ISSN 1022-7954, Russian Journal of Genetics, 2007, Vol. 43, No. 5, pp. 551–562. © Pleiades Publishing, Inc., 2007.
Original Russian Text © V.N. Kharkov, V.A. Stepanov, O.F. Medvedeva, M.G. Spiridonova, M.I. Voevoda, V.N. Tadinova, V.P. Puzyrev, 2007, published in Genetika, 2007, Vol. 43, No. 5, pp. 675–687.

HUMAN GENETICS

Gene Pool Differences between Northern and Southern Altaians Inferred from the Data on Y-Chromosomal Haplogroups

V. N. Kharkov[a], V. A. Stepanov[a], O. F. Medvedeva[a], M. G. Spiridonova[a], M. I. Voevoda[b, c], V. N. Tadinova[d], and V. P. Puzyrev[a]

[a] *Institute of Medical Genetics, Tomsk Scientific Center, Russian Academy of Medical Sciences, Tomsk, 634050 Russia; fax: (3822)51-37-44; e-mail: vadim.stepanov@medgenetics.ru*
[b] *Institute of Internal Medicine, Russian Academy of Medical Sciences, Novosibirsk, 630089 Russia; fax: (3832)64-25-16*
[c] *Institute of Cytology and Genetics, Russian Academy of Sciences, Novosibirsk, 630090 Russia; fax: (383)333-12-78*
[d] *Republican Children Hospital, Gorno-Altaisk, 649002 Altai Republic, Russia; fax: (38822)2-61-86*

Received March 6, 2006

Abstract—Y-chromosomal haplogroups composition and frequencies were analyzed in Northern and Southern Altaians. In the gene pool of Altaians a total of 18 Y-chromosomal haplogroups were identified, including C3xM77, C3c, DxM15, E, F*, J2, I1a, I1b, K*, N*, N2, N3a, O3, P*, Q*, R1*, R1a1, and R1b3. The structuring nature of the Altaic gene pool is determined by the presence of the Caucasoid and Mongoloid components, along with the ancient genetic substratum, marked by the corresponding Western and Eastern Eurasian haplogroups. Haplogroup R1a1 prevailed in both ethnic groups, accounting for about 53 and 38% of paternal lineages in Southern and Northern Altaians, respectively. This haplogroup is thought to be associated with the eastward expansion of early Indo-Europeans, and marks Caucasoid element in the gene pools of South Siberian populations. Similarly to haplogroup K*, the second frequent haplogroup Q* represents paleo-Asiatic marker, probably associated with the Ket and Samoyedic contributions to the Altaic gene pool. The presence of lineages N2 and N3a can be explained as the contribution of Finno–Ugric tribes, assimilated by ancient Turks. The presence of haplogroups C3xM77, C3c, N*, and O3 reflects the contribution of Central Asian Mongoloid groups. These haplogroups, probably, mark the latest movements of Mongolian migrants from the territory of contemporary Tuva and Mongolia. The data of factor analysis, variance analysis, cluster analysis, and phylogenetic analysis point to substantial genetic differentiation of Northern and Southern Altaians. The differences between Northern and Southern Altaians in the haplogroup composition, as well as in the internal haplotype structure were demonstrated.

DOI: 10.1134/S1022795407050110

INTRODUCTION

Ethnic groups of the Altai–Sayan region are especially interesting for ethnogenetic studies. The gene pool of the indigenous populations of Southern Siberia was formed as the result of long and multistage admixture of local gene pools of different Caucasoid and Mongoloid tribes. Mixing of different Turkic, Mongolian, Ket, and Uralian groups at the basis of the genetic substratum of ancient Indo-European tribes resulted in the formation of the varied genetic diversity pattern of the Altaic populations [1]. Anthropologic heterogeneity of South Siberian populations can be traced at the regional, as well as at the intraethnic levels [2, 3].

Altaians represent the indigenous population of the Altai Republic. Linguistically, they belong to the Turkic group of the Altaic linguistic family, and are subdivided by the anthropologists and ethnographers into two isolated groups of Southern and Northern Altaians. Altaians were relatively recently consolidated into one ethnic group. Earlier, they belonged to different tribal and territorial subdivisions. The division of Altaians into Northern and Southern groups was associated with remarkable differences in their ontogeny, as well as with their development at the bases of compositionally different ancestral components. The differences between the groups can be traced at the historical, anthropologic, and linguistic levels [3, 4].

According to the classification of Turkic languages, Southern Altaians were attributed to the Kypchak group. A part of Northern Altaic dialects belongs to the Uyghuric (northeastern) language group [4, 5]. Northern Altaians (including three ethnic groups of Kumandins, Chelkans, and Tubalars) were formed as a result of the interactions between ancient Samoyedic, Ket, Ugric, and Turkic tribes. Southern Altaians (Altai-Kizhi, Telengits, and Teleses) were formed as a result of admixture between Turkic and Mongolian tribes. In respect of anthropology, Southern Altaians belong to South Siberian and Central Siberian Mongoloid types. At the same time, in Northern Altains, Uralian anthropologic type prevails [2, 6, 7].

Despite of intense migrations, structural subdivision of the Altaic ethnic population underwent through no dramatic changes during the last century. Analysis of

the 1897 census showed the isolation of Southern and Northern Altaians, which was manifested not only in the names of seoks (tribes), but also in the tribe settling patterns [6]. Analysis of the tribal composition of Altaians showed that during the last century, their tribal structure at the level of ethnic groups remained unchanged [8, 9]. Genetic demographic analysis of rural populations of the Altai Republic by the surname and tribal structure suggested that the isolation of the two groups of Altaians have been preserved till present [10]. These results, of course, implied the existence of the genetic differentiation of the two groups of Altaians.

The studies of biochemical markers in the indigenous groups of South Siberia, including Altaians, were rather extensive. Analysis of the frequency distributions of the phenotypes and genes of the blood group system, serum proteins, and erythrocytic enzymes revealed the clear subdivision of Altaians into southern and northern groups, as well as the presence of local subgroups within ethnic groups. Taken together, most of the polymorphic systems in Altaians were characterized by the frequencies typical of the Siberian and Central Asian Mongoloids. At the same time, relative to some of the genes examined, Altaians occupied the intermediate position between Caucasoids and Mongoloids, pointing to the presence of ancient Caucasoid component, mostly expressed in Northern Altaians [11–16].

The uniparentally inherited genome components, mitochondrial DNA (mtDNA) and Y chromosome, occupy a special place among the genetic marker systems. Principal importance of the DNA markers of specific human genome regions lies in the fact that they provide individual analysis of maternal and paternal components.

In recent years, mtDNA diversity was extensively studied in the ethnic groups from South Siberia [17–21]. It was demonstrated that mitochondrial gene pools of the populations studied were characterized by different ratios between the Mongoloid and Caucasoid mtDNA lineages. The frequency of Caucasoid lineages declined in the direction from the south to the north, and from the west to the east, which was generally consistent with the anthropological data [18, 19]. Most of these lineages have south Caucasoid origin, while the others belong to east Caucasoid lineages [22].

Studies of the indigenous populations of South Siberia by use Y-chromosomal DNA markers started only recently. Analysis of Y-linked microsatellite markers showed the presence of several major components in the male gene pool of Tuvinians, which were categorized as the Caucasoid and Mongoloid components [23, 24]. Studies of a number of ethnic groups from Altai–Sayan region (including Altaians) also showed the presence of profound paleo-Caucasoid component in their gene pools [25, 26]. These data in general are consistent with those obtained upon the analysis of mtDNA.

The comprehensive studies performed with the use of biallelic Y-chromosomal markers showed that indigenous populations of South Siberia, compared to other indigenous populations of Siberia, possessed the most variable haplogroup composition. Furthermore, Altaians were characterized by a high frequency of haplogroup R1a1, which marked the Caucasoid gene pool component [15, 26, 27]. However, the pattern obtained still rather schematically reflected the actual gene pool structure of Altaians. Effective haplogroup specification requires investigations involving additional samples and possibly larger number of genetic markers.

In the present study, an analysis of the composition and structure of Y-chromosomal haplogroups, distinguished based on the genotyping of 37 biallelic and seven microsatellite markers of its non-recombining part, was performed. The data obtained provided characterization of the gene pool structure in Northern and Southern Altaians. This study is a part of the ongoing project on the analysis of Y-chromosomal gene pool of the indigenous populations of Siberia.

MATERIALS AND METHODS

Total DNA for the experiments was isolated from peripheral blood lymphocytes using standard techniques [28]. Population samples consisting of 146 individuals representing the indigenous population of the Altai Republic were examined. The samples were comprised of males unrelated in at least three generations.

Northern Altaians ($N = 50$) were represented by the samples, formed from the inhabitants of the city of Gorno-Altaisk ($N = 20$) and the settlements of Kurmach-Baigol ($N = 11$) and Turochak ($N = 19$), Turochak raion. Southern Alataians ($N = 96$) were represented by the samples formed from the inhabitants of the settlement of Kulada ($N = 46$), Ongudaisk raion; the settlement of Beshpel'tir ($N = 43$), Chemal'sk raion; and the settlement Kosh-Agach ($N = 7$), Kosh-Agachskii raion of the Altai Republic.

The Y-chromosomal haplogroup composition and structure was examined using two genetic marker systems: biallelic loci, mostly represented by SNPs, and polyallelic, highly variable microsatellites (YSTR).

Using biallelic markers, attribution of the samples to certain haplogroup was performed. Nomenclature of haplogroups is as defined by the Y-Chromosome Consortium [29]. Genotyping was performed using a set of microsatellite markers, and for each sample its individual STR haplotype was determined. Based on the data on haplotype composition within the haplogroups, their internal diversity and detailed phylogenetic relationships were established.

Biallelic markers. Variability of 37 biallelic markers from the nonrecombining portion of Y chromosome were investigated, including *SRY1532*, *SRY2627*, *M1* (*YAP*), *92R7*, *DYF155S2*, *12f2*, *M3* (*DYS199*), *M9*, *M15*, *M17*, *M20*, *M25*, *M46* (*Tat*), *M70*, *M77*, *M89*,

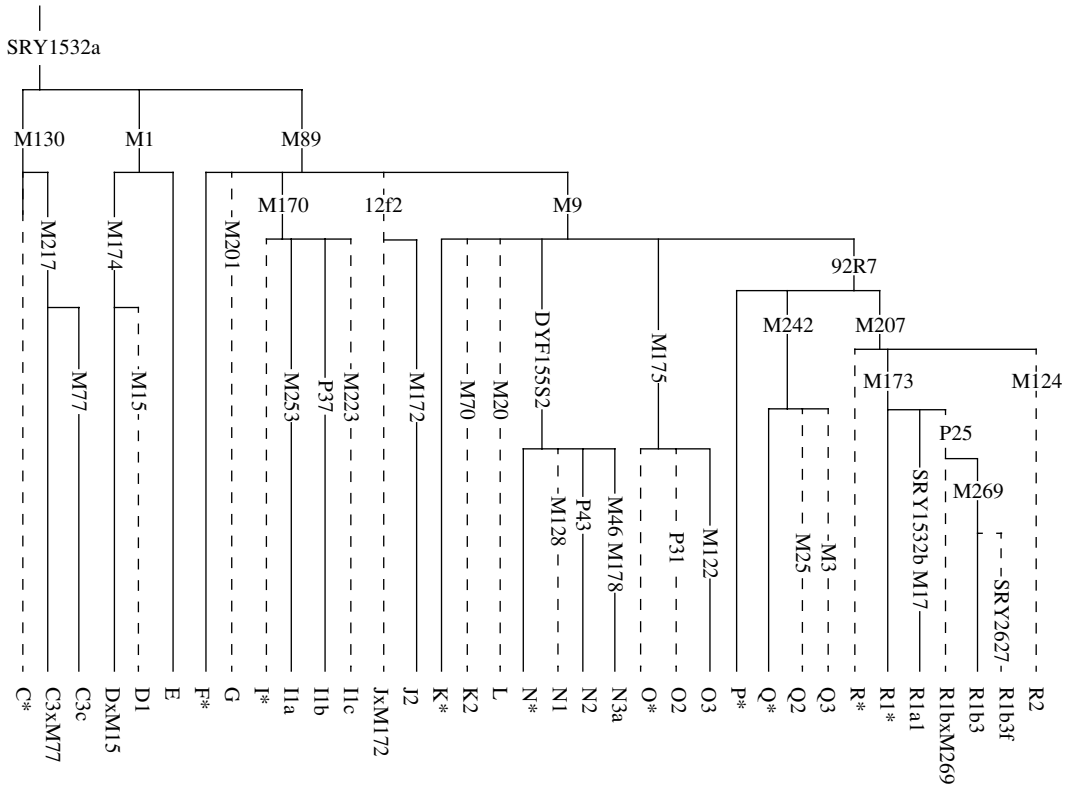

Fig. 1. Phylogenetic tree of Y-chromosomal haplogroups identified in the present study.

M122, *M124*, *M128*, *M130* (*RPS4Y*), *M170*, *M172*, *M173*, *M174*, *M175*, *M178*, *M201*, *M207*, *M217*, *M223*, *M242*, *M253*, *M269*, *P25*, *P31*, *P37*, and *P43* (Fig. 1).

Biallelic markers were genotyped using polymerase chain reaction (PCR) with subsequent analysis of the DNA fragments by means of different methods. Most of the markers were genotyped using RFLP analysis. In this case, the fragment sizes were determined by use of agarose gel electrophoresis, or by direct separation of the products of allele-specific or ordinary PCR, or by sequencing. Most of the primer sequences for biallelic markers, with the exception of the specially mentioned changed variants, were described in the work on the haplogroup system nomenclature [29].

Genotyping of the *YAP*, *92R7*, *DYF155S2*, *12f2*, *Tat*, *M9*, *M17*, *M25*, and *M89* loci was performed as described earlier [30]. Typing of the *SRY2627*, *M124*, *M130*, *M170*, *M172*, *M173*, *M174*, *M178*, *M201*, *M207*, *M242*, *M269*, *P25*, and *P37* loci is described in [31]. Markers *M3* (*DYS199*), *P43*, and *SRY1532* were genotyped as it was suggested earlier [27, 32, 33].

Genotyping of marker *M128* (the 2-bp deletion) was conducted using allele-specific PCR. Each of two reverse primers was strictly complementary to only one of the allelic variants and differed from another primer in its 3'-end structure (M128R1: 5'-CTG TAA ATG AAA ATA ACT GTG AA-3' and M128R2: 5'-CTG TAA ATG AAA ATA ACT GAA A-3'). The 334- or 332-bp fragment was present only in one PCR variant at the annealing temperature of 56°C.

Genotyping of marker *P31* (the C–T transition) was carried out in the same way. Two different forward primers were used: (P31F1: 5'-AAT AAG GTT TTT TTT TGG TTG T-3' and P31F2: 5'-AAT AAG GTT TTT TTT TGG TTG C-3'). The 599-bp fragment was present only in one PCR variant at the annealing temperature of 58°C.

Markers *M70*, *M217*, and *M223* were genotyped, using modified primers, which generated the recognition sites for the appropriate restriction enzymes in case of one of the allelic variants.

Marker *M70*, which was the A–C transversion, was genotyped using PCR with modified forward primer (M70F: 5'-ACT ATA CTT TGG ACT CAT GTC TCC ATG AGG-3'). In these conditions, in case of the allele *C*, there appeared the *Hae*III recognition site, and the 231-bp product was cleaved into the fragments of 201 and 30 bp. Reverse primer was also different from that described in [29] (M70R: 5'-TTT GTC TTG CTG AAA TAT ATT TTA-3').

Genotyping of marker *M217* (the A–C transversion) was conducted using the modified reverse primer (M217R: 5'-TAT GTA TTT TTC CTT CTG AAC AAT T-3'). In case of the allele *C*, the *Mfe*I recognition site was formed, and the 244-bp product was cut into the fragments of 219 and 25 bp.

The modified forward primer for genotyping of marker *M223* (the C–T transition) (M223F: 5'-AGT CTG CAC ATT GAT AAA TTT ACT TAC AAT-3')

generated, in case of allele *T*, the recognition site for the *Mfe*I restriction enzyme, which cleaved the 172-bp amplification product into the fragments of 145 and 27 bp.

Genotyping of markers *M202*, *M77*, *M175*, and *M253* was carried out by use of PCR/RFLP after the literature search for the restriction sites.

The A–G transition (marker *M*20) was identified by the presence of the *Ssp*I restriction site in case of allele *A*. The 413-bp product is cleaved in the fragments of 295 and 118 bp.

Marker *M77* (the C–T transition) was typed by the presence in the PCR product of two *Dra*I restriction sites in case of allele *T*. The initial fragment was cleaved in three parts (243 + 92 + 36 bp). In case of allele *C*, only one restriction site was present (the fragments of 335 and 36 bp).

Marker *M175*, which was the 5-bp deletion, was identified by the *Bst*6I site loss in case of the deleted allele (439 bp), and by the restriction site gain in case of ancestral allele (355 and 89 bp).

Ancestral allele *C* of marker *M253* (the C–T transition) leads to the formation of the *Hin*dII restriction site (fragments of 120 and 180 bp), while the mutant allele is associated with the restriction site loss. The sequences of the primers used are presented in [34].

The restriction enzymes and buffers were produced by SibEnzim (Novosibirsk, Russia) and used in the conditions recommended by the manufacturer. Analysis of amplified DNA fragments was carried out using 2% (for *M20*, *M77*, *M128*, *M175*, *M253*, *SRY1532*, *P31*, and *P43*) or 3% (for *M3*, *M70*, *M223*, and *M217*) agarose gels.

Markers *M15* and *M122* were typed by sequencing of PCR products from the forward primer. PCR product having only one specific band upon analyzing on agarose gel was used as a template. Sequencing primers were the same as used for PCR. Sequencing was performed on the ABI Prism 310 automated sequencer (Perkin-Elmer) using the method of Sanger with BigDye Terminator cycle sequencing kit according to the protocol of the manufacturer. The fragments were separated by means of capillary gel electrophoresis.

Microsatellite markers. STR haplotypes were typed using seven microsatellite markers of nonrecombining portion of Y chromosome (*DYS389I*, *DYS389II*, *DYS390*, *DYS391*, *DYS392*, *DYS393*, and *DYS394* (*DYS19*). One of these markers (*DYS392*) is a trinucleotide repeat, and the others are tetranucleotide repeats. The seven markers used form the so-called minimum haplotype [35]. This YSTR set is the most comprehensively studied in different world populations, and it is the main marker used in population genetic and forensic studies.

Forward primers for each locus were synthesized at the Perkin–Elmer Oligo Factory (Weiterstadt, Germany), and carried the fluorescent label (HEX, FAM, or TET). The markers were genotyped as described earlier [23, 24]. Furthermore, to confirm the coincidence of the labeled fragment sizes observed to the number of tandem repeat components, sequencing of the PCR products of all microsatellite loci with the use of a number of samples analyzed was performed. Sequencing was carried out from the reverse primers carrying no fluorescent label. Allele nomenclature corresponds to the generally accepted (for the *DYS389I* without considering the three-copied TCTG repeat) [36, 37].

Statistical methods. Genetic relationships between the populations were examined using factor analysis and multidimensional scaling. To identify the integral characteristics (factors) determining the population gene frequency variation, the method of principle components was used [38]. Diagrams were constructed and analyzed using the STATISTICA 6.0 (StatSoft, United States) software package. Population genetic diversity, equivalent to the expected heterozygosity for diploid data was calculated as in Nei [39]. Population genetic differentiation was assayed using analysis of molecular variance (AMOVA) [40]. The F_{st} index was used for biallelic haplogroups, and R_{st} index was used with YSTR haplotypes; a total of 10 000 permutations were performed. Statistical significance of the differences between populations in haplogroup and YSTR haplotype frequencies was evaluated using the exact test for population differentiation. Slatkin's matrices of pairwise genetic distances (F_{st} for biallelic haplogroups and R_{st} for YSTR haplotpes) were calculated using 100 permutation steps as implicated in the ARLEQUIN 2.000 software package (http://anthro.unige.ch/arlequin) [41]. Microsatellite tandem repeat number variance was assessed using the MICROSAT software program (http://hpgl.staford.edu/projects/microsat) [42]. Median networks for Y-chromosomal haplotypes were constructed using the Network v. 4.1.1.1 program (Fluxus Technology, www.fluxus-engineering.com) and the median networks method of Bandelt with MJ (median-joining) [43] and RM (reduced median) algorithms (the ε estimator was considered as equal to 0). The principle of maximum parsimony of the mutation event number and tree topology was used. To score different mutation rates upon the networks construction, each STR locus was weighted in accordance with the calculated values (*DYS393* : : *DYS392* : *DYS19* : *DYS389I* : *DYS389II* : *DYS391* : : *DYS390* = 10 : 10 : 5 : 5 : 2 : 2 : 1) [45]. Based on pairwise distances, genetic distance matrices between the populations were constructed. Population phylogenetic trees were constructed by use of neighbor joining algorithm [46] realized in the PHYLIP software package [47].

RESULTS AND DISCUSSION

Frequencies of Y-chromosomal Haplogroups in Northern and Southern Altaians

Analysis of the allele distributions for the marker loci examined in 146 individuals representing the samples of Northern (Kumandins, Chelkans, and Tubalars)

Table 1. Distribution of Y-chromosomal haplogroups in Altaians

Haplogroup	Frequency, %					
	Northern			Southern		
	Gorno-Altaisk ($N = 20$)	Kurmach-Baigol ($N = 11$)	Turochak ($N = 19$)	Beshpel'tir ($N = 43$)	Kulada ($N = 46$)	Kosh-Agach ($N = 7$)
C3xM77	–	–	–	–	2.17 (1)	14.28 (1)
C3c	–	–	–	–	2.17 (1)	–
D × M15	–	–	–	–	10.87 (5)	14.28 (1)
E	–	–	–	2.33 (1)	–	–
F*	–	–	–	6.98 (3)	2.17 (1)	–
I1a	–	–	–	2.33 (1)	–	–
I1b	–	–	–	2.33 (1)	–	–
J2	5.00 (1)	–	–	4.65 (2)	4.35 (2)	–
K*	–	18.18 (2)	10.53 (2)	–	2.17 (1)	–
N*	–	–	–	–	10.87 (5)	–
N2	10.00 (2)	–	–	4.65 (2)	–	28.57 (2)
N3a	–	–	15.79 (3)	2.33 (1)	2.17 (1)	–
O3	–	–	10.53 (2)	9.30 (4)	4.35 (2)	14.28 (1)
P*	–	–	–	2.33 (1)	–	–
Q*	20.00 (4)	63.64 (7)	26.32 (5)	4.65 (2)	4.35 (2)	–
R1*	15.00 (3)	–	–	–	–	–
R1a1	50.00 (10)	18.18 (2)	36.84 (7)	58.14 (25)	52.17 (24)	28.57 (2)
R1b3	–	–	–	–	2.17 (1)	–
H	0.7105 ± 0.0848	0.5818 ± 0.1420 0.7510 ± 0.0406	0.7895 ± 0.0573	0.6545 ± 0.0801	0.7111 ± 0.0682 0.6941 ± 0.0518	0.9048 ± 0.1033
H YSTR	0.9368 ± 0.0354	0.8909 ± 0.0918 0.9281 ± 0.0290	0.9542 ± 0.0301	0.9415 ± 0.0301	0.9691 ± 0.0125 0.9554 ± 0.0123	1.0000 ± 0.0764

Note: *H*, haplogroup-based gene diversity value; *H* YSTR, microsatellite haplotype-based gene diversity value; *N*, sample size.

and Southern (Altai-Kizhi) Altaians showed the presence of 18 haplogroups (Fig. 1, Table 1). The gene pools of the two population samples examined displayed substantial differences relative to the lineages of Y chromosome. In Northern Altaians, only eight haplogroups were identified, while Southern Altaians showed the presence of 17 Y-chromosomal haplogroups. Northern Altaians completely lacked lineages C, D, E, F, I1a, I1b, N*, P*, and R1b3, which accounted for 24% of Y chromosomes in Southern Altaians. On the other hand, Southern Altaians demonstrated the absence of lineage R1*, found in Northern Alataians.

The most frequent variant of Y chromosome among both ethnic groups was haplogroup R1a1, which was the prevailing among Southern Altaians with the total frequency of 53%. This haplogroup is widely distributed on the territory of Eurasia. It is found in the populations of Central and Eastern Europe, Iran, Pakistan, Central Asia, and India [48]. By contrast, in Eastern Asia haplogroup R1a1 is very rare. The appearance of haplogroup R1a1 in South Siberian ethnic groups is thought to be associated with the early stages of their ethnogeny, beginning from the Andronovo cultural association, as a result of early Indo-European movements to the east, and marks the substrate Caucasoid component of their gene pools. Surprisingly, the proportion of this component is higher among Southern Alatians, characterized by more expressed Mongoloid characters. This finding can be, probably, explained in terms of the genetic drift, which substantially affects Y-chromosomal markers due to the low effective population number of Y chromosomes.

Concerning other Western European (Caucasoid) haplogroups, it should be mentioned that in Southern Alatians single cases of haplogroups E, J2, I1a, I1b, and R1b3 were detected. It can be thus concluded that Caucasoid component in Southern Altaians in rather variable. On the contrary, in Northern Altaians this component is represented only by haplogroups R1a1 and J2. Haplogroups E and J2 are thought to be associated with the dispersal of the Middle Eastern farmers during the Neolithic [49–51]. The presence of these lineages on the territory of South Siberia is probably associated with the migratory flows from the Central Asia. The presence of single chromosomes from the Caucasoid haplogroups I1a and I1b, for which it is known that dispersal of the carriers and formation of contemporary distribution area took place during recolonization of Europe from Iberian and Balkan refugees after the Last Glacial Maximum [52] in Altaians, can have different

explanations. On the one hand, they can reflect the recent metisation with Russian immigrants to Siberia. On the other hand, it seems likely that these lineages could appear here as early as in the Neolithic, together with R1a1. The possible confirmation of this idea is the absence of haplogroup R1* (yet undifferentiated lineage R1), found in Northern Altaians, in modern European populations.

Northern and Southern Altaians demonstrate substantial differences relative to the haplogroup Q* frequencies. In the gene pool of Northern Altaians, it is the second most frequent haplogroup after R1a1 (22%); in Southern Altaians, the proportion of this haplogroup is significantly lower (4%). In earlier studies, haplogroup Q* with different frequencies was described in most of Siberian ethnic groups. With low frequency (1 to 17%; with the average of 5%) this haplogroup is found in Middle and Central Asia, India, and Siberia [27, 53]. High frequencies of haplogroup Q* are detected among Native Americans (up to 25%), where it is the second most frequent haplogroup after its derivative, haplogroup Q3 [54]. In Siberia, this haplogroup accounts for the lion's share (66%) of the Selkup gene pool [27], while the highest Q* frequency was detected in Kets (94%). In other parts of the region the Q* frequency varies in the range from 5 to 15%. It is suggested that peak frequency of haplogroup Q* in the region of contemporary residence of Kets and Selkups does not point to this region as to the center of origin of this haplogroup. It is known that presumptive motherland of the ethnic groups mentioned above is located in South Siberia, where their ancestors are thought to have penetrated from Central Asia [55, 56]. These data are consistent with the information on the significant participation of the Samoyedic ethnic groups in the ethnogeny of Northern Altaians.

In Northern and Southern Altaians, the frequencies of haplogroups belonging to clade N are also different. However, the total proportion of these haplogroups in both groups is nearly equal (about 10%). It should be noted, that haplogroups N3a and N2 cannot be unambiguously attributed to either West- or East Eurasian. High frequencies of haplogroup N3a is typical of East Siberian ethnic groups (Yakuts, Buryats, and Evenks). Moreover, in Yakut population, the frequency of this haplogroup constitutes almost 90%. Based on the high N3a frequency in the populations of Siberia, Zerjal et al. [57], who were the first to describe polymorphism of the *Tat* locus, suggested that the origin of this haplogroup was associated with Siberia, and its presence in European populations could be explained by westward migrations of its carriers. However, further investigations demonstrated that diversity of the N3a YSTR haplotypes in Finno–Ugric populations of the Volga region was remarkably higher compared to the populations of Northern Europe and the ethnic groups of Siberia [58, 59]. From here it follows that the most probable place of origin of haplogroup N3a is Eastern Europe, and the migratory roots for the dispersal of this haplogroup are determined as from the west to the east. Higher frequency of N3a observed in some populations of Siberia is the result of genetic drift.

High frequency of haplogroup N2 was observed in Tuvinians and Khants (personal unpublished results). According to the literature data, among the ethnic populations of Siberia, lion share of N2 was typical of Nganasans (92%), Nentsy (57%), and Dolgans (12%) [27]. In the other ethnic groups, the frequency of this lineage is marginal. The second N2 distribution area falls at the Volga–Ural region, where it accounts for 10% of Y-chromosomes in Maris and Chuvashes, 13%, in Komi, and 29%, in Udmurts [60]. Sporadic cases of N2 were recorded among Russians and in Fennoscandia. It can be thus suggested that N2 represents the marker of prehistorical relations between the Siberian and pro-Uralic populations. The recently described mutation *M214* (not genotyped in the present study) was shown to be ancestral relative to calde N, as well as to calde O, typical of Mongoloid populations from the Pacific [61]. The problems of how ancient is this link, where the lineages have diverged, and how the carriers of the undifferentiated N/O appeared in Eastern Europe, the place of origin of N3a, and probably, of N2, remain unclear. The lineage, defined in the present study as K* and accounting for about 8% of the total gene pool of Northern Altaians, is thought to include the undifferentiated group N/O, marked by mutation *M214*.

Haplogroups C3xM77 and C3c, present in Southern Altaians, are among the most frequent in Central Asia [62], and East Siberia [32, 63, 64]. Furthermore, the lion share of calde C from the continental Southeast Asia belongs to this haplogroup. In Siberia, the highest frequency of haplogroup C was observed among Buryats and Oroches. The haplogroup is also found among Tungus–Manchurian populations, namely, Evenks, Evens, and Manchurs. Haplogroups C3xM77 and C3c are thus the markers of the genetic contribution of Mongoloid groups, probably, reflecting the latest contribution of Mongolian immigrants from the territories of contemporary Tuva and Mongolia, beginning from the Hunnu time.

Southern Altians are also characterized by a high frequency of haplogroup O3. This lineage is mostly distributed across Southeast Asia and Oceania, where its carriers penetrated from the continent, and also along the whole pacific coastline of Eurasia. Clade O is generally the most abundant variant of Y chromosome, accounting for almost 80% of all lineages. The highest O3 frequencies were detected among the Chinese from the southern provinces (58%), in Chinese from Taiwan (58%), in Vietnam (45%), Philippines (35%), and Malaysia (up to 30%) [45, 65–67]. In most of the indigenous populations of Siberia the haplogroup is not detected, or its frequency is marginal. The O3 geographic distribution pattern makes it possible to consider it as marking the contribution of the Mongoloid groups to the gene pool of Siberian ethnic groups.

Analysis of the haplogroup and YSTR haplotype distributions showed heterogeneity of the populations examined relative to the level of genetic diversity of their male gene pool (Table 1). The haplogroup genetic diversity in Northern Altaians is several times higher than that in Southern Altaians. This finding can be explained by the fact that the frequency of one lineage (R1a1) in Southern Altaians constitutes more than 50%. The microsatellite haplogroup genetic diversity is somewhat higher in Southern Altaians.

Genetic Relationships between the Populations

Analysis of the data set over the haplogroup frequencies with the help of the method of principle components identified three principle components (PCs), which explained about 95% of the haplotype frequency variation. The two first PCs in total accounted for about 88% of genetic variation of the ethnic groups examined (Fig. 2). Maximum values of factor loadings associated with the first component of the Altaic gene pool variation fell on the west Eurasian haplogroups E, J2, I1a, I1b, and R1a1. Furthermore, genetic closeness of the populations relative to this component was determined by the presence of either higher or lower proportions of Caucasoid component in their gene pools, as well as by the diversity of the gene pool components, reaching maximum values in Southern Altaians from Beshpel'tir and Kulada.

The second PC explained about 20% of the data variation. This component, the higher values of which correlated with an increase of the summarized proportion of Mongoloid haplogroups (maximum values of factor loadings fell on the haplogroups C3xM77, C3c, D, and N*), reaching the peak values in Southern Altaians, effectively differentiated the samples of Northern and Southern Altaians. The data obtained agreed with the anthropological data on the higher proportion of Mongoloid component in the gene pool of Southern Altaians, compared to Northern Altaians.

Another factor, explaining 7.6% of the haplogroup frequency variation (HFV) was the presence of haplogroups Q* and K* in the population Y-chromosomal gene pool. These haplogroups mark the most ancient genetic substratum, N2 and N3a, most likely associated with the Uraloid component. In this case, the loading falling on haplogroup N2 was negative. Thus, the main factors determining the gene pool structure of Altaians are the presence of Caucasoid and Mongoloid components along with the share of Paleolithic/Uraloid genetic substratum.

The multidimensional scaling diagrams of Slatkin's R_{st} distances obtained for both haplogroup and YSTR haplotype frequencies in the samples tested were nearly identical and clearly differentiated the populations of Southern and Northern Altaians. For this reason, only a single diagram, illustrating one of the experiment variants is presented in the figure (Fig. 3). Phylogenetic

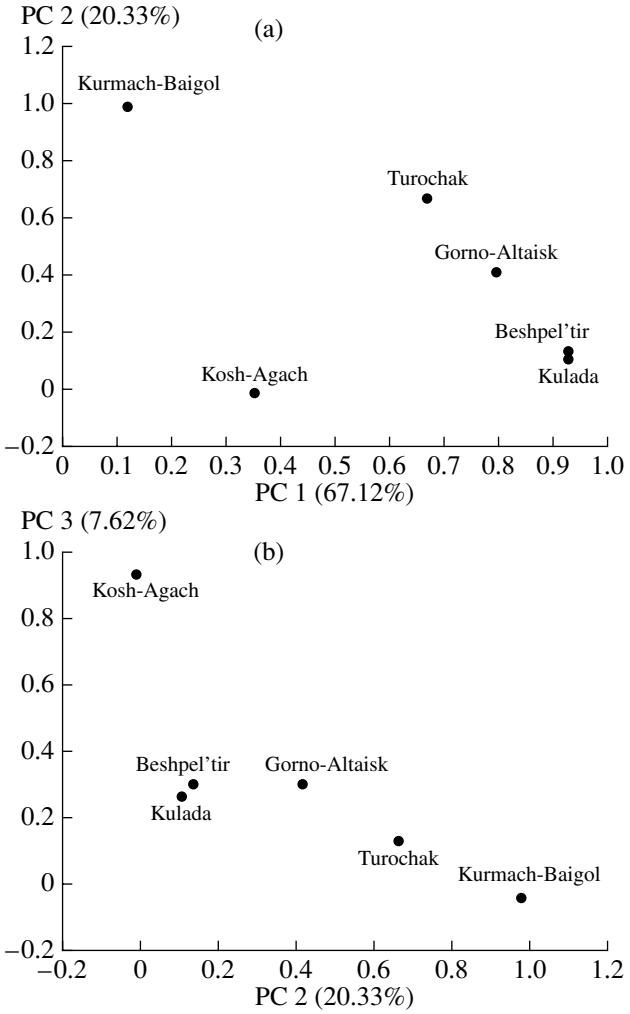

Fig. 2. Position of Altaians in the space of three principle components based on Y-chromosomal haplogroup frequencies. (a) First and second principle components; (b) second and third principle components.

cluster analysis revealed the similar structure of the genetic relationships between the data of populations (Fig. 4). Similarly to the multidimensional scaling data, the samples of Northern and Southern Altaians formed two individual clusters. The only difference between the methods used was that the closest samples on the phylogenetic population tree constructed based on the haplogroup frequencies were those from Kulada and Beshpel'tir, while on the tree based on the YSTR haplotype data, the closest samples were those from Kulada and Kosh–Agach.

Population Genetic Differentiation

Estimates of the significance of among-population differences performed by use of the exact test for the population differentiation, revealed statistically significant differences between the population samples in haplogroup frequencies. Comparisons of Northern

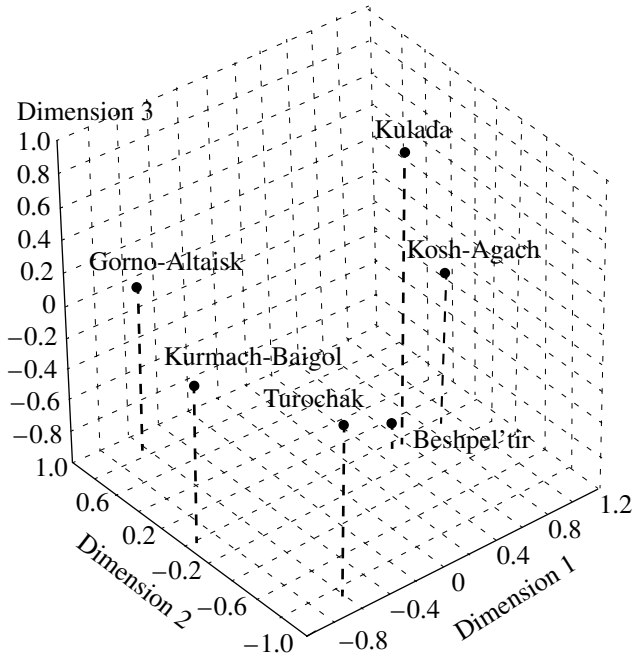

Fig. 3. Multidimensional scaling diagram of Slatkin's R_{st} distances between the populations of Altaians based on the YSTR haplotype frequencies.

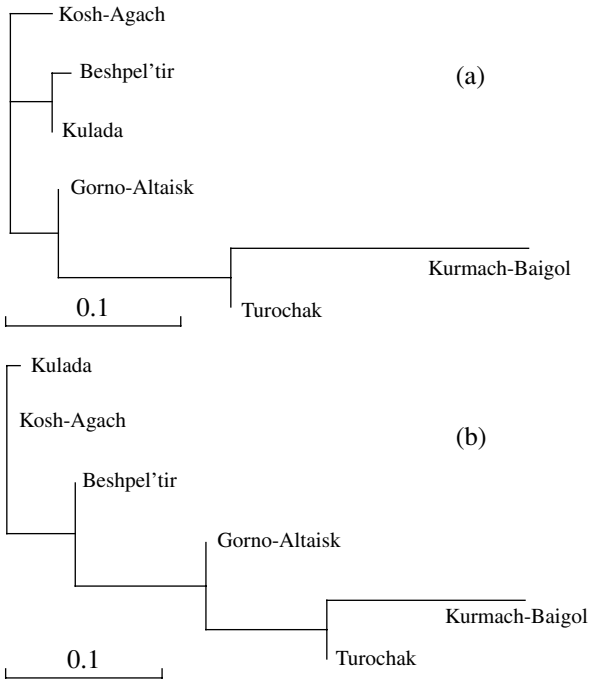

Fig. 4. Dendrogram of genetic relationships between the populations of Altaians based on the genetic distances between (a) the haplogroups and (b) YSTR haplotypes.

Altaic samples revealed no statistically significant differences for the pair of Kurmach–Baigol/Turochak. The sample from Gorno-Altaisk also was not statistically significantly different from the samples from Beshpel'tir and Kosh-Agach. In haplogroup frequencies, all samples of South Altaians were not statistically significantly different from one another. Analogous comparison of all samples relative to the YSTR haplotypes, revealed no statistically significant differences only for the pair of Kyrmach-Baigol/Kosh-Agach, which were the smallest in number.

Genetic differentiation of the samples examined was investigated by analysis of molecular variance (AMOVA) [40], using two marker systems. First, the population differentiation was tested relative to the haplogroup frequencies, and second, this estimate was performed considering molecular diversity of Y-chromosomal microsatellite haplotypes. The AMOVA procedure is based on the analysis of the population genetic structure by means of the analysis of variance [68–70]. AMOVA makes it possible to score the number of mutations between the haplotypes, providing more precise estimation of the repeat number variance. Hierarchical analysis of the population genetic structure consists in the identification of the variance share due to within- and between-group differences at different hierarchical levels. Calculations were made pooling the samples into two groups, Northern and Southern Altaians.

The results of calculations made using two marker systems were noticeably different. Analysis of the haplogroup frequencies showed that the proportion of among-group differences was 5.02%, while the differences between the populations within the groups accounted for 3.72%, and the proportion of total within-population variation was 91.26%. Analogous calculations made using YSTR haplotype data showed the higher share of the among-group differences (8.97%), lower proportion of among-population differences within the groups (1.54%), and the proportion of within-population variation of 89.49%. Thus, analysis of the gene pool structure at the level of Y-chromosomal haplotypes more clearly identified the differentiation of Northern and Southern Altaians and showed lower among-population differences within the groups, compared to the analysis at the level of haplogroup frequencies.

Phylogenetic Analysis of Haplogroup R1a1 in Altaians

Haplogroup R1a1 accounts for most substantial part of Y-chromosomal gene pool in Alatians. Detailed structure of this haplogroup was established through the construction of a phylogenetic tree using the haplotype median network method [43, 44]. Figure 5 shows the phylogenetic tree for the microsatellite haplotypes of haplogroup R1a1. The size of a circle (the tree nod) corresponds to the number of the identified samples, belonging to a haplotype of interest, while the length of a link between the nods corresponds to the number of the mutation steps between the haplotypes. Mutated

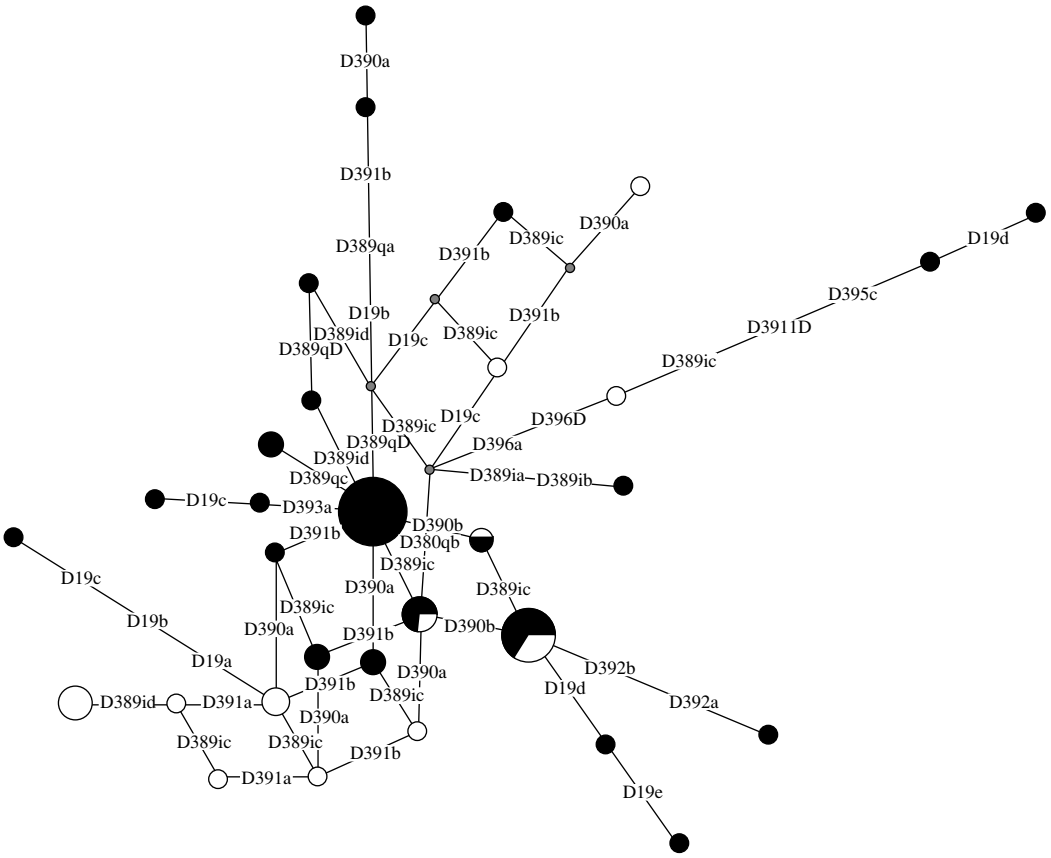

Fig. 5. Phylogenetic tree (median network) of the YSTR haplotypes of haplogroup R1a1 in Altaians. The size of a circle (the tree nod) corresponds to the number of the identified samples, belonging to a haplotype of interest, while the nod color indicates ethnic affiliation of the individual, to whom the sample belongs; Southern Altaians are indicated by black color, Northern Altaians are indicated by white color. The tree branching points not represented by the haplotypes discovered (median vectors) are designated by grey color.

loci are listed along links. The nod color indicates ethnic affiliation of the individual, to whom the sample belongs. The branching points not represented by the haplotypes discovered (median vectors) are designated by small grey circles.

Northern Altaians are characterized by higher genetic diversity and almost identical repeat number variance compared to Southern Altaians (Table 2). However, the mean number of alleles per locus and the repeat number range was remarkably higher in Southern Altaians. It should be noted in this respect that for the STR loci such parameters as variance and the repeat number range, characterizing molecular locus variation, are the more adequate diversity estimates, than gene diversity, based on the allele frequencies. This situation can be explained by the fact that analysis of polyallelic marker systems highly probably can result in the obtaining of the samples biased relative the total population. In addition, calculations are made using the stepwise mutation model, which most adequately corresponds to the mechanism of the haplotypes evolution.

Only three R1a1 haplotypes are common for the two population groups, pointing to the differences between their gene pools not only in the haplogroup composition but also in their internal haplotype structure. In Southern Altaians 15 samples out of 51, carry the most frequent haplotype, while among Northern Altaians, this haplotype was not detected. It is suggested that Southern Alataians demonstrate the founder effect relative to this haplogroup, since their median network is characterized by typical

Table 2. Diversity indices for the haplogroup R1a1 in Altaians

Ethnic group	*H*	Avg Var	Avg All	Avg Ran
Southern Altaians	0.8955	0.292	3.286	2.714
Northern Altaians	0.9560	0.294	2.286	1.286

Note: *H*, gene diversity; Avg Var, averaged repeat number variance; Avg All, averaged number of alleles per locus; Avg Ran, range of the repeat number.

star-like phylogenetic structure with the prevailing frequency of the founder haplotype, with the allelic structure of 16–11–18-25–11–11–13 (*DYS19-DYS389I-DYS389II-DYS390-DYS391-DYS392-DYS393*). Eight haplotypes of those discovered in Southern Alatians stand only one mutation step apart from the founder haplotype. This finding explains the illusive contradiction of the presence of a large proportion of the "Caucasoid" haplogroup R1a1 in Southern Altaians, which according to anthropological classification belong to South Siberian and Central Asian Mongoloid types. There still remains the question of whether haplotype structure of haplogroup R1a1 revealed is typical only for the samples examined, or it is typical to Southern Altaians, or even for the whole South Siberian pool of R1a1. Our data on the R1a1 microsatellite haplotype diversity in Tuvinian populations did not show the high presence of South Altaic founder haplotype (only 1 out 33 Tuvinian samples belonged to that haplogroup) (personal unpublished results). Among ten South Siberian ethnic groups examined in [26], the haplotype of interest was fopund only in Southern Altaians, Teleuts, and Tuvinians. Moreover, in Altaians it also accounted for 30% of the samples belonging to R1a1. The sample examined in the study cited was represented by the inhabitants of Ust'-Kansk raion (the westernmost region of the republic), which was rather geographically distant from the regions examined in the present study. These data allow conclusion that specific structure of haplogropup R1a1 with the prevailing frequency of one founder haplotype, described in the present study, is typical to the whole ethnic group of Southern Altaians, but not to the other South Siberian ethnic groups.

Thus, in the present study a detailed analysis of the gene pools of Northern and Southern Altaians was performed using two systems of the genetic markers on nonrecombining portion of Y chromosome. The results obtained pointed to the multicomponent gene pool of the Altai population. It was demonstrated that Northern and Southern Altaians were characterized by specific haplogroup composition and frequencies, which reflected the contributions of the genetic components of different origin, associated with different migratory flows. The results of the analysis of variance pointed to the higher proportion of the among-group differences compared to the among-population differences within the groups. These results also suggest significant genetic differentiation of the indigenous population of Altai into two population groups, including Northern and Southern Altaians, respectively. Factor analysis, as well as cluster and phylogenetic analysis, and multidimensional scaling also suggested subdivision of the Altaic gene pool. Using haplogroup R1a1 as an example, the differences between Northern and Southern Altaians were estimated, which concerned the haplogroups composition and their internal haplotype structure.

ACKNOWLEDGMENT

This work was supported by the Russian Foundation for Basic Research (grant no. 03-04-49021, V.S.), the Presidential Program of the Russian Federation (grant no. IA-88.2003.04, V.S.), and by the Russian State Program "Investigations and Developments in the Priority Directions of the Development of Science and Technology" (grant nos. 02.442.11.7010, V.S.; 02.445.11.7001, V.P.; and 02.442.11.7074, to V.Kh.).

REFERENCES

1. Kiselev, S.V., *Istoriya Yuzhnoi Sibiri* (History of South Siberia), Moscow: Akad. Nauk SSSR, 1951, pp. 43–48.
2. Alekseev, V.P. and Gokhman, I.I., *Antropologiya asiatskoi chasti SSSR* (Anthropology of the Asian Part of the USSR), Moscow: Nauka, 1984.
3. Potapov, L.P., *Ocherki po istorii altaitsev* (Essays on the History of Altaians), Moscow: Akad. Nauk SSSR, 1953.
4. Potapov, L.P., *Etnicheskii sostav i proiskhozhdenie altaitsev* (Ethnic Composition and Origin of Altaians), Leningrad: Nauka, 1969.
5. Dolgikh, B.O., *Rodovoi i plemennoi sostav narodov Sibiri v XVII v* (Clan and Tribal Composition of Siberian Population in the 17th Century), Moscow, 1960.
6. Alekseev, V.P., Paleoanthropology of Altai–Sayan Upland in Neolith Epoch and in Bronze, in *Antropologicheskii sbornik, III* (Antropological Collection, III), Proceedings of Institute of Ethnography, 1961, vol. 71, pp. 107–206.
7. Katsyuba, D.V. and Nikolaev, R.V., *Etnografiya narodov Sibiri* (Ethnography of the Peoples of Siberia), Kemerovo, 1994.
8. Lotosh, E.A., Kolbasko, A.V., Dranishnikov, A.K., et al., Population, Medical, and Genetic Characteristic of Indigenous People of Altai Mountains, *Vestn. Akad. Med. Nauk SSSR*, 1984, no. 7, pp. 78–81.
9. Luzina, F.A., Hereditary Polymorphism and Genetic Processes among Indigenous Population of Altai Mountains, *Extended Abstract of Cand. Sci. (Biol.) Dissertation*, Moscow State Univ., Moscow, 1987.
10. Kucher, A.N., Tadinova, V.N., and Puzyrev, V.P., Genetic and Demographic Characteristic of Rural Altai Republic Population: Sex and Age Constitution, Surname and Tribal Structure, *Russ. J. Genet.,* 2005, vol. 41, no. 2, pp. 189–194.
11. Rychkov, Yu.G., Serologic Differentiation in Population of Siberia, *Vopr. Antropol.*, 1965, no. 21, p. 18.
12. Sukernik, R.I. and Osipova, L.P., Distribution of Hereditary Variants of Haptoglobin and Transferrin among Some Human Populations in Siberia, *Genetika* (Moscow), 1976, vol. 12, no. 9, p. 139.
13. Sukernik, R.I., Karafet, T.N., Abanina, T.A., et al., Genetic Structure of Two Isolated Ethnic Populations in Siberia (North Altaians) Inferred from Results of Blood Groups and Isozymes Study, *Genetika* (Moscow), 1977, vol. 13, p. 911.
14. Posukh, O.L., Osipova, L.P., Kashinskaya, Yu.O., et al., A Genetic Study of the South Altaian Population of the Mendur-Sokkon Village, Altai Republic, *Russ. J. Genet.,* 1998, vol. 34, no. 1, pp. 92–98.

15. Stepanov, V.A., *Etnogenomika naseleniya Severnoi Evrazii* (Population Ethnogenomics of Northern Eurasia), Tomsk: Pechatnaya Manufaktura, 2002.

16. Khitrinskaya, I.Yu., Genetic Diversity in Ethnic Populations in Siberia and Central Asia Relative to Polymorphic Alu-Insertions, *Extended Abstract of Cand. Sci. (Biol.) Dissertation*, Tomsk, 2003.

17. Salyukov, V.B., Puzyrev, V.P., Golubenko, M.V., and Kutmin, A.I., Polymorphism of Noncoding Regions of the Mitochondrial Genome in the Indigenous Population of Southeastern Tuva Republic, *Russ. J. Genet.,* 1998, vol. 34, no. 1, pp. 1205–1208.

18. Golubenko, M.V., Eremina, E.R., Tadinova, V.N., et al., Territorial Gene Pool Differentiation among Population of Siberia and Central Asia Relative to Restriction Polymorphism of Mitochondrial DNA, *Med. Genet.*, 2002, vol. 1, no. 3, pp. 124–128.

19. Derenko, M.V., Denisova, G.A., Malyarchuk, B.A., et al., The Structure of the Gene Pools Ethnic Population of Altai–Sayan Based on Mitochondrial DNA Polymorphism Data, *Russ. J. Genet.*, 2001, vol. 37, no. 10, pp. 1177–1184.

20. Derenko, M.V., Grzybowski, T., Malyarchuk, B.A., et al., Diversity of Mitochondrial DNA Lineages in South Siberia, *Ann. Hum. Genet.,* 2003, vol. 67, pp. 391–411.

21. Zakharov, I.A., Derenko, M.V., Malyarchuk, B.A., et al., Mitochondrial DNA Variation in the Aboriginal Populations of the Altai–Baikal Region, *Ann. New York Acad. Sci.*, 2004, vol. 1011, pp. 21–35.

22. Derenko, M.V., Malyarchuk, B.A., and Zakharov, I.A., Origin of Caucasoid-Specific Mitochondrial DNA Lineages in the Ethnic Populations of the Altai–Sayan Region, *Russ. J. Genet.*, 2002, vol. 38, no. 9, pp. 1098–1103.

23. Stepanov, V.A. and Puzyrev, V.P., Analysis of the Allele Frequencies of Seven Y-Chromosome Microsatellite Loci in Three Tuvinian Populations, *Russ. J. Genet.,* 2000, vol. 36, no. 2, pp. 179–185.

24. Stepanov, V.A. and Puzyrev, V.P., Y-Chromosome Microsatellite Haplotypes Demonstrate Absence of Subdivision and Presence of Several Components in the Tuvinian Male Gene Pool, *Russ. J. Genet.*, 2000, vol. 36, no. 3, pp. 298–304.

25. Derenko, M.V., Malyarchuk, B.A., Denisova, G.A., et al., Polymorphism of the Y-Chromosome Diallelic Loci in Ethnic Populations of the Altai–Sayan Region, *Russ. J. Genet.*, 2002, vol. 38, no. 3, pp. 309–314.

26. Derenko, M.V., Malyarchuk, B.A., Denisova, G.A., et al., Contrasting Patterns of Y-Chromosome Variation in South Siberian Populations from Baikal and Altai–Sayan Regions, *Hum. Genet.*, 2006, vol. 118, pp. 591–604.

27. Karafet, T.M., Osipova, L.P., Gubina, M.A., et al., High Levels of Y-Chromosome Differentiation among Native Siberian Populations and the Genetic Signature of a Boreal Hunter-Gatherer Way of Life, *Hum. Biol.*, 2002, vol. 74, no. 6, pp. 761–789.

28. Johns, M.B. and Pauls-Thomas, J.E., Purification of Human Genomic DNA from Whole Blood Using Sodium Perchlorate in Place of Phenol, *Anal. Biochem.*, 1989, vol. 180, pp. 276–278.

29. The Y-Chromosome Consortium: A Nomenclature System for the Tree of Human Y-Chromosmal Binary Haplogroups, *Genome Res.*, 2002, vol. 12, pp. 339–348.

30. Kharkov, V.N, Stepanov, V.A., Borinskaya, S.A., et al., Gene Pool Structure of Eastern Ukrainians as Inferred from the Y-Chromosome Haplogroups, *Russ. J. Genet.*, 2004, vol. 40, no. 3, pp. 326–333.

31. Kharkov, V.N., Stepanov, V.A., Feshchenko, S.P., et al., Frequensies of Y Chromosome Binary Haplogroups in Belarussians, *Russ. J. Genet.*, 2005, vol. 41, no. 8, pp. 928–931.

32. Lell, J.T., Brown, M.D., Schurr, T.G., et al., Y Chromosome Polymorphism in Native American and Siberian Populations: Identification of Native American Y Chromosome Haplotypes, *Hum. Genet.*, 1997, vol. 100, pp. 536–543.

33. Kwok, C., Tyler-Smith, C., Mendonca, B.B., et al., Mutation Analysis of the 2 kb 5' to SRY in XY Females and XY Intersex Subjects, *J. Med. Genet.*, 1996, vol. 33, pp. 465–468.

34. Cinnioglu C., King R., Kivisild T., et al., Excavating Y-Chromosome Haplotype Strata in Anatolia, *Hum. Genet.*, 2004, vol. 114, pp. 127–148.

35. De Knijf, P., Kayser, M., Caglia, A., et al., Chromosome Y Microsatellites: Population Genetic and Evolutionary Aspects, *Int. J. Legal Med.,* 1997, vol. 110, pp. 134–140.

36. Kayser, M., Krawczak, M., and Excoffier, L., An Extensive Analysis of Y-Chromosomal Microsatellite Haplotypes in Globally Dispersed Human Populations, *Am. J. Hum. Genet.*, 2001, vol. 68, pp. 990–1018.

37. Kayser, M., Kittler, R., Erler, A., et al., A Comprehensive Survey of Human Y-Chromosomal Microsatellites, *Am. J. Hum. Genet.*, 2004, vol. 74, pp. 1183–1197.

38. Kim, J.O. and Mueller, C.W., *Factor Analysis: Statistical Methods and Practical Issues*, Sage University Paper Series on Quantitative Applications in the Social Sciences, no. 07-014, Newbury Park, CA: Sage, 1978.

39. Nei, M., *Molecular Evolutionary Genetics*, New York: Columbia Univ. Press, 1987.

40. Excoffier, L., Smouse, P., and Quattro, J., Analysis of Molecular Variance Inferred from Metric Distances among DNA Haplotypes: Application to Human Mitochondrial DNA Restriction Data, *Genetics*, 1992, vol. 131, pp. 479–491.

41. Schneider, S., Roessli, D., and Excoffier, L., *Arlequin, Version 2000: A Software for Population Genetics Data Analysis*, 2000.

42. Minch, E., Ruiz-Linares, A., and Goldstein, D., *Microsat, Version 1.5d,* Stanford Univ., 1998.

43. Bandelt, H.-J., Forster, P., and Rohl, A., Median-Joining Networks for Inferring Intraspecific Phylogenies, *Mol. Biol. Evol.*, 1999, vol. 16, pp. 37–48.

44. Bandelt, H.-J., Forster, P., Sykes, B.C., and Richards, M.B., Mitochondrial Portraits of Human Populations Using Median Networks, *Genetics*, 1995, vol. 141, pp. 743–753.

45. Kayser, M., Underhill, P., Shen, P., et al., Extreme Reduction in Y-Chromosome, but Not mtDNA, Diversity in Human Populations from West New Guinea, *Am. J. Hum. Genet.*, 2003, vol. 72, pp. 281–302.

46. Saitou, N. and Nei, M., The Neighbour-Joining Method: a New Method for Reconstructing Phylogenetic Trees, *Mol. Biol. Evol.*, 1987, vol. 4, pp. 406–425.
47. Felsenstein, J., *PHYLIP, Version 3.5*, Seattle Univ., 1993.
48. Quintana-Murci, L., Krausz C., Zerjal E., et al., Y-Chromosome Lineages Trace Diffusion of People and Languages in Southwestern Asia, *Am. J. Hum. Genet.*, 2001, vol. 68, pp. 537–542.
49. Rosser, Z.H., Zerjal T., Hurles M.E., 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.
50. Semino, O., Passarino, G., Oefner, P.J., et al., The Genetic Legacy of Paleolithic *Homo sapiens* in Extant Europeans: A Y-Chromosome Perspective, *Science,* 2000, vol. 290, pp. 1155–1159.
51. Semino, O., Magri, M., Benuzzi, G., et al., Origin, Diffusion, and Differentiation of Y-Chromosome Haplogroups E and J: Inferences on the Neolithization of Europe and Later Migratory Events in the Mediterranean Area, *Am. J. Hum. Genet.,* 2004, vol. 74, pp. 1023–1034.
52. 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.
53. Seielstad, M., Yuldasheva, N., Singh, N., et al., A Novel Y-Chromosome Variant Puts an Upper Limit on the Timing of First Entry into the Americas, *Am. J. Hum. Genet.,* 2003, vol. 73, pp. 700–705.
54. Zegura, S.L., Karafet, T.M., Zhivotovsky, L.A., et al., High-Resolution SNPs and Microsatellite Haplotypes Point to a Single, Recent Entry of Native American Y-Chromosomes into the Americas, *Mol. Biol. Evol.*, 2004, vol. 21, pp. 164–175.
55. Popov, A.A., and Dolgikh, B.O., The Kelts, in *The Peoples of Siberia*, Levin, M.G., Potapov, L.P., et al., Eds., Chicago: Univ. of Chicago Press, 1964, pp. 607–619.
56. Prokofieva, E.D., The Selkups, in *The Peoples of Siberia*, Levin, M.G., Potapov, L.P., et al., Eds., Chicago: Univ. of Chicago Press, 1964, pp. 587–606.
57. Zerjal, T., Dashnymaan, B., Pandya, A., et al., Genetic Relationship of Asians and Northern Europeans, Revealed by Y-Chromosomal DNA Analysis, *Am. J. Hum. Genet.*, 1997, vol. 60, pp. 1174–1183.
58. Villems, R., Adojaan, M., Kivisild, M., et al., Reconstruction of Maternal Lineages of Finno-Ugric Speaking People and Some Remarks on their Paternal Inheritance, in *The Roots of Peoples and Languages of Northern Eurasia: I*, Turku: Societas Historiae Fenno-Ugricae, 1998, p. 241.
59. Rootsi, S., Kivisild, T., Tambets, K., et al., On the Phylogeographic Context of Sex-Specific Genetic Markers of Finno-Ugric Populations, in *The Roots of Peoples and Languages of Northern Eurasia: II–III*, Tartu: Univ. Tartu, 2000, pp. 148–164.
60. Tambets, K., Rootsi, S., Kivisild, T., et al., The Western and Eastern Roots of the Saami: the Story of Genetic "Outliers" Told by Mitochondrial DNA and Y-Chromosomes, *Am. J. Hum. Genet.*, 2004, vol. 74, pp. 661–682.
61. Jobling, M.A. and Tyler-Smith, C., The Human Y Chromosome: An Evolutionary Marker Comes of Age, *Nat. Rev. Genet.,* 2003, vol. 4, pp. 598–612.
62. Zerjal, T., Xue, Y., Bertorelle, G., et al., The Genetic Legacy of the Mongols, *Am. J. Hum. Genet.*, 2003, vol. 72, pp. 717–721.
63. Karafet, T.M., Xu, L., Du, R., et al., Paternal Population History of East Asia: Sources, Patterns, and Microevolutionary Processes, *Am. J. Hum. Genet.*, 2001, vol. 69, pp. 615–628.
64. Lell, J.T., Sukernik, R.I., Starikovskaya, Y.B., et al., The Dual Origin and Siberian Affinities of Native American Y Chromosomes, *Am. J. Hum. Genet.*, 2002, vol. 70, pp. 192–206.
65. Su, B., Xiao, J., Underhill, P.A., et al., Y-Chromosome Evidence for a Northward Migration of Modern Humans into Eastern Asia during the Last Ice Age, *Am. J. Hum. Genet.*, 1999, vol. 65, pp. 1718–1724.
66. Kayser, M., Schiefenhovel, W., Underhill, P.A., et al., Independent Histories of Human Y Chromosomes from Melanesia and Australia, *Am. J. Hum. Genet.*, 2001, vol. 68, pp. 173–190.
67. Xue, Y., Zerjal, T., Bao, W., et al., Recent Spread of a Y-Chromosomal Lineage in Northern China and Mongolia, *Am. J. Hum. Genet.*, 2005, vol. 77, pp. 1112–1116.
68. Cockerham, C.C., Analysis of Gene Frequencies, *Genetics*, 1973, vol. 74, pp. 679–700.
69. Cockerham, C.C., Variance of Gene Frequencies, *Evolution*, 1969, vol. 23, pp. 72–84.
70. Weir, B.S. and Cockerham, C.C., Estimating F-Statistics for the Analysis of Population Structure, *Evolution,* 1984, vol. 38, pp. 1358–1370.