ISSN 1022-7954, Russian Journal of Genetics, 2013, Vol. 49, No. 12, pp. 1236–1244. © Pleiades Publishing, Inc., 2013.
Original Russian Text © V.N. Kharkov, K.V. Khamina, O.F. Medvedeva, K.V. Simonova, I.Yu. Khitrinskaya, V.A. Stepanov, 2013, published in Genetika, 2013, Vol. 49, No. 12, pp. 1416–1425.

HUMAN GENETICS

Gene-Pool Structure of Tuvinians Inferred from Y-Chromosome Marker Data

V. N. Kharkov, K. V. Khamina, O. F. Medvedeva, K. V. Simonova, I. Yu. Khitrinskaya, and V. A. Stepanov

Institute of Medical Genetics, Siberian Branch of the Russian Academy of Medical Sciences, Tomsk, 634050 Russia

e-mail: vadim.stepanov@medgenetics.ru

Received April 22, 2013; in final form, July 22, 2013

Abstract—The gene-pool structure of Tuvinians was examined in terms of the composition and frequency of Y-chromosome haplogroups in five geographically distanct populations. In the Tuvinian gene pool, a total of 22 haplogroups were identified with six of these, which were the most frequent (C3c, C3*, N1b, N1c1, Q1a3, and R1a1a). It was demonstrated that eastern regions of Tuva were most different from the other regions in haplotype frequencies. The evaluation of genetic diversity based on the frequencies of biallelic haplogroups and YSTR haplotypes revealed very high diversity values for all samples. In general, the genetic diversity values identified in Tuvinians were the highest for the indigenous ethnic groups of Siberia. The evaluation of the genetic differentiation of the samples examined using the analysis of molecular variance (AMOVA) showed that the gene pool of Tuvinians was relatively poorly differentiated with respect to haplogroup frequencies. Phylogenetic analysis within haplogroup N1b revealed strong founder effect, i.e., reduced diversity and star-like phylogeny of the median network of haplotypes, which formed a separate subcluster exclusive to Tuvinians. It was demonstrated that, in Tuvinians, haplogroup N1c1 was the most heterogeneous in haplotype profile and consisted of three different haplotype clusters, demonstrating considerable differences of western population from the rest of the Tuva populations. Phylogenetic analysis of haplogroups revealed common components for Tuvinians, Khakasses, Altaians, and Mongols.

DOI: 10.1134/S102279541312003X

INTRODUCTION

In recent years, indigenous ethnic groups of Altai–Sayan region were among the most actively studied by the specialists in different fields of human population genetics and medical genetics. Considerable interest to the problem of identifying the gene pool structure of South Siberian populations is associated with the complexity of their ethnogeny. According to anthropological data, metisation was the leading factor of population differentiation on the territory of steppe and forest steppe parts of South Siberia [1, 2]. The processes of the assimilation and fusion with the participation of Mongoloid and Caucasoid population groups played a major role in the formation of modern Turkic-speaking populations of South Siberia, including Tuvinians. In the Neolithic, Bronze, and Early Iron Ages the territory of Tuva was the part of the habitat of ancient Caucasoid population. Later, on this territory, the cultures of Scythian–Siberian were developed [3]. The penetration of Central Asian Mongoloid component to the territory of South Siberia can be dated back to the 7th–6th age BP. The emergence of forest, taiga Mongoloid component is also attributed to about this time [1, 4]. In the course of time, gradual increase of the Mongoloid component, from the prevalence of Caucasoid component in Scythian time to the formation of modern Central Asian anthropological type of Tuvinians in 13th–14th century AD was observed [3–5]. Genetic relationships between different population groups on the territory of Siberia caused by considerable migrations, which especially intensified during Bronze Age, Early and Late Iron Age, and Middle Ages, resulted in the formation of transitional Mongoloid–Caucasoid populations in the Altai–Sayan region.

At present, Tuvinians are one of the most compact living population groups of Russia. Moreover, the Tuvinian gene pool is relatively isolated because the frequency of intermarriages is very low, even in mixed populations [6, 7]. In addition, the isolation of certain indigenous ethnic groups of Tuva caused by geographical factors and the history of their development should be mentioned. This is because Tuvinians, unlike other populations of Altai–Sayan upland, are characterized by endogamy. At the same time, heterogeneity of tribal composition of Tuvinians was demonstrated [8].

This paper continues a series of studies on the gene pool structure of the indigenous ethnic groups of Siberia [9–13]. The objective of the present study was to describe the gene pool structure of Tuvinians, including its regional and intraethnic subdivision. The investigations were based on the analysis of Y-chromosome haplogroup composition and structure, determined

with the help of genotyping of a wide set of SNP and STR markers from nonrecombining portion of the Ychromosome. Until the present time, the Tuvinian gene pool was poorly studied using Y-chromosome markers. There are no separate articles on Tuvinians. Moreover, although the most detailed study on South Siberian ethnic groups [14] contains information on rather large population samples of Tuvinians (N = 113) and Todja Tuvinians (36), it is focused on a comparative analysis at the interethnic level. The genotyping data for Tuvinian population samples are also reported in the studies aimed at phylogeography and the origin of different Y chromosome haplogroups [15–17]. However, these data refer only to certain components of Tuvinian gene pool. In our earlier study, we examined Y-linked microsatellite markers in three geographically distanced Tuvinian populations. It was demonstrated that the male gene pool of Tuvinians contained several components that were assessed as Caucasoid and Mongoloid [9]. Since then, considerably larger population material, along with a higher number of informative DNA markers became available, which provides the analysis of the Tuvinian gene pool at the new level.

MATERIALS AND METHODS

Study material was represented by total DNA extracted from male peripheral blood leucocytes using standard methods. The population samples with the total number of 419 individual specimens and representing the indigenous population of the Tuva Republic were examined. The samples were collected in the settlement of Teely (west of Tuva) (N = 44), the settlement of Kungurtug (southwest of Tuva) (N = 48), the settlement of Toora-Khem (northeastern part of Tuva) (N = 23), and the city of Kyzyl (N = 296). The Kyzyl samples were assigned to the territorial groups in accordance to the donor birthplaces. The samples were divided into five geographically distanced groups as follows: west (Barun-Khemchinsk, Bai-Taiga, Dzun-Khemchinsk, Sut-Kholsk, and Mongun-Taiga raions) (N = 169), center (Chaa-Kholsk, Tandyn, Kaa-Khemsk, Kyzyul, Ulug-Khemsk, Chedi-Kholsk, Pii-Khemsk, Tes-Khemsk, Ovyursk, and Erzinsk raions) (N = 179), east (N = 71), which included the northeastern (Todzha raion) (N = 23) and southeastern (Tere-Kholsk raion) (N = 48) parts of the republic.

The biallelic loci, mostly represented by SNPs, and polyallelic highly variable microsatellites (YSTRs) were used as markers. Using biallelic markers, the attribution of the samples to certain haplogroup was performed. Nomenclature of haplogroups is as defined by the Y-Chromosome Consortium [18] with further modifications [19, 20]. 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. The haplogroup composition was studied using 60 markers from the nonrecombining portion of Y chromosome, including M1 (YAP), M3 (DYS199), M7, M8, M9, M12, M15, M17, M20, M25, M46 (Tat), M47, M56, M64, M67, M70, M73, M77, M86, M89, M92, M102, M117, M119, M120, M122, M124, M128, M130, M134, M157, M170, M172, M173, M174, M175, M178, M198, M201, M204, M207, M217, M223, M231, M242, M253, M267, M269, M324, M346, M407, M458, SRY1532, 92R7,DYF155S2, 12f2, P25, P31, P37, and P43. Biallelic markers were genotyped using polymerase chain reaction (PCR) with subsequent analysis of the DNA fragments by means of different methods as described earlier [10–13].

Microsatellite markers. Haplotype analysis was carried out using 17 microsatellite markers from the nonrecombining portion of the Y chromosome (DYS19, DYS385a, DYS385b, DYS388, DYS389I, DYS389II, DYS390, DYS391, DYS392, DYS393, DYS426, DYS434, DYS435, DYS436, DYS437, DYS438, and DYS439). Fluorescent-labeled primers (HEX, FAM, TET, and NED dyes used) were produced at Applied Biosystems. Genotyping of microsatellite markers was performed using the ABI Prism 310 and ABI Prism 3130x1 genetic analyzers. Primer sequences for DYS19, DYS389I, DYS389II, DYS390, DYS391, DYS392, DYS393 were described earlier in [9, 23]; sequences for DYS385a, DYS385b, DYS388, DYS426, DYS438 were described in [24]; and sequences for DYS434, DYS435, DYS436, DYS437, DYS439 were described in [25]. The fragment sizes were analyzed using the GeneMapper Software. Allele nomenclature corresponds to the generally accepted (for the DYS389I without taking into consideration the three-copy TCTG repeat and for DYS437 without terminal [TCTG]2–[TCTA]4 tandems).

Statistical methods. Genetic relationships among the populations were analyze using factor analysis (the method of principal components). Computations and graph construction were conducted using a STATISTICA 7.0 software package. Genetic diversity was evaluated using the formula suggested by Nei [26]. The genetic differentiation of the population was assayed using the analysis of molecular variance (AMOVA) [27]. Using the F_{st} coefficient, 10 000 permutations of the initial data was performed. The statistical significance of the differences between the populations in haplogroup and YSTR haplotype frequencies was evaluated using the exact test for population differentiation (the number of Markov chain steps was equal to 10000, the number of steps not taken into consideration was 1000, and the significance level was equal to 0.05). The matrices of Slatkin's pairwise distances (F_{st}) were calculated using 100 permutations of the initial data. The calculations were per-

formed in the ARLEQUIN 3.11 software package (http://cmpg.unibe.ch/software/arlequin3) [28]. Median networks for Y-chromosome haplotypes were constructed in the Network v. 4.6.1.1 software program (www.fluxus-engineering.com) using the median network method of Bandelt with the sequential use of the RM (reduced median) and MJ (median-joining) algorithms (the ε estimator was considered to be equal to 0) [29, 30]. To score different mutation rates upon the networks constructions, each STR locus was given a weight proportional to its variability within the haplotype array examined.

RESULTS AND DISCUSSION

Frequencies of Y-Chromosome Haplogroups in Population Samples of Tuvinians

Genotyping of the biallelic marker loci examined showed the presence of 22 Y-chromosome haplogroups in the gene pool of Tuvinians. Only six haplogroups (C3c, C3*, N1b, N1c1, Q1a3, and R1a1a) were identified in all five sample groups. The frequency of these haplogroups in the total sample of Tuvinians constituted more than 5%, and they represented 85% of Y-chromosome gene pool of Tuvinians. The highest number of haplogroups was identified in the population samples from the group representing the center of Tuva (20 haplogroups). In the west of Tuva, as well as in local samples from the south and southeast of the republic, considerably lower number of haplogroups was identified. However, most of Y-chromosome lineages characterized by a high frequency in the central group of Tuvinian samples were identified in the populations from other groups. In the population sample of Todja Tuvinians, only seven haplogroups were identified (table). On the one hand, these data suggest more intensive migratory flows from the outside to the steppe territory of Tuva compared to mountain–taiga regions of northeast and southeast of the republic. On the other hand, many rare haplogroups are represented by single samples, and their identification can be associated with the larger sizes of western and central groups of samples compared to the three other groups. The northeast of Tuva is represented by the least number of haplogroups, which can probably be associated with the historically formed isolated position of this region.

In general, population samples of Tuvinians showed that their gene pool contained far more haplogroups than in neighboring Altaians [10], Teleuts [12], Khakasses [13], and Buryats [31]. The assessment of the differences between the populations in the haplotype frequencies using hte exact test for population differentiation revealed statistically significant differences between western, central, and eastern groups in all pairwise comparisons.

In Tuvinians, the most frequent Y-chromosome variant was haplogroup N1b, which constituted 24% in the total gene pool. This lineage was the most abundant in the groups from the center and south of the republic. Moreover, in the latter group, almost 39% of all Y chromosomes belonged to this lineage. In this respect, it should be noted that no other haplogroup with such a high frequency was detected in any population. It seems likely that high frequency of this haplogroup in modern Tuvinians reflects the contribution the Samoyedic ethnic groups that previously populated the territory of Tuva to their gene pool.

Haplogroup N1c1 is the second highly frequent haplogroup in Tuvinians (19% out of all samples). In the total sample, the frequency of this haplogroup is only 5% lower than that of haplogroup N1b. Moreover, in the west of Tuva, Nic1 was present in about 30% of the samples examined. The distinguishing feature of the populations from the west and northeast of the republic is that the frequency of haplogroup N1c1 in these populations is higher than that of N1b. At the same time, in other parts of Tuva, the frequency of haplogroup N1b is orders of magnitude higher than that of N1c1. The lowest frequency of haplogroup N1c1 was observed in the south and southeast of Tuva. Thus, haplogroups of N group displayed a gradient of decreasing frequencies, from west to east (for haplogroup N1b) and northwest to southeast (for haplogroup N1c1).

It seems likely that the presence of haplogroup N1c1 in the gene pool of Tuvinians cannot be associated with any one particular ethnic element. However, similar to haplogroup N1b, some proportion of this lineage should reflect the Samoyedic ethnic substratum. It can be suggested that, in the gene pool of Tuvinians, this lineage also marks the contribution of Ugric tribes.

Haplogroup Q1a3 was found in 14% of all Tuvinian specimens. This haplogroup was detected with the highest frequency in the eastern samples (25%), while it was absent in the samples from the south of the republic. Q1a3 demonstrated the east–west gradient of decreasing frequencies. The highest for Tuvinians frequency of haplogroup Q1a3, observed in mountain taiga and southeastern regions, as well as in Todja, probably results from geographic inaccessibility of these regions and, as a consequence, from relative genetic isolation of local populations. Under these conditions, in the gene pool of the Tuva population, the highest proportion of ancient component was preserved. Alternatively, this situation can be associated with the fact that the gene pool of Todja Tuvinians was formed with more active participation of Kets, who are characterized by the high frequency of haplogroup Q1a3.

The proportion of haplogroup R1a1a in the total gene pool of Tuvinians is somewhat lower (12%). This lineage is characterized by high frequencies in the populations of South Siberia. Furthermore, R1a1a dominates in Altaians, constituting 60% [10]. In Khakasses, this haplogroup is the second most frequent

Distribution of Y-chromosome haplogroups in Tuvinians

Haplogroup	Frequency, % (*N*)						
	West (*N* = 169)	Center (*N* = 179)	East (*N* = 71)	Total (*N* = 419)	Northeast (*N* = 23)	Southeast (*N* = 48)	South (*N* = 27)
C3* (xM48, M77, M86)	3.6 (6)	7.3 (13)	5.6 (4)	5.48 (23)	4.3 (1)	6.3 (3)	3.6 (1)
C3c (M48, M77, M86)	8.9 (15)	11.2 (20)	4.2 (3)	8.7 (38)	4.3 (1)	4.2 (2)	14.3 (4)
C3d (M407)	–	2.2 (4)	5.6 (4)	1.9 (8)	–	8.4 (4)	3.6 (1)
D (M15)	–	0.6 (1)	–	0.2 (1)	–	–	3.6 (1)
E (M1)	–	0.6 (1)	1.4 (1)	0.5 (2)	–	2.1 (1)	–
I1 (M253)	1.8 (3)	0.6 (1)	–	0.9 (4)	–	–	–
I2a1 (P37)	–	0.6 (1)	–	0.2 (1)	–	–	–
J* (xM172, M287)	–	0.6 (1)	–	0.2 (1)	–	–	–
J2a1b1 (M92)	–	0.6 (1)	–	0.2 (1)	–	–	–
J2* (xM47, M67)	0.6 (1)	–	–	0.2 (1)	–	–	–
N* (M231, xP43, M46, M178)	1.2 (2)	1.7 (3)	1.4 (1)	1.4 (6)	–	2.1 (1)	–
N1b (P43)	24.9 (42)	27.9 (50)	18.3 (13)	24.3 (105)	13.0 (3)	20.8 (10)	39.3 (10)
N1c1 (M178)	28.4 (48)	12.8 (23)	8.5 (6)	18.9 (77)	17.4 (4)	4.2 (2)	7.1 (2)
O* (xM122, P31)	–	0.6 (1)	4.2 (3)	0.9 (4)	–	6.3 (3)	–
O2 (P31)	–	0.6 (1)	–	0.2 (1)	–	–	3.6 (1)
O3a* (M122, xM134)	4.7 (8)	1.7 (3)	2.8 (2)	3.1 (13)	–	4.2 (2)	7.1 (2)
O3a3c1 (M117)	3.0 (5)	2.2 (4)	–	2.1 (9)	–	–	3.6 (1)
O3a3c* (M134, xM117)	1.2 (2)	–	1.4 (1)	0.7 (3)	4.3 (1)	–	–
Q1a3 (M346)	11.8 (20)	9.4 (17)	25.6 (18)	13.9 (55)	30.4 (7)	22.9 (11)	–
R1b* (M269, xM73)	–	0.6 (1)	1.4 (1)	0.5 (2)	–	4.2 (1)	–
R1a1a (M17)	8.3 (15)	13.9 (25)	18.3 (13)	12.3 (53)	26.1 (6)	14.6 (7)	10.7 (3)
R1b1b1 (M73)	1.2 (2)	4.4 (8)	1.4 (1)	2.8 (11)	–	4.2 (1)	3.6 (1)
H HG	0.8276 ± 0.0157	0.8601 ± 0.0140	0.8620 ± 0.0212	0.8579 ± 0.0080	0.8221 ± 0.0447	0.8794 ± 0.0247	0.8405 ± 0.0576

(28%) [13]. Interestingly, in western Truvinian sample, the frequency of haplogroup R1a1a was considerably lower than in the central sample. Based on the closeness of the Altai, which is populated by the representatives of a more Caucasoid South-Siberian racial type, it would be reasonable to expect the west–east decrease of the R1a1a frequency on the territory of Tuva. However, this was not observed, and the change of the haplogroup frequency was rather the opposite, as the eastern samples demonstrated maximum frequency of this haplogroup. At the first glance, the result obtained is paradoxical. Specifically, in terms of anthropology, the most Caucasoid population of the western parts of Tuva displays the minimum of haplogroup R1a1a, while in the most Mongoloid population of Todja, the maximum of this haplogroup is observed.

In Tuvinians, single samples of the other seven west Eurasian (Caucasoid) haplogroups were observed (E, I1, I2a1, J*, J2*, J2a1b1, and R1b*). Thus, in Tuvinians, the Caucasoid component is rather diverse, and probably has east European, as well as Central Asian origin. The five geographical groups of Tuivinians demonstrated no statistically significant differences in the summarized frequencies of west Eurasian haplogroups, which can be associated with the antiquity of the distribution of the Caucasoid population over the territory of formation of the Tuvinian gene pool.

Two Central Asian haplogroups, C3* and C3c, are the least frequent among Tuvians. The proportion of haplogrpoup C3* is the highest in the central sample (7%) and of haplogroup C3c in the southern group (15%). It has been suggested that, in Tuvinians, haplogroups C3* and C3c mostly mark the genetic contribution of late Mongoloid migrants, which reflects the contribution of Hun and Mongolian immigrants. In the Tuvinian gene pool, the southeast-to-west clinal decrease of the frequencies of these haplogroups was observed. Mongolian nomads penetrated into the territory of Tuva from the south, gradually spreading to northern territories. Accordingly, the Mongolizatiion of the population of Tuva was most expressed in southern regions, which indicated the paleontological data [3], the anthropological characteristics of the modern population [32], and the linguistic data. The latter characterize the southeastern dialect formed as a result of the considerable influence of the Mongolian language.

Two Eastern Eurasian haplogroups, O3a* and O3a3s1, in the total sampling have a frequency of 3.1 and 2.1%, and in the southern regions—7.1 and 3.6%, respectively. In Tuvinians, D, E, I1, I2a1, J*, J2a1b1, J2*, N*, O*, O2, O3a3c*, and R1b*, which are represented by single samples, haplogroups are rare and their proportion in total sample constitutes less than 2%. Among these, only haplogroup N* was detected in two different samples. Haplogroups O3a3c1 and R1b* have frequencies higher than 2% and belong to samples that represent three different regional groups.

Population Genetic Diversity

The genetic diversity estimates based on the frequencies of biallelic haplogroups and YSTR haplotypes showed very high diversity values in all samples at both marker systems. The *H* values obtained (more than 0.8, for haplogroups, and more than 0.9, for microsatellite haplotypes of Y-chromosome) were higher than those observed in all indigenous ethnic populations of Siberia examined in our studies and elsewhere. For instance, in neighboring Altaians, the value of this index was 0.5–0.7 [10]; in Teleuts, it was 0.7 [12]; in Khakasses, it was 0.2–0.7 [13]; in Buryats, it was 0.5 [14]; in Shorians, it was 0.6 [14]. Thus, Tuvinians, despite their current endogamy, are characterized by the highest level of genetic diversity level among Siberian ethnic groups within the marker system chosen. These findings are probably associated with the fact that multicomponent composition of Y-chromosome haplogroups in the the Tuvinial gene pool was formed prior to the formation of endogamous marriage traditions.

An analysis of the haplogroup distributions showed no noticeable heterogeneity of the samples examined in the levels of genetic diversity of their male gene pools. The difference in the genetic diversity values was small (table). The highest diversity in the composition of the haplogroup in local samples was observed in the southeast of Tuva, in small population from hardly accessible mountain territory. The lowest genetic diversity value was obtained for the western samples, which is also obviously associated with the high frequency of haplogroups N1b and N1c in the northeast of Tuva (Todja).

Previously, using these populations, it was demonstrated that Tuvinians were characterized by the heterozygote deficiency at the Alu repeat loci, which probably reflected the relatively high level of inbreeding in these populations due to their small number and high degree of isolation [33]. The genetic demographic analysis of Truvinians showed that they are characterized by a very high value of the inbreeding coefficient calculated based on isonymous marriages, the frequency of which in Tuva is an order of magnitude higher than in other previously examined indigenous populations of Siberia examined [6, 7, 34]. The highest level of inbreeding was observed in the western sample form the Bai-Taiga region, which was completely consistent with the relatively low diversity at Y-chromosome markers.

At the same time, the high level of genetic diversity in Tuvinian populations was observed by analyzing classic markers [7], autosome microsatellites, and Alu repeats [33]. In these conditions, the sample from the southeast of Tuva often demonstrated the highest level of genetic diversity.

Genetic diversity of Tuvinians inferred from the mtDNA haplogroup data was also very high (about 0.94–0.96) [35]. This concordance of the results

obtained using two uniparentally inherited marker systems deserves special interest and is a reflection of complex processes of the formation and development of Tuvinian ethnos, as well as of the specific features of the sex–age and marriage population structure.

Genetic Differentiation of the Population

The genetic differentiation of the samples of the examined population was performed using the analysis of molecular variance using two independent marker systems (haplogroup frequencies and the frequencies of Y-chromosome microsatellite haplotypes). Calculations within two marker systems produced almost the same results. An analysis of the haplogroup frequencies showed that the differences between the center, west and east of Tuva constituted 2.9%, and the differences among five samples constituted 2.1%. Calculations performed using YSTR haplotype data yield somewhat lower differences 1.5%). These values were an order of magnitude lower than those obtained for the neighboring ethnic groups (Altais and Buryats) [10, 31]. The within-population differences of the five Tuvinian samples constituted 97.9% (by haplogroup frequencies) and 98.5% (by haplotypes) of total genetic variability.

Thus, the AMOVA results indicated the absence of considerable genetic differences among the three territorial groups of Tuvinians. This means that, despite the information on substantial anthropological differentiation of the indigenous population of Tuva since Scythian times, contemporary language differences between the ethnoterritorial groups of Tuvinians, the differences in anthropological characters, tribal composition, and some others, the male part of the Tuvinian gene pool is very low differentiated.

This situation can probably be explained in terms of the long period of time since the last intensive migrations to Tuva from outside. Tuvinian ethnic group has long existed as a relatively closed population system. Specific features of the economic activity left their marks on the differences in migration activity between Tuvinian men and women, and despite the Tuvinian-typical endogamy, the differences in the Y-chromosome gene pools between distant from each other Tuvinian populations were considerably smoothed out. It can be said with confidence that migration was the main population dynamics factor, which played the key role in this process.

An analysis of the same populations with the use of mtDNA markers demonstrated that mitochondrial gene pool of Tuvinians was also characterized by low intraethnic differentiation and could be considered as a single system [35, 36]. An analysis of autosomal Alu repeat variations also suggests the integrity of the gene pool of modern Tuvinians and showed that the population of the west of Tuva was genetically statistically significantly different from other Tuvinian populations

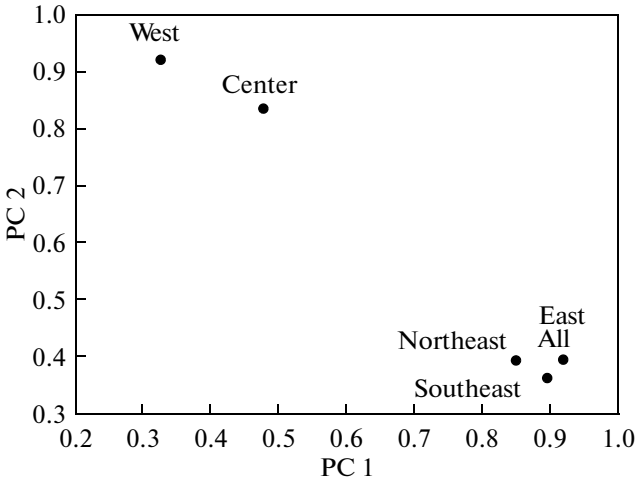

Fig. 1. Position of Tuvinian populations in the space of principal components based on Y-chromosome haplogroup frequencies.

[33]. In turn, an analysis of autosomal STR loci revealed genetic differences between the northeastern Todja Tuvinians and the rest of the Tuva population [33].

Genetic Relationships between Populations

An analysis of the data set over the haplogroup frequencies with the help of factor analysis (Fig. 1) showed that the samples from western and central parts of Tuva were the genetically closest to each other, which agreed with the data on the genetic closeness of the population of these two regions with respect to a complex of marker systems (classic markers, mtDNA, Alu insertion polymorphism) [7]. The samples from the northeast and southeast of Tuva were considerably distant from the first samples. The first principal component (PC1), which explained 82% of the total variability, effectively differentiated all population samples. In this case, maximum factor loadings fell on the frequent haplogroups, C3*, C3c, N1b, Q1a3, R1a1a, as well as on a number of minor haplogroups. In the space of PC2 (12% of total variability), the samples from the west and east of the republic were those, maximally distant from each other. The highest factor loadings in this case fell on haplogroups N1c1 and C3d.

The isolation of southeast and northeast of Tuva (Todja) from the rest of the republic established is probably the consequence of geographical isolation of this region. These data are in good agreement with anthropological [1–5] and linguistic [6, 7] data.

Phylogenetic Analysis of Haplogroup N1b in Tuvinians

An analysis of the YSTR haplotypes structure and diversity within the most frequent in Tuvinians haplogroup N1b showed that this haplogroup contained two individual clusters (Fig. 2). The smaller cluster (I)

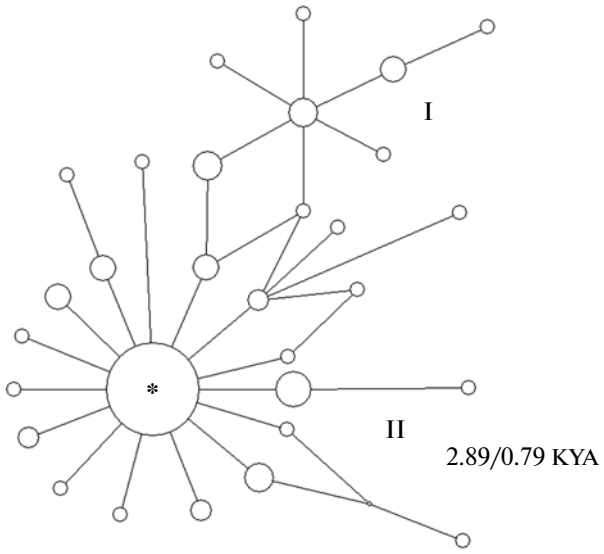

Fig. 2. Median network of the YSTR haplotypes of haplogroup N1b in Tuvinians. The size of a node corresponds to the number of identified specimens belonging to a given haplotype. The haplotype-founder of the Tuvanian cluster is marked with an asterisk. I, II—see text.

consisted of the N1b ancestral haplotype and a number of its derivatives (11 specimens in total). Only few of the Tuvinian specimens belonged to this cluster. The second cluster of haplotypes (II) is spaced from the first cluster by three mutation steps. This cluster is characterized by starlike phylogeny, which indicates the strong local founder effect in this Tuvinian lineage (a total of 118 specimens). Allelic structure of modal haplotype from the main Tuvinian cluster is 15-12-13-12-10-16-23-11-14-13-11-8-11-12-8-10-11 (marked by asterisk in the figure).

Haplotypes from this cluster were exclusive to Tuvinians, and they were not detected upon analysis of the N1b lineage in other ethnic groups, including South Siberian Altaians, Teleuts, Khakasses, Shorians, and Buryats (personal and literature data available on 17 YSTR markers). The F_{st} value calculated for this haplotype cluster was low ($F_{st} = 4.13\%$). The exact test for population differentiation showed no statistically significant differences between different geographical groups of Tuvinians with respect to haplogroup N1b. It can be concluded that, in the Tuvinian ethnic group, the genetic component marked by this haplogroup is homogeneous in structure and allele frequencies. The age estimates for the main Tuvinian cluster range from 2.89 ± 1.24 to 0.79 ± 0.16 thousand years (KYA). Most of the N1b haplotypes found in Tuvinians do not overlap with the haplotypes of other South Siberian ethnic groups. It can be suggested that this component of the Tuvinian gene pool is associated with the expansion of Samoyedic tribes on the territory of Tuva.

Phylogenetic Analysis of Haplogroup N1c1 in Tuvinians

The structure of haplotype median network obtained upon phylogenetic analysis of haplogroup N1c1 in Tuvinians was quite different from the first one. Haplotypes of this lineage were clearly subdivided into three haplotype clusters that were approximately equal in the number of specimens (Fig. 3). The first of these clusters (I), similar to the smaller cluster of haplogroup N1b, is the closest to the founder haplotype of the entire haplogroup (14 samples), and the other two represent considerable ethnic and geographic specificity. Both of them are relatively young and, like N1b, are a prime example of a "star" phylogeny coupled with the recent founder effect. The second cluster (II) (26 samples) combines Tuvinians, Khakasses, and Shorians [13] and is not found in other Siberian ethnic groups. The generation time of diversity in Tuvinians is $2.59 \pm 0.60/0.78 \pm 0.17$ thousand years. Founder haplotype of the cluster (15-12-12-12-11-16-23-11-14-13-11-8-11-12-8-10-11) stands four mutation steps apart from the founder haplotype of the whole haplogroup. The first and second haplotype clusters demonstrate no statistically significant frequency differences in different Tuvinian samples.

The third cluster (III) carries the haplotypes, which are shared by Tuvinians, Altaians, Buryats, and a considerable part of the Mongolian sample from the SMGF database [37]. Age estimates for this cluster range from 1.69 ± 0.48 to 0.46 ± 0.13 KYA. This cluster (modal haplotype 14-11-11-12-11-16-23-11-14-14-11-8-11-12-8-10-10) is distributed mainly in the western sample (23 specimens out of 34). Interestingly, in Khakasses, no haplotypes of this cluster were detected [13]. Obviously, the population expansion, associated with the distribution of this Y-chromosome lineage occurred much later than the dispersal of the carries of Tuvinian–Shorian–Khakass N1c1 branch. The phylogeography of this branch of N1c1 haplotypes can be a consequence of the appearance of the Mongolian ethnic component in Tuva, Buryatia, and Altai. In this case, higher frequency of this cluster observed in western Tuvinian sample in comparison to the samples from other regions, can be explained in terms of local founder effect (similarly to the extremely high frequency of N1b).

Statistically significant differences in the frequencies of the N1c1 haplotypes were observed between the western and all other samples, except Todja. The extent of genetic differentiation among five geographical groups relative to haplogroup N1c1 was almost four times higher than the values obtained for haplogroup N1b (F_{st} constituted 16.89%). This is the most heterogeneous component of the Tuvinian gene pool.

Thus, in the present study, a detailed analysis of the Tuvinian gene pool was conducted using genetic markers from nonrecombining portion of Y-chromosome. The data obtained indicate the multicomponent composition of the gene pools of the Tuvinian popula-

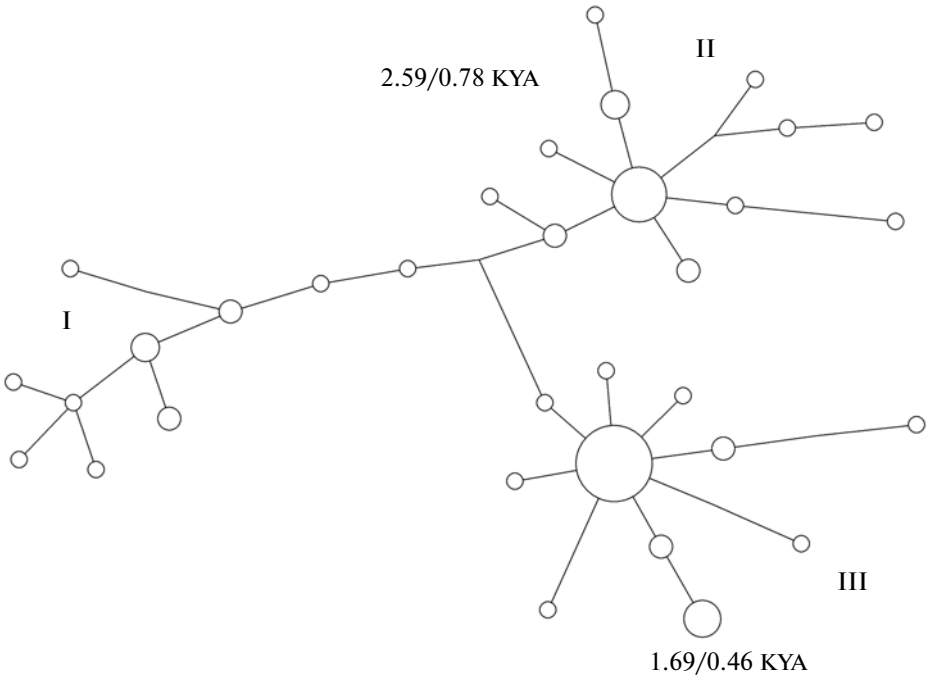

Fig. 3. Median network of the YSTR haplotypes of haplogroup N1c1 in Tuvinians. I, the main haplotype cluster; II, Tuvinian–Khakass cluster; III, Tuvinian–Buryat–Mongolian cluster.

tions, as well as to the very high genetic diversity of the general Tuvinian gene pool. In Tuvinian ethnic group, statistically significant differences between the samples from western, central, and eastern parts of Tuva were observed. An analysis of molecular variance revealed the relatively low proportion of the between-population differences in the total variability of the Tuvinian gene pool. In addition, the results of this analysis indicated the low genetic differentiation between the indigenous populations from western and central Tuva relative to the markers from nonrecombining portion of the Y chromosome. At the same time, large genetic distance between the samples from eastern regions of Tuva and the other samples was demonstrated along with the differences in the degree of heterogeneity and haplotype profiles within the most frequent N1b and N1c1 lineages.

ACKNOWLEDGMENTS

This work was supported by the Federal Targeted Programs: Scientific and Scientific-Pedagogical Personnel of Innovative Russia (grant no. 11.519.11.2036) and Research and Development in Priority Fields of Science and Technology Complex (State Contract no. 8042), and by the Russian Foundation for Basic Research (grant no. 12-04-00595a).

REFERENCES

1. Alekseev, V.P. and Gokhman, I.I., *Antropologiya aziatskoi chasti SSSR* (Anthropology of the Asian Part of the Soviet Union), Moscow, 1984.
2. Moiseev, V.G., Anthropological content of the Turkic speaking peoples on the West-Siberian Plain inferred from craniological data, in *Vtoroi mezhdunarodnyi kongress etnografov i antropologov: rezyume dokladov i soobshchenii* (The Second International Congress of Ethnographers and Anthropologists: Abstracts of Reports and Communications), Ufa: Vostochnyi Univ., 1997, part 1, pp. 57–58.
3. Alekseev, V.P., A brief review of palaeoanthropology of Tuva in relation to historical questions, in *Antropologicheskie issledovaniya v Tuve* (Anthropological Research in Tuva), Moscow: Nauka, 1984, pp. 17–28.
4. Kiselev, S.V., *Istoriya Yuzhnoi Sibiri* (History of South Siberia), Moscow: Akad. Nauk SSSR, 1951.
5. Debets, G.F., *Paleoantropologiya SSSR* (Palaeoanthropology of the Soviet Union), vol. 4 of *Trudy instituta etnografii AN SSSR* (Transections of the Institute of Ethnography, Academy of Sciences of the Soviet Union), Moscow: Akad. Nauk SSSR, 1948.
6. Puzyrev, V.P., Erdynieva, L.S., Kucher, A.N., and Nazarenko, L.P., *Genetiko-epidemiologicheskoe issledovanie naseleniya Tuvy* (Genetic Epidemiologic Study of the Tuva Population), Tomsk: STT, 1999.
7. Kucher, A.N., Ondar, E.A., Stepanov, V.A., et al., *Tuvintsy: geny, demografiya, zdorov'e* (Tuvinians: Genes, Demography, Health), Tomsk: Pechatnaya Manufaktura, 2003.

8. Potapov, L.P., *Ocherki narodnogo byta tuvintsev* (Essays on the Traditional Life of Tuvinians), Moscow: Nauka, 1969.
9. 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.
10. Khar'kov, V.N, Stepanov, V.A., Medvedeva, O.F., et al., Gene pool differences between Northern and Southern Altaians inferred from the data on Y-chromosomal haplogroups, *Russ. J. Genet.*, 2007, vol. 43, no. 5, pp. 551–562.
11. Khar'kov, V.N., Stepanov, V.A., Medvedeva, O.F., et al., The origin of Yakuts: analysis of the Y-chromosome haplotypes, *Mol. Biol.* (Moscow), 2008, vol. 42, no. 2, pp. 198–237.
12. Khar'kov, V.N., Medvedeva, O.F., Luzina, F.A., et al., Comparative characteristics of the gene pool of Teleuts inferred from Y-chromosomal marker data, *Russ. J. Genet.*, 2009, vol. 45, no. 8, pp. 994–1003.
13. Khar'kov, V.N., Khamina, K.V., Medvedeva, O.F., et al., Genetic diversity of the Khakass gene pool: subethnic differentiation and the structure of Y-chromosome haplogroups, *Mol. Biol.* (Moscow), 2011, vol. 45, no. 3, pp. 404–416.
14. Derenko, M., Malyarchuk, B., Denisova, G., et al., Contrasting patterns of Y-chromosome variation in South Siberian population from Baikal and Altai–Sayan regions, *Hum. Genet.*, 2006, vol. 118, pp. 591–604.
15. Derenko, M., Malyarchuk, B., Denisova, G., et al., Y Chromosome haplogroup N dispersals from South Siberia to Europe, *J. Hum. Genet.*, 2007, vol. 52, no. 9, pp. 763–770.
16. Malyarchuk, B., Derenko, M., and Denisova, G., Phylogeography of the Y-chromosome haplogroup C in northern Eurasia, *Ann. Hum. Genet.*, 2010, vol. 74, pp. 539–546.
17. Malyarchuk, B., Derenko, M., and Denisova, G., Ancient links between Siberians and native Americans revealed by subtyping the Y chromosome haplogroup Q1a, *J. Hum. Genet.*, 2011, vol. 56, pp. 583–586.
18. The Y-Chromosome Consortium: A nomenclature system for the tree of human Y-chromosmal binary haplogroups, *Genome Res.*, 2002, vol. 12, pp. 339–348.
19. Karafet, T.M., Mendez, F.L., Meilerman, M.B., et al., New binary polymorphisms reshape and increase resolution of the human Y chromosomal haplogroup tree, *Genome Res.*, 2008, vol. 18, pp. 830–838.
20. http://www.isogg.org/
21. Khar'kov, 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. 326–331.
22. Khar'kov, V.N, Stepanov, V.A., Feshchenko, S.P., et al., Frequencies of Y chromosome binary haplogroups in Belarussians, *Russ. J. Genet.*, 2005, vol. 41, pp. 928–931.
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. Butler, J.M., Schoske, R., Vallone, P.M., et al., A novel multiplex for simultaneous amplification of 20 Y chromosome STR markers, *Forensic Sci. Int.*, 2002, vol. 129, no. 1, pp. 10–24.
25. Ayub, Q., Mohyuddin, A., Qamar, R., et al., Identification and characterization of novel human Y-chromosomal microsatellites from sequence database information. *Nucleic Acids Res.*, 2000, vol. 28, no. 2, p. e8.
26. Nei, M., *Molecular Evolutionary Genetics*, New York: Columbia Univ. Press, 1987.
27. 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.
28. 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.
29. 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.
30. Bandelt, H.-J., Forster, P., and Rohl, A., Median-joining networks for inferring intraspecific phylogenies, *Mol. Biol. Evol.*, 1999, vol. 16, pp. 37–48.
31. Khar'kov, V.N. and Stepanov, V.A., Gene pool structure of Buryats as inferred from the Y-chromosome haplogroups: clinal variation and territorial partitioning, in *Osob' i populyatsiya—strategii zhizni* (Individual and Population—Life Strategies), (Proc. 9th All-Russian Population Seminar), Ufa, 2006, pp. 520–523.
32. Bogdanova, V.I., Some questions of the anthropological content formation of modern Tuvinians, *Sov. Etnograf.*, 1978, no. 6, pp. 46–60.
33. Stepanov, V.A., *Etnogenomika naseleniya Severnoi Evrazii* (Ethnogenomics of the Northern Eurasia Population), Tomsk: Pechatnaya Manufaktura, 2002.
34. Kucher, A.N., Puzyrev, V.P., Sanchat, N.O., and Erdynieva, L.S., Genetic demographic characteristics of the rural population of the Tuva Republic: ethnic and tribal composition and sex and age structure, *Russ. J. Genet.*, 1999, vol. 35, no. 5, pp. 580–585.
35. Golubenko, M.V., Eremina, E.R., Tadinova, V.N., et al., Distribution of Europeoid and Mongoloid mitochondrial DNA haplogroups among populations of Siberia and Central Asia, in *Genofond naseleniya Sibiri* (Gene Pool of Siberian Population), Konenkov, V.I., Puzyrev, V.P., and Voevody, M.I., Eds., Novosibirsk: Institute Arkheologii i Etnografii Sibirskogo Otdeleniya Rossyskoy Akademii Nauk, 2003, pp. 9–13.
36. Golubenko, M.V., Puzyrev, V.P., Salyukov, V.B., et al., Distribution of "Mongoloid" haplogroups of mitochondrial DNA among indigenous population of the Tuva Republic, *Russ. J. Genet.*, 2001, vol. 37, no. 6, pp. 683–692.
37. http://www.smgf.org/

Translated by N. Maleeva