Russian Journal of Genetics, Vol. 40, No. 11, 2004, pp. 1292–1299. Translated from Genetika, Vol. 40, No. 11, 2004, pp. 1562–1570.

HUMAN GENETICS

Restriction Polymorphism of Mitochondrial DNA in Koreans and Mongolians

M. V. Derenko[1], A. V. Lunkina[1], B. A. Malyarchuk[1], I. A. Zakharov[2], Ts. Tsedev[3], K. S. Park[4], Y. M. Cho[4], H. K. Lee[4], and Ch. H. Chu[5]

[1] *Institute of the Biological Problems of the North, Russian Academy of Sciences, Magadan, 685000 Russia; fax: (41322) 344-63; e-mail: mderenko@ibpn.kolyma.ru*
[2] *Vavilov Institute of General Genetics, Russian Academy of Sciences, Moscow, 119991 Russia*
[3] *Institute of Biological Sciences, Mongolian Academy of Sciences, Ulanbatar, 510351 Mongolia*
[4] *Seoul National University, College of Medicine, Seoul, 151-742 Korea*
[5] *Department of History, Kangwon National University, Chunchon, 200-701 Korea*

Received February 12, 2004; in final form, July 1, 2004

Abstract—Using the data on mitochondrial DNA (mtDNA) restriction polymorphism, the gene pools of Koreans ($N = 164$) and Mongolians ($N = 48$) were characterized. It was demonstrated that the gene pools were represented by the common set of mtDNA haplogroups of East Asian origin (M*, M7, M8a, M10, C, D4, G*, G2, A, B*, B5, F1, and N*). In addition to this set, mtDNA haplogroups D5 and Y were identified in Koreans while Mongolians possessed haplogroup Z. Only in Mongolians, a European component with the frequency of 10.4% and represented by the mtDNA types belonging to haplogroups K, U4, and N1, was identified. Phylogenetic and statistical analyses of the data on mtDNA variation in the populations of South Siberia, Central, and East Asia suggested the existence of interpopulation differentiation within these regions, the main role in which was played by the geographical and linguistic factors. Analysis of the pairwise F_{ST} distances demonstrated close genetic similarity of Koreans to Northern Chinese, which in turn, were clearly different from Southern Chinese populations. Mongolians occupied an intermediate position between the ethnic groups of South Siberia and Central/East Asia.

INTRODUCTION

East Asia is one of several world regions, where the first stages of the development of anatomically modern humans took place. According to paleontological data, primary eastern race-forming area was developed on the territory of the Yellow River and the Yangtze River basins as early as during Lower Paleolithic [1]. It is suggested that later (during Upper Paleolithic) the development of Mongoloid characters, differentially expressed in the present-day populations inhabiting the vast territory of Asia began on this area. Analysis of paleontological data showed that the Central Asian race-forming area was developed later (during the Neolithic) within the territory, limited by southern steppe regions of Transbaikalia, central and eastern regions of Mongolia, and, probably, by some regions of Northern China. Beginning from the Bronze Age, Caucasoid groups started colonization of the territories of Central Asia (till Western Mongolia) and the south of Siberia. The result of this was the appearance (at the turn of the Common Era) of the South Siberian race-forming area on the territories of South Siberian steppe and the north of Central Asia. Within this area, intensive mixing between Caucasoids and Mongoloids took place. One of such local population groups has emerged in the Altai–Sayan Region [1–3].

At the same time, on the territory of China, the most important region relative to the understanding of the early stages of the Mongoloid race development, there are still no paleontological findings, covering a large time interval (from 40 to 100 thousand years ago) [4]. This is one of the reasons for such great interest to the results of molecular genetic investigations, permitting reconstruction of prehistoric (racial genetic and demographic) events based on the genetic analysis of the present-day and ancient populations. The studies of mitochondrial DNA (mtDNA) and Y-chromosome variability in East Asian populations showed that the primary colonization of this region by modern humans happened about 60 thousand years ago, during migrations of ancestral populations from Africa to the south of East Asia (in accordance with the "southern wave" model) [4, 5]. The data obtained suggested that colonization of East Asia occurred from the south to the north and that only during Upper Paleolithic these populations have reached the north of China, Mongolia, and Siberia. The existing genetic differences between southern and northern East Asian populations were probably caused by the loss of genetic diversity during the process of migration of southern populations to the north of Asia [4]. Thus, one of the main tasks of ethnic genetics is investigation of genetic differentiation of the populations of East and Central Asia with the purpose of determining the levels of the regional gene pools dif-

ferentiation. In future, this will enable localization of the areas of the formation of long-term genetic diversity, corresponding to different stages of East Asian population differentiation. Also, it will be possible to reconstruct the main stages in colonization of Central and Northern Asia.

The data on the mtDNA variation in the populations of China suggest the existence of genetic differentiation between the northern and southern Chinese populations. Southern and southeastern Chinese populations demonstrate higher diversity of the mtDNA types compared to the populations of Central, East, and Northeast China. The age of a number of mtDNA haplogroups described in the populations of China is estimated as more than 50 000 years, confirming the hypothesis on the colonization of the territory of China about 60 000 years ago [6].

At the same time, many populations of East and Central Asia, for example, Koreans and Mongolians, remain poorly studied in respect of the mtDNA markers distribution. Investigation of classical genetic markers showed that Koreans displayed maximum genetic similarity to Mongolians [7, 8]. On the contrary, the data on mtDNA variation pointed to genetic closeness of Koreans to Chinese and Japanese [9, 10]. Recent studies of the Y-chromosome markers polymorphism showed that the gene pool of Koreans contained the lineages prevalent both in Northeast and Southeast Asia. Moreover, relative to the male gene pool structure, Koreans displayed maximum genetic similarity to Manchurians, Southern Chinese from the province of Yunnan, and Vietnamese [11]. It can be thus suggested that the population of Korea was formed as a result of multiple migration events with the substantial genetic contribution from the northern Asian groups, and also from the populations inhabiting the south and north of China [11, 12]. It is also suggested that the ancestors of Koreans could have common origin with the ethnic groups of Northeast Asia, which inhabited vast territories of the Altai–Sayan and Baikal regions of Southeast Siberia [13].

Mongolians inhabit the territories proposed to represent the origin of Central Asian anthropological type, prevalent in addition to Mongols, among the majority of South and East Siberian ethnic populations [3]. Investigations of the mtDNA variation revealed higher level of the mitochondrial gene pool diversity in the populations of Mongolia compared to the populations of Siberia and Southeast Asia [14]. In the meantime, the data on biochemical markers and nuclear loci variation along with those on the polymorphism of the main noncoding region of mtDNA, point to an extremely low level of genetic differentiation characteristic of Mongolian populations [14–16]. The low level of interpopulation differentiation in Mongolians can be probably explained by specific features of their economic and cultural life style, and also by active integration and assimilation processes, which accompanied the formation of the Mongolian ethnic groups. Archaeological and anthropological data suggest that the territory of Mongolia could have been already colonized during Upper Paleolithic. These data also point to the succession of archaeological cultures observed within this region till the early Middle Ages [3, 5]. Mongolian population was formed under the influence of different groups (Hunnu, Syanbi, Jujan, Uighurs, Kidans, etc.), ethnolinguistic and anthropological affiliation of which still remains controversial. For instance, recent investigations of skeletal remains from the necropolis in the Egyin Gol Valley (Northern Mongolia) showed that the population of Xiongnu period (3rd century B.P.–2nd century A.D.) was characterized by mixed anthropological composition, since 11% of the mitochondrial lineages, identified in the study cited, belonged to western Eurasian mtDNA haplogroups [17]. Their genetic succession, however, cannot be established due to the lack of data on mtDNA polymorphism among the ancient and present-day population of Mongolia and adjacent territories.

The present study, thereby, was focused on the analysis of the structure and diversity of the mitochondrial gene pools of Mongolians and Koreans, as well as on the evaluation of the population differentiation patterns of the ethnic populations from South Siberia, Central, and East Asia.

MATERIALS AND METHODS

The sample of Mongolians (48 individuals) was formed from the inhabitants of different aimaks of Mongolia. The sample of Koreans (164 individuals) was comprised of the inhabitants of different provinces of Korea. Only individuals who were apparently unrelated at least for three generations participated in the study. Genomic DNA used as the study material was isolated from the biological tissue samples (blood and hair bulbs) using standard techniques [18].

Screening for polymorphic sites determining the main haplogroups of mtDNA types prevalent in the populations of Eurasia (Table 1) was conducted using restriction endonuclease analysis of mtDNA fragments amplified in polymerase chain reaction with the primers suggested in [19, 20]. Restriction fragments were fractionated by electrophoresis in 8% polyacrylamide gels. Gels were stained with ethidium bromide and DNA fragments were visualized in the UV light.

Polymorphism was scored by the presence (+) or absence (–) of the restriction endonuclease recognition sites. The mtDNA types were identified in accordance with the existing classification of the haplogroups of mtDNA types in the populations of Eurasia [6, 10, 21, 22]. Haplogroups and subhaplogroups of mtDNA types are defined as monophyletic clusters marked by certain polymorphism variants (Table 1). In accordance with the classification, mtDNA haplogroups are denoted by single Roman letters (excluding haplogroup HV), and

Table 1. Polymorphic restriction variants determining haplogroups of mtDNA types in the populations of Eurasia

MtDNA haplogroup	Key restriction variants
M*	–12704 *Mbo*II; +10394 *Dde*I; +10397 *Alu*I
C	–12704 *Mbo*II; +10394 *Dde*I; +10397 *Alu*I; –13259 *Hinc*II/+13262 *Alu*I
D*	–12704 *Mbo*II; +10394 *Dde*I; +10397 *Alu*I; –5176 *Alu*I
D4	–12704 *Mbo*II; +10394 *Dde*I; +10397 *Alu*I; –5176 *Alu*I; –3010 *Fnu*DII
D5	–12704 *Mbo*II; –10394 *Dde*I; –10397 *Alu*I; –5176 *Alu*I
E	–12704 *Mbo*II; +10394 *Dde*I; +10397 *Alu*I; –7598 *Hha*I
G*	–12704 *Mbo*II; +10394 *Dde*I; +10397 *Alu*I; +4830 *Hae*II/+4831 *Hha*I
G2	–12704 *Mbo*II; +10394 *Dde*I; +10397 *Alu*I; +4830 *Hae*II/+4831 *Hha*I; –7598 *Hha*I
M7	–12704 *Mbo*II; +10394 *Dde*I; +10397 *Alu*I; +9820 *Hin*fI
M8a	–12704 *Mbo*II; +10394 *Dde*I; +10397 *Alu*I; –16298 *Mse*I; +14465 *Acc*I
M10	–12704 *Mbo*II; +10394 *Dde*I; +10397 *Alu*I; +10646 *Rsa*I
Z	–12704 *Mbo*II; +10394 *Dde*I; +10397 *Alu*I; –16298 *Mse*I
A	–12704 *Mbo*II; +663 *Hae*III
Y	–12704 *Mbo*II; +10394 *Dde*I; +7933 *Mbo*I
N*	–12704 *Mbo*II
N1	–12704 *Mbo*II; –12498 *Nla*III
I	–12704 *Mbo*II; +10394 *Dde*I; –12498 *Nla*III; +10032 *Alu*I; +16389 *Bam*HI
W	–12704 *Mbo*II; –8994 *Hae*III
X	–12704 *Mbo*II; +14465 *Acc*I
R*	+12704 *Mbo*II
B*	+12704 *Mbo*II; 9-bp deletion in region V
B5	+12704 *Mbo*II; +10394 *Dde*I; 9-bp deletion in region V
F1	+12704 *Mbo*II; –12406 *Hpa*I/*Hinc*II
U*	+12704 *Mbo*II; +12308 *Hin*fI
U4	+12704 *Mbo*II; +12308 *Hin*fI; +4646 *Rsa*I
K	+12704 *Mbo*II; +10394 *Dde*I; +12308 *Hin*fI; –9052 *Hae*II
HV	+12704 *Mbo*II; –14766 *Mse*I
V	+12704 *Mbo*II; –14766 *Mse*I; +4577 *Nla*III
H	+12704 M*bo*II; –14766 *Mse*I; –7025 *Alu*I
T*	+12704 *Mbo*II; +15606 *Alu*I
T1	+12704 *Mbo*II; +15606 *Alu*I; –12629 *Ava*II
J	+12704 *Mbo*II; +10394 *Dde*I; –13704 *Bst*NI

Note: Haplogroups of mtDNA types are designated according to classification from [6, 10, 21, 22].

subhaplogroups within the haplogroups, by the digits added to the letter code of the group (Table 1). The mtDNA types that belong to a particular haplogroup, but yet cannot be attributed to any of the known subhaplogroups, are designated by asterisks.

Statistical significance of interpopulation differences relative to the frequencies of mtDNA haplogroups were evaluated using an exact test of population differentiation [23]. The indices of mtDNA population diversity, as well as *F*-statistics values were calculated using ARLEQUIN 2.000 software package [24]. Correlation between the genetic, geographical, and linguistic distances was evaluated using Mantel's test (100 permutations). Geographical distances were calculated using Great Circle Distance program (http://www.mercury.demon.co.uk/dist/dodist.html) through the comparison of geographical coordinates. The matrix of interpopulation linguistic distances was constructed in accordance with the methods suggested in [25].

For comparative analysis, the data on the frequencies of mtDNA haplogroups in Southern Altaians (Altai-Kizhi), Khakassians, Buryats, Sojots, Todjins, Tuvinians, Tofalars, and Western Evenks, reported by us earlier [26, 27], as well as the data for Hun Chinese

from northern (Shandong, Liaoning, and Xinjiang [6]) and southern (Yunnan, Hubei, and Guangdong (the cities of Guangzhou [10] and Zhanjiang [6]) provinces of China were used. The matrix of pairwise F_{ST} distances was used for the presentation of phylogenetic relationships between the populations by means of multivariate scaling methods realized in STATISTICA 5.0 (StatSoft, Inc. Tulsa, OK, United States) software package.

RESULTS AND DISCUSSION

The present study continues description of the mitochondrial gene pool structures in ethnic populations of Northern Asia. Earlier, based on the hypervariable segment 1 (HVS1) sequence variation data and/or restriction analysis of the mtDNA coding sequences, a comprehensive characterization of the mitochondrial gene pools of ten ethnic population groups from South and East Siberia, including Southern Altaians (Altai-Kizhi), Khakassians, Tuvinians, Todjins, Tofalars, Buryats, Sojots, Western Evenks, Shorians, and Yakuts, was carried out [26, 27]. The total size of the sample examined was 663 individuals. In the present study, using high-resolution restriction analysis of mtDNA, which included screening of 32 polymorphic sites, determining 32 haplogroups and subhaplogroups of mtDNA, prevalent in the populations of Eurasia (Table 1), mitochondrial genomes of 212 individuals representing two ethnic groups of Central Asia (Mongolians) and East Asia (Koreans) were characterized.

The data on the prevalence of mtDNA haplogroups and subhaplogroups among Koreans and Mongolians are demonstrated in Table 2. As shown in the table, the populations examined were characterized by similar composition of mtDNA haplogroups, which was also typical of the other Mongoloid populations of Asia [6, 10, 26, 27, 29]. More than the half of mtDNA lineages in each of the populations (58.5% in Koreans and 54.2% in Mongolians) belonged to different haplogroups of Asian macrohaplogroup M, which included haplogroups C, D, E, G, M7, M8, M9, M10, M*, and Z. It should be noted in this respect that in Koreans, high frequency (34.76%) of mtDNA lineages belonging to haplogroup D along with the absence of the lineages belonging to haplogroup C was observed. At the same time, the gene pool of Mongolians was characterized by relatively high frequencies of haplogroups C and D (16.67 and 10.42%, respectively). Note that in Mongolians only D4 mtDNA types were identified, while in Koreans haplogroup D also included mtDNA variants of subhaplogroup D5, which was mostly distributed among the populations of China (with the frequencies ranging from 4.8 to 10%), and with the low frequencies (1.8–2.2%) in the populations of South Siberia [6, 27]. Single mtDNA variants belonging to haplogroup Z, mostly prevalent in Northeast Asia, were detected only in the Mongolian sample. Both samples were characterized by low frequencies (1 to 2%) of mtDNA types from haplogroup M10, which was earlier detected in the populations of China with similarly low frequencies (1.3 to 5.9%) [6]. Single variants of mtDNA lineages belonging to haplogroup M10 were earlier described in the gene pools of Altaians and Buryats [27].

Table 2. Prevalence of mtDNA haplogroups in Koreans and Mongolians

MtDNA haplogroup	Koreans (N = 164)		Mongolians (N = 48)	
	N	%	N	%
M*	3	1.83	2	4.17
M7	13	7.93	2	4.17
M8a	9	5.49	2	4.17
M10	2	1.22	1	2.08
C	1	0.61	8	16.67
D4	46	28.04	5	10.42
D5	11	6.71	0	0
G*	7	4.27	1	2.08
G2	4	2.44	4	8.33
Z	0	0	1	2.08
A	15	9.15	6	12.50
Y	2	1.22	0	0
B*	17	10.36	5	10.42
B5	17	10.36	2	4.17
F1	8	4.88	3	6.25
N*	9	5.49	1	2.08
N1	0	0	1	2.08
U4	0	0	3	6.25
K	0	0	1	2.08
$h \pm$ s.e.	0.873 ± 0.015		0.930 ± 0.016	

The frequencies of haplogroup G mtDNA lineages in the gene pools of Koreans and Mongolians were 6.71 and 10.42%, respectively. Note that haplogroup G in Koreans was mostly composed of the mtDNA types belonging to haplogroup G* (4.27%), while higher frequency of haplogroup G2 mtDNA variants was observed in Mongolians (8.33%). Earlier, it was demonstrated that subhaplogroup G2 was mostly prevalent in Central Asia (8.8% as reported in [10]), while it was practically absent in China and Southeast Asia. In South Siberia, maximum frequencies of mtDNA types from haplogroup G2 were observed in Mongolian-speaking Buryats and also in genetically close to them Sojots (14.3 and 6.7%, respectively) [27]. Thus, the origin and distribution of mitochondrial subhaplogroup G2 could be associated with the Mongolian-speaking populations of Central Asia.

MtDNA lineages belonging to East Asian haplogroup M7 were present in the gene pools of Koreans and Mongolians with the frequencies of 7.93 and 4.17%, respectively. Earlier, it was shown that haplogroup M7 prevailed in China, Vietnam, the Korean Pen-

insula, and in the islands of Southeast Asia, while it was absent in Northeast and Central Asia [10]. The coalescence time for different subhaplogroups of haplogroup M7 is estimated as ranging between 6000 and 18 000 years, which corresponds to the time of the settlement process in the area around the southern coast of Japanese Sea after the Last Glacial Maximum and before the onset of the Jomon culture [10].

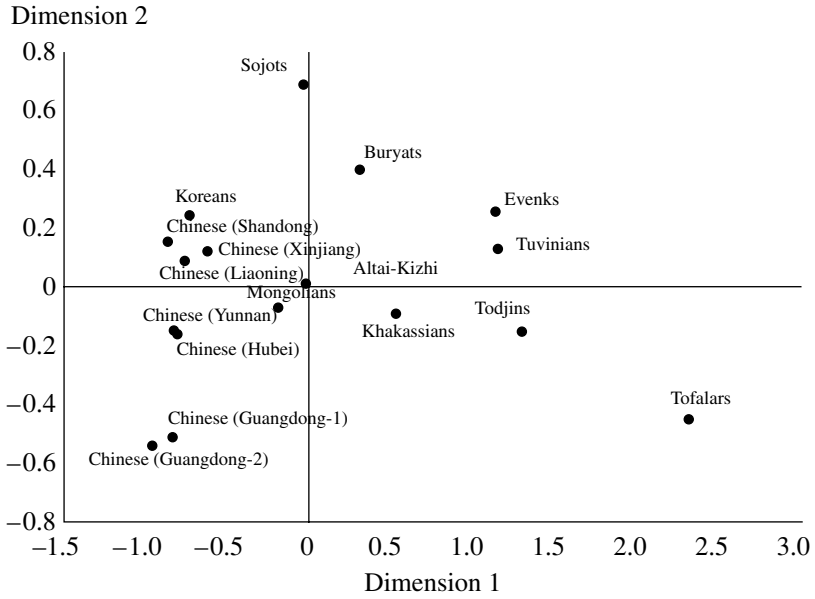
In Korean and Mongolian samples the frequency of mtDNA lineages belonging to three other haplogroups (A, B, and F) of macrohaplogroup R and distributed predominantly among Mongoloid populations, constituted 34.8 and 33.3%, respectively. Relatively high frequencies of haplogroup A, observed in Koreans (9.15%) and Mongolians (12.5%) bring these populations close to the populations of China, where the frequencies of mtDNA variants belonging to this haplogroup vary from 4.7 to 16.7% [6]. Haplogroup A was found in the populations of South Siberia with the averaged frequency of 2.7%, and reaching its maximum values in Sojots (10%) [27].

The gene pools of both, Koreans and Mongolians, were characterized by high prevalence of mtDNA types belonging to haplogroup B, distinguished by the presence of the 9-bp deletion in the mtDNA region V. The frequency of B lineages in Koreans was 20.7%, and 14.6% in Mongolians. Note that high frequencies of haplogroup B, varying in the range from 12 to 30.4%, were found in all samples of Northern and Southern Chinese examined [6]. The frequency of mtDNA B types in the Mongolian sample examined earlier was 9.7% [14]. In the populations of South Siberia the frequencies of haplogroup B did not exceed the maximum value of 7.8% observed in Tuvinians [26]. Haplogroup B in the populations studied was represented by two subhaplogroups, B* and B5. In addition to the 9-bp deletion in the region V, subhaplogroup B5 is characterized by the presence of the +10 394 *Dde*I variant. In the gene pool of Koreans, both subgroups were represented with equal frequencies (10.37%), while in Mongolians, mtDNA variants belonging to subhaplogroup B* prevailed (10.42%). Predominant distribution of subhaplogroup B*, represented by the B4, B4a, and B4b variants, was observed earlier in the populations of Northern and Southern Chinese. The exclusion is the Chinese population from the province Xinjiang, in the gene pool of which the frequency of the mtDNA B5 lineages exceeds the frequency of the B* lineages (6.4 versus 2.1%) [6].

Note that B* mtDNA types are nearly equally frequent both in the south and east of Siberia, while the presence of subhaplogroup B5 (with the frequency lower than 2.5%) is typical only of the gene pools of Altaians, Buryats, and Sojots [26, 27]. Earlier, it was demonstrated that mitochondrial haplogroup B represented the most ancient component of the Asian gene pools (with the evolutionary age of about 56 000 years), with the possible place of origin and diversification lying in South China and/or Southeast Asia [6]. Thus, the high frequency and the marked structure of haplogroup B in the populations of Koreans and Mongolians indicate a high diversity level and the ancient age of this gene pool component. This, in turn, favors the earlier suggestion on the existence of a mitochondrial gene pool diversity-forming center of Mongoloid populations from Northern Asia on the territories of Mongolia, Korea, and China [28].

Mitochondrial haplogroup F1, represented in the populations of Siberia, Central, East, and Southeast Asia by subhaplogroups F1a, F1b, and F1c occurred detected in the gene pools of the Korean and Mongolian populations at frequencies of 4.88 and 6.25%, respectively. Earlier, it was demonstrated that F1a lineages prevailed in the populations of Southeast Asia, while F1b mtDNA types were more frequently observed among the populations of Central and East Asia [10]. In Chinese, the frequency of F1 variants increased southward, from the minimum values (4%) in the province of Liaoning to the maximum values (16.3 and 20.2%, respectively) in the populations of Yunnan and Guangdong (Guangzhou) [6]. In the populations of South Siberia, F1 mtDNA lineages had high frequencies (22 and 43%, respectively) in the gene pools of Khakassians and Shorians, while in Altaians their frequency slightly exceeded 8% [26, 27, 29].

In the present study, Caucasoid mtDNA lineages belonging to haplogroups U and N1 were detected only in Mongolians. The frequency of U4 mtDNA variants in the gene pool of Mongolians was 6.25%. These variants were earlier observed among the populations of Altai–Sayan Region, Altaians and Khakassians, with the frequencies of 5.5 and 9.4%, respectively [27]. In addition, a single mtDNA variant belonging to haplogroup K was detected in Mongolians. This haplogroup was earlier found in some Siberian populations, Shorians, Western Evenks, and Altai-Kizhi, with the frequencies of 2.4, 2.5, and 6.5%, respectively [26, 29]. Haplogroup N1, detected in Mongolians with the frequency of 2.08%, was earlier found only in Altaians with similar frequency of 2.7% [26]. In general, it should be noted that Caucasoid component of the Altai–Sayan gene pools displays substantial diversity, since it is represented by mtDNA lineages belonging to haplogroups and subhaplogroups H, U2, U3, U4, U5a, U5b, K, J, T*, T1, I, X, R*, N1a, and HV* [26]. Caucasoid lineages belonging to haplogroups H, V, J, T*, T1, U2, U4, U5a, and U7 are also characteristic of the gene pools of Northwest Siberian populations, Mansi, Nganasans, and Kets [30, 31]. It should be noted, that excluding some common mtDNA types (belonging to haplogroups U4, U5a, K, and H), Caucasoid components of the Altai–Sayan and West Siberian gene pools are substantially different, probably, due to more pronounced influence on the latter from the side of East European populations. Thus, high frequencies and diversity of Caucasoid mtDNA lineages in South Siberian gene pools support the hypothesis (based on paleontological



Multivariate scaling of F_{ST}-distances between the populations of South Siberia, Central, and East Asia. The samples of Southern Chinese from Guangzhou and Zhanjiang, province Guangdong are designated as "Guangdong-1" and "Guangdong-2", respectively.

data) on the early (during the Neolithic, Bronze Age, and the early Iron Age) penetration of Caucasoids into South Siberia [2, 3]. In turn, the presence of Caucasoid U4 lineages in the Mongolian mitochondrial gene pool enables suggestion that Caucasoids have also penetrated into Mongolia.

Paleoanthropological findings of Pre-Scythian time from the things-free burial sites in western Mongolia, and also the findings of later periods, which are characterized by Caucasoid and/or mixed characters enabled substantial expansion of the Caucasoid race area, including in it not only the Caucasus, Central Asia, and South Siberian steppe, but also West Mongolia [3]. For these reasons, taking into consideration the absence of Caucasoid mtDNA lineages in the gene pools of Northern and Southern Chinese examined so far [6, 10], along with the absence of these lineages in Koreans and Japanese [9], it can be considered that the territory of Mongolia represents the eastern border of the maternal Caucasoid lineages distribution in East Eurasia.

Mitochondrial gene pools of the populations examined were characterized by similar diversity values, the h values for Koreans and Mongolians constituted 0.873 and 0.930, respectively. These mtDNA diversity values are within the range of h values (0.603 to 0.930) observed in the populations of South Siberia, and Central and East Asia. Moreover, Mongolians, similarly to Liaoning (h = 0.929) and Hubei (h = 0.925) Chinese, and also Altai-Kizhi (h = 0.924) were characterized by the maximum diversity values among the populations examined so far.

Comparative analysis of the mtDNA haplogroup and subhaplogroup frequency distribution patterns in Koreans, Mongolians, Northern and Southern Chinese, and in the ethnic groups of South Siberia revealed high level of interpopulation differentiation (F_{ST} = 8.05%; $P < 0.001$). Moreover, despite the fact that in the populations examined similar sets of mtDNA haplogroups were observed, statistically significant differences ($P < 0.05$) were observed between almost all population pairs compared. For instance, Koreans were statistically significantly different from Mongolians, and also from all South Siberian and Chinese populations, excluding the sample of Northern Chinese from Liaoning. Mongolians were statistically significantly different from almost all other populations used in comparative analysis, excluding Southern Chinese from Hubei and Northern Chinese from Liaoning. In turn, Southern and Northern Chinese were statistically significantly different from each other. The exclusion was two samples from the northern provinces of Xinjiang and Liaoning, which displayed no genetic differences from each other, and also from Northern Chinese from Shandong, as well as from Southern Chinese from Yunnan and Hubei. Taken together, the data obtained indicate that Central and East Asian populations are characterized by the high levels of genetic diversity along with substantial gene pool differentiation. High degree of interpopulation differentiation is also typical of the indigenous populations of Siberia [27].

The data on mtDNA haplogroup and subhaplogroup frequency distribution in 17 populations of South Siberia, Central and East Asia, including Korean and Mongolian populations examined in the present study, were used for the investigation of the evolutionary relationships between the ethnic groups of these regions. The data on multivariate scaling of the pairwise F_{ST}-distance matrix are presented in the figure. As can be seen,

central position in the space of two dimensions is occupied by Mongolians and Altai-Kizhi, who are characterized by maximum genetic diversity values. Tofalars and Sojots, demonstrating minimum levels of mtDNA gene pool diversity among the populations compared occupy the marginal position on the map. The position of South Siberian populations reveals no distinct clusters, while in general, in the first dimension they relatively clearly differentiate form the Central and East Asian ethnic groups. The second dimension, in turn, clearly differentiates Northern and Southern Chinese populations. The Korean sample tested forms a separate cluster together with Chinese from northern provinces of Shandong, Xinjiang, and Liaoning. Southern Chinese populations form two distinct clusters, one of which includes Shandong, Xinjiang, and Liaoning samples, while another one contains two samples from the southernmost province of Guangdong. In general, the population distribution follows a geographical, rather than linguistic or anthropological, pattern. Because of this, it seems reasonable to carry out further investigations of the relationships between the genetic diversity and geographical and linguistic characteristics of the populations in the region examined.

As follows from molecular variance analysis (AMOVA) of mtDNA haplogroups characterizing the gene pools of the populations examined, the main role in the observed differentiation of South Siberian, Central and East Asian populations was played by geographical and linguistic factors. Combination of the populations into regional groups showed that the proportion of intergroup variation was remarkably higher compared to interpopulation differences within the groups (6.72 versus 2.76%). Clustering of the populations according to linguistic principle showed that the proportion of the observed intergroup differences was also higher compared to the values of the interpopulation differences within the groups (6.12 versus 4% for linguistic families, and 4.89 versus 4.06% for the groups within linguistic families). Analysis of the correlation between genetic distance matrices (F_{ST}-distances) and the matrices of geographical and linguistic distances by use of Mantel's test pointed to the existence of statistically significant correlation between the genetic and geographical distances ($r = 0.429$; $P < 0.001$). The partial coefficient of correlation between the genetic and geographical distances (with constant linguistic differences) was also statistically significant, but at lower significance level ($r = 0.241$; $P = 0.02$). There was also statistically significant correlation between the genetic and linguistic distances ($r = 0.387$; $P = 0.003$). However, statistically significant correlation between genetic and linguistic data is not revealed in case of geographical distance constancy ($r = 0.138$; $P = 0.087$), thereby indicating the determining role of the geographical factors in the development of genetic diversity of the indigenous populations of this region of Asia.

In conclusion, the approach applied in this study provided detailed characterization of the mtDNA gene pool structures in Koreans and Mongolians, and also enabled determination of their phylogenetic position among South Siberian, Central, and East Asian ethnic groups. It was demonstrated that mitochondrial gene pool of Koreans was represented exclusively by East Asian mtDNA lineages. At the same time, the gene pool of Mongolians displayed the presence of Caucasoid mtDNA lineages with the frequency of 10.4%, favoring the hypothesis on early penetration of Caucasoids into the territory of Central Asia. Comprehensive phylogeographic sequence analysis of mitochondrial lineages revealed among the populations of Central and East Asia will provide future determination of the origins of all gene pool components along with the adequate description of the genetic processes in the populations studied, as well as the identification of the factors (evolutionary, population, and historical) that had the most important influence on the formation of the populations of different regions of Northern Eurasia.

ACKNOWLEDGMENTS

This work was supported by the Russian Foundation for Basic Research (grant no. 04-04-48746) and the Russian State Program "Research and Developments within the Priority Directions of Science and Technologies," section "Technology of the Living Systems," subsection "Biology," theme "Human, Animal, Plant, and Microbial Genome Databases."

REFERENCES

1. Alekseev, V.P., *Geograficheskie ochagi formirovaniya chelovecheskikh ras* (Geographical Foci of the Formation of Human Races), Moscow: Mysl', 1985.
2. Alekseev, V.P., *Istoricheskaya antropologiya i antropogenez* (Historical Anthropology and Anthropogenesis), Moscow: Nauka, 1989.
3. Alekseev, V.M. and Gokhman, I.I., *Antropologiya aziatskoi chasti SSSR* (Anthropology of the Asian Part of the Soviet Union), Moscow: Nauka, 1984.
4. Jin, L. and Su, B., Natives or Immigrants: Modern Human Origin in East Asia, *Nat. Rev. Genet.*, 2000, vol. 1, pp. 126–133.
5. Cavalli-Sforza, L.L., Menozzi, P., and Piazza, A., *The History and Geography of Human Genes*, Princeton, New York: Princeton Univ. Press, 1994, pp. 195–252.
6. Yao, Y.-G., Kong, Q.-P., Bandelt, H.-J., *et al.*, Phylogeographic Differentiation of Mitochondrial DNA in Han Chinese, *Am. J. Hum. Genet.*, 2002, vol. 70, pp. 635–651.
7. Saha, N. and Tay, J.S.H., Origin of Koreans: A Population Genetic Study, *Am. J. Phys. Anthropol.*, 1992, vol. 88, pp. 27–36.
8. Hong, S.S., Suh, J.A., Chae, J.J., *et al.*, Frequency Distribution of Alleles at D1S80 and Apo-B 3'-VNTR Loci in Korean Population, *Mol. Cell*, 1993, vol. 3, pp. 457–463.

9. Horai, S., Murayama, K., Hayasaka, K., *et al.*, mtDNA Polymorphism in East Asian Populations, with Special Reference to the Peopling of Japan, *Am. J. Hum. Genet.,* 1996, vol. 59, pp. 579–590.
10. Kivisild, T., Tolk, H.-V., Parik, J., *et al.*, The Emerging Limbs and Twigs of the East Asian mtDNA Tree, *Mol. Biol. Evol.*, 2002, vol. 19, pp. 1737–1751.
11. Jin, H.-J., Kwak, K.-D., Hammer, M., *et al.*, Y-Chromosomal DNA Haplogroups and Their Implications for the Dual Origin of the Koreans, *Hum. Genet.*, 2003, vol. 114, pp. 27–35.
12. Karafet, T.M., Xu, R., 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.
13. Kim, K., People and Language, in *Korea—Its People and Cultures*, Kim, K., Ed., Seoul: Hakwonsa, 1970, pp. 10–20.
14. Kolman, C.J., Sambuughin, N., and Bermingham, E., Mitochondrial DNA Analysis of Mongolian Populations and Implications for the Origin of New World Founders, *Genetics*, 1996, vol. 142, pp. 1321–1334.
15. Sambuughin, N., Petrishchev, V.N., and Rychkov, Yu.G., DNA Polymorphism in the Mongolian Population: RFLP Analysis of Mitochondrial DNA, *Genetika* (Moscow), 1991, vol. 27, no. 11, pp. 2143–2151.
16. Sambuughin, N., Petrishchev, V.N., and Rychkov, Yu.G., DNA Polymorphism in the Mongolian Population: RFLP Analysis of Seven Loci of the Nuclear Genome, *Genetika* (Moscow), 1991, vol. 27, no. 11, pp. 2137–2142.
17. Keyser-Tracqui, C., Crubezy, E., and Ludes, B., Nuclear and Mitochondrial DNA Analysis of a 2000-Year-Old Necropolis in the Egyin Gol Valley of Mongolia, *Am. J. Hum. Genet.,* 2003, vol. 73, pp. 247–260.
18. Walsh, P.S., Metzger, D.A., and Higuchi, R., Research Report: Chelex 100 as a Medium for Simple Extraction of DNA for PCR-Based Typing from Forensic Material, *BioTechniques*, 1991, vol. 10, pp. 506–513.
19. Finnila, S., Hassinen, I., Ala-Kokko, L., and Majamaa, K., Phylogenetic Network of the mtDNA Haplogroup U in Northern Finland Based on Sequence Analysis of the Complete Coding Region by Conformation-Sensitive Gel Electrophoresis, *Am. J. Hum. Genet.*, 2000, vol. 66, pp. 1017–1026.
20. Torroni, A., Huoponen, K., Francalacci, P., *et al.*, Classification of European mtDNAs from an Analysis of Three European Populations, *Genetics*, 1996, vol. 144, pp. 1835–1850.
21. Macaulay, V., Richards, M., Hickey, E., *et al.*, The Emerging Tree of West Eurasian mtDNAs: A Synthesis of Control-Region Sequences and RFLPs, *Am. J. Hum. Genet.*, 1999, vol. 64, pp. 232–249.
22. Kong, Q.-P., Yao, Y.-G., Sun, C., *et al.*, Phylogeny of East Asian Mitochondrial DNA Lineages Inferred from Complete Sequences, *Am. J. Hum. Genet.*, 2003, vol. 73, pp. 671–676.
23. Raymond, M. and Rousset, F., An Exact Test of Population Differentiation, *Evolution*, 1995, vol. 49, pp. 1280–1283.
24. Schneider, S., Roessl, D., and Excoffier, L., *Arlequin V. 2.000: A Software for Population Genetics Data Analysis*, Geneva: Univ. of Geneva, 2000.
25. Poloni, E.S., Semino, O., Passarino, G., *et al.*, Human Genetic Affinities for Y-Chromosome P49a,f/*Taq*I Haplotypes Show Strong Correspondence with Linguistics, *Am. J. Hum. Genet.*, 1997, vol. 61, pp. 1015–1035.
26. Derenko, M.V., Malyarchuk, B.A., Denisova, G.A., *et al.*, Molecular Genetic Differentiation of the Ethnic Populations of South and East Siberia Based on Mitochondrial DNA Polymorphism, *Russ. J. Genet.*, 2002, vol. 38, no. 10, pp. 1196–1202.
27. 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.
28. Derenko, M.V. and Malyarchuk, B.A., Gradient of Mitochondrial Genome Diversity among Mongoloid Populations of Asia Inferred from mtDNA Hypervariable Region I Variability, *Mol. Biol.* (Moscow), 1998, vol. 32, no. 5, pp. 651–657.
29. Derenko, M.V., Denisova, G.A., Malyarchuk, B.A., *et al.*, The Structure of the Gene Pools of the Ethnic Populations of Altai–Sayan Region Based on Mitochondrial DNA Polymorphism Data, *Russ. J. Genet.*, 2001, vol. 37, no. 10, pp. 1177–1185.
30. Derbeneva, O.A., Starikovskaya, E.B., Volodko, N.V., *et al.*, Mitochondrial DNA Variation in the Kets and Nganasans and Its Implications for the Initial Peopling of Northern Eurasia, *Russ. J. Genet.*, 2002, vol. 38, no. 11, pp. 1316–1321.
31. Derbeneva, O.A., Starikovskaya, E.B., Wallace, D.C., and Sukernik, R.I., Traces of Early Eurasians in the Mansi of Northwest Siberia Revealed by Mitochondrial DNA Analysis, *Am. J. Hum. Genet.*, 2002, vol. 70, pp. 1009–1014.