Genetic Evidence for the Mongolian Ancestry of Kalmyks

Ivan Nasidze,[1]* Dominique Quinque,[1] Isabelle Dupanloup,[2] Richard Cordaux,[1] Lyudmila Kokshunova,[3] and Mark Stoneking[1]

[1]*Department of Evolutionary Genetics, Max Planck Institute for Evolutionary Anthropology, D-04103 Leipzig, Germany*
[2]*Center of Integrative Genomics, University of Lausanne, CH-1015 Lausanne Dorigny, Switzerland*
[3]*Department of Human and Animal Physiology, Kalmyk State University, Elista 358000, Republic of Kalmykia, Russian Federation*

KEY WORDS Kalmyks; Y chromosome; mtDNA

ABSTRACT The Kalmyks are an ethnic group along the lower Volga River in Russia who are thought to have migrated there from Mongolia about 300 years ago. To investigate their origins, we studied mtDNA and Y-chromosome variation in 99 Kalmyks. Both mtDNA HV1 sequences and Y-chromosome SNP haplogroups indicate a close relationship of Kalmyks with Mongolians. In addition, genetic diversity for both mtDNA and the Y chromosome are comparable in Kalmyks, Mongolians, and other Central Asian groups, indicating that the Kalmyk migration was not associated with a substantial bottleneck. The so-called "Genghis Khan" Y-chromosome short tandem repeat (STR) haplotype was found in high frequency (31.3%) among Kalmyks, further supporting a strong genetic connection between Kalmyks and Mongolians. Genetic analyses of even recent, relatively well-documented migrations such as of the Kalmyks can therefore lead to new insights concerning such migrations. Am J Phys Anthropol 128:846–854, 2005.

The Kalmyks (also known as the Khalmag) currently live along the western bank of the lower Volga River but are thought to be the descendants of Oyrats, originating from west Mongolia (Jungaria). During the late 16th and early 17th centuries, the deficit of pasture lands and feudal internecine dissension led the large Oyrat tribal unions of Torgouts and Derbets to migrate to the steppes of western Siberia (Erdeniev, 1985). After the Yermak Expedition (1579–1584), these territories came under the control of Russia, and in 1608 and 1609 the Oyrats gave their oath of allegiance to the Russian czar. Their descendants settled in the territory circumscribed by the Ural and Volga Rivers (Fig. 1). In the second half of the 17th century, they formed the Kalmyk Khanate in the Lower Volga and laid the foundation for the new Mongolian-speaking ethnic group, the Kalmyks (Erdeniev, 1985). There is no evidence from the historical record for any subsequent migration from Mongolia to this region; thus the Kalmyks have been isolated for some 300 years from their presumed parental population.

Little population genetic information is available concerning the Kalmyks. An analysis of classical genetic markers (blood groups ABO and RH(D), serum proteins HP, TF, and GC, and red cell enzymes ACP1, PGM1, ESD, GLO1, and SOD-A) suggested a genetic resemblance of Kalmyks with the contemporary Buryats of the Baikal region of southeastern Siberia and with Mongolians (Galushkin et al., 2002). Polymorphisms of glutathione S-transferases M1 and T1 (GSTM1 and GSTT1) also suggest similarities between Kalmyks and Buryats (Popova et al., 2002).

In order to test whether the putative history of a recent Mongolian origin of the Kalmyks is reflected in their mtDNA and Y-chromosome gene pools, as well as to assess whether a bottleneck was associated with the Kalmyk migration, and the extent of possible subsequent maternal or paternal admixture between the Kalmyks and surrounding populations, we analyzed mitochondrial DNA HV1 sequence variability, screened the 9-bp deletion between the mitochondrial *tRNAlys* and *COII* genes, and genotyped 13 biallelic markers and nine short tandem repeat (STR) loci on the Y chromosome in 99 Kalmyks. We compared the patterns of mtDNA and Y-chromosome variation in the Kalmyks with data from their geographic neighbors and from their putative source population.

MATERIALS AND METHODS

Samples and DNA extraction

In total, 99 cheek-cell samples from unrelated male individuals were collected in Elista (Republic of Kalmykia, Russian Federation). Informed consent and informa-

Grant sponsor: Max Planck Society, Germany; Grant sponsor: Italian National Research Council; Grant sponsor: European Science Foundation; Grant number: JA03-B02.

Current address of Richard Cordaux: Biological Computation and Visualization Center, Department of Biological Sciences, Louisiana State University, 202 Life Sciences Building, Baton Rouge, LA 70803.

*Correspondence to: Ivan Nasidze, Department of Evolutionary Genetics, Max Plank Institute for Evolutionary Anthropology, Deutscher Platz 6, 04103 Leipzig, Germany.
E-mail: nasidze@eva.mpg.de

Received 23 February 2004; accepted 3 August 2004.

DOI 10.1002/ajpa.20159
Published online 18 July 2005 in Wiley InterScience (www.interscience.wiley.com).

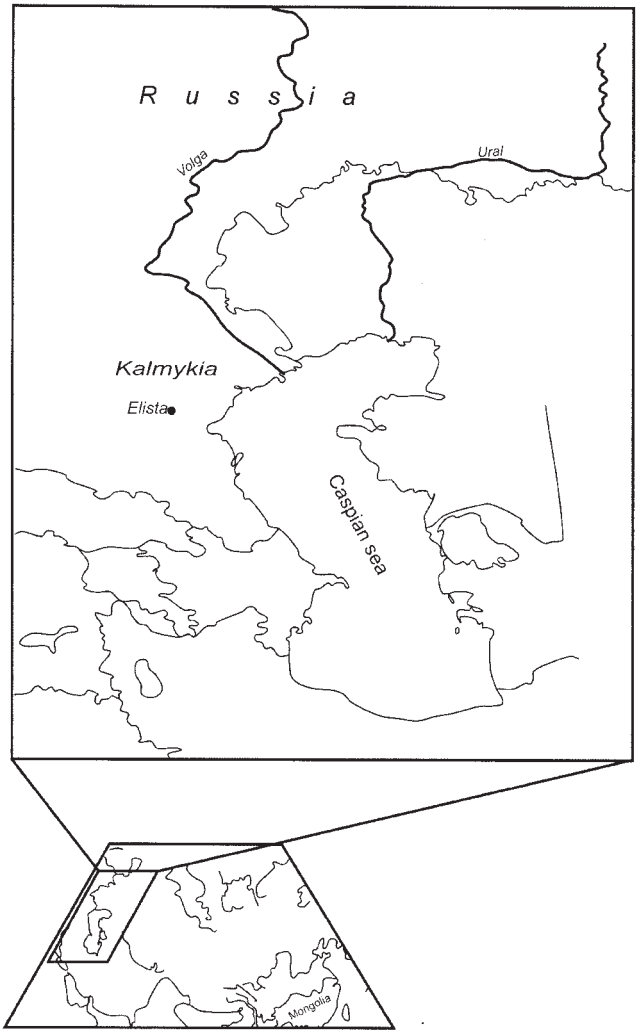

Fig. 1. Map of Russia (above) and Eurasia (below), showing present location of Kalmyks and area of their putative origin.

tion about birthplace, parents, and grandparents were obtained from all donors.

Genomic DNA from cheek-cell swabs was extracted using a salting-out procedure (Miller et al., 1988).

Mitochondrial DNA

The first hypervariable segment of the noncoding mtDNA control region (HVI) was amplified with primers L15996 and H16410 (Vigilant et al., 1989), as described previously (Redd et al., 1995). The nested primers L16001 (Cordaux et al., 2003) and H16401 (Vigilant et al., 1989) were then used to determine sequences for both strands of PCR products with the DNA Sequencing Kit (Perkin-Elmer), following the protocol recommended by the supplier, and with an ABI 3700 automated DNA sequencer. Sequences with a C at position 16189 (Anderson et al., 1981) usually terminated prematurely at the "C-stretch" region (positions 16184–16193); these were sequenced again in each direction, so that each base was determined twice.

Published mtDNA HV1 sequences were also used from 139 Sakha and 56 Evens (Pakendorf et al., 2003), 46 Itel'men (Schurr et al., 1999), 36 Tuvinians and 76 Buryats (Derenko et al., 2002), 58 Evenks (Torroni et al., 1993), 84 Xi'an and 82 Changsha Han Chinese (Oota et al., 2002), 89 Japanese (Horai et al., 1996), 102 Russians (Orekhov et al., 1999), 18 Slavs (Maliarchuk et al., 1995), 55 Kazakhs and 94 Kyrgiz (Comas et al., 1998), 103 Mongolians (Khalkha and Daridanga) (Kolman et al., 1996), 17 Altaians (Shields et al., 1993), and 13 Mari (Sajantila et al., 1995).

The 9-bp deletion in the COII-tRNA^{lys} intergenic region was screened in all samples, as described elsewhere (Redd et al., 1995). For comparative analyses, published data for this marker were used from Siberia, Central Asia, and East Asia (Sambungyin et al., 1991; Harihara et al., 1992; Petrischev et al., 1993).

Y chromosome

Twelve Y-chromosomal SNP markers (M130 (RPS4Y), M48, M9, M46 (TAT C), M89, M124, M45, M173, M17, M201, M170, and M172) (Underhill et al., 2000; Zerjal et al., 1997) and the YAP *Alu* insertion polymorphism were typed (Hammer and Horai, 1995). The markers M9, M46, and RPS4Y were typed by means of PCR-RFLP assays, as described elsewhere (Kayser et al., 2000; Zerjal et al., 1997). The markers M17, M124, M170, M172, and M201 were typed using primer-introduced restriction analysis (PIRA)-PCR assays (Yoshimoto et al., 1993), as described previously (Cordaux et al., 2004). M89 was typed by a PIRA-PCR assay, as described previously (Kayser et al., 2000a). New PIRA-PCR assays were designed for M173, M48, and M45 (Table 1). The YAP *Alu* insertion was typed as described previously (Hammer et al., 1995). Samples were genotyped according to hierarchical order of the markers (Underhill et al., 2000). The Y-SNP haplogroup nomenclature used here is according to the recommendations of the Y Chromosome Consortium (2002).

Published Y-SNP data from East Europe and Central and East Asia were used from 42 Tuvinians, 24 Mongolians, 45 Koreans, 44 Karakalpak, 56 Uzbek (Bukhara), 68 Uzbek (Surkhandarya), 70 Uzbek (Khorezm), 43 Uzbek (Tashkent), 63 Uzbek (Fergana Valley), 45 Uzbek (Samarkand), 25 Ishkashimi, 30 Bartangi, 44 Shugrian, 31 Yagnobi, 22 Tajik (Koiant), 16 Tajik (Dushambe), 52 Kyrgiz, 40 Dungan, 54 Kazakh, 41 Uigur, 28 Pomor, 49 Russians (North), 89 Russians (Tashkent), and 38 Kazan Tatar (Wells et al., 2001).

Nine Y-chromosome short tandem repeat (Y-STR) loci were analyzed: DYS19 (synonymous with DYS394, amplified with a different primer set as described elsewhere) (Kayser et al., 2001), DYS385a, DYS385b, DYS389I, DYS389II, DYS390, DYS391, DYS392, and DYS393. These loci were amplified in pentaplex and quadruplex PCRs, or alternatively in a single nanoplex PCR, and analyzed on an ABI PRISM 377 DNA sequencer (Applied Biosystems), as described elsewhere (Kayser et al., 1997). In order to distinguish genotypes at the duplicated DYS385a and DYS385b loci, an additional PCR was carried out (Kittler et al., 2003) and analyzed on an ABI PRISM 377 DNA sequencer (Applied Biosystems).

Statistical analysis

Basic parameters of molecular diversity and population genetic structure, including analyses of molecular variance (AMOVA) and a minimum spanning network

TABLE 1. PIRA-PCR assays for Y-chromosome SNPs M45, M48, and M173 Y chromosome SNPs[1]

Marker	Primers	Restriction enzymes	Sizes of digested PCR products
M45	For-AATTGGCAGTGAAAAATTATAGCTA Rev-AACTCTCCTACTCTGATGAGCA	*Nla*IV	128 bp + 23 bp
M48	For-AATTGGCAGTGAAAAATTATAGCTA Rev-TCAATGTAAATGTTAGTATAAGGATG	*Bst*F5I	60 bp + 30 bp
M173	For-TTACAATTCAAGGGCATTTAGGA Rev-AGGTGTATCTGGCATCCGTTA	*Nla*IV	129 bp + 23 bp

[1] Derived state of each SNP is digested by restriction enzyme.

TABLE 2. Parameters summarizing some characteristics of mtDNA HV1 sequence variability in Kalmykians[1]

Population	N	No. of haplotypes	Haplotype diversity ± SE	MPD	Tajima's D	Source
Kalmykians	**99**	**85**	0.996 ± 0.002	**6.80**	**−1.83**	**Present study**
Tuvinians	36	28	0.978 ± 0.015	6.89	−1.14	Derenko et al., 2002
Sakha	139	44	0.961 ± 0.007	6.47	−1.20	Pakendorf et al., 2003
Itel'men	46	18	0.929 ± 0.022	4.14	−0.57	Schurr et al., 1999
Evens	56	25	0.949 ±0.016	5.59	−0.95	Pakendorf et al., 2003
Evenks	58	23	0.936 ± 0.017	6.09	−0.98	Torroni et al., 1993
Xi'an (Han Chinese)	84	76	0.997 ± 0.002	5.82	−1.77	Oota et al., 2002
Changsha (Han Chinese)	82	70	0.995 ± 0.003	6.22	−1.87	Oota et al., 2002
Japanese	89	62	0.982 ± 0.006	5.01	−1.99	Horai et al., 1996
Buryats	76	58	0.992 ± 0.004	7.14	−1.83	Derenko et al., 2002
Russians	102	63	0.964 ± 0.012	4.22	−1.99	Orekhov et al., 1999
Slavs	18	18	1.000 ± 0.019	4.41	−1.35	Maliarchuk et al., 1995
Kazakh	55	45	0.990 ± 0.006	6.64	−1.81	Comas et al., 1998
Kirgiz	94	69	0.989 ± 0.004	6.25	−1.95	Comas et al., 1998
Mongolians	103	82	0.995 ± 0.002	6.51	−1.82	Kolman et al., 1996
Mari	13	10	0.949 ± 0.051	4.13	−1.22	Sajantila et al., 1995
Altaians	17	16	0.993 ± 0.023	5.52	−1.25	Shields et al., 1993

[1] Data from some Siberian and Central and East Asian populations are given for comparison. N, sample size; MPD, mean number of pairwise differences.

for Y-STR haplotypes, were calculated using the software package Arlequin 2.000 (Schneider et al., 2000). F_{st} values were calculated based on the number of pairwise differences between HV1 sequences or Y-SNP haplotypes; the statistical significance of F_{st} values was estimated by permutation analysis, using 10,000 permutations. The statistical significance of the correlation between genetic distance matrices based on mtDNA HV1 sequences and Y-chromosome SNP-haplogroups was evaluated by the Mantel test with 10,000 permutations. The STATISTICA package (StatSoft, Inc.) was used for multidimensional scaling (MDS) analysis. Network analysis for mtDNA HVI sequence data was carried out using the software package NETWORK version 3.1 (Bandelt et al., 1999), and a neighbor-joining tree of the HV1 sequences was constructed using PYLIP (Felsenstein, 1993).

RESULTS

MtDNA variation

In total, 377 bp of the mtDNA HV1 region, comprising nucleotide positions 16024–16400 (Anderson et al., 1981), were determined for 99 Kalmyks. For the purposes of comparing the sequences reported here with published data, further analyses were restricted to 365 bp (nucleotide positions 16024–16388) of HV1. As a check on the accuracy of the HV1 sequences, we used the network method to search for so-called "phantom" mutations (Bandelt et al., 2002). No such artifacts were found in the Kalmykian HVI sequences (analysis not shown). The sequences will be deposited in the HVRbase database (www.HVRbase.de) at time of publication.

Parameters summarizing some characteristics of mtDNA HV1 sequence variability in Kalmyks and additional Central and East Asian and East European populations are presented in Table 2. The haplotype diversity was 0.996, among the highest values observed in groups from this region (Table 2). The mean number of pairwise nucleotide differences (MPD) was 6.80, which exceeds the upper limit of the range of MPD values in European groups (3.15–5.03) (Comas et al., 1997), but is comparable to the MPD values for Central Asian groups (5.91–6.64) (Comas et al., 1998).

Pairwise F_{st} values indicate that the lowest values are found between Kalmyks and Mongolians ($F_{st} = 0.007$, $P = 0.029$), and then between these two groups and Central and East Asians. An MDS plot (Fig. 2A) based on the individual pairwise F_{st} values further confirms these observations: the Kalmyks cluster with Mongolians, Kazakh, Kyrgiz, Buryats, and some other Central and East Asian populations. The Eastern European groups cluster separately from the Kalmyks.

Although these analyses indicate that Kalmyk HV1 sequences as a whole group with Mongolian sequences, it may be that some individual Kalmyk HV1 sequences are of east European (or other) origin. We therefore constructed a neighbor-joining tree (based on p-distances, i.e., the number of pairwise differences) of the 304 individual mtDNA HVI sequences from Russians, Kalmyks, and Mongolians. The majority of Kalmyk HV1 sequences (more then 80%) clustered with Mongolian sequences, whereas Russian sequences formed separate branches on the tree, with only few of the Kalmyk sequences clustered with the Russian sequences (tree not shown).

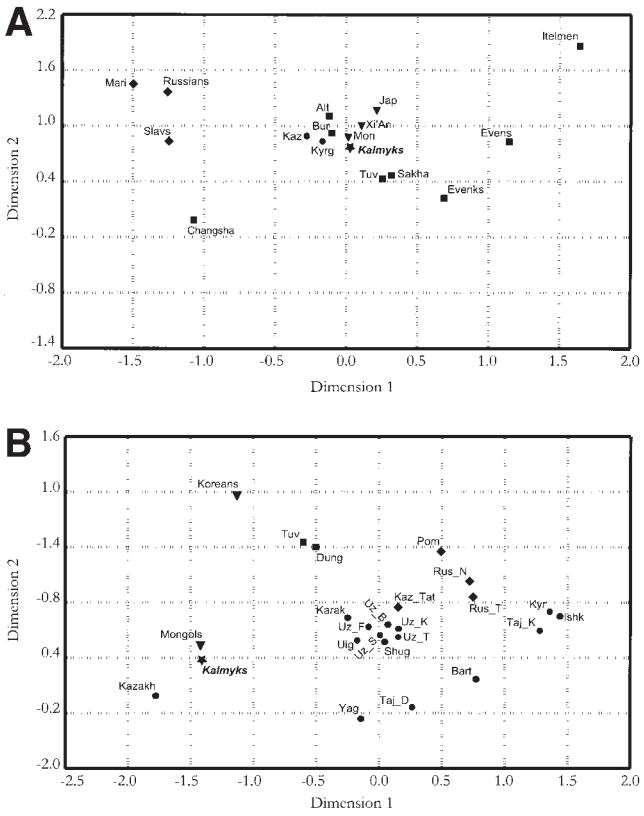

Fig. 2. MDS plots based on pairwise F_{st} values, showing relationships among Kalmykian, Eastern European, Siberian, and Central and East Asian populations. Kalmyks are represented by a star; Siberian groups by squares; Central Asian groups by circles; Eastern European groups by diamonds; and East Asian groups by triangles. **A:** MtDNA HVI sequence data. Stress value for MDS plot is 0.092. Names of populations are abbreviated as follows: Alt, Altaians; Jap, Japanese; Kaz, Kazakhs; Mon, Mongolians; Tuv, Tuvinians; Kyrg, Kyrgiz; Bur, Buryats. **B:** Y-chromosome SNP data. Stress value for MDS plot is 0.117. Names of populations are abbreviated as follows: Tuv, Tuvinians; Dung, Dungans; Pom, Pomors; Rus_N, Russians (North); Rus_T, Russians (Tashkent); Kyr, Kyrgiz; Taj_K, Tajiks (Koinat); Taj_D, Tajiks (Dushambe); Yag, Yagnobi; Kaz_Tat, Kazan Tatars; Karak, Karakalpak; Uig, Uigurs; Uz_F, Uzbeks (Fergana); Uz_S, Uzbek (Surkhadarya); Uz_K, Uzbek (Khorezm); Uz_T, Uzbek (Tashkent); Uz_B, Uzbeks (Bukhara); Shug, Shugrian; Ishk, Ishkashimi.

The frequency of the COII-tRNA^{lys} intergenic 9-bp deletion was about 7% in Kalmyks, similar to the frequency observed in Koreans, Mongolians, and Buryats (Table 3). Since the 9-bp deletion is virtually absent in Eastern Europe (Table 3), these results further support a Central-East Asian origin of Kalmyks.

Y-SNP haplogroups

Nine Y-SNP haplogroups were found in Kalmyks (Table 4). The most frequent haplogroups were C* (RPS4Y) and C3c (M48), followed by K* (M9) and P* (M45); together, the frequency of these haplogroups was 0.858. Haplogroup C* is common in Central Asian and Mongol populations, but absent in Eastern Europe (Table 4). Haplogroup C3c occurs at appreciable frequencies only in Mongolians and Kazakhs, while haplogroup K* is widely distributed in virtually all Eastern European and Central and East Asian groups (Table 4). Haplogroup P* is absent or present in very low frequencies in almost all Eastern European and Central and East Asian groups. The common Y-chromosome haplogroup N3* in Eastern European groups was found only in one Kalmykian individual out of 99 (Table 4). The other Kamlykian Y-haplogroups occurred at low frequencies (less than 10%). Haplogroup diversity in Kalmyks (0.773) is as high as in Central Asian populations (average 0.762) and higher than in Eastern European groups (average 0.755).

Pairwise F_{st} comparisons showed a close relationship of Kalmyks with Mongolians ($F_{st} = 0.010$, $P = 0.221$); by contrast, Eastern European and Central Asian groups were more distant from the Kalmyks (average $F_{st} = 0.224$ and 0.164, respectively). The MDS plot (Fig. 2B) similarly groups Kalmyks with Mongolians, with Kazakhs as the next most similar group.

The Y-SNP M-201, which distinguishes haplogroup G* from haplogroup F*, was not analyzed by Wells at al. (2001) in the populations used for comparison. In our analyses, these individuals were classified as haplogroup F*, although some unknown proportion could in fact be haplogroup G*. To determine if this inability to distinguish between haplogroups F* and G* for some groups influences the results of MDS and F_{st} analyses, we classified all haplogroup G* Kalmykian individuals as haplogroup F* and repeated the analyses. The results (not shown) were essentially identical; thus the inability to distinguish between haplogroups F* and G* in some groups does not influence our conclusions.

The sample of Mongolians on which we base our conclusion of a close genetic relationship between Kalmyks and Mongolians is rather small, with only 24 individuals (Wells et al., 2001). Data on Y-SNP haplogroups are available for a larger sample of Mongolians (147 individuals) in Karafet et al. (2001). However, that study did not distinguish between haplogroups C* (RPS4Y) and C3c (M48), which have frequencies of 0.24 and 0.37, respectively, in Kalmyks. Since failure to distinguish between these haplogroups could lead to inaccurate conclusions, in the above analyses we used the smaller Mongolian data set from Wells et al. (2001), in which haplogroups C* and C3c were distinguished. Nevertheless, to investigate the influence of sample size on our conclusions, we included the data from Karafet et al. (2001) by combining haplogroups C* and C3c; the resulting MDS plot (not shown) groups Mongolians from both studies with Kalmyks.

Y-STR haplotypes

Two male-specific alleles were observed at DYS19 in 31 Kalmyks, indicating a regional Y-chromosome duplication with subsequent microsatellite mutation. Fourteen haplotypes were observed (Table 5), all on the background of haplogroup C3c. Duplication of DYS19 was previously observed in Mongolian, Kazakh, and Kyrghiz Y chromosomes, albeit at low frequency in Kyrghiz and at higher frequencies in Mongolians and Kazakhs (C. Tyler-Smith, personal communication). Moreover, the duplication in Kazakhs mostly involves alleles 15 and 17, whereas both Mongolians and Kalymks mostly exhibit alleles 16 and 17 (C. Tyler-Smith, personal communication), further strengthening the connection between Mongolian and Kalmykian Y chromosomes.

TABLE 3. Polymorphism of mtDNA 9-bp deletion in Kalmyks and comparative data from some Siberian, Central and East Asian, and European populations[1]

Population	N	9-bp deletion n	9-bp deletion %	Source
Kalmykians	**99**	**7**	**7.07**	**Present study**
North Mongolians	292	23	7.9	Sambungyin et al., 1991
South Mongolians	278	23	8.1	Sambungyin et al., 1991
Japanese	116	18	16.0	Harihara et al., 1992
Koreans	64	5	7.8	Harihara et al., 1992
Buryats	65	5	7.7	Petrischev et al., 1993
Northern Altaians	127	13	10.2	Petrischev et al., 1993
Mansi	75	0	0.0	Petrischev et al., 1993
Eastern Slavs	108	0	0.0	Maliarchuk et al., 1995
Finns	32	0	0.0	Lahermo et al., 1996
Saami	129	0	0.0	Lahermo et al., 1996

[1] N, sample size; n, absolute number of 9-bp deletions.

TABLE 4. Y-chromosome haplogroup frequencies in Kalmykians and additional populations from Eastern Europe and Central and East Asia (Wells et al., 2002)[1]

		Haplogroups													
Population	N	E* YAP	C* RPS4Y	C3c M48	K* M9	N3* M46	P1 M124	P* M45	R1* M173	R1a1* M17	F* M89	G* M201	J2* M172	I* M170	HD
Kalmykians	**99**	**0.01**	**0.24**	**0.37**	**0.13**	**0.01**	**0.06**	**0.11**	**0.0**	**0.0**	**0.05**	**0.01**	**0.0**	**0.0**	**0.773**
Tuvinian	42	0.00	0.10	0.07	0.47	0.02	0.0	0.17	0.02	0.14	0.0		0.0	0.0	0.727
Mongolian	24	0.04	0.13	0.46	0.25	0.0	0.0	0.0	0.0	0.04	0.08		0.0	0.0	0.732
Korean	45	0.07	0.16	0.0	0.69	0.0	0.0	0.06	0.0	0.0	0.02		0.0	0.0	0.503
Karakalpak	44	0.0	0.20	0.02	0.25	0.0	0.07	0.0	0.09	0.18	0.09		0.09	0.0	0.852
Uzbek/Bukhara	56	0.02	0.09	0.0	0.19	0.0	0.02	0.02	0.07	0.25	0.18		0.16	0.0	0.842
Uzbek/ Surkhandarya	68	0.08	0.12	0.0	0.17	0.04	0.01	0.04	0.06	0.29	0.03		0.16	0.0	0.846
Uzbek/Khorezm	70	0.07	0.06	0.04	0.14	0.01	0.01	0.09	0.09	0.30	0.07		0.11	0.01	0.862
Uzbek/Tashkent	43	0.05	0.07	0.0	0.09	0.02	0.02	0.05	0.12	0.28	0.09		0.14	0.07	0.877
Uzbek/Fergana Valley	63	0.06	0.13	0.0	0.13	0.0	0.05	0.05	0.13	0.22	0.07		0.1	0.05	0.885
Uzbek/ Samarkand	45	0.02	0.16	0.02	0.15	0.0	0.02	0.07	0.11	0.13	0.13		0.16	0.02	0.893
Ishkashimi	25	0.0	0.0	0.0	0.12	0.0	0.08	0.0	0.04	0.68	0.08		0.0	0.0	0.530
Bartangi	30	0.0	0.0	0.0	0.0	0.0	0.17	0.13	0.03	0.4	0.23		0.03	0.0	0.763
Shugrian	44	0.11	0.02	0.0	0.16	0.0	0.0	0.14	0.07	0.23	0.16		0.11	0.0	0.868
Yagnobi	31	0.0	0.03	0.0	0.13	0.0	0.0	0.03	0.32	0.16	0.0		0.32	0.0	0.772
Tajik/Koiant	22	0.0	0.05	0.05	0.05	0.0	0.09	0.0	0.0	0.64	0.03		0.09	0.0	0.597
Tajik/Dushambe	16	0.06	0.00	0.0	0.0	0.0	0.06	0.0	0.0	0.19	0.19		0.31	0.0	0.795
Kyrgyz	52	0.00	0.08	0.08	0.10	0.02	0.0	0.02	0.02	0.63	0.02		0.02	0.02	0.585
Dungan	40	0.00	0.03	0.0	0.48	0.03	0.05	0.06	0.05	0.1	0.05		0.13	0.03	0.742
Kazakh	54	0.02	0.09	0.57	0.11	0.02	0.02	0.06	0.06	0.04	0.02		0.0	0.0	0.653
Uigur	41	0.00	0.15	0.0	0.19	0.02	0.0	0.07	0.0	0.22	0.12		0.2	0.02	0.854
Pomor	28	0.00	0.0	0.0	0.02	0.43	0.0	0.0	0.0	0.36	0.04		0.04	0.11	0.698
Russian/North	49	0.02	0.0	0.0	0.02	0.20	0.0	0.0	0.0	0.43	0.02		0.04	0.27	0.716
Russian/ Tashkent	89	0.03	0.0	0.0	0.06	0.13	0.0	0.0	0.07	0.47	0.03		0.08	0.12	0.736
Kazan Tatar	38	0.03	0.0	0.0	0.09	0.13	0.0	0.05	0.03	0.24	0.13		0.11	0.18	0.869

[1] HD, haplotype diversity.

The Y-STR haplotypes In the haplogroup C3c background exhibit a "star-like" configuration in Kalmyks (Fig. 3), as observed previously for Y-STR haplotypes In the haplogroup C3c background in Mongolians (Zerjal et al., 2003). We estimated a time to most recent common ancestor (TMRCA) for the Kalmyk Y-STR haplotypes in the haplogroup C3c background, using the ρ estimate (Morral et al., 1994). Assuming a generation time of 30 years, the TMRCA estimate is 855 years (95% confidence interval limits, ~560–1,160 years), which is in remarkable agreement with the TMRCA estimate of 860 years for this haplogroup in Mongolians (Zerjal et al., 2003).

Comparison of mtDNA and Y-chromosome data

With respect to within-population diversity, the Kalmyks exhibit levels of Y-SNP haplogroup and mtDNA haplotype diversity that are comparable to East and Central Asian and Eastern European groups (Tables 2 and 5). The pairwise F_{st} value based on Y-SNP haplogroups is lowest between Kalymks and Mongolians ($F_{st} = 0.010$), followed by Central Asian groups (average $F_{st} = 0.164$) and then Eastern European groups (average $F_{st} = 0.224$). Similarly, based on mtDNA HV1 sequences, the lowest F_{st} value is between Kalmyks and Mongolians ($F_{st} = 0.007$), followed

TABLE 5. Y-STR haplotypes in background of Y-SNP C3c haplogroup

Haplotypes	Y-STR loci									Number of individuals
	DYS393	DYS390	DYS394	DYS391	DYS392	DYS385a	DYS385b	DYS389I	DYS389II	
1	13	24	16/17	9	11	12	12	11	17	13
2	13	24	16/17	9	11	12	13	11	17	4
3	13	25	16/17	9	11	12	12	11	17	2
4	13	24	16/17	9	11	12	12	12	17	1
5	13	24	16/17	9	11	12	13	12	17	1
6	13	24	16/17	9	11	12	13	11	18	1
7	13	24	16/17	9	11	12	12	11	18	1
8	12	25	16/17	9	11	12	12	11	17	1
9	13	23	16/17	9	11	12	12	12	16	1
10	13	24	16/17	9	11	12	12	11	16	1
11	13	24	17/18	9	11	12	12	11	17	2
12	13	24	15/17	9	11	12	12	11	17	1
13	13	24	15/16	9	11	12	12	11	17	1
14	13	24	15/17	9	11	12	12	11	18	1
15	13	24	15	9	11	12	13	11	17	1
16	13	24	15	9	11	12	13	10	18	1
17	13	24	17	9	11	12	12	11	17	3

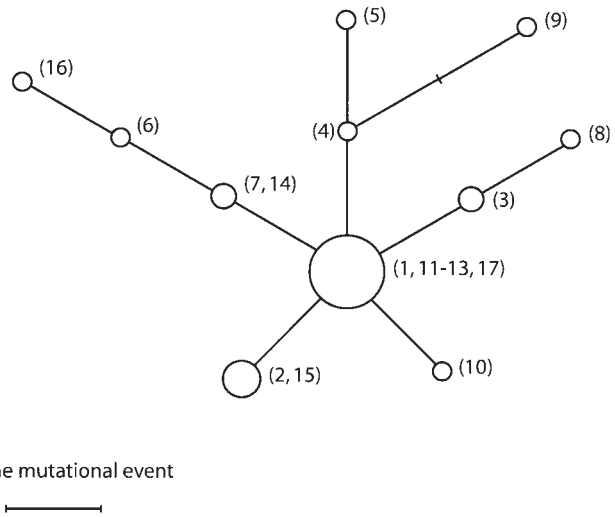

Fig. 3. Minimum spanning network of Y-chromosome STR haplotypes in background of haplogroup C3c-M48. DYS19 was excluded from network analysis because it is duplicated in haplogroup C3c. Numbers in parentheses correspond to haplotypes in Table 5.

again by Central Asian groups (average $F_{st} = 0.05$) and then by East Asian groups (average $F_{st} = 0.058$). However, the correlation between pairwise F_{st} distances among pairs of Kalmyks and Eastern European and Central Asian groups, based on mtDNA and the Y chromosome, was not statistically significant (Mantel test, $Z = 0.168$, $P = 0.334$), suggesting some differences in the genetic structure of these groups based on mtDNA vs. the Y chromosome.

The geographic and linguistic structure of Kalmykian, Eastern European, and Central Asian groups, assessed by mtDNA and Y-chromosome variation, was further investigated by the AMOVA procedure. For mtDNA, both geographic and linguistic groupings gave similar results (Table 6). This is probably because the two groupings are quite similar, differing only in that geographically Kalmyks group with Kazakhs, apart from Altaians, Kyrgiz, and Mongolians, whereas linguistically Kalmyks group with Mongolians, and Kazakhs group with Altaians and Kirghiz (Table 6). However, for Y-SNP haplogroups, the level of between-population differentiation was much higher than for mtDNA; moreover, classifying groups on the basis of geography gave a much better fit to the data than the linguistic classification, in that the among-group component was much bigger than the among-populations-within-groups component for the geographic classification (Table 6).

DISCUSSION

Both the mtDNA and the Y-chromosome results indicate that Kalmyks are most closely related to Mongolians. This is entirely in accordance with their history, which indicates that in the early 17th century, large groups of Oyrat tribes from western Mongolia moved to their current home in the Volga region (Erdeniev, 1985). The Derbets were the second largest tribal subdivision of the Oyrats, followed by Khoshuts. There were also smaller tribal groups such as the Zungars, Khoits, and Tsaatans. They migrated to the left bank of the lower Volga River and are found now in Russian towns from Astrakhan to Samara.

The genetic results do not merely confirm the historical record, but also add additional insights into the Kalmykian migration. Levels of genetic diversity for both the mtDNA and Y chromosome are comparable in Kalmyks and Mongolians, indicating that there was no loss of diversity, and hence no bottleneck, associated with this migration. Even the Y-STR diversity associated with the C3c haplogroup in Kalmyks is virtually identical to that reported previously in Mongolians (Zerjal et al., 2003). Thus, the number of both males and females involved in the migration of Kalmyks must have been substantial. This is in contrast with other human migrations for which bottlenecks have been postulated for either mtDNA or the Y chromosome (or both), such as the colonization of Polynesia (Kayser et al., 2000a), the migration of Turkish-speaking Yakuts to Siberia (Pakendorf et al., 2002), and the settlement of the New World (Torroni et al., 1993; Starikovskaya et al., 1998; Bortolini et al., 2003). Comparing the Kalmyk migration with such other migrations may therefore shed light on the social or other circumstances that influence the number of migrating individuals.

TABLE 6. AMOVA results according to different classifications[1]

	mtDNA			Y-SNP		
Classifications	Among groups	Among populations within groups	Within populations	Among groups	Among populations within groups	Within populations
Geography	4.40	0.88	94.72	12.22	6.53	81.25
Linguistic	4.54	0.74	94.72	6.62	11.17	82.21

[1] Geography: Eastern Europe (Russians and Slavs), Central Asia (Mongolians, Kyrgiz, Altaians), and Kalmyks and Kazakhs for mtDNA. Eastern Europe (Russians/North and Russians/Tashkent), Central Asia (Kyrgiz, Mongolians, and Uzbek), and Kazakh and Kalmykians for Y-SNPs. Linguistic: Turkic (Altaians, Kyrgiz, and Kazakh), Mongol-Langam (Mongolians and Kalmyks), and Indo-European (Russians and Slavs) for mtDNA. Turkic (Kazakh, Kyrgiz, and Uzbek), Mongol-Langam (Mongolians and Kalmyks), and Indo-European (Russian/Tashkent and Russians/North) for Y-SNPs (according to Ethnologue: http://www.ethnologue.com/).

The genetic results also indicate that there has been no substantial admixture with Russians, along either paternal or maternal lines, during the 300 years that the Kalmyks have been living in close proximity to Russians. The lack of detectable similarity between Russians and Kalmyks does not simply reflect insufficient time for such admixture to have occurred, since European admixture can be readily detected in African-Americans (Kayser et al., 2003), whose time-depth in North America is comparable to that of Kalmyks in Russia. Other migrations, such as those to Polynesia (Kayser et al., 2000a), were accompanied by substantial admixture, so it appears that there are particular social circumstances that either promote or inhibit admixture. The Kalmyks differ from Russians in language, religion, and subsistence, but it is not clear if these factors alone are sufficient to inhibit admixture, as there is evidence for Russian admixture with Yakuts (Pakendorf et al., 2002), who also differ in language, religion, and subsistence. It may be that the substantial size of the Kalmyk migration (as evidenced by the lack of reduced diversity for either mtDNA or the Y chromosome) inhibited admixture, but further comparison of the Kalmyk migration with other migrations is needed to understand the factors that influence admixture.

The mtDNA and Y-chromosome results are also consistent in indicating some degree of genetic similarity between Kazakhs and Kalmyks. This contrast between Kazakh and Russian admixture with Kalmyks leads to the prediction that the Kalmyk language might show a greater impact (i.e., borrowing) from Kazakh than from Russian, or that the Kazakh language might show an influence of the Kalmykian language. Indeed, despite the fact that Kazakhs speak a Turkic language, there are many Mongolian loan words in Kazakh (L. Johanson, personal communication), although further work is needed to demonstrate a specifically Kalmykian origin of these loan words.

An unusual genetic feature of the Kalmyks is the high frequency of duplicate alleles for the DYS19 locus. While duplicate alleles at DYS19 were found in Mongolians (C. Tyler-Smith, personal communication), they are at a much lower frequency (10.8%) than in Kalmyks (31.3%). In other populations, duplication of DYS19 is extremely rare, occurring at an overall frequency of 0.12% based on 7,772 individuals (Kayser et al., 2000b). This unusual duplication further points to a close relationship of Kalmyks with Mongolians.

CONCLUSIONS

The genetic results support the historical record in that they indicate a close relationship between Kalmyks and Mongolians. Moreover, the genetic results indicate that the Kalmyk migration involved substantial numbers of individuals, and that Kalmyks have not experienced detectable admixture with Russians. Thus, genetic studies of even such recent, relatively well-documented migrations as the Kalmyks can provide additional insights into the circumstances surrounding such migrations. Moreover, contrasting the Kalmyk migration with other migrations that differ in one or more aspects (such as the colonizations of Polynesia or the New World, both of which involved substantial bottlenecks) can shed light on the consequences of human migrations.

ACKNOWLEDGMENTS

We thank Manfred Kayser and Brigitte Pakendorf for useful discussions. This research was supported by the Max Planck Society, Germany. I.D. was supported by grants from the Italian National Research Council (CNR), within the European Science Foundation Euro Cores Program "The Origins of Man, Language, and Languages" (Project JA03-B02).

LITERATURE CITED

Anderson S, Bankier AT, Barrell BG, et al. 1981. Sequence and organization of the human mitochondrial genome. Nature 290: 457–465.

Bandelt HJ, Forster P, Röhl A. 1999. Median-joining networks for inferring intraspecific phylogenies. Mol Biol Evol 16:37–48.

Bandelt H-J, Quintana-Murci L, Salas A, Macaulay V. 2002. The fingerprint of phantom mutations in mitochondrial DNA data. Am J Hum Genet 71:1150–1160.

Bortolini M-C, Salzano FM, Thomas MG, Stuart S, Nasanen SP, Bau CH, Hutz MH, Layrisse Z, Petzl-Erler ML, Tsuneto LT, Hill K, Hurtado AM, Castro-de-Guerra D, Torres MM, Groot H, Michalski R, Nymadawa P, Bedoya G, Bradman N, Labuda D, Ruiz-Linares A. 2003. Y-chromosome evidence for differing ancient demographic histories in the Americas. Am J Hum Genet 73:524–539.

Comas D, Calafell F, Mateu E, Pérez-Lezaun A, Bosch E, Bertranpetit J. 1997. Mitochondrial DNA variation and the origin of the Europeans. Hum Genet 99:443–449.

Comas D, Calafell F, Mateu E, Pérez-Lezaun A, Bosch E, Martinez-Arias R, Clarimon J, Facchini F, Fiori G, Luiselli D, Pettener D, Bertranpetit J. 1998. Trading genes along the Silk Road: mtDNA sequences and the origin of Central Asian populations. Am J Hum Genet 63:1824–1838.

Cordaux R, Saha N, Bentley GR, Aunger R, Sirajuddin SM, Stoneking M. 2003. Mitochondrial DNA analysis reveals diverse histories of tribal populations from India. Eur J Hum Genet 11:253–264.

Cordaux R, Aunger R, Bentley G, Nasidze I, Saha N, Sirajuddin SM, Stoneking M. 2004. Independent origins of Indian caste and tribal paternal lineages. Curr Biol 14:231–235.

Derenko MV, Maliarchuk BA, Denisova GA, Dambueva IK, Kakpakov VT, Dorzhu CM, Luzina FA, Lotosh EA, Ondar UN, Kaplina MI, Zakharov IA. 2002. Molecular genetic differentiation of ethnic populations in southern and eastern Siberia based on mitochondrial DNA polymorphism. Russ J Genet 38:1409–1416.

Erdeniev UE. 1985. Kalmyks. Moscow: Nauka.

Felsenstein J. 1993. PYLIP (phylogeny inference package) version 3.5c. (http:evolution.genetics.washington.edu/phylip.html)

Galushkin SK, Spitsyn VA, Crawford MH. 2002. Genetic structure of Mongolic-speaking Kalmyks. Hum Biol 73:823–834.

Hammer M, Horai S. 1995. Y chromosomal DNA variation and the peopling of Japan. Am J Hum Genet 56:951–962.

Harihara S, Hirai M, Suutou Y, Shimizu K, Omoto K. 1992. Frequency of a 9-bp deletion in the mitochondrial DNA among Asian populations. Hum Biol 64:161–166.

Horai S, Murayama K, Hayasaka K, Matsubayashi S, Hattori Y, Fucharoen G, Harihara S, Park KS, Omoto K, Pan IH. 1996. MtDNA polymorphism in East Asian populations, with special reference to the peopling of Japan. Am J Hum Genet 59:579–590.

Karafet T, Xu L, Du R, Wang W, Feng S, Wells RS, Redd AJ, Zegura SL, Hammer MF. 2001. Paternal population history of East Asia: sources, patterns, and microevolutionary processes. Am J Hum Genet 69:615–628.

Kayser M, Caglia A, Corach D, Fretwell N, Gehrig C, Graziosi G, Heidorn F, Herrmann S, Herzog B, Hidding M, Honda K, Jobling M, Krawczak M, Leim K, Meuser S, Meyer E, Oesterreich W, Pandya A, Parson W, Penacino G, Perez-Lezaun A, Piccinini A, Prinz M, Schmitt C, Roewer L. 1997. Evaluation of Y chromosomal STRs: a multicenter study. Int J Legal Med 110:125–133.

Kayser M, Brauer S, Weiss G, Underhill PA, Roewer L, Schiefenhovel W, Stoneking M. 2000a. Melanesian origin of Polynesian Y chromosome. Curr Biol 10:1237–1246.

Kayser M, Roewer L, Hedman M, Henke L, Henke J, Brauer S, Kruger C, Krawczak M, Nagy M, Dobosz T, Szibor R, de Knijff P, Stoneking M, Sajantila A. 2000b. Characteristics and frequency of germline mutations at microsatellite loci from the human Y chromosome, as revealed by direct observation in father/son pairs. Am J Hum Genet 66:1580–1588.

Kayser M, Krawczak M, Excoffier L, Dieltjes P, Corach D, Pascali V, Gehrig C, Bernini LF, Jespersen J, Bakker E, Roewer L, de Knijff P. 2001. An extensive analysis of Y-chromosomal microsatellite haplotypes in globally dispersed human populations. Am J Hum Genet 68:990–1018.

Kayser M, Brauer S, Schadlich H, Prinz M, Batzer MA, Zimmerman PA, Boatin BA, Stoneking M. 2003. Y chromosome STR haplotypes and the genetic structure of U.S. populations of African, European, and Hispanic ancestry. Genome Res 4: 624–634.

Kittler R, Erler A, Brauer S, Stoneking M, Kayser M. 2003. Apparent intrachromosomal exchange on the human Y chromosome explained by population history. Eur J Hum Genet 11:304–314.

Kolman CJ, Sambuughin N, Bermingham E. 1996. Mitochondrial DNA analysis of Mongolian populations and implications for the origin of New World founders. Genetics 142:1321–1334.

Lahermo P, Sajantila A, Sistonen P, Lukka M, Aula P, Peltonen L, Savontaus ML. 1996. The genetic relationship between the Finns and the Finnish Saami (Lapps): analysis of nuclear DNA and mtDNA. Am J Hum Genet 6:1309–1322.

Maliarchuk BA, Derenko MV, Solovenchuk LL. 1995. Types of regulatory regions in mitochondrial DNA in eastern Slavs. Russ J Genet 31:846–851.

Miller SA, Dykes DD, Polesky HF. 1988. A simple salting out procedure for extracting DNA from human nucleated cells. Nucleic Acids Res 16:1215.

Morral N, Bertranpetit J, Estivill X, Nunes V, Casals T, Gimenez J, Reis A, Varon-Mateeva R, Macek M Jr, Kalaydjieva L, Angelichev D, Dancheva R, Romeo G, Russo MP, Garnerone S, Restagno G, Ferrari M, Magnani C, Claustres M, Desgeorges M, Schwarz M, Dallapiccola B, Novelli G, Ferec C, Dearce M, Nemeti M, Kere T, Anvret M, Dahl N, Kadasi L. 1994. The origin of the major cystic fibrosis mutation (ΔF508) in European populations. Nat Genet 7:169–175.

Oota H, Kitano T, Jin F, Yuasa I, Wang L, Ueda S, Saitou N, Stoneking M. 2002. Extreme mtDNA homogeneity in continental Asian populations. Am J Phys Anthropol 118:146–153.

Orekhov V, Poltoraus A, Zhivotovsky LA, Spitsyn V, Ivanov P, Yankovsky N. 1999. Mitochondrial DNA sequence diversity in Russians. FEBS Lett 445:197–201.

Pakendorf B, Morar LA, Tarskaia M, Kayser M, Soodyall H, Rodewald A, Stoneking M. 2002. Y-chromosomal evidence for a severe reduction in male population size of Yakuts. Hum Genet 110:198–200.

Pakendorf B, Wiebe V, Tarskaia LA, Spitsyn VA, Soodyall H, Rodewald A, Stoneking M. 2003. Mitochondrial DNA evidence for admixed origins of central Siberian populations. Am J Phys Anthropol 120:211–224.

Petrischev VN, Kutueva AB, Rychkov YG. 1993. Deletion-insertion polymorphism in ten mongoloid populations of Siberia. Frequency of deletion correlates with the geographical coordinates of the area. Russ J Genet 29:1196–1204.

Popova SN, Slominskii PA, Galushkin SN, et al. 2002. Polymorphism of glutathione S-transferases M1 and T1 in several populations of Russia. Russ J Genet 38:281–284.

Redd AJ, Takezaki N, Sherry ST, et al. 1995. Evolutionary history of the COII/tRNA^{Lys} intergenic 9 base pair deletion in human mitochondrial DNAs from Pasific. Mol Biol Evol 12: 604–615.

Sajantila A, Lahermo P, Anttinen T, Lukka M, Sistonen P, Savontaus ML, Aula P, Beckman L, Tranebjaerg L, Gedde-Dahl T, Issel-Tarver L, DiRienzo A, Pääbo S. 1995. Genes and languages in Europe: an analysis of mitochondrial lineages. Genome Res 5:42–52.

Sambungyin N, Petrischev VN, Rychkov YG. 1991. DNA polymorphism in Mongols. Mitochondrial DNA RFLP analysis. Russ J Genet 27:2143–2151.

Schneider S, Roessli D, Excoffier L. 2000. Arlequin version 2.000: a software for population genetics data analysis. Geneva: Genetics and Biometry Laboratory, University of Geneva.

Schurr TG, Sukernik RI, Starikovskaya YB, Wallace DC. 1999. Mitochondrial DNA variation in Koryaks and Itel'men: population replacement in the Okhotsk Sea-Bering Sea region during the Neolithic. Am J Phys Anthropol 108:1–39.

Shields GF, Schmiechen AM, Frazier BL, Redd A, Voevoda MI, Reed JK, Ward RH. 1993. MtDNA sequences suggest a recent evolutionary divergence for Beringian and northern North American populations. Am Hum Genet 53:549–562.

Starikovskaya YB, Sukernik RI, Schurr TG, Kogelnik AM, Wallace DC. 1998. MtDNA diversity in Chukchi and Siberian Eskimos: of ancient Beringia and the peopling of the New World—implications for the genetic history. Am J Hum Genet 5:1473–1491.

Torroni A, Sukernik RI, Schurr TG, Starikorskaya YB, Cabell MF, Crawford MH, Comuzzie AG, Wallace DC. 1993. MtDNA variation of aboriginal Siberians reveals distinct genetic affinities with Native Americans. Am J Hum Genet 53: 591–608.

Underhill PA, Shen P, Lin AA, Jin L, Passarino G, Yang WH, Kauffman E, Bonne-Tamir B, Bertranpetit J, Francalacci P, Ibrahim M, Jenkins T, Kidd JR, Mehdi SQ, Seielstad MT, Wells RS, Piazza A, Davis RW, Feldman MW, Cavalli-Sforza LL, Oefner PJ. 2000. Y chromosome sequence variation and the history of human populations. Nat Genet 26:358–361.

Vigilant L, Pennington R, Harpending H, Kocher TD, Wilson AC. 1989. Mitochondrial DNA sequences in single hairs from a southern African population. Proc Natl Acad Sci USA 86: 9350–9354.

Wells RS, Yuldasheva N, Ruzibakiev R, Underhill PA, Evseeva I, Blue-Smith J, Jin L, Su B, Pitchappan R, Shanmugalakshmi S, Balakrishnan K, Read M, Pearson NM, Zerjal T, Webster MT, Zholoshvili I, Jamarjashvili E, Gambarov S, Nikbin B, Dostiev A, Aknazarov O, Zalloua P, Tsoy I, Kitaev M, Mirrakhimov M, Chariev A, Bodmer WF. 2001. The Eurasian heartland: a continental perspective on Y-chromosome diversity. Proc Natl Acad Sci USA 98:10244–10249.

Y Chromosome Consortium. 2002. A nomenclature system for the tree of human Y- chromosomal binary haplogroups. Genome Res 2:339–348.

Yoshimoto K, Iwahana H, Fukuda A, Sano T, Itakura M. 1993. Rare mutations of the Gs alpha subunit gene in human endocrine tumors. Mutation detection by polymerase chain reaction-primer-introduced restriction analysis. Cancer 72:1386–1393.

Zerjal T, Dashnyam B, Pandya A, Kayser M, Roewer L, Santos FR, Schiefenhovel W, Fretwell N, Jobling MA, Harihara S, Shimizu K, Semjidmaa D, Sajantila A, Salo P, Crawford MH, Ginter EK, Evgrafov OV, Tyler-Smith C. 1997. Genetic relationships of Asians and Northern Europeans, revealed by Y-chromosomal DNA analysis. Am J Hum Genet 5:1174–1183.

Zerjal T, Xue Y, Bertorelle G, Wells RS, Bao W, Zhu S, Qamar R, Ayub Q, Mohyuddin A, Fu S, Li P, Yuldasheva N, Ruzibakiev R, Xu J, Shu Q, Du R, Yang H, Hurles ME, Robinson E, Gerelsaikhan T, Dashnyam B, Mehdi SQ, Tyler-Smith C. 2003. The genetic legacy of the Mongols. Am J Hum Genet 72:717–723.