, Y-chromosome, haplogroups, Slavic expansion, Central Russia, Volga–Oka region

**Funding:** the study was supported by the RSF grant No. 25-28-01594.

**Acknowledgements:** the authors would like to thank all participants of the expedition survey, who provided their biological samples for the study, to Administration and employees of the Ministry of Healthcare of the Ryazan Region and Republic of Mordovia for institutional support and assistance in conducting expeditions, as well as to Biobank of North Eurasia for access to DNA collections.

**Author contribution:** Ponomarev GYu — genotyping and Y-SNP marker analysis, study design; Shlykov AG — manager of the expedition survey of the gene pool of the Ryazan Region and Mordovia; Ponomarev GYu, Voronina MM, Petrov VA — expedition members, questionnaire survey data analysis; Adamov DS, Potanina AYu, Gorin IO — statistical analysis; Koshel SM — cartographic analysis; Adamov DS, Balanovska EV — 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). The data were acquired after obtaining the written informed consent from the assessed individuals and anonymised.

✉ **Correspondence should be addressed:** Georgy Yu. Ponomarev  
Moskvorechye, 1, 115522, Moscow, Russia; st261900@gmail.com

**Received:** 16.10.2025 **Accepted:** 22.11.2025 **Published online:** 01.12.2025

**DOI:** 10.24075/brsmu.2025.062

**Copyright:** © 2025 by the authors. **Licensee:** Pirogov University. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

## ГЕНЕТИЧЕСКИЕ ПОРТРЕТЫ ВОЛГО-ОКСКОГО РЕГИОНА В КОНТЕКСТЕ ГЕНОФОНДА ЦЕНТРАЛЬНОЙ РОССИИ (Y-SNP ПОЛИМОРФИЗМ)

Г. Ю. Пономарев<sup>1</sup>✉, А. Г. Шлыков<sup>1</sup>, И. О. Горин<sup>1</sup>, М. М. Воронина<sup>1</sup>, А. Ю. Потанина<sup>1</sup>, В. А. Петров<sup>1</sup>, С. М. Кошель<sup>2</sup>, Д. С. Адамов<sup>1</sup>, Е. В. Балановская<sup>1</sup>

<sup>1</sup> Медико-генетический научный центр, Москва, Россия

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

Актуальную проблему роли славянской экспансии в формировании генофондов народонаселения Европы активно анализируют с применением разных методов изучения древних и современных популяций. Но ее решение затрудняет отсутствие сводных данных о гаплогруппах Y-хромосомы в славяно-, финно- и тюркоязычных популяциях европейской части России. Цель работы — рассмотреть генетические портреты популяций Мордовии и русских Рязанской области в сравнении с окружающими популяциями коренного народонаселения Центральной России. Y-генофонд Волго-Окского региона (10 популяций,  $n = 1136$  индивидов) изучен в широком контексте собственных данных о генофондах европейской части России (по единой панели 35 Y-гаплогрупп, 80 популяций,  $n = 9712$ ). Анализ проведен методами многомерного шкалирования (MDS) и компьютерной картографии (GeneGeo). Созданная серия 35 карт Y-генофонда Центральной России и карт генетических расстояний от народов Волго-Окского региона детально описывает генеогеографический ландшафт региона. Показано, что все изученные русские популяции входят в единый «славянский» кластер, включающий также представителей западных славян. К русским популяциям Центральной России наиболее генетически близким оказался кластер популяций Волго-Уральского региона, включающий три финноязычные популяции Мордовии (эрзя, мокша, шокша) и семь тюркоязычных популяций чувашей и мишарей. Выведена гипотеза, что этот массив популяций восходит к генофонду древнего автохтонного финноязычного населения Волго-Окского региона.

**Ключевые слова:** генофонд, генеогеография, Y-хромосома, гаплогруппы, славянская экспансия, Центральная Россия, Волго-Окский регион

**Финансирование:** исследование выполнено при поддержке гранта РНФ № 25-28-01594.

**Благодарности:** авторы благодарны всем участникам экспедиционного обследования, предоставившим свои биологические образцы для исследования, администрации и сотрудникам Министерства здравоохранения Рязанской области и Республики Мордовия за организационную поддержку и содействие в проведении экспедиций, а также Биобанку Северной Евразии за предоставление коллекций ДНК.

**Вклад авторов:** Г. Ю. Пономарев — генотипирование и анализ Y-SNP маркеров, дизайн исследования; А. Г. Шлыков — организатор экспедиционного обследования генофонда Рязанской области и Мордовии; Г. Ю. Пономарев, М. М. Воронина, В. А. Петров — участники экспедиций, анализ анкетных данных; Д. С. Адамов, А. Ю. Потанина, И. О. Горин — статистический анализ; С. М. Кошель — картографический анализ; Д. С. Адамов, Е. В. Балановская — дизайн исследования и подготовка рукописи.

**Соблюдение этических стандартов:** исследование одобрено этическим комитетом Медико-генетического научного центра имени Н. П. Бочкова (протокол № 1 от 29 июня 2020 г.). Данные получены при добровольном письменном информированном согласии участников обследования и деперсонифицированы.

✉ **Для корреспонденции:** Георгий Юрьевич Пономарев  
ул. Москворечье, д. 1, 115522, г. Москва, Россия; st261900@gmail.com

**Статья получена:** 16.10.2025 **Статья принята к печати:** 22.11.2025 **Опубликована онлайн:** 01.12.2025

**DOI:** 10.24075/vrgmu.2025.062

**Авторские права:** © 2025 принадлежат авторам. **Лицензиат:** РНИМУ им. Н. И. Пирогова. Статья размещена в открытом доступе и распространяется на условиях лицензии Creative Commons Attribution (CC BY) (<https://creativecommons.org/licenses/by/4.0/>).

In the recent decade, the research focused on the origin of modern Slavic-speaking peoples has risen to a new level due to the studies of modern and ancient DNA. The genetic data on the autosomal DNA, Y-chromosome, mitochondrial DNA provide an opportunity to see various aspects of communication between pre-Slavic and Slavic populations during massive expansion of Slavic tribes in the large territories of Europe in the 6<sup>th</sup>–12<sup>th</sup> century A.D. One of the main issues of this communication is the relationship between two processes: population change (complete substitution of the indigenous people's gene pool with the newcomer Slavic tribes) and assimilation of the pre-Slavic population ("slavicisation" — interpenetration of gene pools of Slavic and pre-Slavic populations with the adoption of Slavic culture). With the historical, archaeological, linguistic, and genetic data accumulated to date, a clear answer to the question cannot be given.

Synthesis of data on the autosomal genome, Y-chromosome, and mtDNA of the modern Slavic population has shown that most of the Baltic-Slavic genetic variation is related mainly to assimilation of pre-Slavic gene pools that were different in Western, Eastern and Southern Slavs [1].

In 2025, two important reports were published, the authors of which analyzed communication between the pre-Slavic population and Slavic tribes based on various systems of genetic data of modern populations and ancient DNA. Autosomal genomes of ancient samples ( $n = 555$ ) from burials in the eastern Germany, northwestern Balkans, Poland (and northwestern Ukraine) dating to the period before (6–7<sup>th</sup> century) and after the Slavic expansion were assessed [2]. In addition to these three major territorial clusters, the data for the Volga–Oka region are also provided. To calculate the extent, to which the indigenous population was substituted, samples from the earliest Slavic inhumations that replaced the cremation ceremony (Powiat Hrubieszowski, Poland, 600–900 A.D.) were used. DNA of these samples was used to calculate ancestral genetic components (qpAdm) as a reference. Based on the calculation results, the authors came to a conclusion about the dominance of the process of intense population substitution in the territories of Slavic expansion: the Slavs substituted  $82 \pm 1\%$  of the local gene pools in the northwestern Balkans,  $83 \pm 6\%$  in the eastern Germany,  $93 \pm 3\%$  in Poland and Ukraine. The values reported for the Volga–Oka region are much lower:  $65 \pm 4\%$  of the indigenous population was substituted by the newcomer Slavic migrants [2].

The second study performed by our team based on the Y-chromosome data was limited to the Slavic expansion northeastern periphery: the Volga–Oka region [3]. Haplogroup R1a predominates in modern Y gene pools of the population of the Volga–Oka interfluvium ( $n = 935$ ): 56% in Russians of the Ryazan Region and 44% in the indigenous population of Mordovia. That is why the authors focused mainly on the comprehensive analysis of the detailed data on the Y-SNP and Y-STR polymorphisms of haplogroups R1a-CTS1211 and R1a-Z92, along with five ancient samples dating to the 6<sup>th</sup>–12<sup>th</sup> century A.D. Calculation of the time to the most recent common ancestor (TMRCA) by two independent methods revealed 10 informative clusters in phylogenetic networks of the 37-marker Y-STR haplotypes aged 1600–2900 years, i.e. before the beginning of Slavic expansion. About a half of Russians of the Ryazan Region (carriers of the studied R1a branches) can be male descendants of the pre-Slavic population of the region, presumably the ancient indigenous Finnish-speaking tribes. In contrast to the findings of the previous paper [2], we concluded that the Russian population of the Volga–Oka interfluvium was shaped largely through cultural assimilation,

not complete substitution of the pre-Slavic population by the Slavs. This important conclusion is supported by the earlier studies conducted by our team that were focused on the Y-chromosome [1, 4–9] and the autosomal gene pool [1, 10].

The conclusions about substitution of about 65% of the gene pool of the pre-Slavic population of the Volga–Oka region [2] are based on the data on the ancient DNA of the Suzdal Opole ( $n = 31$ ) acquired by another team, which recorded the gene pool changes associated with the spread of the Slavs and was more careful when drawing conclusions about the scope of substitution [11]. These authors report that the medieval Suzdal principality was populated by various ethnic groups that eventually yielded a hybrid, but totally Russian-speaking population inhabiting this region by today.

It is clear that further research is necessary to clarify an important question: what is the relationship between the process of intense substitution of the indigenous population with the Slavs and the process of interpenetration of their gene pools. The question is difficult to answer due to the lack of consolidated data on Y-haplogroups in various Slavic-, Finnish-, and Turkic-speaking populations of European Russia. Such a gap in our knowledge prevents further solution to the problem of the nature of communication between Slavic and pre-Slavic populations in Russia. The study aimed to perform comparative analysis of the major Y-haplogroups in the context of the gene pool of central part of European Russia, which would help fill the gap.

## METHODS

The data on the present-day indigenous population of the Volga–Oka region were collected during expedition surveys in 2005–2023 ( $n = 1136$  in total): in the Ryazan Region ( $n = 497$ ; Kadomsky, Kasimovsky, Mikhaylovsky, Sapozhkovsky, Sarayevsky, Spassky, Shilovsky districts) and the Republic of Mordovia ( $n = 639$ ; Ardatovsky, Insarsky, Ichalkovsky, Krasnoslobodsky, Lukyanovsky, Ruzayevsky, Tengushevsky, Torbeyevsky, Chamzinsky districts). Inclusion criteria: samples exclusively from unrelated males, whose ethnicity was confirmed at least three generations deep.

DNA was isolated by nucleic acid purification on magnetic particles in the QIAAsymphony system (Qiagen; Germany). The Y-SNP marker genotyping was performed by the TaqMan Open Array method in the QuantStudio 12K Flex thermocycler (Thermo Fisher Scientific; USA) using custom plates.

The data on the gene pool of 80 indigenous populations of Central Russia ( $n = 9712$ ) provided by the Biobank of North Eurasia [12] were assessed based on the single panel of 35 haplogroups: C-M217, D-M174, E-M96, G1-M285, G2-P15, G-M201(xM285,P15), H-M69, I-M170(xM253,M223,P37), I-M223, I-M253, I-P37, J1-M267, J2-M172, L-M20, N2-P43, N3a1-B211, N3a2-M2118, N3a3-CTS10760, N3a4-Z1936, N3a5a-F4205, N-M231(xP43,M178), O-M175, Q-M242, R1a-M198(xM458,CTS1211,Z92,Z93), R1a-CTS1211, R1a-M458, R1a-Z92, R1a-Z93, R1b-M343(xM269,M73), R1b-M269 (xL51,Z2105), R1b-L51, R1b-M73, R1b-Z2105, R2-M124, T-M70. These were complemented with the published data on 10 populations of Latvia, Lithuania, Poland, Finland, Sweden, and Estonia [13–15]. Frequencies of 35 Y-haplogroups were used to calculate a pairwise matrix of Nei's genetic distances between 90 populations (Djgenetics [4]); the multidimensional scaling (MDS) plots were created in Statistica version 7.1 (TIBCO Software, USA).

Maps of the Y-haplogroup spread and maps of genetic distances were created in the GeneGeo 2.8 software package

**Table.** Frequencies of Y-chromosome haplogroups (%) in the indigenous population of the Ryazan Region and Mordovia

| Population              | Consolidated data on the populations |                               |          | Russians from 7 districts of the Ryazan Region |            |           |                             | Populations of Mordovia |        |          |
|-------------------------|--------------------------------------|-------------------------------|----------|--|------------|-----------|-----------------------------|-------------------------|--------|----------|
| Y-haplogroups           | Volga–Oka region                     | Russians of the Ryazan Region | Mordovia | Kadomsky                                       | Kasimovsky | Shilovsky | Total for 4 other districts | Erzya                   | Moksha | Shoksha* |
| Number of samples       | 1136                                 | 497                           | 639      | 79   | 110        | 91        | 217                         | 322                     | 252    | 65       |
| R1a (total)             | 50.7                                 | 57.3                          | 45.5     | 58.2   | 63.6       | 58.2      | 53.5                        | 55.3                    | 33.3   | 44.6     |
| R1a-CTS1211             | 31.9                                 | 28.4                          | 34.6     | 20.3   | 44.5       | 35.2      | 20.4                        | 41.6                    | 24.2   | 40.0     |
| R1a-Z92                 | 9.8                                  | 17.3                          | 3.9      | 17.7   | 6.4        | 13.2      | 24.5                        | 4.0                     | 4.4    | 1.5      |
| R1a-PF6202 (eq.M458)    | 4.0                                  | 7.2                           | 1.4      | 5.1  | 9.1        | 8.8       | 6.5                         | 2.5                     | 0.4    | –        |
| R1a-Z93                 | 2.9                                  | 1.9                           | 3.8      | 3.8  | 1.8        | –         | 2.0                         | 6.5                     | 1.2    | –        |
| R1a-M198*               | 2.1                                  | 2.5                           | 1.8      | 11.3   | 1.8        | 1.0       | 0.1                         | 0.7                     | 3.1    | 3.1      |
| N3 (total)              | 14.1                                 | 11.3                          | 16.3     | 11.4   | 16.4       | 5.5       | 11.1                        | 13.0                    | 13.5   | 43.1     |
| N3a1-B211               | 5.7                                  | 0.2                           | 10.0     | –  | –          | 1.1       | –                           | 8.7                     | 8.3    | 23.1     |
| N3a3-CTS10760 (eq.VL29) | 4.7                                  | 6.6                           | 3.1      | 6.3  | 8.2        | 4.4       | 6.9                         | 0.6                     | 4.8    | 9.2      |
| N3a4-Z1936              | 3.7                                  | 4.4                           | 3.1      | 5.1  | 8.2        | –         | 4.1                         | 3.7                     | 0.4    | 10.8     |
| Haplogroups (xR1a, N3)  | 35.2                                 | 31.4                          | 38.2     | 30.4   | 20.0       | 36.3      | 35.5                        | 31.7                    | 53.2   | 12.3     |
| E-M78                   | 6.7                                  | 4.6                           | 8.3      | 1.3  | 0.9        | 5.5       | 7.4                         | 2.8                     | 16.7   | 3.1      |
| J2-M172                 | 6.1                                  | 2.0                           | 9.2      | –  | 2.7        | 1.1       | 2.8                         | 5.6                     | 15.9   | 1.5      |
| I1-M253                 | 4.4                                  | 4.0                           | 4.7      | 7.6  | 0.9        | 3.3       | 4.6                         | 8.7                     | 0.8    | –        |
| I2-P37.2                | 4.0                                  | 8.5                           | 0.6      | 8.9  | 8.2        | 11.0      | 7.4                         | 0.9                     | 0.4    | –        |
| G2-P15                  | 3.8                                  | 1.4                           | 5.6      | –  | 0.9        | 4.4       | 0.9                         | 4.0                     | 9.1    | –        |
| R1b-L51                 | 3.7                                  | 1.8                           | 5.2      | 2.5  | 2.7        | 1.1       | 1.4                         | 3.4                     | 7.1    | 6.2      |
| R1b-Z2105               | 2.2                                  | 1.6                           | 2.7      | 5.1  | 1.8        | 2.2       | –                           | 5.0                     | 0.4    | –        |
| Other                   | 4.3                                  | 7.4                           | 1.9      | 5.1  | 1.8        | 7.7       | 11.1                        | 1.2                     | 2.8    | 1.5      |

**Note:** The red background marks the values corresponding to the 5% polymorphism criterion. Haplogroups are presented in the descending order of their frequencies in the gene pool of the Volga–Oka region. \* Shoksha — is not a generally accepted term for the population of Erzya of the Tengushevsky district in northwestern Mordovia

[16] by the weighted average interpolation method with an influence radius of 500 km and the weight function decrement 3.

## RESULTS

### Volga–Oka region gene pool diversity

Among 28 Y-haplogroups found in the gene pool of the studied populations of the Volga–Oka region, the macro-haplogroup R1a predominates, accounting for on average 51% (Table). However, populations of the region show considerable differences in its frequency: minimum 33% in Moksha, maximum 64% in Russians of the Kasimovsky district, Ryazan Region. The majority of other haplogroups vary much more (Fig. 1). Thus, the haplogroup N3 frequency varies between 5.5% in Russians of the Shilovsky district and 43% in Shoksha. The “Other” category (Table) includes haplogroups with the frequencies 5% in all studied populations (C-M217, G1-M285, I2-M223, J1-M267, L-M20, N2-P43, N-M231\*, Q-M242, R1a-M198\*, R1b-M269\*, R1b-M73, R2a-M124, T1a-M70).

In the gene pool of Russians of the Ryazan Region, haplogroups are distributed by descending order of frequency as follows: R1a-CTS1211 (28%), R1a-Z92 (17%), I2a-P37.2 (9%), R1a-M458 (7%). This set of conditionally “Slavic” haplogroups is typical for the Russian populations of central and southern Russia in general [4]. Among the conditionally “Uralic” haplogroups the most common are N3a3 (7%) and N3a4 (4%).

The Erzya gene pool is close to the Russian populations of the Ryazan Region: R1a shows high frequency (55%), and the branch R1a-CTS1211 (44%) is much more common in

Erzya, than R1a-Z92 (4%). In the Erzya gene pool, haplogroups R1a-Z93 (6.5%), N3a1 (9%), and I1-M253 (9%) are more frequent, than in Russian populations.

In Shoksha (Tengushevsky Erzya), the macrohaplogroup N3 (43%) predominates, which makes Shoksha dramatically different from Erzya (13%). All three N3 branches in the Shoksha gene pool reach their maximum in the Volga–Oka region: N3a1 (23%), N3a4 (11%), N3a3 (9%). Based on the frequency of R1a-CTS1211 (40%) Shoksha are close to Erzya (44%).

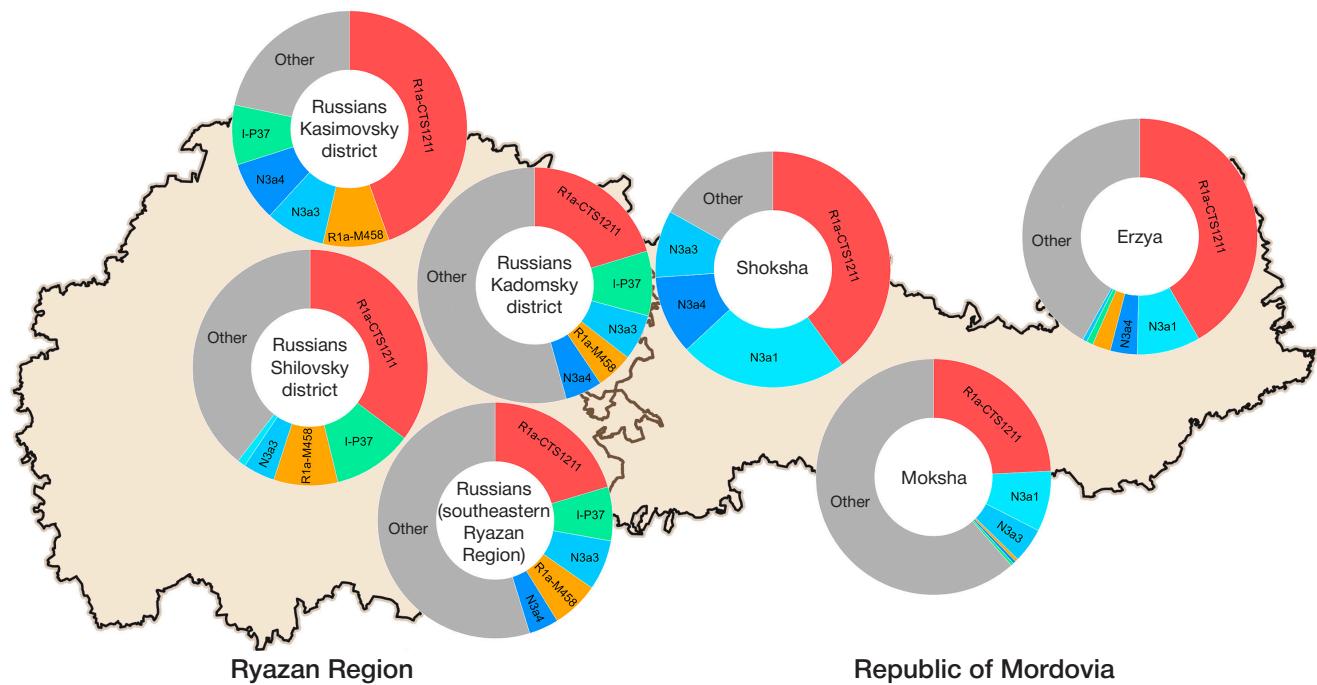
The Moksha gene pool uniqueness is represented not only by the reduced frequency of R1a (33%), but also by the increased frequencies of haplogroups E-M78 (17%), J2-M172 (16%), G2-P15 (9%), R1b-L51 (7%).

Frequencies of Y-haplogroups revealed significant genetic differences between populations of Mordovia, which suggests the need to consider these as separate populations, not allowing us to combine these into a single Mordva group.

### Volga–Oka region gene pool position among populations of Central Russia

The fact that the detailed information about gene pools of the indigenous population of European Russia is available for the team allows us to consider the major patterns of the gene pool variation and assess the position of the Volga–Oka region gene pool in it.

Frequencies of 35 Y-haplogroups were used to calculate a pairwise matrix of Nei's genetic distances between 90 populations of Russia and the Baltics, as well as to create a multidimensional scaling (MDS) plot. Populations of Central Russia formed three major clusters in its genetic space (Fig. 2): Slavic, Western Finnish, Ural–Volga.

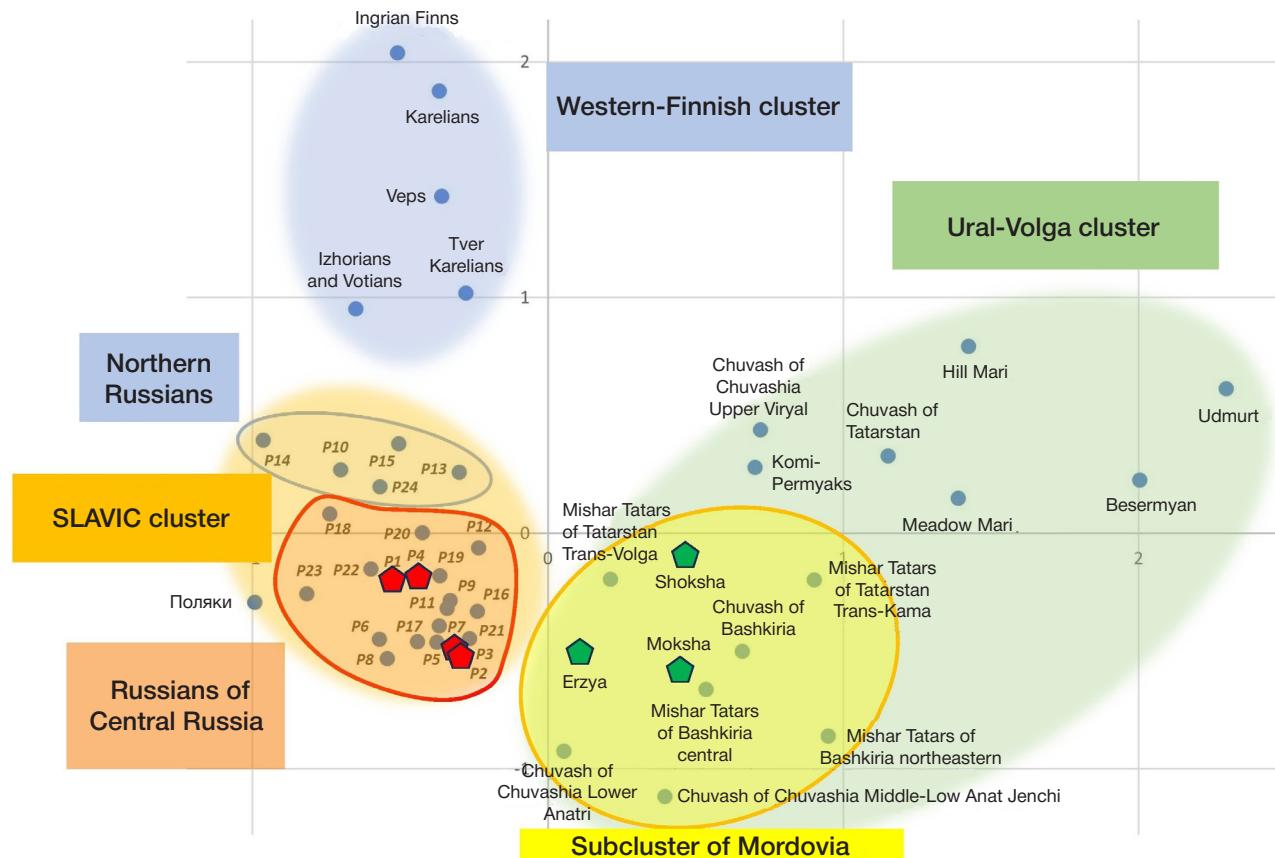


**Fig. 1.** Genetic portraits of populations of the Volga-Oka region (within the borders of the Ryazan Region and Mordovia). The data on the conditionally “Slavic” haplogroups R1a-CTS1211, R1a-M458, I2a-P37.2 and the conditionally “Uralic” haplogroups N3a1, N3a3, N3a4 are provided for all populations; the set of other haplogroups is referred to as “Other”

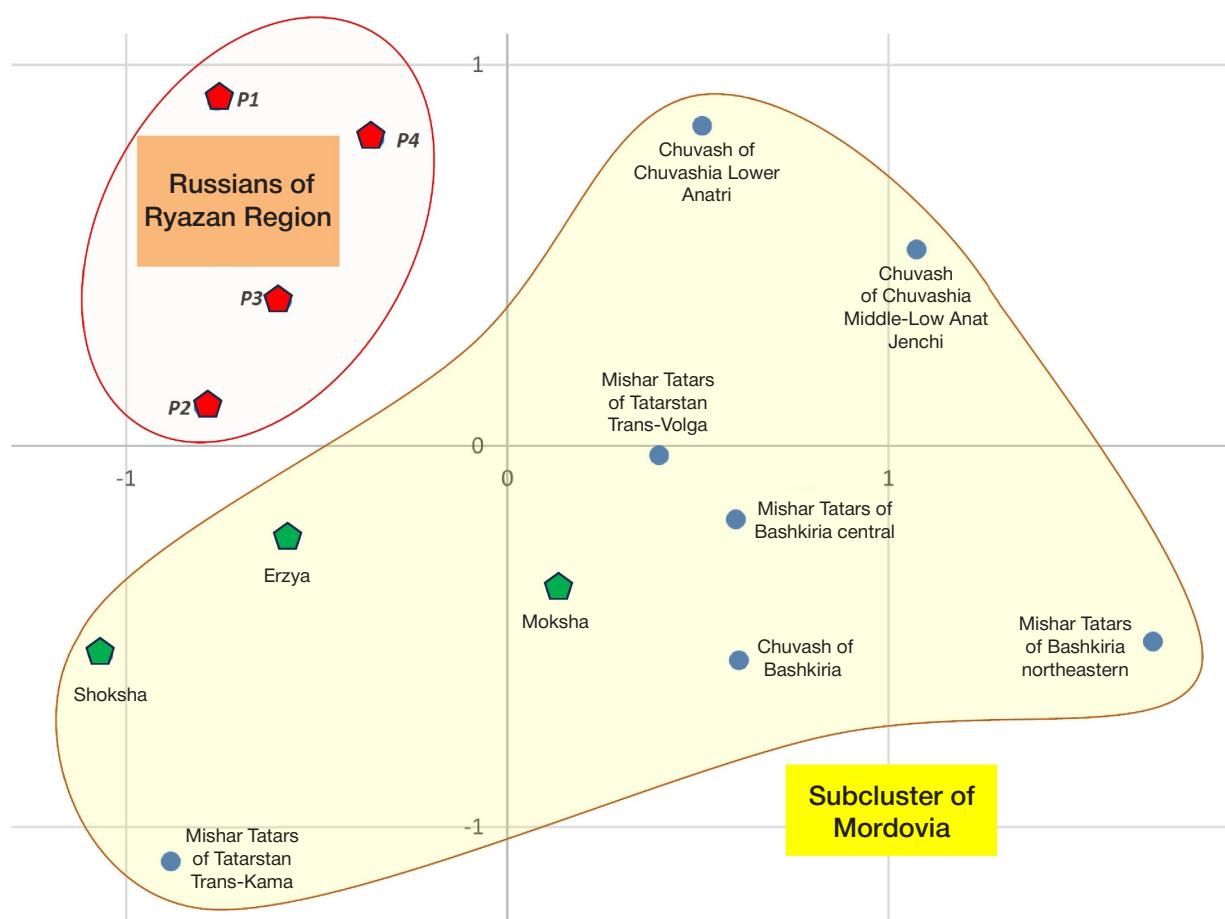
The Slavic cluster is divided into two subclusters: the first one includes the northernmost population of Central Russia, the second one includes the main bulk of the Central Russia's populations.

The subcluster of Slavs including 17 Russian populations of central and southern Russia turned out to be the most

dense and compact in the genetic space. It is significant that Poles, representatives of Western Slavs, are very close to it ( $n = 598$ ) [14]. This suggests a large geographic range of the populations included in the central Slavic subcluster, as well as considerable similarity of their gene pools. High frequencies of haplogroups R1a-CTS1211,



**Fig. 2.** Position of populations of European Russia in the genetic space based on the pairwise analysis of genetic distances (stress coefficient = 0.16; alienation coefficient = 0.17)



**Fig. 3.** Position of populations of the Slavic and Ural-Volga clusters relative to each other on the pairwise analysis of genetic distances (stress coefficient = 0.09; alienation coefficient = 0.13)

R1a-Z92, R1a-M458, I2a-P37.2, and N3a3 are the subcluster characteristic features.

The northern cluster of Central Russia's Slavs combined five populations of the Kostroma, Novgorod, and Yaroslavl regions [4, 8, 9]. Uniqueness of its gene pool is manifested by higher frequencies of haplogroups E and R1a-M458 and proximity to the gene pool of Finnish-speaking populations of northwestern Russia due to haplogroups I1-M253, I2a-P37.2, N3a4-Z1936, N2-P43, R1a-CTS1211, R1b-L51.

The Western Finnish cluster included all the studied populations of Finnish-speaking residents of northwestern Russia: Veps, Votians, Izhorians, Karelia's Karelians, Tver Karelians, Ingrian Finns (the relationship between those and the gene pool of the Russian North was reported in [8, 17]). The cluster gene pool combines increased frequencies of haplogroups of the western (I1-M253, R1b-L51) and Uralic (N3a4, N2) origin.

The Ural-Volga cluster combined Finnish-speaking and Turkic-speaking peoples of the region. Its range in the genetic space is the largest, which suggests considerable differences between gene pools. Positions of populations on the plot do not depend on the linguistic identity, but reflect their location in the geographic space. The gene pools of those combine the "Uralic" haplogroups N3a1, N2 and the conditionally "steppe" ones E, G2-P15, J, R1a-Z93, R1b-M73, R1b-Z2105 in different proportions. It is important to highlight two features of the Volga-Oka populations. First, populations of Mordovia (Erzya, Moksha, Shoksha) are genetically most close to the cluster of Russian populations, which confirms the conclusions of our previous study considering significant pre-Slavic substrate in the gene pool of the central and southern parts of European Russia [3]. Second, the Mishar Tatars and Chuvash are close

to the Volga-Oka populations (Fig. 3), which requires special discussion.

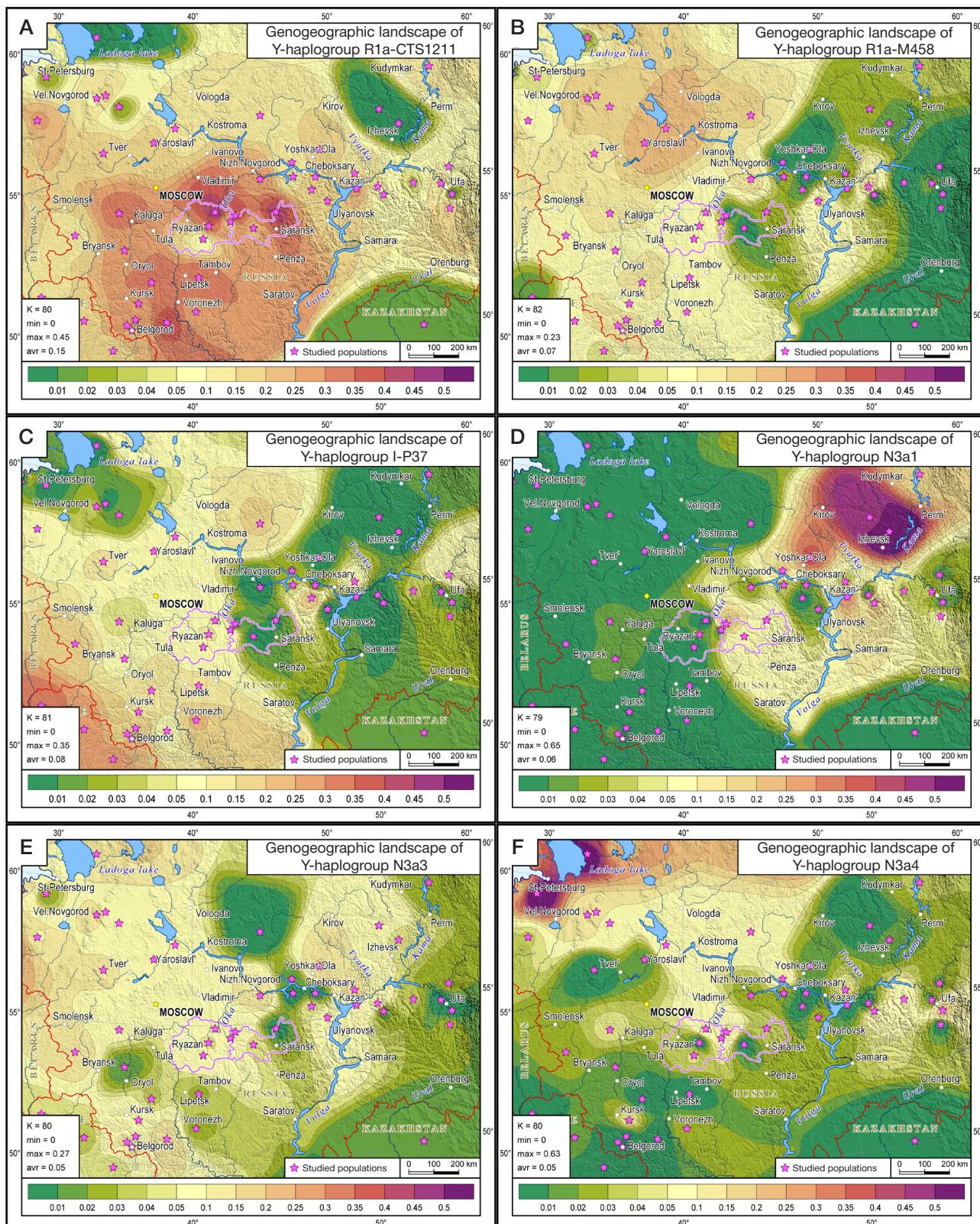
#### Gene-geographic landscape of Y-haplogroups in the indigenous population of Central Russia

The most important objective of the study is the analysis of the gene-geographic landscape of the indigenous peoples of European population providing valuable information about the nature and intensity of communication between the indigenous population and the newcomer Slavic tribes.

Being unable to provide all 35 maps of spatial variation of Y-haplogroup frequencies, we selected six most typical maps (Fig. 4): maps of "Slavic" (R1a-CTS1211, R1a-M458, and I2a-P37.2) and "Uralic" (N3a1, N3a3, N3a4) haplogroups. Both terms are very relative, since the maps themselves demonstrate how big the ranges are. However, the term "Uralic" also suggests both linguistic affiliation to the Uralic-speaking peoples and the geographic range, including the populations of Turkic-speaking peoples of the Ural-Volga region. Borders of the Ryazan Region and Mordovia allowing one to see the position of the Volga-Oka region in the gene-geographic landscape of Central Russia are provided in the maps.

Haplogroup R1a-CTS1211 (Fig. 4A) reaches its maximum frequencies (shades of red) not only in Russians of Central and Southern Russia, but also in Erzya and Moksha of Mordovia [3]. Its average frequencies (shades of yellow) cover all the periphery of the maps, except for minima (shades of green) in Kazakhstan, Udmurtia and Northwestern Finns.

Haplogroup R1a-M458 (Fig. 4B) is typical for the more western portion of the range compared to R1a-CTS1211.



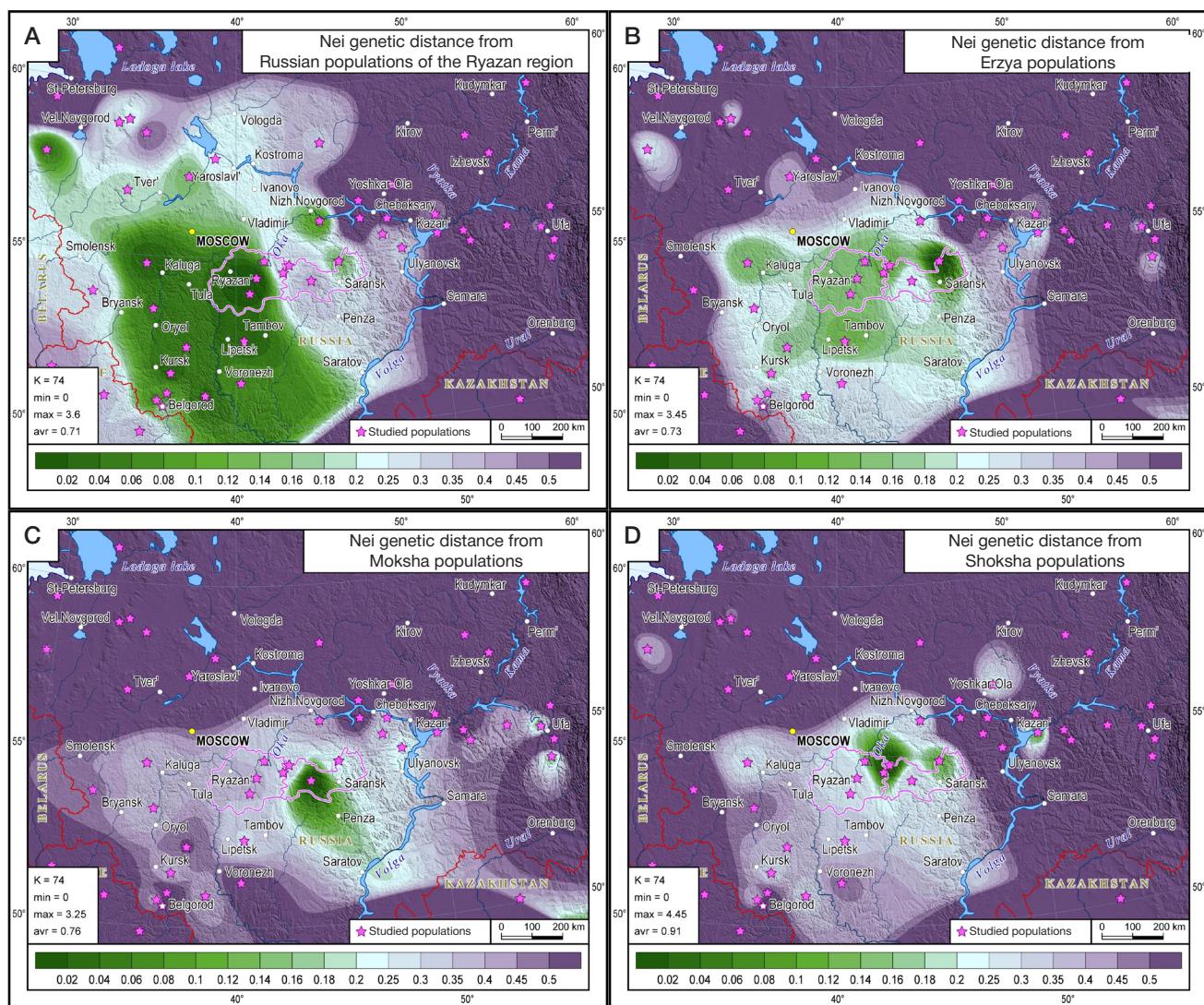
**Fig. 4.** Geographic distribution and frequencies of the major haplogroups of the indigenous population of European Russia: haplogroup R1a-CTS1211 (A); haplogroup R1a-M458 (B); haplogroup I-P37 (C); haplogroup N3a1 (D); haplogroup N3a3 (E); haplogroup N3a4 (F). Shades of green reflect the minimum haplogroup frequency; shades of yellow — medium frequencies; orange and red — maximum frequencies

The maximum frequencies are found in the populations of the northern subcluster of the MDS plot. The low-frequency zone (shades of green) stretches to Mordovia and Mari El.

Haplogroup I2a-P37.2 (Fig. 4C) having the same average gene-geographic landscape frequency ( $0.07 \leq q \leq 0.08$ ), as R1a-M458, is characterized by the larger scale of frequencies,

clearer transition boundaries, and the position of maximum frequencies in the southwest of European Russia.

Haplogroup N3a1 (Fig. 4D), as a truly "Uralic" one, is almost an alternative to the previous map of the "Slavic" I2a-P37.2. An absolute maximum (65%) is located in the north-east of Udmurtia, the boundary of average frequencies embraces



**Fig. 5.** Maps of Nei's genetic distances from the studied populations of the Volga-Oka region: from the total sample of Russian populations of the Ryazan Region (A); from populations of Erzya (B); from populations of Moksha (C); from populations of Shoksha (D). Shades of green reflect the minimum genetic distances (maximum similarity of gene pools); shades of purple and blue — medium genetic distances; shades of violet — maximum genetic distances (considerable differences between gene pools)

Mordovia (with the local maximum in Shoksha), but Russian populations are located in the low-frequency zone (shades of green).

Haplogroup N3a3 (Fig. 4E) is traditionally affiliated as “Uralic”, since it is typical for the Uralic-speaking northwestern peoples. The frequency maximum is located in the Baltics (28–58%), Fennoscandia, and adjacent regions of the Russian North [18], but the local maximum is also found in the Volga Region in Mari (15%). N3a3 with the medium frequencies is spread across almost all European Russia, which reflects ancient “Baltic” influences.

Haplogroup N3a4 (Fig. 4F) is distinguished by a clear gradient of the frequency decrease from north to south. Its maxima are located in Fennoscandia [15], but very high frequencies (up to 60%) are found in the northwestern Finnish-speaking peoples of Russia and in the Russian North. Furthermore, the medium frequency zone (shades of yellow) extends to the south to Ryazan and to the east to the foothills of the Urals (17% N3a4 in Bashkir).

#### Volga-Oka region gene pool position in the gene-geographic landscape of Central Russia

Visual assessment of the Volga-Oka region genetic similarity to various populations of European Russia is provided by maps of Nei's genetic distances (dNei) (Fig. 5).

Genetic distances from Russians of the Ryazan Region (average values for seven districts; Fig. 5A) confirm the detected (MDS; Fig. 2) similarity of those to the gene pools of almost all Russian populations of Central Russia (shades of green). A green spot of maximum similarity is also found in Mordovia, in Erzya. The zone of medium genetic distances (shades of purple and blue) covers the northern Russian populations, eastern Slavs, and the entire Volga right bank.

However, the pattern of genetic distances from Erzya is surprising (Fig. 5B): the “green” zone of high similarity covers only Russian populations, stretching to the Kaluga and Kursk regions. The “violet” zone of medium values is even broader: this stretches to the Belgorod, Bryansk, Nizhny Novgorod regions, but of Uralic populations includes only Mishar Tatars of the Trans-Kama region (Alexeyevsky, Alkeyevsky, Spassky, Chistopolsky districts of Tatarstan).

The unique nature of the Moksha Y gene pool manifested by the lack of populations being on the small genetic distance from Moksha was mentioned above (Fig. 5C). However, the zone of medium values is rather large: along with Erzya and Shoksha, this includes the Mishar Tatars of Bashkiria and the Trans-Volga region (Drozhzhanyovsky, Apastovsky districts of Tatarstan), Chuvash of Bashkiria, Anatri Chuvash of Chuvashia.

All these populations are located in the Mordovia subcluster on the plot (MDS, Fig. 3). The Shoksha gene pool (Fig. 5D) turned out to be the most close to that of the Mishar Tatars of the Trans-Kama region and Meadow Mari.

## DISCUSSION

The analysis of all Slavic populations [1] showed a strong correlation of the data on the autosomal genome and Y chromosome. Such a correlation was confirmed in the earlier report: the autosomal gene pool substitution during the Slavic expansion was accompanied by the no less dramatic change of Y-haplogroups that resulted in predominance of haplogroups R1a-CTS1211, R1a-M458, and I2a-P37.2 [2].

Therefore, in this study special attention is paid to the geographic landscape of the conditionally "Slavic" haplogroups. The total frequency of haplogroups R1a-CTS1211, R1a-Z92, R1a-M458, R1a-M198\*, and I2a-P37.2 calculated based on the set of 19 populations of Russians and 15 populations of Belarusians and Ukrainians constituted 67% of the gene pool of Eastern Slavs. This value is fully consistent with the total frequency of these "Slavic" haplogroups in Russians of Central Russia (66%) and is 3 times higher, than in Finnish-speaking and Turkic-speaking peoples of the Ural–Volga region (21%). However, this parameter is 2 times higher in populations of Mordovia (42%), than in the region (21%).

The in-depth phylogenetic study of modern and ancient carriers of haplogroups R1a-CTS1211 and R1a-Z92 in the Volga–Oka region has shown that about a half of the R1a-CTS1211 and R1a-Z92 haplogroup lineages on the Slavic expansion eastern periphery (Ryazan Region) are of pre-Slavic origin [3]. Frequency of these two haplogroups in Russians of the Ryazan Region eastern districts (46%) corresponds to their frequency in all Russian populations of central Russia (46%). This allows us to extend the lower estimate of the contribution of pre-Slavic population to the Y gene pool of Russians of central and southern Russia to all populations of Russians in this large region: this is about 43%. This estimate suggests the lower bound of the indigenous population contribution to the Slavic gene pool, since it was obtained specifically for the haplogroups most typical for the Slavs. Further in-depth phylogenetic study of modern and ancient carriers of "Uralic" haplogroups in the Volga–Oka region will show, how much higher could be the estimated contribution of pre-Slavic population to the Russian gene pool.

Among the conditionally "Uralic" haplogroups, haplogroups N3a1, N3a3, and N3a4 turned out to be the most common in the Volga–Oka region. However, we would like to emphasize again that the term "Uralic" is in this case both geographic and linguistic, it suggests the spread of these haplogroups across both studied peoples speaking Uralic languages (both northwestern — Veps, Votians, Izhorians, Karelians, Ingrian Finns; and eastern — Besermyan, Mari, Moksha, Udmurt, Shoksha, Erzya) and the considered Turkic-speaking populations of the Ural–Volga region (Bashkir, Kazan Tatars, Kryashens, Mishar Tatars, Chuvash). The total frequencies of the N3a1, N3a3, and N3a4 "Uralic" haplogroups in these three groups of populations are distributed in the following descending order: northwestern Finnish-speaking peoples (48%), eastern Finnish-speaking peoples (37%), Turks of the Ural–Volga region (12%). Based on the set of "Uralic" haplogroups, populations of Mordovia (23%) occupy an intermediate position in the Ural–Volga region, getting closer to Russians of the Ryazan Region

(11%) and Russian populations of Central Russia in general (11%) in this parameter.

The first discovered "kinship" of the Mordovia's population Y gene pools, all four studied populations of the Mishar Tatars of Tatarstan and Bashkiria, and the majority of Chuvash populations (Antri, Anat Jenchi, and Chuvash of Bashkiria) was an important finding of the analysis of the gene-geographic landscape of Central Russia. All these gene pools of the set of Finnish-speaking and Turkic-speaking populations form a common subcluster, which suggests their close genetic kinship. In the genetic space, this subcluster turned out to be close not only to the Ryazan Russians, but also to the entire set of Russian populations of Central and Southern Russia, based on the ratio of "Slavic" (35%) and "Uralic" haplogroups (14%) as well.

The fact, that all populations of Mordovia together with the Mishar Tatars and Chuvash form a common genetic set, allows us to hypothesize that all the populations of this cluster can be traced back to the common ancient Finnish-speaking population of the region. The sources of such unity can be different. It can be traced back to the early medieval Imenkovo archaeological culture or be related to migration of ancient Balts [19–21]. However, the Meschera hypothesis of the Mishar Tatar origin should be tested once again. According to anthropology data, [22] the Mishar Tatars are closer to Erzya, than to Meschera, but the new data on the Y gene pool suggest that not only Erzya and Mishar Tatars, but also all populations of Mordovia together with the Mishar Tatars and Chuvash form the common genetic set, which can result from the common nature of the tribes that included Meschera and the related tribes before Slavic expansion. These hypotheses can be tested based on the more thorough assessment of appropriate gene pools using the broad panels of Y-STR markers [3] and the ADMIXTURE analysis of the autosomal gene pool ancestral components.

## CONCLUSIONS

1. The findings of the gene-geographic analysis of variation of 35 Y-haplogroups in 80 populations of indigenous peoples of Central Russia (our own data) and 10 populations of the North European countries [13–15] are presented as multidimensional scaling plots and a series of the gene-geographic landscape maps.
2. It has been shown that the gene pool of Russian populations of the Volga–Oka region (seven districts of the Ryazan Region) is completely inside the compact cluster of populations of Central Russia, it belongs to the common gene pool of the Eastern and Western Slavs. The map of genetic distances from Russians of the Ryazan Region and all the studied populations of indigenous peoples of Central Russia has fully confirmed this conclusion.
3. The gene pool of Finnish-speaking peoples of the Volga–Oka region (Moksha, Erzya, Shoksha) belongs to the large cluster of the Ural–Volga populations, but it shows maximum genetic proximity to the Russian populations of Central Russia.
4. It has been found that the subcluster that is close to the Russian gene pool includes the majority of Chuvash populations and all the studied populations of the Mishar Tatars of Tatarstan and Bashkiria, along with the Mordovia's populations. It has been hypothesized that this set of Finnish-speaking and Turkic-speaking populations can be traced back to the gene pool of the ancient indigenous Finnish-speaking population of the Volga–Oka region.

## References

1. Kushniarevich A, Utevska O, Chuhryaeva M, Agdzhoyan A, Dibirova K, Uktveryte I, et al. Genetic heritage of the Balto-Slavic speaking populations: a synthesis of autosomal, mitochondrial and Y-chromosomal data. *PLoS ONE*. 2015; 10 (9): e0135820.
2. Gretzinger J, Biermann F, Mager H, King B, Zlámalová D, Traverso L, et al. Ancient DNA connects large-scale migration with the spread of Slavs. *Nature*. 2025. Available from: <https://doi.org/10.1038/s41586-025-09437-6>.
3. Adamov D, Shlykov A, Potanina A, Voronina M, Gorin I, Ponomarev G, et al. Pre-Slavic and Slavic Interaction at Eastern Periphery of Slavic Expansion in Northeastern Europe (Y-Gene Pools of Volga-Oka Region). *Genes*. 2025; 16 (10): 1149.
4. Balanovsky O, Roots S, Pshenichnov A, Kivisild T, Churnosov M, Evseeva I, et al. Two sources of the Russian patrilineal heritage in their Eurasian context. *American Journal of Human Genetics*. 2008; 82 (1): 236–50.
5. Balanovska EV, Balanovsky OP. Russkij genofond na Russkoj ravnine. M.: Luch, 2007; 416 s. Russian.
6. Okovancev VS, Ponomarev GYu, Agdzhoyan AT, Agdzhoyan AT, Pylyov VYu, Balanovska EV. Svoeobrazie pomorov Onezhskogo poluostrova i Zimnego berega v geneticheskem kontekste severa Evropy. *Vestnik Moskovskogo universiteta. Seriya 23: Antropologiya*. 2022; 5: 5–14. Russian.
7. Balanovska EV, Pezhemskij DV, Romanov AG, Baranova EE, Romashkina MV, Agdzhoyan AT i dr. Genofond Russkogo Severa: slavyane? finny? paleoevropcy? *Vestnik Moskovskogo universiteta. Seriya 23: Antropologiya*. 2011; 3: 27–58. Russian.
8. Balanovska EV, Agdzhoyan AT, Skhalyaho RA, Balaganska OA, Frejdin GS, Chernevskij KG i dr. Genofond novgorodcev: mezhdu severom i yugom. *Genetika*. 2017; 53: 1259–71. Russian.
9. Chuhryaeva MI, Pavlova ES, Napol'skih VV, Garin EV, Klopo AS, Temnyatkin SN i dr. Sohranilis' li sledy finno-ugorskogo vliyaniya v genofonde russkogo naseleniya Yaroslavskoj oblasti? Svidetel'stva Y-hromosomy. *Genetika*. 2017; 53: 388–99. Russian.
10. Balanovska EV, Gorin IO, Ponomarev GYu, Pylyov VYu, Petrushenko VS, Markina NV i dr. Sledy vzaimodejstviya finnoyazychnogo, slavyanskogo i tyurkoyazychnogo naseleniya v sovremenном genofonde i ih otrazhenie v farmakogenetike. *Vestnik RGMU*. 2022; №2: 18–27. Russian.
11. Peltola S, Majander K, Makarov N, Dobrovolskaya M, Nordqvist K, Salmela E, et al. Genetic admixture and language shift in the medieval Volga-Oka interfluvie. *Current Biology*. 2023; 33 (1): 174–82.
12. Balanovska EV, Zhabagin MK, Agdzhoyan AT, Chuhryaeva MI, Markina NV, Balaganskaya OA i dr. Populyacionnye biobanki: principy organizacii i perspektivy primeneniya v genogeografi i personalizirovannoj medicine. *Genetika*. 2016; 52 (12): 1371–87. Russian.
13. Ilumäe AM, Post H, Flores R, Karmin M, Sahakyan H, Mondal M, et al. Phylogenetic history of patrilineages rare in northern and eastern Europe from large-scale re-sequencing of human Y-chromosomes. *Eur J Hum Genet*. 2021; 29: 1510–19. Available from: <https://doi.org/10.1038/s41431-021-00897-8>.
14. Milewski M, Dawidziuk M. Most males in modern Poland carry Y-chromosomal lineages from clades that have recently expanded over Central, Eastern and SouthEastern Europe. *Human Genetics*. 2025; Available from: <https://doi.org/10.1007/s00439-025-02781-7>.
15. Preussner A, Leinonen J, Riikinen J, Pirinen M, Tukiainen T. Y chromosome sequencing data suggest dual paths of haplogroup N1a1 into Finland. *Eur J Hum Genet*. 2025; 33: 89–97. Available from: <https://doi.org/10.1038/s41431-024-01707-7>.
16. Koshel SM. Geoinformacionnye tekhnologii v genogeografi. V knige: I. K. Lur'e, V. I. Kravcova, redaktory. Sovremennaya geograficheskaya kartografiya. M.: Data+, 2012; s. 158–66. Russian.
17. Agdzhoyan A, Ponomarev G, Pylev V, (Kagazezheva) Zh, Gorin I, Esvyukov I, et al. The Finnic Peoples of Russia: Genetic Structure Inferred from Genome-Wide and Y-Chromosome Data. *Genes*. 2024; 15: 1610. Available from: <https://doi.org/10.3390/genes15121610>.
18. Ilumäe A M, Reidla M, Chuhryaeva M, Järve M, Post H, Karmin M, et al. Human Y Chromosome Haplogroup N: A Non-trivial TimeResolved Phylogeography that Cuts across Language Families. *The American Journal of Human Genetics*. 2016; 99 (1): 163–73.
19. Matveeva GI. O proiskhozhdenii imen'kovskoj kul'tury. Drevnie i srednevekovye kul'tury Povolzh'ya. Kujbyshev, 1981. Russian.
20. Sedov VV. Ocherki po arheologii slavyan. Moskva. 1994. 128 s. Russian.
21. Napol'skih VV. Balto-slavyanskiy yazykovoj komponent v Nizhnem Prikam'e v ser. I tys. n. e. Slavyanovedenie. 2006; 2: 3–19. Russian.
22. Alekseeva TI, Vasilev BA. K voprosu o geneticheskem rodstve russkoj meshchery i tatar-misharej. *Kratkie soobshcheniya Instituta etnografii*. 1959; 31: 3–13. Russian.

## Литература

1. Kushniarevich A, Utevska O, Chuhryaeva M, Agdzhoyan A, Dibirova K, Uktveryte I, et al. Genetic heritage of the Balto-Slavic speaking populations: a synthesis of autosomal, mitochondrial and Y-chromosomal data. *PLoS ONE*. 2015; 10 (9): e0135820.
2. Gretzinger J, Biermann F, Mager H, King B, Zlámalová D, Traverso L, et al. Ancient DNA connects large-scale migration with the spread of Slavs. *Nature*. 2025. Available from: <https://doi.org/10.1038/s41586-025-09437-6>.
3. Adamov D, Shlykov A, Potanina A, Voronina M, Gorin I, Ponomarev G, et al. Pre-Slavic and Slavic Interaction at Eastern Periphery of Slavic Expansion in Northeastern Europe (Y-Gene Pools of Volga-Oka Region). *Genes*. 2025; 16 (10): 1149.
4. Balanovsky O, Roots S, Pshenichnov A, Kivisild T, Churnosov M, Evseeva I, et al. Two sources of the Russian patrilineal heritage in their Eurasian context. *American Journal of Human Genetics*. 2008; 82 (1): 236–50.
5. Балановская Е. В., Балановский О. П. Русский генофонд на Русской равнине. М.: Луч, 2007; 416 с.
6. Окованцев В. С., Пономарев Г. Ю., Агджоян А. Т., Агджоян В. Ю., Балановская Е. В. Свообразие поморов Онежского полуострова и Зимнего берега в генетическом контексте севера Европы. *Вестник Московского университета. Серия 23: Антропология*. 2022; 5: 5–14.
7. Балановская Е. В., Лежемский Д. В., Романов А. Г., Баранова Е. Е., Ромашкина М. В., Агджоян А. Т. и др. Генофонд Русского Севера: славяне? финны? палеоевропейцы? *Вестник Московского*
- университета. Серия 23: Антропология. 2011; 3: 27–58.
8. Балановская Е. В., Агджоян А. Т., Схалъо Р. А., Балаганская О. А., Фрейдин Г. С., Черневский К. Г. и др. Генофонд новгородцев: между севером и югом. *Генетика*. 2017; 53: 1259–71.
9. Чухряева М. И., Павлова Е. С., Напольских В. В., Гарин Э. В., Клопов А. С., Темняткин С. Н. и др. Сохранились ли следы финно-угорского влияния в генофонде русского населения Ярославской области? Свидетельства Y-хромосомы. *Генетика*. 2017; 53: 388–99.
10. 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: 18–27.
11. Peltola S, Majander K, Makarov N, Dobrovolskaya M, Nordqvist K, Salmela E, et al. Genetic admixture and language shift in the medieval Volga-Oka interfluvie. *Current Biology*. 2023; 33 (1): 174–82.
12. Балановская Е. В., Жабагин 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.
13. Ilumäe AM, Post H, Flores R, Karmin M, Sahakyan H, Mondal M, et al. Phylogenetic history of patrilineages rare in northern and eastern Europe from large-scale re-sequencing of human

Y-chromosomes. *Eur J Hum Genet.* 2021; 29: 1510–19. Available from: <https://doi.org/10.1038/s41431-021-00897-8>.

14. Milewski M, Dawidziuk M. Most males in modern Poland carry Y-chromosomal lineages from clades that have recently expanded over Central, Eastern and SouthEastern Europe. *Human Genetics.* 2025; Available from: <https://doi.org/10.1007/s00439-025-02781-7>.
15. Preussner A, Leinonen J, Riikonen J, Pirinen M, Tukiainen T. Y chromosome sequencing data suggest dual paths of haplogroup N1a1 into Finland. *Eur J Hum Genet.* 2025; 33: 89–97. Available from: <https://doi.org/10.1038/s41431-024-01707-7>.
16. 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.
17. Agdzhoyan A, Ponomarev G, Pylev V, (Kagazezheva) Zh, Gorin I, Esvyukov I, et al. The Finnic Peoples of Russia: Genetic Structure Inferred from Genome-Wide and Y-Chromosome Data. *Genes.* 2024; 15: 1610. Available from: <https://doi.org/10.3390/genes15121610>.
18. Ilumäe A M, Reidla M, Chukhryaeva M, Järve M, Post H, Karmin M, et al. Human Y Chromosome Haplogroup N: A Non-trivial TimeResolved Phylogeography that Cuts across Language Families. *The American Journal of Human Genetics.* 2016; 99 (1): 163–73.
19. Матвеева Г. И. О происхождении именьковской культуры. Древние и средневековые культуры Поволжья. Куйбышев, 1981.
20. Седов В. В. Очерки по археологии славян. Москва. 128 с, 1994.
21. Напольских В. В. Балто-славянский языковой компонент в Нижнем Прикамье в сер. I тыс. н. э. *Славяноведение.* 2006; 2: 3–19.
22. Алексеева Т. И., Васильев Б. А. К вопросу о генетическом родстве русской мещеры и татар-мишарей. *Краткие сообщения Института этнографии.* 1959; 31: 3–13.