

## ORIGINAL ARTICLE

# Ancient links between Siberians and Native Americans revealed by subtyping the Y chromosome haplogroup Q1a

Boris Malyarchuk<sup>1</sup>, Miroslava Derenko<sup>1</sup>, Galina Denisova<sup>1</sup>, Arkady Maksimov<sup>2</sup>, Marcin Wozniak<sup>3</sup>, Tomasz Grzybowski<sup>3</sup>, Irina Dambueva<sup>4</sup> and Ilya Zakharov<sup>5</sup>

To investigate the structure of Y chromosome haplogroups R-M207 and Q-M242 in human populations of North Asia, we have performed high-resolution genotyping using both single nucleotide polymorphisms and short tandem repeat (STR)-based approaches of 121 M207- and M242-derived samples from 885 males of 16 ethnic groups of Siberia and East Asia.

As a result, the following Y chromosome haplogroups were revealed: R1b1b1-M73 (2.0%), R1b1b2-M269 (0.7%), R2-M124 (1.1%), Q1a\*-MEH2 (0.5%), Q1a2-M25 (0.1%), Q1a3\*-M346 (9.2%) and Q1a3a-M3 (0.2%). Despite the low coalescence age of haplogroup Q1a3\*-M346, which is estimated in South Siberia as about  $4.5 \pm 1.5$  thousand years ago (Ka), divergence time between these Q1a3\*-M346 haplotypes and Amerindian-specific haplogroup Q1a3a-M3 is equal to  $13.8 \pm 3.9$  Ka, pointing to a relatively recent entry date to America. In addition, unique cluster of haplotypes belonging to Q1a\*-MEH2 was found in Koryaks inhabiting the Sea of Okhotsk coast (at a frequency of 10.3%). Although the level of STR diversity associated with Q1a\*-MEH2 is very low, this lineage appears to be closest to the extinct Palaeo-Eskimo individuals belonging to the Saqqaq culture arisen in the New World Arctic about 5.5 Ka. This finding suggests that Q1a\*-MEH2 likely traces a population migration originating in Northeast Siberia across the Bering Strait.

*Journal of Human Genetics* (2011) 56, 583–588; doi:10.1038/jhg.2011.64; published online 16 June 2011

**Keywords:** haplogroup Q; haplogroup R; human migrations; Siberia; Y chromosome

## INTRODUCTION

Recent studies of genome-wide autosomal single nucleotide polymorphisms (SNPs) showed that Siberian populations are characterized by significant amount of West Eurasian admixture.<sup>1</sup> This is consistent with results of mitochondrial DNA (mtDNA), and Y chromosome studies showing that North-Asian populations have a closer genetic relationship with Central Asian and West-Eurasian populations.<sup>2–7</sup> The presence of West-Eurasian genetic components in North and East Asia can be interpreted either as the evidence of ancient migration via the northern route<sup>8</sup> or simply as the reflection of recent population admixture.<sup>9,10</sup> However, results of mtDNA analysis in northern Eurasian populations do not support a northern Asian migration route out of Africa, but they do predict that there were at least two migrations into South Siberia, one from East Asia and one from West Eurasia.<sup>6</sup> West Eurasian mtDNA haplogroups found in gene pools of South Siberians (for example, N1e, I4, J1b2, N1a, U4 and X2e) demonstrate an obvious link between populations of Siberia and those of West Asia, the Caucasus and East Europe. It is

noteworthy that complete mtDNA genome-based coalescence times for haplogroups X2e, J1b2 and N1a suggest their postglacial flows (about 12–15 Ka) from the west. Similarly, evidence of geographic distribution and Y chromosome short tandem repeat (Y-STR) diversity indicates that haplogroup Q1a3\*-M346 (sister haplogroup of the Native American-specific haplogroup Q1a3a-M3), R1b-M343 and R1a1\*-M17 probably migrated into North Asia and northern East Asia from the west.<sup>7</sup> The age estimation of Y-STR variation within these Y chromosome haplogroups suggests the existence of such migrations in postglacial time (about 15–18 Ka). In addition, ancient DNA studies demonstrate that West Eurasian admixture was present in South Siberia and Northwest China as early as the early Bronze Age, thus indicating that West Eurasian genetic input occurred earlier than the proposed relatively recent population admixture.<sup>11–14</sup>

It is assumed that haplogroup P-92R7 consisting of subclades Q-M242 and R-M207 originated in Central Asia about 40 Ka.<sup>15</sup> Haplogroup Q-M242 is present in Eurasia, but it is most frequent in North Asia.<sup>7,16–19</sup> It was suggested also that Q-M242 carriers

<sup>1</sup>Genetics Laboratory, Institute of Biological Problems of the North, Russian Academy of Sciences, Magadan, Russia; <sup>2</sup>Laboratory for Physiology of Extreme Conditions, Scientific Research Center 'Arktika', Russian Academy of Sciences, Magadan, Russia; <sup>3</sup>Forensic Medicine Institute, The Ludwik Rydygier Medical College, The Nicolaus Copernicus University in Torun, Bydgoszcz, Poland; <sup>4</sup>Animal Ecology Laboratory, Institute of General and Experimental Biology, Russian Academy of Sciences, Ulan-Ude, Russia and <sup>5</sup>Animal Comparative Genetics Laboratory, N.I. Vavilov Institute of General Genetics, Russian Academy of Sciences, Moscow, Russia

Correspondence: Dr BA Malyarchuk, Genetics Laboratory, Institute of Biological Problems of the North, Russian Academy of Sciences, Portovaya street, 18, 685000 Magadan, Russia.

E-mail: malyarchuk@ibpn.ru

Received 30 March 2011; revised 22 April 2011; accepted 24 April 2011; published online 16 June 2011

migrated through the Altai/Baikal region of Siberia into the Americas.<sup>16,17,20</sup> Haplogroup R-M207 consisting of two currently defined subclades R1-M173 and R2-M124 occurs at high frequencies in some regions of Eurasia. Haplogroup R1-M173 is estimated to have arisen during the Last Glacial Maximum, most likely in Southwest Asia.<sup>8,15</sup> This haplogroup includes two main subclades, R1a-M420 and R1b-M343, with distinctive Eurasian distribution.<sup>18,21–23</sup> R1a-M420 is most frequently observed in East Europe, Altai region of Siberia and Southwest Asia.<sup>2,24,25</sup> R1b-M343 originated in West Asia consists of two most frequent subclades, R1b1b1-M73 and R1b1b2-M269.<sup>8,18,23,26</sup> Haplogroup R1b1b1-M73 is observed mainly in Asia, and haplogroup R1b1b2-M269 is frequent in Europe, especially on the west, but it is present in West Asia as well.<sup>7,23</sup> Moreover, a recent study by Balaresque *et al.*<sup>27</sup> suggests a West Asian origin of R1b1b2-M269 and its Neolithic expansion. Haplogroup R2-M124 is most often observed in Asia, especially in South and Central Asia.<sup>7,8,15,25</sup>

Although the presence of haplogroups R-M207 and Q-M242 in South Siberia and neighboring regions has been reported earlier,<sup>16,17,20,28,29</sup> there has not been detailed examination of their substructure in Siberians, although this might be very important for providing new insights into the early migrations into the Americas. Therefore, in this study, we have performed phylogeographic analysis of subclades within haplogroups R-M207 and Q-M242 in different populations of Siberia and northern East Asia, based on high-resolution genotyping of Y chromosomes using both SNPs and STR-based approaches.

## MATERIALS AND METHODS

### Subjects and DNA typing

A total of 885 samples (whole blood and hair roots specimens) from unrelated males were collected in populations of South Siberia (Altaians, Teleuts, Khakassians, Shors, Tuvinians, Todjins, Tofalars, Sojots, Buryats), Central and East Siberia (Evenks, Evens, Yakuts and Koryaks), East Asia (Mongols and Koreans) and East Europe (Kalmyks) (Table 1). As the Mongolic-speaking Kalmyks appear to be descendants of western Mongolians (Oirats) migrated to the Caspian region in the 17th century from Central Asia and taking into account the considerable genetic similarity between Kalmyks, Buryats and Mongols demonstrated both by Y chromosome and mtDNA variability data,<sup>6,30</sup> the Kalmyks are considered here as 'North Asians'. All samples studied were collected with appropriate ethics approval and informed consent.

Haplogroup P markers 92R7 (for the whole haplogroup P), M242 (for Q), MEH2 (for Q1a), M120 (for Q1a1), M25 (for Q1a2), M346 (for Q1a3), M3 (for Q1a3a), P48 (for Q1a4), P89 (for Q1a5), M323 (for Q1a6), M173 (for R1), M343 (for R1b), M73 (for R1b1b1), M269 (for R1b1b2) and M124 (for R2) were assayed using PCR primers summarized in Karafet *et al.*<sup>31</sup> All polymorphisms were typed by means of DNA sequencing on ABI 3130 Genetic Analyzer (Applied Biosystems, Foster City, CA, USA). The Y-SNP haplogroup nomenclature used here is according to the recommendations of the Y Chromosome Consortium.<sup>31</sup>

A total of 121 samples belonging to haplogroups Q (88 samples), R1b (23 samples) and R2 (10 samples) were analyzed at 12 STR loci (DYS19, DYS385a, DYS385b, DYS389I, DYS389II, DYS390, DYS391, DYS392, DYS393, DYS437, DYS438, DYS439) using PowerPlex Y System (Promega, Madison, WI, USA). Electrophoresis results were analyzed using Genscan v. 3.7 and Genotyper v. 3.7 software (Applied Biosystems).

### Data analysis

Median joining networks of STR-haplotypes were constructed using the Network 4.6 program (<http://fluxus-engineering.com>). For the network construction, STR variants were weighted (with a weight assigned to a range of variance values) following the distribution of the number of mutations at character.<sup>32</sup> Loci DYS385a and DYS385b were excluded from the median joining network analysis, because an unambiguous assignment of the alleles to these loci is impossible without their separate typing. In the network construction and the age calculation procedures, the allele sizes for DYS389II were determined with the subtraction of DYS389I.

The age of STR variation within haplogroups was estimated as the average squared difference in the number of repeats between all current chromosomes and the founder haplotype (formed by the median values of the repeat scores at each STR locus within the haplogroup), averaged over STR loci and divided by means of a mutation rate.<sup>25,33</sup> The evolutionary effective mutation rate of  $6.9 \times 10^{-4}$  per 25 years based on STR variation within Y chromosome haplogroups in the populations with documented short-term histories was used.<sup>33</sup> The upper bound for divergence time of two groups of haplotypes was calculated as divergence time estimate, assuming STR variance in repeat number at the beginning of population subdivision ( $V_0$ ) equal to zero.<sup>34</sup> In the age calculation procedures, only tri- and tetranucleotide markers were used, so, besides the ambiguous loci DYS385a and DYS385b, locus DYS438 with pentanucleotide repeats was excluded from the calculations.

## RESULTS

SNP analysis of 885 Y chromosomes from 16 ethnic groups representing populations of northern East Asia demonstrates that haplogroup

**Table 1** Haplogroups Q, R2 and R1b distribution (no. of individuals and % values in parenthesis) in populations studied

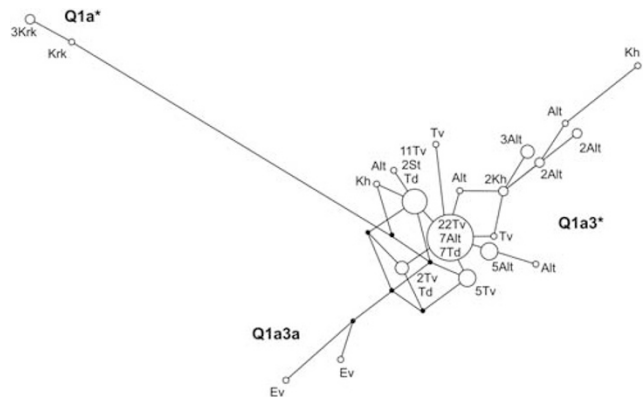
Population	N	Linguistic affiliation	Q1a*-MEH2	Q1a2-M25	Q1a3*-M346	Q1a3a-M3	R2-M124	R1b1b1-M73	R1b1b2-M269
Altaians	89	Turkic	0	0	23 (25.8)	0	0	1 (1.1)	0
Teleuts	44	Turkic	0	0	0	0	0	5 (11.4)	0
Khakassians	64	Turkic	0	0	4 (6.3)	0	0	2 (3.2)	0
Shors	38	Turkic	0	0	0	0	0	5 (13.2)	0
Todjins	26	Turkic	0	0	10 (38.5)	0	0	0	0
Tuvinians	108	Turkic	0	0	41 (38.0)	0	0	2 (1.9)	0
Yakuts	10	Turkic	0	0	0	0	0	0	0
Sojots	28	Turkic	0	0	2 (7.1)	0	0	0	0
Tofalars	30	Turkic	0	0	0	0	0	0	2 (6.7)
Mongols	46	Mongolic	0	0	0	0	0	0	2 (4.3)
Buryats	147	Mongolic	0	0	0	0	4 (2.7)	0	1 (0.7)
Kalmyks	60	Mongolic	0	1 (1.7)	1 (1.7)	0	6 (10.0)	2 (3.3)	0
Evenks	41	Tungusic	0	0	0	0	0	0	1 (2.4)
Evens	63	Tungusic	0	0	0	2 (3.2)	0	0	0
Koryaks	39	Chukotko-Kamchatkan	4 (10.3)	0	0	0	0	0	0
Koreans	52	Korean	0	0	0	0	0	0	0

Q is frequent in some Siberian populations studied (Table 1). This haplogroup consists of four subgroups: Q1a\*-MEH2, Q1a2-M25, Q1a3\*-M346 and Q1a3a-M3. Haplogroup Q1a3\*-M346 was found in Turkic-speaking Tuvians (38%), Todjins (38.5%), Altaians (25.8%), Sojots (7.1%) and Khakassians (6.3%), and only once in Mongolic-speaking Kalmyks (1.1%). Rare haplogroup Q1a2-M25 previously detected mostly in Iranians, Turks, Uygurs, Uzbeks and Han<sup>7,18,28,35</sup> was found also in Kalmyks (1.1%). Amerindian-specific haplogroup Q1a3a-M3 was present at a frequency of 3.2% only in Tungusic-speaking Evens inhabiting the Sea of Okhotsk coast. Four Koryak individuals (10.3%) from the same region belong to paralog Q1a\* (xQ1a1, Q1a2, Q1a3, Q1a4, Q1a5, Q1a6). They are all characterized by similar Y-STR profiles (Supplementary Table S1) and probably belong to a certain haplogroup, yet, unidentified by SNP marker.

Typing of Y chromosome marker M173 allowed us to define R1-structure in geographic region under investigation. We reported earlier phylogeographic pattern of haplogroup R1a1-M17 in Siberian populations, indicating that Shors and Teleuts had the highest frequency of this haplogroup (more than 50%), and that R1a1-M17 generally was more frequent in populations from Altai and eastern Sayan region than in the adjacent areas.<sup>24</sup> Subsequent SNP analysis demonstrates that two R1b haplogroups—R1b1b1-M73 and R1b1b2-M269—are present in some Siberian populations (Table 1). R1b1b2-M269 that is frequent in Europe is rarely observed in diverse set of Siberian populations: Evenks (2.4%), Buryats (0.7%), Mongols (4.3%) and Tofalars (6.7%). However, more interesting fact is the presence of haplogroup R1b1b1-M73 in the whole series of Turkic-speaking populations—Shors (13.2%), Teleuts (11.4%), Khakassians (3.2%), Tuvians (1.9%), Altaians (1.1%), as well as in Mongolic-speaking Kalmyks (2.2%). Quite the contrary, the remaining R-haplogroup, R2-M124, is present only in Mongolic-speaking Buryats (2.7%) and Kalmyks (6.6%) (Table 1). Haplogroup R2-M124 has a predominant distribution in South Asia (in India and Pakistan),<sup>25</sup> but it is also found in China (in Uygurs, Han, Hui)<sup>7</sup> and Central Asia (in Tajiks and Kyrgyz),<sup>28</sup> thus, its presence in some Siberian populations is likely a result of male influence from Central/East Asia.

To obtain a better resolution of phylogenetic relationships between Y chromosomes, we have analyzed haplotypes at 12 STR loci (Supplementary Table S1). Median joining network of haplogroup Q1a demonstrates subdivision of STR-haplotypes in accordance with their haplogroup affiliation into three clusters corresponding to haplogroups Q1a3\*-M346, Q1a3a-M3 and Q1a\*-MEH2 (Figure 1). Coalescence age of South Siberian Q1a3\*-M346, based on the average squared difference in the number of tri- and tetranucleotide repeats, is about  $4.03 \pm 1.25$  Ka, while the age of the Koryak Q1a\*-MEH2 appears to be only  $1.0 \pm 1.0$  Ka.

Median network of haplogroup R1b1b1-M73 shows that there are two subclusters of haplotypes in Siberian populations studied. One of them (designated as A in Figure 2) is determined by median haplotype 14-13-16-13-17-22-11-13-13-15-10-13 (for all loci studied: DYS19, DYS385a, DYS385b, DYS389I, DYS389II, DYS390, DYS391, DYS392, DYS393, DYS437, DYS438, DYS439) and another one (B) by haplotype 14-13-13-14-16-19-11-13-13-15-10-13 (Figure 2). The coalescence age of R1b1b1-M73 in South Siberia, based on tri- and tetranucleotide repeats, is estimated as about  $18.2 \pm 10.5$  Ka. The ages of subclusters A and B are amounted to  $4.4 \pm 1.5$  and  $5.6 \pm 4.0$  Ka, respectively. Analysis of published data demonstrates that these two subclusters are present simultaneously not only in Siberia but also in different ethnic populations of China<sup>7</sup> and the Caucasus.<sup>23,36</sup> In addition, in the Caucasus populations, a third subcluster of



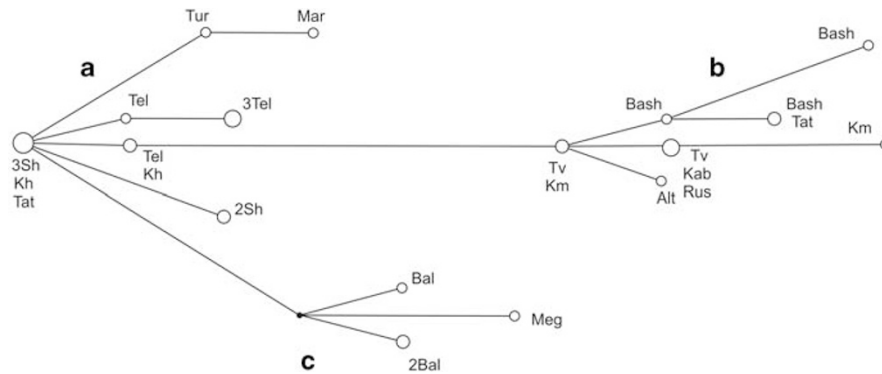
**Figure 1** Median joining network of haplogroups Q1a\*-MEH2, Q1a3\*-M346 and Q1a3a-M3 based on STR loci DYS19, DYS389I, DYS389II, DYS390, DYS391, DYS392, DYS393, DYS437, DYS438 and DYS439 in populations of North Asia. Each circle represents a haplotype, defined by a combination of STR markers. Circle size is shown proportional to haplotype frequency and the smallest circle represents one haplotype. The lines between circles represent mutational distance, the shortest distance being a single mutational step. Median vectors are indicated by black points. Haplotypes are labeled as follows: Alt—Altaians, Ev—Evens, Kh—Khakassians, Krk—Koryaks, St—Sojots, Td—Todjins, Tv—Tuvians.

R1b1b1-M73 indicated here as 'C' may be recognized (Figure 2). To infer migration history of haplogroup R1b1b1-M73, a further search of the subcluster-defining SNPs is required. In addition, more Y-STR data are needed for R1b1b1-M73 to make clear the origin of this haplogroup in Siberians.

## DISCUSSION

Previous studies have shown that haplogroup R1a1-M17 is present in South Siberia from the Holocene period ( $11.3 \pm 3.2$  Ka),<sup>24</sup> while the oldest age estimates dating back to Mesolithic times (approximately 18 Ka) were obtained for this haplogroup in South India and South Pakistan.<sup>37</sup> Similarly, high estimation of the age has been received for haplogroup R1b1b1-M73 ( $22.9 \pm 9.3$  Ka), based on variation of tri- and tetranucleotide repeats in populations of the Caucasus, South Ural and East Europe.<sup>36</sup> The age of this haplogroup in South Siberia is also high, about  $18.2 \pm 10.5$  Ka. However, a much younger age was found for haplogroup Q1a3\*-M346 in South Siberia ( $4.0 \pm 1.25$  Ka). One should note that the age of Q1a3\*-M346 calculated from haplotypes detected in populations of China and North Pakistan appears to be much higher— $17.8 \pm 4.1$  Ka according to Zhong *et al.*<sup>7</sup>

To compare STR diversity of haplogroups Q1a3\*-M346 and R1b1b1-M73 in different geographic regions we have formed a set of SNP-typed haplotypes limited by seven STR markers (DYS19, DYS389I, DYS389II, DYS390, DYS391, DYS392, DYS393) and retrieved from different population studies (Table 2). Haplogroup R1b1b1-M73 is characterized by the high intra-haplogroup variability and appears to be very ancient, being dated at about 40 Ka. We should note, however, that the calculated age of R1b1b1-M73 is so high probably due to very high repeat variance at locus DYS390 (4.31 versus 0.25–0.93 at the remaining six loci). Coalescence time estimate for this haplogroup in South Siberia corresponds to about 20 Ka, although its two phylogenetic STR-clusters (A and B) appear to be young ( $<4$  Ka), suggesting that their entrance in South Siberia occurred in relatively recent historical times or that evolution of this Y chromosome lineages in South Siberia was associated with recent population bottlenecks. It seems that appropriate ancient DNA samples should be studied to solve this problem.



**Figure 2** Median joining network of haplogroup R1b1b1-M73 based on STR loci DYS19, DYS389I, DYS389II, DYS390, DYS391, DYS392, DYS393, DYS437, DYS438 and DYS439 in populations of northern Eurasia. Designations are as in Figure 1. Haplotypes are labeled as follows: Alt—Altaians, Kh—Khakassians, Km—Kalmyks, Sh—Shors, Tel—Teleuts, Tv—Tuvinians. The network includes also haplotypes which were reported elsewhere;<sup>36</sup> Bal—Balkars, Bash—Bashkirs, Kab—Kabardinians, Mar—Mari, Meg—Megreles, Rus—Russians, Tat—Tatars, Tur—Turks. STR-based subclusters are designated by capital letters a, b and c.

**Table 2** Coalescence age estimates and median haplotypes of Y chromosome haplogroups

Haplogroup (geographic region)	Sample size	Coalescence age estimate (in Ka)	Median haplotype (for loci DYS19-DYS389I-DYS389II-DYS390- DYS391- DYS392- DYS393)
Q1a3*-M346 (China and North Pakistan) <sup>a</sup>	13	20.31 ± 4.58	13-13-16-23-10-14-13
Q1a3*-M346 (South Siberia) <sup>b</sup>	80	4.53 ± 1.49	13-13-18-23-10-14-13
Q1a3*-M346 (pooled) <sup>c</sup>	93	7.85 ± 2.54	13-13-18-23-10-14-13
R1b1b1-M73 (South Siberia) <sup>b</sup>	17	20.4 ± 13.45	14-13-17-22-11-13-13
R1b1b1-M73 (pooled) <sup>d</sup>	63	39.6 ± 21.25	14-13-17-22-11-13-13

<sup>a</sup>Including individuals from North Pakistan<sup>25</sup> and China (Uygurs, Han, Tibetans, Hui).<sup>7</sup>

<sup>b</sup>Including individuals from South Siberia (present study).

<sup>c</sup>Including individuals from North Pakistan, China and South Siberia.

<sup>d</sup>Including individuals from South Siberia (present study), China (Han, Tibetans, Mongolians, Uygurs, Kyrgyz, Naxi),<sup>7,25</sup> Turkey,<sup>18</sup> North Pakistan,<sup>25</sup> the Caucasus (Balkars, Kabardinians, Megreles),<sup>36</sup> East Europe (Mari, Tatars, Bashkirs, Russians).<sup>36</sup>

Similar problem exists with haplogroup Q1a3\*-M346, which is dated in South Siberia at 4.5 Ka. This value is much lower than that obtained in previous studies, although those estimations were based solely on Q\*(xQ1a3a-M3) STR diversity as the Y chromosome data have lacked enough resolution. For instance, Q\* has been dated to 17.7 Ka in South Siberia<sup>2</sup> or either 15.4 Ka in Mongolia.<sup>16</sup> According to study by Zegura *et al.*,<sup>20</sup> divergence time estimate for the Altaian and North Asian Q\* haplotypes versus the Native American Q1a3a-M3 haplotypes was  $17.2 \pm 4.6$  Ka. We have also calculated divergence time estimate between Q1a3\*-M346 haplotypes found in South Siberians ( $n=80$ ) and Q1a3a haplotypes revealed in northern Native Americans ( $n=184$ ; retrieved from Malhi *et al.*<sup>38</sup>). As a result, we found that, despite the low current diversity of haplogroup Q1a3\*-M346 in South Siberia, divergence time between Q1a3\*-M346 and Q1a3a-M3 is equal to  $13.81 \pm 3.88$  Ka, thereby pointing to a relatively recent entry date to America. One should note that similar divergence time was estimated for the other Amerindian founder haplogroup C3b-P39— $13.9 \pm 3.2$  Ka.<sup>20</sup> In addition, complete mtDNA genome studies have shown that all pan-American mtDNA haplogroups reveal entry times of 15–18 Ka, which are suggestive of a concomitant post-Last Glacial Maximum arrival from Beringia with early Amerindians.<sup>39–42</sup> As for the question whether the Q1a3\*-M346 was present in South Siberia in the post-Last Glacial Maximum, we can report that divergence time estimate for South Siberian versus Chinese/North Pakistan Q1a3\*-M346 haplotypes is  $15.29 \pm 5.49$  Ka, thus supporting the idea for the real presence of Q1a3\*-M346 in South Siberia in those times.

As for the presence of Amerindian-specific haplogroup Q1a3a-M3 in Evens, this finding is consistent with previous studies of Northeast Siberian populations. Earlier, haplogroup Q1a3a has been discovered at a low frequency in Siberian Eskimos, Chukchi and Evens.<sup>43,44</sup> Karafet *et al.*<sup>43</sup> have hypothesized that this haplogroup may have originated in the New World/eastern Beringia and that its infrequent presence in Northeast Siberian populations can be explained by back-migration from Alaska to Siberia. A search of Q1a3a-M3 STR-haplotypes revealed in Evens in the YHRD 3.0 database (<http://www.yhrd.org>; release 36 built on 15 April 2011; 91 493 haplotypes within 686 world populations) has shown that matching haplotypes (for loci DYS19, DYS385a, DYS385b, DYS389I, DYS389II, DYS390, DYS391, DYS392 and DYS393) are rarely present in modern Native American populations. Thus, this favors back-migration hypothesis for the presence of Q1a3a-M3 haplotypes in Northeast Siberia.

An interesting cluster of Q1a\*-MEH2 haplotypes was detected in Koryaks inhabiting the Sea of Okhotsk coast. Although such haplotypes are relatively frequent, exclusively in Koryaks (10.3%), they are characterized by low age, which corresponds to about 1.0 Ka. Nevertheless, it seems probable that this Q1a\*-MEH2 cluster appears to be much older taking into account that Q1a\* haplotype (xQ1a1, Q1a3, Q1a4, Q1a5, Q1a6) has been detected in extinct Palaeo-Eskimo individuals belonging to the Saqqaq culture (dated approximately 4.75–2.5<sub>14C</sub> ky).<sup>1</sup> It is noteworthy that according to results of SNP genotyping, the populations closest to the Saqqaq individuals are Koryaks and Chukchi.<sup>1</sup> As that study suggested that the ancestral



Saqqaq peoples separated from their Old World relatives about 5.5 Ka, it is possible that Q1a\*-MEH2 lineage detected in Koryaks represents an ancient genetic component, which in the past united the peoples of Northeast Asia, North America and Greenland. In order to determine the range of Q1a\*-MEH2 haplotypes detected in Koryaks we have performed a search for similar STR haplotypes in the YHRD 3.0 database. As a result, we did not find identical haplotypes among 91 493 haplotypes within 686 world populations, but the only similar 9-marker haplotype (loci DYS19, DYS385a, DYS385b, DYS389I, DYS389II, DYS390, DYS391, DYS392 and DYS393) was found in Yukaghirs from Northeast Siberia (Yakutia). According to Pakendorf *et al.*,<sup>29</sup> this haplotype (13-11-19-15-16-25-11-15-13; one-step differences from the Koryak Q1a\*-MEH2 haplotypes are shown in bold) was detected in 4 out of 13 Yukaghirs studied and belongs to haplogroup Q-P36. Therefore, it is likely that the range of Q1a\*-MEH2 may cover a distance of about 1000 km between the coasts of East Siberian Sea and Sea of Okhotsk. Coalescence age of the Koryak/Yukaghir Q1a\*-MEH2 haplotypes is about  $3.5 \pm 1.5$  Ka, that is within the bounds of the Saqqaq culture dating.

To sum up, our study demonstrates ancient genetic links between modern Siberians and Native Americans and supports a single-migration model, with a common genetic source for Native Americans in the Altai-Sayan region of South Siberia.<sup>20</sup> In addition, we found Y chromosome evidence demonstrating the close relationship between extinct Palaeo-Eskimo individual belonging to Saqqaq culture and modern Koryaks from Northeast Siberia, thus indicating that this region was a starting point of additional migrations across the Bering Strait.

## ACKNOWLEDGEMENTS

We are grateful to all the voluntary donors of DNA samples used in this study. This study was supported by the Program of Presidium of Russian Academy of Sciences 'Biodiversity and Gene Pools' and the Russian Foundation for Basic Research (11-04-00620, 11-04-98508).

- 1 Rasmussen, M., Li, Y., Lindgreen, S., Pedersen, J. S., Albrechtsen, A., Moltke, I. *et al.* Ancient human genome sequence of an extinct Palaeo-Eskimo. *Nature* **463**, 757–762 (2010).
- 2 Karafet, T. M., Osipova, L. P., Gubina, M. A., Posukh, O. L., Zegura, S. L. & Hammer, M. F. High levels of Y chromosome differentiation among native Siberian populations and the genetic signature of a boreal hunter-gatherer way of life. *Hum. Biol.* **74**, 761–789 (2002).
- 3 Derenko, M. V., Grzybowski, T., Malyarchuk, B. A., Dambueva, I. K., Denisova, G. A., Czarny, J. *et al.* Diversity of mitochondrial DNA lineages in South Siberia. *Ann. Hum. Genet.* **67**, 391–411 (2003).
- 4 Yao, Y. G., Kong, Q. P., Wang, C. Y., Zhu, C. L. & Zhang, Y. P. Different matrilineal contributions to genetic structure of ethnic groups in the Silk road region in China. *Mol. Biol. Evol.* **21**, 2265–2280 (2004).
- 5 Xue, Y., Zerjal, T., Bao, W., Zhu, S., Shu, Q., Xu, J. *et al.* Male demography in East Asia: a north-south contrast in human population expansion times. *Genetics* **172**, 2431–2439 (2006).
- 6 Derenko, M., Malyarchuk, B., Grzybowski, T., Denisova, G., Dambueva, I., Perkova, M. *et al.* Phylogeographic analysis of mitochondrial DNA in northern Asian populations. *Am. J. Hum. Genet.* **81**, 1025–1041 (2007).
- 7 Zhong, H., Shi, H., Qi, X. B., Duan, Z. Y., Tan, P. P., Jin, L. *et al.* Extended Y chromosome investigation suggests postglacial migrations of modern humans into East Asia via the northern route. *Mol. Biol. Evol.* **28**, 717–727 (2011).
- 8 Underhill, P. A., Passarino, G., Lin, A. A., Shen, P., Mirazon Lahr, M., Foley, R. A. *et al.* The phylogeography of Y chromosome binary haplotypes and the origins of modern human populations. *Ann. Hum. Genet.* **65** (Part 1), 43–62 (2001).
- 9 Comas, D., Calafell, F., Mateu, E., Perez-Lezaun, A., Bosch, E., Martínez-Arias, R. *et al.* Trading genes along the silk road: mtDNA sequences and the origin of central Asian populations. *Am. J. Hum. Genet.* **63**, 1824–1838 (1998).
- 10 Xu, S., Huang, W., Qian, J. & Jin, L. Analysis of genomic admixture in Uyghur and its implication in mapping strategy. *Am. J. Hum. Genet.* **82**, 883–894 (2008).
- 11 Lalueza-Fox, C., Sampietro, M. L., Gilbert, M. T., Castri, L., Facchini, F., Pettener, D. *et al.* Unravelling migrations in the steppe: mitochondrial DNA sequences from ancient central Asians. *Proc. Biol. Sci.* **271**, 941–947 (2004).
- 12 Moeder, K. P., Schurr, T. G., Bamforth, F. J., Bazaliiski, V. I. & Savel'ev, N. A. Population affinities of Neolithic Siberians: a snapshot from prehistoric Lake Baikal. *Am. J. Phys. Anthropol.* **129**, 349–361 (2006).
- 13 Keyser, C., Bouakaze, C., Crubézy, E., Nikolaev, V. G., Montagnon, D., Reis, T. *et al.* Ancient DNA provides new insights into the history of south Siberian Kurgan people. *Hum. Genet.* **126**, 395–410 (2009).
- 14 Li, C., Li, H., Cui, Y., Xie, C., Cai, D., Li, W. *et al.* Evidence that a West-East admixed population lived in the Tarim Basin as early as the early Bronze Age. *BMC Biol.* **8**, 15 (2010).
- 15 Wells, R. S., Yuldasheva, N., Ruzibakiev, R., Underhill, P. A., Evseeva, I., Blue-Smith, J. *et al.* The Eurasian heartland: a continental perspective on Y chromosome diversity. *Proc. Natl Acad. Sci. USA* **98**, 10244–10249 (2001).
- 16 Bortolini, M. C., Salzano, F. M., Thomas, M. G., Stuart, S., Nasanen, S. P., Bau, C. H. *et al.* Y-chromosome evidence for differing ancient demographic histories in the Americas. *Am. J. Hum. Genet.* **73**, 524–539 (2003).
- 17 Seielstad, M., Yuldasheva, N., Singh, N., Underhill, P., Oefner, P., Shen, P. *et al.* A novel Y-chromosome variant puts an upper limit on the timing of first entry into the Americas. *Am. J. Hum. Genet.* **73**, 700–705 (2003).
- 18 Cinnioglu, C., King, R., Kivisild, T., Kalfoglu, E., Atasoy, S., Cavalleri, G. L. *et al.* Excavating Y-chromosome haplotype strata in Anatolia. *Hum. Genet.* **114**, 127–148 (2004).
- 19 Pakendorf, B., Novgorodov, I. N., Osakovskij, V. L. & Stoneking, M. Mating patterns amongst Siberian reindeer herders: inferences from mtDNA and Y-chromosomal analyses. *Am. J. Phys. Anthropol.* **133**, 1013–1027 (2007).
- 20 Zegura, S. L., Karafet, T. M., Zhivotovsky, L. A. & Hammer, M. F. High-resolution SNPs and microsatellite haplotypes point to a single, recent entry of Native American Y chromosomes into the Americas. *Mol. Biol. Evol.* **21**, 164–175 (2004).
- 21 Semino, O., Passarino, G., Oefner, P. J., Lin, A. A., Arbuzova, S., Beckman, L. E. *et al.* The genetic legacy of Paleolithic *Homo sapiens* sapiens in extant Europeans: a Y chromosome perspective. *Science* **290**, 1155–1159 (2000).
- 22 Underhill, P. A., Myres, N. M., Rootsi, S., Metspalu, M., Zhivotovsky, L. A., King, R. J. *et al.* Separating the post-Glacial coancestry of European and Asian Y chromosomes within haplogroup R1a. *Eur. J. Hum. Genet.* **18**, 479–484 (2010).
- 23 Myres, N. M., Rootsi, S., Lin, A. A., Järve, M., King, R. J., Kutuev, I. *et al.* A major Y-chromosome haplogroup R1b Holocene era founder effect in Central and Western Europe. *Eur. J. Hum. Genet.* **19**, 95–101 (2011).
- 24 Derenko, M., Malyarchuk, B., Denisova, G., Wozniak, M., Dambueva, I., Dorzh, C. *et al.* Contrasting patterns of Y chromosome variation in South Siberian populations from Baikal and Altai-Sayan regions. *Hum. Genet.* **118**, 591–604 (2006).
- 25 Sengupta, S., Zhivotovsky, L. A., King, R., Mehdi, S. Q., Edmonds, C. A., Chow, C. E. *et al.* Polarity and temporality of high-resolution Y chromosome distributions in India identify both indigenous and exogenous expansions and reveal minor genetic influence of central Asian pastoralists. *Am. J. Hum. Genet.* **78**, 202–221 (2006).
- 26 Cruciani, F., Santolamazza, P., Shen, P., Macaulay, V., Moral, P., Olckers, A. *et al.* A back migration from Asia to sub-Saharan Africa is supported by high-resolution analysis of human Y-chromosome haplotypes. *Am. J. Hum. Genet.* **70**, 1197–1214 (2002).
- 27 Balareshque, P., Bowden, G. R., Adams, S. M., Leung, H. Y., King, T. E., Rosser, Z. H. *et al.* A predominantly neolithic origin for European paternal lineages. *PLoS Biol.* **8**, e1000285 (2010).
- 28 Kharkov, V. N. *Y-Chromosome Structure in Populations of Siberia*. PhD thesis (Research Institute of Medical Genetics, Tomsk, 2005) (in Russian).
- 29 Pakendorf, B., Novgorodov, I. N., Osakovskij, V. L., Danilova, A. P., Protod'jakonov, A. P. & Stoneking, M. Investigating the effects of prehistoric migrations in Siberia: genetic variation and the origins of Yakuts. *Hum. Genet.* **120**, 334–353 (2006).
- 30 Nasidze, I., Quinque, D., Dupanloup, I., Cordaux, R., Kokshunova, L. & Stoneking, M. Genetic evidence for the Mongolian ancestry of Kalmyks. *Am. J. Phys. Anthropol.* **128**, 846–854 (2005).
- 31 Karafet, T. M., Mendez, F. L., Meilerman, M. B., Underhill, P. A., Zegura, S. L. & Hammer, M. F. New binary polymorphisms reshape and increase resolution of the human Y-chromosomal haplogroup tree. *Genome Res.* **18**, 830–838 (2008).
- 32 Bandelt, H.-J., Macaulay, V. & Richards, M. Median networks: speedy construction and greedy reduction, one simulation, and two case studies from human mtDNA. *Mol. Phylogenet. Evol.* **16**, 8–28 (2000).
- 33 Zhivotovsky, L. A., Underhill, P. A., Cinnioglu, C., Kayser, M., Morar, B., Kivisild, T. *et al.* The effective mutation rate at Y chromosome short tandem repeats, with application to human population-divergence time. *Am. J. Hum. Genet.* **74**, 50–61 (2004).
- 34 Zhivotovsky, L. A. Estimating divergence time with the use of microsatellite genetic distances: impacts of population growth and gene flow. *Mol. Biol. Evol.* **18**, 700–709 (2001).
- 35 Regueiro, M., Cadenas, A. M., Gayden, T., Underhill, P. A. & Herrera, R. J. Iran: tricontinental nexus for Y-chromosome driven migration. *Hum. Hered.* **61**, 132–143 (2006).
- 36 Järve, M., Zhivotovsky, L. A., Rootsi, S., Help, H., Rogaev, E. I., Khushnutdinova, E. K. *et al.* Decreased rate of evolution in Y chromosome STR loci of increased size of the repeat unit. *PLoS One* **4**, e7276 (2009).
- 37 Mirabal, S., Regueiro, M., Cadenas, A. M., Cavalli-Sforza, L. L., Underhill, P. A., Verbenko, D. A. *et al.* Y-chromosome distribution within the geo-linguistic landscape of northwestern Russia. *Eur. J. Hum. Genet.* **17**, 1260–1273 (2009).
- 38 Malhi, R. S., Gonzalez-Oliver, A., Schroeder, K. B., Kemp, B. M., Greenberg, J. A., Dobrowski, S. Z. *et al.* Distribution of Y chromosomes among native North Americans: a study of Athapaskan population history. *Am. J. Phys. Anthropol.* **137**, 412–424 (2008).

- 39 Tamm, E., Kivisild, T., Reidla, M., Metspalu, M., Smith, D. G., Mulligan, C. J. *et al*. Beringian standstill and spread of Native American founders. *PLoS One*. **2**, e829 (2007).
- 40 Fagundes, N. J., Kanitz, R. & Bonatto, S. L. A reevaluation of the Native American mtDNA genome diversity and its bearing on the models of early colonization of Beringia. *PLoS One*. **3**, e3157 (2008).
- 41 Perego, U. A., Achilli, A., Angerhofer, N., Accetturo, M., Pala, M., Olivieri, A. *et al*. Distinctive Paleo-Indian migration routes from Beringia marked by two rare mtDNA haplogroups. *Curr. Biol*. **19**, 1–8 (2009).
- 42 Perego, U. A., Angerhofer, N., Pala, M., Olivieri, A., Lancioni, H., Kashani, B. H. *et al*. The initial peopling of the Americas: a growing number of founding mitochondrial genomes from Beringia. *Genome Res*. **20**, 1174–1179 (2010).
- 43 Karafet, T., Zegura, S. L., Vuturo-Brady, J., Posukh, O., Osipova, L., Wiebe, V. *et al*. Y chromosome markers and trans-Bering Strait dispersals. *Am. J. Phys. Anthropol.* **102**, 301–314 (1997).
- 44 Lell, J. T., Brown, M. D., Schurr, T. G., Sukernik, R. I., Starikovskaya, Y. B., Torroni, A. *et al*. Y chromosome polymorphisms in Native American and Siberian populations: identification of Native American Y chromosome haplotypes. *Hum. Genet.* **100**, 536–543 (1997).

Supplementary Information accompanies the paper on Journal of Human Genetics website (<http://www.nature.com/jhg>)