
HUMAN
GENETICS

Gene Pool Structure of Russian Populations from the European Part of Russia Inferred from the Data on Y Chromosome Haplogroups Distribution

B. A. Malyarchuk and M. V. Derenko

*Institute of Biological Problems of the North, Russian Academy of Sciences, Magadan, 685000 Russia;
e-mail: malyarchuk@ibpn.ru*

Received November 8, 2006

Abstract—Population structure of Russian population from the European part of Russia was investigated by analyzing the distribution of 23 SNP makers of Y chromosome in Russian populations from Kaluga, Yaroslavl', Vladimir, Nizhni Novgorod, Pskov, Tula, Belgorod, and Novgorod oblasts. In the populations studied a total of 14 Y-chromosome haplogroups (E, F*, I, J, K*, N3a, N2, P*, R1*, R1a1, C3, G, H, and A) were discovered, of which haplogroups R1a1, I, and N3a were the prevailing. Analysis of Φ statistics in the populations grouped in accordance to the dialect subdivision of the Russian language, showed the absence of statistically significant differences between Russian population groups. Analysis of the Y-chromosome markers distribution patterns among Russian population (10 population groups) in comparison with the population of Germany (11 population groups) and Poland (8 population groups) revealed statistically significant differences between the gene pools of Slavs (Russians and Poles) and Teutons (Germans).

DOI: 10.1134/S1022795408020105

INTRODUCTION

In recent years, analysis of the variation of nonrecombining part of Y chromosome in human populations attracted increasing attention in association with investigations in the fields of evolutionary and population genetics, as well as in forensic medicine and crime detection [1]. This interest is mostly determined by different evolutionary rates of the main markers of Y chromosome. These markers are represented by the SNP loci characterized by single-nucleotide substitutions with low accumulation rate during the evolution, and the STR loci, characterized by extremely high microsatellite instability [2]. At present, combination of these approaches for detection of Y chromosome variation is an effective tool for studying Y chromosome evolution, as well as for the reconstruction of the ancient history of humans and timing of its main episodes, and for the analysis of rather recent population differentiation events.

Investigations of Y chromosome variation in Russian populations from Eastern Europe showed that Russian gene pool contained a set of DNA lineages belonging to Y chromosome groups typical of many populations of northern and eastern parts of Europe [3–9]. The data obtained showed that only some Russian populations (from Pskov and Pomor'e) have demonstrated substantial similarity to Finno-Ugric and Baltic populations of Northern and Eastern Europe. However, the majority of Russian populations cluster together with Slavic populations (Poles, Ukrainians, and Belorus-

sians) [6]. The latter finding is explained in terms of specific features of the population genetic structures examined. The majority of Russian populations, similarly to Ukrainians and Belarusians, are characterized by the high frequency of haplogroup R1a1 along with moderate or low frequency of haplogroup N3. At the same time, some populations, specifically, from Pskov and Pomor'e, similarly to Finno-Ugric and Baltic populations, show the increased frequency of haplogroup N3 along with the reduced frequency of R1a1. Although the between-population differentiation of Russian populations from Eastern Europe has been characterized (based on polymorphism data at 12 biallelic loci of Y chromosome [6]), extension of the spectrum of SNP markers involved in the analysis seems to be reasonable for further characterization of the genetic structure of Russian population. In this study, the data on the distribution of Y chromosome haplogroups inferred from the analysis of polymorphism at 23 SNP loci in ten Russian populations from the European part of Russia are presented.

MATERIALS AND METHODS

Distribution of Y chromosome haplogroups was examined in Russian populations from Kaluga, Yaroslavl', Vladimir, Nizhni Novgorod, Pskov, Tula, Orel, Belgorod, and Novgorod oblasts. The total size of the samples tested in the present study constituted 414 unrelated male individuals.

Table 1. Prevalence of Y-chromosome haplogroups in the Russian populations from the European part of Russia

Population	<i>n</i>	E	F*	I	J	K*	N3a	N2	P*	R1*	R1a1	C3	G	H	A
Nizhni Novgorod	58	5 (8.6)	2 (3.4)	11 (19.0)	0	0	8 (13.8)	0	5 (8.6)	6 (10.3)	21 (36.2)	0	0	0	0
Kaluga	42	2 (4.8)	3 (7.1)	8 (19.0)	0	1 (2.4)	2 (4.8)	0	0	4 (9.5)	21 (50.0)	0	0	0	1 (2.4)
Tula	43	1 (2.3)	2 (4.7)	8 (18.6)	0	0	5 (11.6)	0	0	3 (7.0)	22 (51.2)	0	1 (2.3)	1 (2.3)	0
Volot	40	2 (5.0)	0	6 (15.0)	0	1 (2.5)	6 (15.0)	0	0	2 (5.0)	20 (50.0)	0	0	3 (7.5)	0
Veliki Novgorod	37	3 (8.1)	0	4 (10.8)	0	0	5 (13.5)	0	1 (2.7)	3 (8.1)	20 (54.1)	0	1 (2.7)	0	0
Orel	36	0	1 (2.8)	4 (11.1)	0	0	7 (19.4)	0	1 (2.8)	4 (11.1)	19 (52.8)	0	0	0	0
Belgorod	44	3 (6.8)	0	8 (18.2)	6 (13.6)	2 (4.5)	3 (6.8)	0	0	1 (2.3)	20 (45.5)	1 (2.3)	0	0	0
Vladimir	51	1 (2.0)	0	9 (17.6)	0	0	6 (11.8)	1 (2.0)	0	3 (5.9)	29 (56.9)	0	2 (3.9)	0	0
Yaroslavl	23	0	0	1 (4.3)	0	3 (13.0)	2 (8.7)	0	2 (8.7)	1 (4.3)	14 (60.9)	0	0	0	0
Pskov	40	3 (7.5)	0	7 (17.5)	0	0	14 (35.0)	0	0	1 (2.5)	14 (35.0)	0	1 (2.5)	0	0
Russians altogether	414	20 (4.8)	8 (1.9)	66 (15.9)	6 (1.4)	7 (1.7)	58 (14.0)	1 (0.2)	9 (2.2)	28 (6.8)	200 (48.3)	1 (0.2)	5 (1.2)	4 (1.0)	1 (0.2)

In the populations of Novgorod oblast (two samples, the city of Veliki Novgorod and the settlement of Volot) distribution of 23 Y chromosome markers was examined, including M173, M17, SRY-10831, RPS4Y, M217, M8, M38, SRY-8299, M89, 12f2, M9, M207, M168, M20, 92R7, Tat, LLY22g, M178, P43, M170, M52, and M201. Distribution of the *Alu* insertions (YAP in the locus DYS287), which determine the gene pool structure in the populations of Northern Eurasia [10, 11] was analyzed. In the populations from Kaluga, Yaroslavl, Vladimir, Nizhny Novgorod, Pskov, Tula, Orel, and Belgorod oblasts, polymorphism at 12 loci (M8, M17, M38, P43, M52, M168, M170, M173, M178, M201, M207, and M217) was studied, since other loci were already analyzed in the previous study [6]. Polymorphism was determined in accordance with the scheme of the analysis suggested in [3, 7, 11–17] using the method of DNA analysis in 8% polyacrylamide gels.

The levels of the population genetic differentiation were evaluated using analysis of Φ statistics (AMOVA, included into the ARLEQUIN 3.0 software package [18]). The degree of the population partitioning was analyzed with the help of Φ statistics, evaluating genetic differences among the population groups (Φ_{CT}),

among the populations within the groups (Φ_{SC}), and total genetic differentiation (Φ_{ST}) [18]. Statistical significance of the differences between the values of Φ statistics upon the pairwise population comparisons was tested using nonparametric permutation approach (a total of 10 100 permutations) [18]. The matrices of the pairwise among-population Φ_{ST} distances were used for the analysis of spatial population distribution with the help of the method of multidimensional scaling (STATISTICA/w 5.0). For comparative analysis, the data on the distribution of Y chromosome haplogroups in the populations of Germany and Poland were used [19].

MATERIALS AND METHODS

Analysis of 23 Y chromosome markers in 10 Russian populations from the European part of Russia revealed the presence of 14 Y chromosome haplogroups, including E, F*, I, J, K*, N3a, N2, P*, R1*, R1a1, C3, G, H, and A (Table 1). Among the haplogroups tested, haplogroups R1a1, I, and N3a were considered as most prevalent. In Russians, the summarized frequency of these haplogroups constituted 78.2%, on average. Similar frequencies of these three haplogroups were observed previously in the gene pools of other

Table 2. The among-population differentiation of Russian populations from the European part of Russia, based on the Φ_{ST} distances

Population	1	2	3	4	5	6	7	8	9
1. Kaluga									
2. Yaroslavl	0.016								
3. Vladimir	-0.003	0.007							
4. Nizhni Novgorod	0.006	0.044*	0.024						
5. Pskov	0.062*	0.090*	0.053*	0.018					
6. Tula	-0.016	0.012	-0.017	0.005	0.036*				
7. Orel	0.001	0.001	-0.008	0.009	0.026	-0.013			
8. Belgorod	0.002	0.025	0.009	0.012	0.049*	0.000	0.017		
9. Veliki Novgorod	-0.005	-0.005	-0.014	0.007	0.037	-0.015	-0.018	0.005	
10. Volot	-0.004	0.006	-0.016	0.005	0.021	-0.019	-0.013	-0.002	-0.020

Note: Statistically significant among-population differences ($P < 0.05$) are designated by the asterisks.

Eastern Slavic populations, Belarusians and Eastern Ukrainians [8].

The data of the AMOVA analysis of among-population differentiation (Table 2) point to the low level of among-population differentiation in Russians in terms of Y chromosome lineages. Only the structure of the population from Pskov was statistically significantly different from the other Russian populations (from Kaluga, Yaroslavl, Vladimir, Belgorod, and Tula oblasts) (Table 2). Analysis of the among-population differences, performed with the help of the exact test, also pointed to the difference of Pskov and Nizhny Novgorod populations from other Russian populations examined ($P < 0.05$). Based on Y chromosome haplogroup distribution patterns in different Russian populations, it can be concluded that the among-population differences observed were mainly due to the reduced frequency of haplogroup R1a1 (down to 35%, approximately) in Pskov and Nizhni Novgorod samples (Table 1). In addition, compared to all other Russian populations, the sample of Pskov oblast was characterized by the increased frequency of haplogroup N3a (up to 35%). It seems likely, that these features determine genetic specificity of Pskov population (Table 1). It should be noted, that Novgorod populations (the city of Veliki Novgorod and the settlement of Volot), despite their geographical closeness to Pskov population, are more distant from this population (Φ_{ST} values constituted 3.7 and 2.1%, respectively) than from other Russian populations (Table 2). Similar conclusion on the genetic position of Novgorod population was made by

Khrunin et al. [20]. The authors, analyzing polymorphism at five microsatellite loci of Y chromosome, have demonstrated that Russians from Novgorod oblast were close to Russian populations from Kursk oblast, Ukrainians, and Belarusians, and at the same time, distant from Russians of Arkhangel'sk oblast. The latter population is characterized by genetic similarity to Finno-Ugric populations (Finns and Saami).

The works focused on differentiation of Y chromosome lineages in different world populations are at their beginning. Recently, the data of such analysis were published for German and Polish populations [19]. In the study cited, substantial level of genetic differences between the neighboring populations ($\Phi_{ST} = 14.2\%$) was reported. The differences observed were mainly due to the low frequencies of some Y chromosome haplogroups (for example, R1a1 and I1b), typical of Slavic populations, in Germans. Comparative analysis of the main Y chromosome haplogroups (P*, R1*, R1a1, I, E, F*, N*, N3, J, and K*) in German, Polish, and Russian populations (at the level of three population groups), performed in the present study, showed the high level of the among-group differentiation: $\Phi_{CT} = 11.68\%$, $\Phi_{ST} = 12.6\%$, $\Phi_{SC} = 0.88\%$ (all differences were statistically significant). The populations groups were different in terms of the levels of among-population differentiation. The most homogenous population was that of Poles ($\Phi_{ST} = 0.3\%$), while German populations were the most diverse ($\Phi_{ST} = 1.4\%$), and Russian populations occupied the intermediate position ($\Phi_{ST} = 0.9\%$) (Table 3). Analysis of the among-population differences (the val-

Table 3. indices of diversity and among-population differentiation in Germans, Poles, and Russians

Population	No. of populations	Φ_{ST}	h (range of values)
Germans	11	0.0143 ($P = 0.00$)	0.745 ± 0.025 (0.65–0.79)
Poles	8	0.003 ($P = 0.13$)	0.629 ± 0.044 (0.56–0.72)
Russians	10	0.0093 ($P = 0.07$)	0.703 ± 0.061 (0.62–0.80)

Note: The mean values of the genetic diversity index h and their errors are presented in accordance with the AMOVA results [18].

ues of Φ_{ST} distances and the results of their multidimensional scaling) showed that Russian and Polish populations did not differentiate from one another and cluster apart from German populations (Table 4, figure).

It was demonstrated [19] that populations from eastern and western parts of Germany were statistically significantly different in terms of Y chromosome haplogroup composition ($\Phi_{ST} = 1.2\%$; $P = 0$). At the same

Table 4. Comparative analysis of the among-population Φ_{ST} distances in the populations of Germans, Poles, and Russians

Population	Range of values for the among-population Φ_{ST} distances		
	Germans	Poles	Russians
Germans	0–0.084	0.045–0.290	0.034–0.290
Poles		0–0.018	0–0.132
Russians			0–0.090

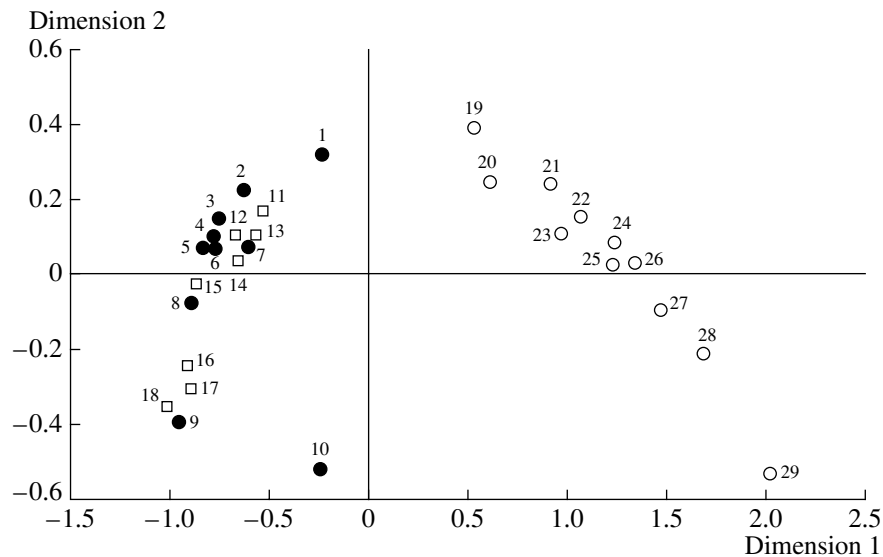
Table 5. Genetic differences between Russian populations grouped in accordance with linguistic (I and II) and anthropological (III) data

Population groups	Differences, %		
	between the groups	between the populations within the groups	within populations
I	0.32	0.74	98.94
II	–0.22	1.09	99.13
III	–0.53	1.42	99.11

Note: I, populations are grouped relative to subdialects and patois of Russian language [21]: southern patois (Belgorod, Orel, Kaluga, and Tula), northern patois (Yaroslavl), Middle-Russian patois (Veliki Novgorod, Volot, Pskov, Vladimir, and Nizhni Novgorod). II, populations are grouped by dialect zones [21]: southwestern zone (Belgorod, Orel, Kaluga, and Tula), northwestern zone (Veliki Novgorod, Volot, and Pskov), and northeastern zone (Vladimir, Nizhni Novgorod, and Yaroslavl). III, populations are grouped according to anthropological data [22]: the presumptive descendants of Novgorod Slovans (Veliki Novgorod, Volot, and Pskov), Krivichi (Vladimir, Nizhni Novgorod, and Yaroslavl), Vyatichi (Kaluga and Tula), and Severyane (Orel and Belgorod).

time, no genetic differences were observed between the populations from eastern and western parts of Poland ($\Phi_{ST} = 0.09\%$; $P = 0.22$). To reveal possible genetic differences between Russian populations from Eastern Europe, an analysis of the Φ statistics was carried out in the populations grouped in accordance with the dialect subdivision of Russian language in the European part of Russia. It should be noted in this context that dialect groupings are believed to play an important role in within-ethnic differentiation of Russians [21]. The analysis showed that despite higher value of the among-group differences obtained upon grouping of the populations in accordance with the distribution of subdialects and patois of the Russian language ($\Phi_{CT} = 0.0032$, $\Phi_{SC} = 0.0074$, $\Phi_{ST} = 0.0106$) in comparison with the value obtained using the map of dialect zones ($\Phi_{CT} = -0.0022$, $\Phi_{SC} = 0.0109$, $\Phi_{ST} = 0.0087$), in both cases the among-population differences were not statistically significant ($P > 0.07$) (Table 5). These findings suggest the absence of genetic differences in terms of the distribution of Y chromosome lineages in Russian population groups, distinguished based on the linguistic data. Analogous conclusion ($P = 0.083$) follows from the analysis of the differentiation of Russian populations grouped based on anthropological data ($\Phi_{CT} = -0.0053$, $\Phi_{SC} = 0.0142$, $\Phi_{ST} = 0.0089$) (Table 5). In this case populations were grouped in accordance with ethnic territorial subdivision of Russian population, based on the idea on the reflection of anthropological characteristics of annalistic Slavic tribes in ethnic territorial groups of contemporary Russian population [22]. The placement of Pskov population, dramatically contrasting with the other Russian populations (figure), into a separate group did not result in the increase of the level of genetic differentiation ($\Phi_{CT} = 0.0023$, $\Phi_{SC} = 0.0073$, $\Phi_{ST} = 0.0096$, $P > 0.07$).

Thus, analysis of Y chromosome variation shows that among-population differentiation of Russians from the European part of Russia is either absent (based on the analysis of the SNP distribution patterns), or is very small ($\Phi_{ST} = 1.09\%$ ($P = 0.0003$) for 12 microsatellite loci of Y chromosome [23]). However, as it was demonstrated in the present study, structural similarity of the gene pools of Russian and Polish populations is so high, that on the interethnic level of comparisons it is



Positions of Russian, German, and Polish populations within two-dimensional space inferred from the multidimensional scaling of pairwise among-population Φ_{ST} distances, based on the Y chromosome haplogroup distribution patterns in the populations. Russian populations (black circles): 1, Nizhni Novgorod; 2, Kaluga; 3, Tula; 4, Volot; 5, Veliki Novgorod; 6, Orel; 7, Belgorod; 8, Vladimir; 9, Yaroslavl; 10, Pskov. Polish populations (pale squares): 11, Wroclaw; 12, Szczecin; 13, Warsaw; 14, Bydgoszcz; 15, Suwalki; 16, Gdansk; 17, Lublin; 18, Krakow. German populations (pale circles): 19, Rostock; 20, Berlin; 21, Magdeburg; 22, Greifswald; 23, Leipzig; 24, Kolon; 25, Hamburg; 26, Munich; 27, Munster; 28, Meinz; 29, Freiburg.

possible to distinguish between the gene pools of Slavs (Russians and Poles) and Teutons (Germans).

ACKNOWLEDGMENTS

B.A. Malyarchuk is thankful to the Foundation for Promotion of National Science.

This work was supported by the Program of the Presidium of Russian Academy of Sciences "Dynamics of Gene Pools and Biodiversity", the Program of the Far East Branch of the Russian Academy of Sciences (grant no. 06-I-P11-033), and the Russian Foundation for Basic Research (grant no. 06-04-48136).

REFERENCES

1. Jobling, M. and Tyler-Smith, C., The Human Y Chromosome: An Evolutionary Marker Comes of Age, *Nat. Rev. Genet.*, 2003, vol. 4, pp. 598–612.
2. Kayser, M., Roewer, L., Hedman, M., et al., Characteristics and Frequency of Germline Mutations at Microsatellite Loci from the Human Y Chromosome, as Revealed by Direct Observation in Father/Son Pairs, *Am. J. Hum. Genet.*, 2000, vol. 66, pp. 1580–1588.
3. Rosser, Z.H., Zerjal, T., Hurles, M., et al., Y-Chromosomal Diversity in Europe Is Clinal and Influenced Primarily by Geography, Rather than by Language, *Am. J. Hum. Genet.*, 2000, vol. 67, pp. 1526–1543.
4. Semino, O., Passarino, G., Oefner, P.J., et al., The Genetic Legacy of Paleolithic *Homo sapiens sapiens* in Extant Europeans: A Y Chromosome Perspective, *Science*, 2000, vol. 290, pp. 1155–1159.
5. Zerjal, T., Beckman, L., Beckman, G., et al., Geographical, Linguistic, and Cultural Influences on Genetic Diversity: Y-Chromosomal Distribution in Northern European Populations, *Mol. Biol. Evol.*, 2001, vol. 18, pp. 1077–1087.
6. Malyarchuk, B., Derenko, M., Grzybowski, T., et al., Differentiation of Mitochondrial DNA and Y Chromosome in Russian Populations, *Hum. Biol.*, 2004, vol. 76, pp. 877–900.
7. Rootsi, S., Magri, C., Kivisild, T., et al., Phylogeography of Y-Chromosome Haplogroup I Reveals Distinct Domains of Prehistoric Gene Flow in Europe, *Am. J. Hum. Genet.*, 2004, vol. 75, pp. 128–137.
8. Kharkov, V.N., Stepanov, V.A., Feshchenko, S.P., et al., Frequencies of Y Chromosome Binary Haplogroups in Belarussians, *Russ. J. Genet.*, 2005, vol. 41, no. 8, pp. 928–931.
9. Underhill, P.A., Passarino, G., Lin, A.A., et al., The Phylogeography of Y-Chromosome Binary Haplotypes and the Origins of Modern Human Populations, *Ann. Hum. Genet.*, 2001, vol. 65, pp. 43–62.
10. Y Chromosome Consortium: A Nomenclature System for the Tree of Human Y-Chromosomal Binary Haplogroups, *Genome Res.*, 2002, vol. 12, pp. 339–348.
11. Hurles, M.E., Irvén, C., Nicholson, J., et al., European Y-Chromosomal Lineages in Polynesia: A Contrast to the Population Structure Revealed by mtDNA, *Am. J. Hum. Genet.*, 1998, vol. 63, pp. 1793–1806.
12. Hammer, M.F. and Horai, S., Y Chromosomal DNA Variation and the Peopling of Japan, *Am. J. Hum. Genet.*, 1995, vol. 56, pp. 951–962.
13. Kwok, C., Tyler-Smith, C., Medonca, B.B., et al., Mutation Analysis of 2 kb 5' to SRY in XY Females and XX

- Intersex Subjects, *J. Med. Genet.*, 1996, vol. 33, pp. 465–468.
14. Hurles, M.E., Veitia, R., Arroyo, E., et al., Recent Male-Mediated Gene Flow Over a Linguistic Barrier in Iberia, Suggested by Analysis of a Y-Chromosomal DNA Polymorphism, *Am. J. Hum. Genet.*, 1999, vol. 65, pp. 1437–1448.
 15. Zerjal, T., Dashnyam, B., Pandya, A., et al., Genetic Relationships of Asians and Northern Europeans, Revealed by Y-Chromosomal DNA Analysis, *Am. J. Hum. Genet.*, 1997, vol. 60, pp. 1174–1183.
 16. Underhill, P.A., Shen, P., Lin, A.A., et al., Y-Chromosome Sequence Variation and the History of Human Populations, *Nat. Genet.*, 2000, vol. 26, pp. 358–361.
 17. Ke, Y., Su, B., Song, X., et al., African Origin of Modern Humans in East Asia: A Tale of 12,000 Y Chromosomes, *Science*, 2001, vol. 292, pp. 1151–1152.
 18. Excoffier, L., Laval, G., and Schneider, S., Arlequin ver. 3.0: An Integrated Software Package for Population Genetics Data Analysis, *Evol. Bioinform. Online*, 2005, vol. 1, pp. 47–50.
 19. Kayser, M., Lao, O., Anslinger, K., et al., Significant Genetic Differentiation between Poland and Germany Follows Present-Day Political Borders, As Revealed by Y-Chromosome Analysis, *Hum. Genet.*, 2005, vol. 117, pp. 428–443.
 20. Khrunin, A.V., Bebyakova, N.A., Ivanov, V.P., et al., Polymorphism of Y-Chromosomal Microsatellites in Russian Populations from the Northern and Southern Russia as Exemplified by the Populations of Kursk and Arkhangel'sk Oblasts, *Russ. J. Genet.*, 2005, vol. 41, no. 8, pp. 922–927.
 21. Zakharova, K.F. and Orlova, V.G., *Dialektnoe chlenenie russkogo yazyka. Izd. 2* (Dialect Subdivision of Russian Language), Moscow: URSS, 2004.
 22. Alekseeva, T.I., *Etnogenez vostochnykh slavyan* (Ethnogeny of the Eastern Slavs), Moscow: Mosk. Gos. Univ., 1973.
 23. Perkova, M.A., Derenko, M.V., and Malyarchuk, B.A., Distribution of Anthropologically Informative Genetic Markers among the East European Russian Populations, *Genetika v Rossii i mire* (Genetics in Russia and in the World), *Proc. Int. Conf. Dedicated to 40th Anniversary of Vavilov Institute of General Genetics, Russian Academy of Sciences*, Moscow, 2006, p. 152.