Analysis of Polymorphism of Uniparental Markers in Reindeer-Herding Populations: The Tozhu Tuvans of Russia and The Tsaatans Of Mongolia

Natalia Balinova¹, Irina Khomyakova², Elena Ayyzhi³, Murat Dzhaubermezov⁴, Sergey Litvinov⁴, Elza Khusnutdinoiva³, Galina El'chinova¹, Rena Zinchenko⁴,⁵, Nailya Spitsyna¹, Jüri Parik⁸, Maere Reidla⁸, Siiri Rootsi⁸

¹Research Centre for Medical Genetics, Moscow, Russian Federation
²Anuchin Research Institute and Museum of Anthropology, Moscow State University, Moscow, Russian Federation
³Tuvan State University, Tuva Republic, Russian Federation
⁴Department of Genetics and Fundamental Medicine, Bashkir State University, Ufa, Republic of Bashkortostan, Russian Federation
⁵Institute of Biochemistry and Genetics – Subdivision of the Ufa Federal Research Centre of the Russian Academy of Sciences, Ufa, Republic of Bashkortostan, Russian Federation
⁶N.A. Semashko National Research Institute of Public Health, Moscow, Russian Federation
⁷N.N. Miklukho-Maklai Institute of Ethnology and Anthropology RAS, Moscow, Russian Federation
⁸University of Tartu Institute of Genomics, Tartu, Estonia

ABSTRACT

We analyzed the data on the variability of the Y chromosome and mitochondrial DNA (mtDNA) in populations of the Tsaatans of Mongolia and the Tozhu Tuvans of Russia. The populations studied are characterized by low genetic diversity for both marker systems. The analysis of Y chromosome haplogroups in the Tsaatan and Tozhu revealed three haplogroups in the Tsaatan and seven haplogroups in the Tozhu. The composition of the haplogroups is coherent to literature data on the Tuvans, which is explained by common origin. According to the data on mitochondrial DNA variability, 12 haplogroups were determined in 46 Tozhus, of which C4b (30.43%) and F1b1b (23.91%) are major haplogroups. According to the HVS–1 (HyperVariable Segment) data, 15 haplotypes were found in the Tozhu Tuvans and the diversity coefficient of 0.8677 turned out to be much lower than among the Torghut of Mongolia (0.9857). In 23 Tsaatans, 14 haplogroups were determined; the most common of which are C4b (22.73%) and C5a1 (18.18%). According to HVS-1, 14 haplotypes were revealed in the Tsaatan, the diversity is 0.9486. The data obtained on uniparental marker systems in the Tozhus and Tsaatans are due to the isolated and inaccessible taiga region and the manifestation of the “founder effect”. The Tsaatans are less polymorphic in terms of the variety of Y chromosome haplogroups, while the Tozhus are less polymorphic in terms of mitochondrial DNA, which is probably a consequence of a high rate of endogamic marriages in the populations studied.

Key words: anthropology, reindeer herders, uniparental markers, haplogroups, mitochondrial DNA (mtDNA), Y chromosome

Introduction

The Tuvans or Tuvinians (endonym Tyva; obsolete names Tannu-Tuvans, Soyots, Uriankhais) are Turkic-speaking Buddhist Lamaists; historically and culturally they are divided into western (mountain-steppe regions of western, central and southern Tuva) and eastern, or Tozhu Tuvans (mountain-taiga part of northeast and southeast Tuva). The total number of Tuvans in Russia is 263,934 people. They live mainly in the Tyva Republic, where they represent a numerical majority of 82.04% (according to the 2010 census)³.

The Tozhu are an ethnic group of Tuvans living compactly in the Tozhu District in north-eastern Tyva. They are engaged in reindeer herding and speak the Tozhu dialect of Tuvan. The Tozhu belong to the indigenous small-numbered peoples. According to the 2010 census, their number is 1856¹. Their housekeeping practices are similar to those of the Tsaatans, with many families still...
living in the taiga in skin tents (chums) and maintaining reindeer herds.

The Tsaatan are a small tribe in north-western Mongolia who live off the reindeer. According to the 2010 census, their number is 282. Their name comes from the Mongolian word tsaa, meaning ‘reindeer’, although they call themselves Tuba or Tukha. They speak the Tozhu dialect of Tuvan. Due to the small number of Tsaatans and almost inaccessible reindeer pastures, we surveyed those who had left their camps to participate in elections and receive social benefits payable to small-numbered peoples.

The Tuvans have been thoroughly studied in different time periods under all classical anthropological programmes\textsuperscript{2–18}. The ethnic groups of the Altai-Sayan region were selected as model groups in the discussion of fundamental issues of ethnic anthropology of contact zones between representatives of the Central Asian Mongolid type and carriers of mixed South Siberian complexes. The issue of participation of the ancient Caucasian population in their formation and the manifestation of borders between the Turkic-speaking and Mongolian-speaking peoples of Asia were considered both by paleoanthropologists and geneticists. The presence of a Caucasian admixture was confirmed not only by paleoanthropologists\textsuperscript{12} but also by morphologists\textsuperscript{19} and geneticists\textsuperscript{14}.

The last decades of the twentieth century are characterized by a rapid development of population research of highly polymorphic genetic systems in solving ethnogenetic problems. The uniparentally inherited genome components, mitochondrial DNA (mtDNA) and Y chromosome, provide individual analysis of maternal and paternal components. In recent years, mtDNA diversity was extensively studied in the ethnic groups from South Siberia\textsuperscript{20–23}. It was demonstrated that mitochondrial gene pools of the populations studied were characterized by different ratios between the Mongolid and Caucasian mtDNA lineages\textsuperscript{21}. The frequency of Caucasian 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. Most of these lineages have south Caucasian origin, while the others belong to east Caucasian lineages\textsuperscript{24}.

Studies of Y chromosome in ethnic groups from Altai-Sayan region also showed the presence of profound paleo-Caucasoid component in their gene pools\textsuperscript{25, 26}. These data are consistent with those obtained earlier upon the analysis of mtDNA: the ethnic groups 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 Caucasian gene pool component\textsuperscript{27–29}.

Materials and Methods

All procedures performed in studies involving human participants were in accordance with the ethical standards of the institutional and/or national research committee and with the 1964 Helsinki declaration and its later amendments or comparable ethical standards. Informed consent was obtained from all individual participants included in the study. This genetic and epidemiological study was approved by the Research Ethics Committee of the Research Centre for Medical Genetics (protocol No. 5/2 dated 9 February 2015).

We studied mitochondrial DNA (mtDNA) and Y chromosome in 23 Tsaatans and 46 Tozhu Tuvans. The sample is representative because of very small size of total population. All respondents included into the study were interviewed about unrelated relatives up to 3 generations. DNA was extracted from peripheral blood leukocytes by the standard chlorphenol method. The haplogroup composition was studied using markers from the nonrecombining portion of Y chromosome: M9, M130, M217, M77, F6379, B469, B80, B90, L1373, M407, M217, M48, M401, M207, page07, Z93, Z95, Z2125, M343, M269, M412, Z2105, M458, M558, M478, M124, M175, M122, M119, M268, M95, P203, P201, M7, M231, P43, TAT, M128, F2930, F4205, CTS6967, B478, B525, B187 M2118, YAP, M170, M423, 12f2, M410, Page 08, M35, M174, M201, P15, M242, M25, M346. Genotyping was performed using PCR and subsequent fragment analysis by electrophoretic separation or Sanger sequencing. Mitochondrial diversity was studied using HVS-1, and then analyzed to confirm haplogroups: HVS-2, 663, 827, 1005, 1048, 1107, 1119, 1382, 1391, 1719.1G, 1719, 1811, 3010, 3206, 3552, 3759, 3816, 3826, 3970, 4491, 4646, 4715, 4833, 4850, 4904, 5021, 5093, 5108, 5153, 5178, 5231, 5360, 5442, 5465, 5601, 5821, 6023, 6047, 6024, 6212, 6392, 6962, 7028, 7100, 7476, 7768, 7828, 8020, 8273, 8414, 8479, 8485, 8563, 8701, 8703, 8762, 8793, 8964, 8994, 9120, 9181, 9449, 10084, 10208, 10398, 10400, 10427, 11215, 11467, 11696, 12672, 12705, 13104, 13942, 14233, 14766, 15071, 15080, 15204, 15326, 15475, 15487, 15884. Statistics

Excel 2016 was used to visualize the results. For cluster analysis and multivariate scaling, the Statistica software version 12.6 was used.

We used R to perform the principal component analysis (PCA). AMOVA analysis, values of Tajima’s D and Fu’s Fs tests for neutrality were calculated in Arlequin 3.5.

Results

The analysis of Y-haplogroups in the Tsaatan and Tozhu was previously studied during the analysis of the Oirat ethnogenesis\textsuperscript{30}. The distribution of haplogroups is presented in Table 1 and Figure 1. The low diversity of
haplogroups in the populations studied is explained by the geographical isolation of these groups. The “founder effect” is clearly represented in the Tsaatan population, since the genealogy is traced along two male lineages, haplogroups N3a5-F4205 (52.2%) and Q1a1b-M25 (43.5%) constitute the main diversity. According to the analysis of marriage records, the Tsaatans show a high endogamy rate since marriage relationships with Tozhu Tuvans ceased after

| Sample / haplogroup | Mongols Torguhs (N=47) | Mongols Khoshuts (N=18) | Mongols Derbets (N=40) | Mongols Tsaatans (N=23) | Tozhu Tuvans (N=46) |
|---------------------|------------------------|-------------------------|------------------------|-------------------------|---------------------|
| C3 (XC3c1, M407)    | 2.1                    | 0.0                     | 2.5                    | 0.0                     | 0.0                 |
| C3c-M48             | 4.3                    | 11.1                    | 2.5                    | 0.0                     | 0.0                 |
| C3c1a-Z40439        | 0.0                    | 0.0                     | 0.0                    | 4.3                     | 0.0                 |
| C3c1b-F6379         | 29.8                   | 66.7                    | 50.0                   | 4.3                     | 0.0                 |
| C3f1-M407           | 14.9                   | 0.0                     | 7.5                    | 0.0                     | 0.0                 |
| C3h-F914 (M401)     | 14.9                   | 0.0                     | 7.5                    | 0.0                     | 4.3                 |
| C3c1a-B469          | 0.0                    | 0.0                     | 2.5                    | 0.0                     | 0.0                 |
| G2a-P15             | 2.1                    | 0.0                     | 0.0                    | 0.0                     | 0.0                 |
| J2a-M410            | 4.3                    | 5.6                     | 2.5                    | 0.0                     | 0.0                 |
| N2a1-B478           | 2.1                    | 0.0                     | 0.0                    | 0.0                     | 13.0                |
| N2a1-B525           | 2.1                    | 0.0                     | 0.0                    | 0.0                     | 0.0                 |
| N3a5a-vF4205        | 10.6                   | 5.6                     | 12.5                   | 52.2                    | 15.2                |
| N3b-B187            | 0.0                    | 0.0                     | 0.0                    | 0.0                     | 2.2                 |
| O-M175              | 2.1                    | 0.0                     | 0.0                    | 0.0                     | 0.0                 |
| O2a2-P201           | 2.1                    | 5.6                     | 7.5                    | 0.0                     | 0.0                 |
| O37-M122            | 4.3                    | 0.0                     | 0.0                    | 0.0                     | 0.0                 |
| Q-M242              | 0.0                    | 5.6                     | 0.0                    | 0.0                     | 0.0                 |
| Q2a-M25             | 0.0                    | 0.0                     | 0.0                    | 43.5                    | 50.0                |
| R1a1c-M558          | 0.0                    | 0.0                     | 0.0                    | 0.0                     | 2.2                 |
| R1a2-Z2125          | 2.1                    | 0.0                     | 0.0                    | 0.0                     | 0.0                 |
| R1a2-Z39            | 2.1                    | 0.0                     | 0.0                    | 0.0                     | 8.7                 |
| R2a-M124            | 0.0                    | 0.0                     | 5.0                    | 0.0                     | 0.0                 |

Fig. 1. Bar chart of Y-chromosomal haplogroup distribution among the studied populations.
C4a2a1, C4b, C5a1, C5b1a1, C5d1, D4, D4o2a, F1b1b, G1a1, G2, G2a1e), of which C4b (30.43%) and F1b1b (23.91%) are major haplogroups (Table 2, Figure 2). According to the HVS-1 data, 15 haplotypes were found and the diversity coefficient of 0.8677 turned out to be much lower than among the Torghut of Mongolia (0.9857). In 23 Tsaatans, 14 haplogroups were determined (C4a1, C4a1a, C4b, C5a1, C5b1a1, D4, F1a1, F1b1b, G2, H20a, M7c1, T2, Y1), with C4b (22.73%) and C5a1 (18.18%) being the most common. According to HVS-1, 14 haplotypes were revealed, and the diversity is 0.9486.

### Table 2

| Haplogroups | Tsaatan N = 23 | Tozhu N = 46 |
|-------------|----------------|--------------|
| C4a1        | 4.55           | 2.17         |
| C4a1a       | 4.55           | 0.00         |
| C4a2a1      | 0.00           | 8.70         |
| C4b         | 22.73          | 30.43        |
| C5a1        | 18.18          | 6.52         |
| C5b1a1      | 4.55           | 6.52         |
| C5d1        | 0.00           | 2.17         |
| D4          | 9.09           | 2.17         |
| D4o2a       | 0.00           | 2.17         |
| F1a1        | 4.55           | 0.00         |
| F1b1b       | 4.55           | 23.91        |
| G1a1        | 0.00           | 4.35         |
| G2          | 4.55           | 8.70         |
| G2a1e       | 0.00           | 2.17         |
| H20a        | 4.55           | 0.00         |
| M7c1        | 4.55           | 0.00         |
| T2          | 4.55           | 0.00         |
| Y1          | 9.09           | 0.00         |

### Discussion

The presented diversity indicates the similarity of the gene pools of the Tsaatans and Tozhus with the results obtained for Tuva populations, where a greater number of haplogroups were found, and the most frequent haplogroups are N1b (N3), N1c (N2), Q1a3 (Q1a1b)\(^3\). The phylogeography of haplogroup N is well structured: the subclade N3b-B187 being specific to southern Siberia and Mongolia is found in Kazakhs, Altaians, and Tuvans; N3a5 is restricted to eastern Eurasia, N3a5-F4205 is prominent around Lake Baikal among Mongolic-speaking Buryats and Mongols; the N2a1-B478 group covers the western and southern parts of Siberia, the Taimyr Peninsula and the Volga-Uralic region and does not extend to eastern Siberia\(^3,4\).

Haplogroup Q1a1b-M25 was frequent in Turkic-speaking populations, and its time estimates were at approximately 3–5 Thousand Years Ago\(^3\). It had spread from Central Asia to Western Asia and Central Europe. The results coincided with the Turkic nomadic migrations from Southern Siberia and Mongolia to Central and Western Asia, the Caucasus and Eastern Europe\(^4\). The high level of this contribution to the gene pool of the Tsaatans and Tozhus is evidence of an ancient component that has been preserved in remote taiga regions.

The low frequency of haplogroups C3 and C2, that are widespread in the steppe area of Mongolia, indicates that waves of population displacement associated with nomadic migrations of Mongols penetrated the taiga zone insignificantly (Figure 1).

The haplogroup R1a in southern Siberia demonstrates the geographical variation in occurrence, reaching a maximum of more than 50% in the south of Altai, and up to 18% in Tuva populations\(^3,6\).

The genetic diversity was assessed based on the frequencies of diallelic haplogroups in populations. The minimum diversity in the frequencies of haplogroups was found in the Tozhus of Tuva (0.66) and the Tsaatans of Mongolia (0.56), which is apparently due to the inaccessibility and isolation of small mountain populations.

**Fig. 2.** Mitochondrial DNA haplogroups (%) in the populations of the Tsaatans of Mongolia (a) and the Tozhus of Tuva (b).
Position of the studied populations among the geographical neighbors shows that the Tsaatans of Mongolia and Tozhu are close to the populations of reindeer herders who are at a considerable distance, which indicates a common origin (Table 3).

That is why for a comparative analysis, literature data on geographically neighboring and Northern reindeer herders were taken. Cluster analysis was carried out for a complex of haplogroup frequencies in each studied sample. Haplogroups with the frequency of at least 25% in all cohorts (19 in total) were selected. For clustering, the Ward's method was used. An important property of the Ward procedure is the robustness to the influence of various kinds of statistically random interference. As a result, clusters are obtained for which the internal variability of features is close to the least.

An interesting fact is the unification of the Tsaatans of Mongolia with the Khants (Figure 4), which are located at a significant geographic distance. We can assume that N3a5a and Q2a haplogroups were associated with the resettlement of reindeer herders and further with the Finno-Ugric peoples. The subclade N1a1a1a1a-L392 (N3a) was determined in ancient samples BOO002 and BOO004 (aged about 3500 years ago) from the Oleneostrovsky burial ground in the Murmansk region (Bolshoi Oleny Island, Kola Bay). The oldest sample N1a (subclade N1a1a1a1a-L1026 / Z1973, L392 or N3a3′5) from the Estonian Iron Age dated back to 770–430 BC. Q1a1b-YP1500 / B143>YP1500 * was determined for sample Kolyma1 from the Duvanny Yar site (9769 years ago).

Based on the data on the nucleotide variability of mtDNA HVSI, in samples from populations of Torghut, Derbet, Tozhu and Tsaatan, we determined the following indices: gene diversity (H), absolute number of detected haplotypes (K), number of polymorphic sites (S), average number of pairwise nucleotide differences (Pi), as well as Tajima’s D and Fu’s Fs neutrality tests (Table 4). The highest level of genetic diversity in the populations studied is observed in the Torghuts of Mongolia (0.9857), which also corresponds to the high diversity of haplotypes found in them (43 at N=60). The lowest value is found in the Tsaatan population (0.8677), the mitochondrial diversity of which is represented by haplogroups C4b (22.73%) and C5a1 (18.18%). Statistically significant Fu’s Fs values were obtained for all populations studied. The high reliably negative value of the Fu’s Fs test for selective neutrality in populations indicates exponential growth in population numbers. Negative values of Tajima’s neutrality tests were also obtained for all populations except the Tozhu population, where the value is close to 0, which indicates selective neutrality and a constant population size. Inconsistent neutrality test values for the Tozhu population may be a result of heterogeneous mutation rates in the mtDNA control region, which could mask the signals of population expansion.

The results of the analysis of the Y-chromosome haplogroups suggest that the Tsaatans of Mongolia and the

### Table 3

| Source of variation | % variation | F-statistics | P    |
|---------------------|-------------|--------------|------|
| Among geographic groups | 1.77 | F<sub>CT</sub>=0.01771 | 0.13196 |
| Among populations within groups | 12.89 | F<sub>SC</sub>=0.13119 | 0   |
| Within populations | 85.34 | F<sub>ST</sub>=0.14658 | 0   |
| Among groups according to CA<sup>a</sup> | 15.67 | F<sub>CT</sub>=0.15670 | 0.00293 |
| Among populations within groups | 5.20 | F<sub>SC</sub>=0.06168 | 0   |
| Within populations | 79.13 | F<sub>ST</sub>=0.20872 | 0   |

<sup>a</sup> Two groups: Europe (Kalmyk Derbet, Kalmyk Khoshut, Kalmyk Torghut, Kalmyk Buzava, Xinjiang Kalmyk) and Asia (Mongol Derbet, Mongol Khoshut, Mongol Torghut, Sart Kalmak, Mongol Tsaatan, Tozhu Tuvan).

<sup>b</sup> Three groups according to correspondence analysis results based on haplogroup frequencies: (1) Kalmyk Derbet, Kalmyk Khoshut, Kalmyk Torghut, Kalmyk Buzava, and Tozhu Tuvan; (2) Mongol Torghut, Sart Kalmak, Mongol Khoshut, Mongol Tsaatan, Tozhu Tuvan and (3) Sart Kalmak.

---

**Fig. 3.** Position of the studied populations among geographical neighbors.
Tozhus of Tuva, engaged in reindeer herding, appeared to have common paternal lines with other reindeer breeders (Khanty). Close ethnocultural contacts with Tuvans demonstrate some paternal genetic similarity, which could indicate a common origin. The anthropological research program conducted by us came to similar conclusions. The data obtained on the variability of the Y chromosome and mtDNA in the Tozhus and Tsaatans are due to the isolated and inaccessible taiga region and the manifestation of the "founder effect". The Tsaatans are less polymorphic in terms of the variety of Y chromosome haplogroups, while the Tozhus are less diverse in terms of mitochondrial DNA HVS-1 haplotypes, which is probably a consequence of a high rate of endogamic marriages in the populations studied.

Acknowledgements

This research was partially funded by the RFBR grant № 18-09-00258 and was carried out as part of the State Assignment for the Research Centre of Medical Genetics. The study was also carried out in the framework of research project № AAAA-A19-119013090163-2 – “Anthropology of Eurasian populations (biological aspects)”.

---

**TABLE 4**

| Ethnic groups          | N  | H     | K     | S     | Pi    | Tajima’s D | Fu’s Fs  | East Eurasian mtDNA types, % | West Eurasian mtDNA types, % |
|------------------------|----|-------|-------|-------|-------|------------|---------|------------------------------|-------------------------------|
| Mongols Torghts        | 60 | 0.9857| 43    | 73    | 8.46060 | -1.55987   | -24.78814| 0.8                          | 0.2                           |
| Mongols Derbets        | 38 | 0.9832| 30    | 56    | 7.02091 | -1.63832   | -25.06715| 0.74                         | 0.26                          |
| Tozhus of Tuva         | 47 | 0.8677| 13    | 32    | 6.69843 | -0.25480   | -25.11964| 100                          | 0                             |
| Tsaatans of Mongolia   | 22 | 0.9526| 15    | 33    | 5.78656 | -1.35028   | -22.32441| 0.91                         | 0.09                          |

1 N — sample size; H — genetic diversity; K — number of revealed haplotypes; S — number of polymorphic sites; Pi — average number of nucleotide differences based on paired alignments; Tajima’s D — Tajima’s D test for neutrality (p<0.05); Fu’s Fs — Fu’s Fs test for neutrality (p<0.02).
ANALIZA POLIMORFIZMA UNIPARENTALNIH BILJEGA U POPULACIJAMA KOJE SE BAVE UZGOJEM SOBOVA: TOZHU TUVANCI IZ RUSIJE I TSAATANI IZ MONGOLIJE

SAŽETAK

Analizirali smo podatke o varijabilnosti Y kromosoma i mitohondrijske DNK (mtDNA) u populaciji Tsaatan u Mongoliji i populaciji Tozhu Tuvana u Rusiji. Ispitane populacije karakterizira niska genska raznolikost za oba markera. Analiza haplogrupa Y kromosoma otkrila je tri haplogrupe u populaciji Tsaatan i sedam haplogrupa u skupini Tuvanaca. Sastav haplogrupa podudara se s podacima o Tuvancima iz literature, što se objašnjava zajedničkim podrijetlom. Prema podacima o varijabilnosti mitohondrijske DNK, u 46 Tuvanaca utvrđeno je 12 haplogrupa od kojih su glavne C4b (30,43%) i F1b1b (23,91%). Prema podacima HVS (HyperVariable Segment) –1, 15 haplotipova pronađeno je kod Tozhu Tuvanaca, a koeficijent raznolikosti od 0,8677 pokazao se mnogo nižim nego kod Torghuta iz Mongolije (0,9857). U 23 Tsaatana otkriveno je 14 haplogrupa; od kojih su najčešći C4b (22,73%) i C5a1 (18,18%). Prema HVS–1, kod Tsaatana je otkriveno 14 haplotipova čija raznolikost je 0,9486. Podaci dobiveni o uniparentalnim markercima sustavima Tuvanaca i Tsaatana posljedica su izolirane i nepristupačne regije tajge i manifestacije efekta utemeljitelja. Tsaatani su manje polimorfni u pogledu raznolikosti haplogrupa Y kromosoma, dok su Tuvanci manje polimorfni u pogledu mitohondrijske DNK, što je vjerojatno posljedica visoke stope endogamnih brakova u proučavanim populacijama.