

# Variation of HLA class II genes in the Nganasan and Ket, two aboriginal Siberian populations

T. S. Uinuk-ool,\* ‡ N. Takezaki,\* O. A. Derbeneva, † N. V. Volodko † & R. I. Sukernik †

## Summary

Allelic frequencies at the three most polymorphic loci of the HLA class II region (DRB1, DQA1 and DQB1) were determined in the Nganasan and Ket, the remnants of the two most ancient groups in the Lower Yenisey River/Taimyr Peninsula region in northern Siberia. By single-stranded conformational polymorphism typing, verified by sequencing, 19 HLA-DRB1-DQA1-DQB1 haplotypes and 15 HLA-DRB1, seven DQA1 and 11 DQB1 alleles were found. The most frequent alleles were DRB1\*1301 (23.5%), DQA1\*0103 (29.4%), \*0501/03/05 (29.4%), and DQB1\*0301/09 (32.4%) in the Ket, and DRB1\*0901 (25%), DQA1\*0301 (39.6%), and DQB1\*0301/09 (37.5%) in the Nganasan. The distribution patterns and comprehensive phylogenetic analysis based on the haplotype frequencies of 17 Siberian populations suggest that the founders of both the Ket and the Nganasan came from Palaeolithic populations in the Altai-Sayan Upland.

## Introduction

The archaeological record of northern Eurasia suggests that southern Siberia was populated by modern humans at least 40 000 years before present (YBP) (Vasiliev *et al.*, 2002). However, it then took millennia to develop the biological and cultural adaptations necessary for the occupation of, and survival in, the Siberian Arctic zone, including the Lower Yenisey region and the Taimyr Peninsula. The populations of the Taimyr Peninsula appear to be related to those occupying the Altai-Sayan Upland, where the earliest modern Siberians underwent profound differentiation fomented by relatively moderate climatic conditions

and rich and diverse natural resources. This affinity is indicated by archaeological, anthropological and genetic evidence (Simchenko, 1976; Khlobystin, 1998; Derbeneva *et al.*, 2002).

Extensive studies of global populations have shown that common HLA haplotypes shared by populations on different continents and in different geographic regions arose, if not before the emergence of modern humans, then at least immediately after the African exodus (Clayton & Lonjou, 1997; Uinuk-ool *et al.*, 2003). These ancient (Dawkins *et al.*, 1991) or 'frozen' (Klein *et al.*, 1991) HLA haplotypes have presumably remained untouched by recombination since the time of their emergence. The distribution of haplotype frequencies in anthropological isolates has been largely determined by lineal dispersion followed by genetic drift (Takahata, 1991). Consequently, the frequencies can be used to elucidate the genetic history of human populations.

In our previous studies (Grahovac *et al.*, 1998; Uinuk-ool *et al.*, 2002), we determined the frequencies of HLA class II genes in 15 indigenous Siberian populations and demonstrated, by comprehensive phylogenetic analysis, that Asians separated into two clusters, one of which encompassed nearly all Siberians and all Native Americans, while the other consisted of the remaining Asian populations. The divergence time of these two clusters was estimated as 24 000–21 000 YBP. This estimate is at odds with archaeological data, which indicate that the initial human occupation of Siberia by modern humans occurred 45 000–40 000 YBP (Goebel & Aksenov, 1995; Goebel, 1999). The discrepancy could be explained by assuming that part of the HLA diversity originally present in the previously studied, relatively large populations might have been lost. This diversity could, however, still be present in small tribal populations of presumably very ancient origin. A few such populations still remain in Siberia, and they are largely restricted to the Arctic and Subarctic regions. In this paper we have therefore examined the variation at HLA class II loci (DRB1, DQA1, and DQB1) in the remnants of two ancient groups, the Ket in the Lower Yenisey River Basin and the Nganasan in the adjacent Taimyr Peninsula. The frequencies of the alleles obtained were compared with those characteristic of the Ket in the Middle Yenisey region, and with those of the indigenous populations inhabiting western, southern, and eastern parts of Siberia. Together, these data provide

\* Department of Immunogenetics, Max Planck Institute of Biology, Tübingen, Germany, and † Laboratory of Human Molecular Genetics, Institute of Cytology and Genetics, Siberian Division, Russian Academy of Sciences, Novosibirsk, Russia.

‡ Present address: Max-Planck-Institut für Infektionsbiologie, Abteilung Molekulare Biologie, Schumannstr. 21/22, D-10117 Berlin, Germany.

Received 6 August 2003; revised 28 October 2003; accepted 31 December 2003

Correspondence: Tatiana Uinuk-ool, Max-Planck-Institut für Infektionsbiologie, Abteilung Molekulare Biologie, Schumannstr. 21/22, D-10117 Berlin, Germany. Tel.: + 49 30 28460 0; Fax: + 49 30 28460 111; E-mail: tanya@tuebingen.mpg.de



**Figure 1.** Map of Siberia showing the location of populations from which blood samples were collected.

additional insights into the genetic relationships among the Siberian populations and between the linguistic groups of indigenous Siberians.

## Materials and methods

### Populations

Traditionally, the Ket comprised a few hunting and gathering tribes inhabiting the boreal forest zone of the Yenisey River basin (Aleksenko, 1967; Dolgikh, 1960). At present, they number several hundred individuals, most of whom have been assimilated by Russian settlers or merged with their western neighbours, the Selikups. The latter have a similar culture to the Ket but a different language. The language family of the Ket remains undefined, although some similarities to Tibetan languages have been noted (Ruhlen, 1988; Shimkin, 1994). In this study, we used blood samples drawn from 17 unrelated Ket currently residing in Turukhansk but representing the Lower Yenisey territorial group by their origin or place of birth (Fig. 1) (Derbeneva *et al.*, 2002). The remnants of the Southern (Upper) Ket, whose ancestors are believed to have originated from the Altai-Sayan region, have been almost totally integrated into the surrounding unrelated populations. The only exception is a small settlement, Sulamai, in which an original Ket population has survived into the 21st century. The data set obtained by analysis of

the HLA class II loci in 22 Ket from Sulamai has been published elsewhere (Grahovac *et al.*, 1998).

The Nganasan are direct descendants of the reindeer-hunting tribes scattered across northern Siberia since the early Neolithic (Simchenko, 1976; Khlobystin, 1998). Their language is part of the Uralic linguistic family. Throughout the 19th and 20th centuries the Nganasan exchanged marital partners with their western neighbours, the Entsi (former Yenisey Samoyeds; Dolgikh, 1952; Simchenko, 1976). The Entsi, now on the brink of extinction, originally inhabited the mouth of the Yenisey River. The current population size of the Nganasan, including those who merged with the Entsi, does not exceed 500 individuals. In this study we used 24 blood samples drawn from unrelated Nganasan and Entsi residing in the town of Dudinka (Derbeneva *et al.*, 2002).

### DNA extraction and polymerase chain reaction (PCR)

Venous blood was collected in vacutainers containing citric acid-dextrose as an anticoagulant. Genomic DNA was extracted from the buffy coat using the Qiagen DNA extraction kit (Qiagen, Hilden, Germany). PCR amplifications were carried out with the primers used by Uinuk-ool *et al.* (2002). The reaction was set up in 1 × reaction buffer (Qiagen), 2 mM of each of the four deoxynucleoside triphosphates, 2 μM of each of the two primers, 200 ng of genomic DNA template, and 2.5 units of HotStar

*Taq* DNA polymerase (Qiagen). The amplification programme consisted of 15 min of initial denaturation at 95 °C, followed by 35 amplification cycles of denaturation at 94 °C for 1 min, primer annealing at 54–62 °C (primer dependent), and primer extension for 1 min at 72 °C. The reaction was completed by a primer extension step at 72 °C for 10 min. The amplifications were carried out in the PTC-100 and PTC-200 Programmable Thermal Controllers (MJ Research, Biozym, Hessisch Oldendorf, Germany).

#### Single-stranded conformation polymorphism (SSCP) analysis

SSCP analysis (Orita *et al.*, 1989) was performed using the GeneGel Excel 12.5/24 Kit (Amersham Biosciences, Freiburg, Germany) under the conditions recommended by the manufacturer. Six-microlitre aliquots containing the specific PCR product were mixed with 6 µl of formamide-dye solution (88% formamide, 10 mM EDTA, 0.01% bromphenol-blue, and 0.01% xylene cyanol), heat-denatured at 95 °C for 5 min, and cooled rapidly in ice water for 5 min. Eight microlitres of the denatured DNA solution was loaded onto GeneGel Excel 12.5/24 gels. The electrophoresis conditions were 10 min at 200 V, 10 mA, 5 W and then 3 h at 375 V, 15 mA, 10 W at 15 °C (Mohabeer *et al.*, 1991; Kimura & Sasazuki, 1992; Kimura *et al.*, 1992). The DNA was visualized using the DNA Silver Staining Kit (Amersham Biosciences).

#### Subcloning of PCR products and sequencing

The PCR products were separated by 1.5% low-melting-point agarose gel electrophoresis. PCR fragments of the expected size were extracted from the gel with the help of the QIAquick gel extraction kit (Qiagen) and subcloned into the pGEM-T vector using the pGEM-T Easy Vector System (Promega, Mannheim, Germany). Plasmid DNA was isolated using the Qiagen Plasmid isolation kit (Qiagen). DNA was sequenced using the ThermoSequenase Primer Cycle Sequencing kit (Amersham Biosciences) with reverse and universal primers annealing to the multiple cloning sites of the pGEM-T Easy vector. Sequencing was carried out on the LI-COR Long ReadIR 4200 DNA sequencer (LI-COR Biosciences, Bad Homburg, Germany). The sequences were aligned with the help of the SEQPUP program (Gilbert 1995).

#### Statistical and phylogenetic analyses

Allele frequencies were obtained by direct counting. Three-loci haplotypes were either identified by family studies (using 16 samples from the Tuva population; data not shown) or inferred indirectly from homozygotes at one, two or all three loci, taking into account allele frequencies at the heterozygous loci.

Heterozygosities ( $H$ ) were calculated as

$$H = \frac{2n-1}{2n} \left(1 - \sum x_i^2\right)$$

where  $2n$  is the number of chromosomes examined in the population and  $x_i$  is the frequency of the  $i$ th allele (Nei, 1987). The expected homozygosity was computed as  $1 - H$ .

Dendrograms of populations were computed from allele frequency data by the neighbour-joining method (Saitou & Nei, 1987) using  $D_A$  distances (Nei *et al.*, 1983), which were obtained by the angular transformation of two points in a multidimensional space representing two populations. The distances were computed using the formula

$$D_A = 1 - \frac{1}{r} \sum_j^r \sum_i^{m_j} \sqrt{x_{ij}y_{ij}}$$

where  $r$  is the number of loci,  $m_j$  the number of alleles at the  $j$ th locus, and  $x_{ij}$  and  $y_{ij}$  the frequencies of the  $i$ th allele at the  $j$ th locus in populations  $X$  and  $Y$ , respectively. In comparative computer simulations, the  $D_A$  distance was shown to be more efficient in obtaining a correct tree topology than other distance measures, particularly when a small number of loci were examined, apparently because of the smaller coefficient of variation (Nei *et al.*, 1983; Takezaki & Nei, 1996).

## Results and Discussion

Each genomic DNA sample extracted from the Ket and Nganasan blood samples was used to amplify by PCR exon 2 of three HLA class II loci: DRB1, DQA1 and DQB1. PCR products were analysed by the SSCP method and samples showing different patterns were sequenced. In this manner, all the alleles of all the samples were identified on the basis of their sequences and are referred to by the designations of the HLA Nomenclature Committee (IMGT/HLA Sequence Database: <http://www.ebi.ac.uk/imgt/hla/index.html>). Frequencies of the alleles were computed and used for phylogenetic analyses. In all cases, the observed frequencies of homozygotes did not deviate significantly from those expected under Hardy–Weinberg equilibrium (data not shown).

In the populations tested, 15 DRB1, seven DQA1 and 11 DQB1 alleles were found (Tables 1–3). At the DRB1 locus, 10 and 13 alleles were found in the Ket and Nganasan populations, respectively. All these alleles were previously found in other Siberian populations (Grahovac *et al.*, 1998; Uinuk-ool *et al.*, 2002). Most of the shared alleles at the DRB1 locus in the Ket and Nganasan samples (\*0701, \*0801, \*1101, \*1201/06 and \*1403) were also common in the populations tested in the other two studies of Siberian populations (Grahovac *et al.*, 1998; Uinuk-ool *et al.*, 2002). One of these alleles (DRB1\*1403) is specific for central and east Asia; the others are present mostly in Old World populations and are absent or rare in Amerinds (for references, see Fig. 2 legend). Alleles DRB1\*1106, \*1401, \*0101, \*0401 and \*0403 were found only in the Nganasan and alleles DRB1\*0404 and 0408 only in the Ket.

The sample of 17 Ket individuals showed a similar distribution of allelic frequencies to the sample of 22 individuals of a different Ket population typed in our previous study (Grahovac *et al.*, 1998). Of the 10 alleles found in

**Table 1.** DRB1 allele frequencies (%) in two Siberian populations

Allele	Ket		Nganasan	
	<i>n</i>	<i>F</i> (%) (2 <i>n</i> = 34)	<i>n</i>	<i>F</i> (%) (2 <i>n</i> = 48)
0101	–	–	1	2.1
0401	–	–	1	2.1
0403	–	–	6	12.5
0404	1	2.9	–	–
0408	2	6	–	–
0701	4	11.8	5	10.4
0801	4	11.8	1	2.1
0901	1	2.9	12	25
1101	3	8.8	8	16.7
1106	–	–	1	2.1
1201/06 <sup>a</sup>	6	17.7	6	12.5
1301	8	23.5	2	4.2
1401	–	–	2	4.2
1403	2	5.9	2	4.2
1501	3	8.8	1	2.1

*n*, number of alleles; *F*, frequency. <sup>a</sup> These alleles could not be distinguished by the methods used.

**Table 2.** DQA1 allele frequencies (%) in two Siberian populations

Allele	Ket		Nganasan	
	<i>n</i>	<i>F</i> (%) (2 <i>n</i> = 34)	<i>n</i>	<i>F</i> (%) (2 <i>n</i> = 48)
0101/04/05 <sup>a</sup>	–	–	3	6.3
0102	1	2.9	1	2.1
0103	10	29.4	2	4.2
0201	4	11.8	5	10.4
0301/02/03 <sup>a</sup>	5	14.7	19	39.6
0401	4	11.8	1	2.1
0501/03/05 <sup>a</sup>	10	29.4	17	35.4

*n*, number of alleles; *F*, frequency. <sup>a</sup> These alleles could not be distinguished by the methods used.

each of these two populations, eight were shared and two were different. The frequencies of the latter two alleles were low and it is therefore possible that the alleles might have been missed in one or the other population. The DRB1\*1201/06 allele was found at a high frequency in both populations, but the frequency of the DRB1\*1301 allele was higher (23.5%) in the present population compared to that studied previously (5%). By contrast, the frequency of the DRB1\*1403 allele was lower in the present population (5.9%) than in the population studied earlier (30%). If not a result of a sampling error, this difference could indicate divergence of the two populations.

Of the 13 alleles found in the Nganasan, the most frequent was DRB1\*0901 (25%), which occurs in south-east Siberia at higher frequencies than in other parts of northern Asia (Uinuk-ool *et al.*, 2002). Moreover, the Nganasan also had a low frequency of the DRB1\*1106 allele (2.1%), which is otherwise present only in south-east Siberia. This finding could indicate that gene flow from

**Table 3.** DQB1 allele frequencies (%) in two Siberian populations

Allele	Ket		Nganasan	
	<i>n</i>	<i>F</i> (%) (2 <i>n</i> = 34)	<i>n</i>	<i>F</i> (%) (2 <i>n</i> = 48)
0201/02 <sup>a</sup>	4	11.8	5	10.4
0301/09 <sup>a</sup>	11	32.4	18	37.5
0302	3	8.8	6	12.5
03032	1	2.9	12	25
0401	4	11.8	1	2.1
0501	–	–	1	2.1
0502	–	–	1	2.1
0503	–	–	1	2.1
0601	1	2.9	–	–
0602/06111 <sup>a</sup>	10	29.4	2	4.2
0603/14	–	–	1	2.1

*n*, number of alleles; *F*, frequency. <sup>a</sup> These alleles could not be distinguished by the methods used.

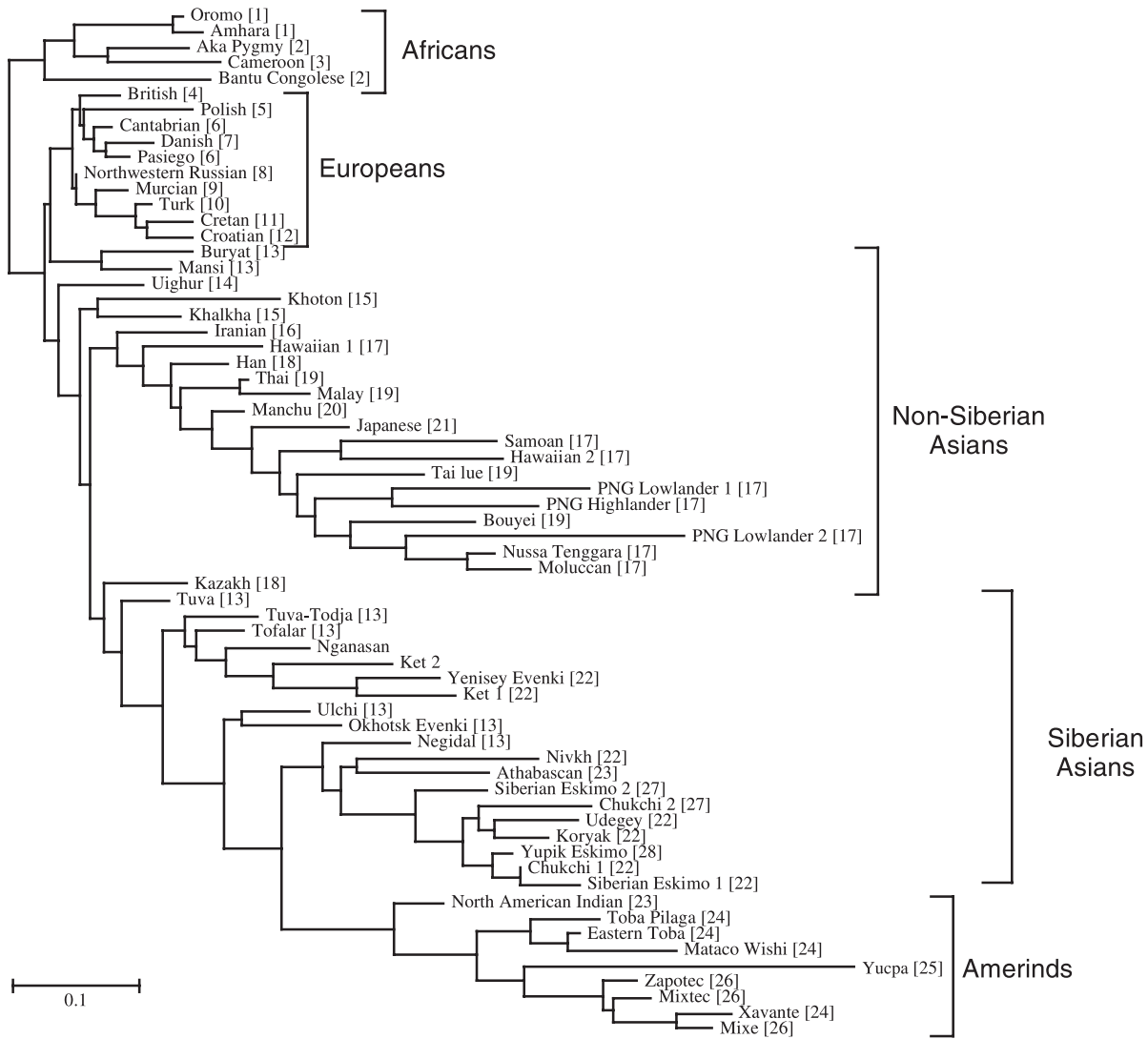
south-east Siberia might have occurred during the formation of the Nganasan gene pool.

At the DQA1 locus, six alleles were found in the Ket and seven alleles in the Nganasan populations (Table 2). The most frequent alleles were DQA1\*0301/02/03 (39.6%) and DQA1\*0501/03/05 (35.4%) in the Nganasan, and DQA1\*0103 and \*0501/03/05 (29.4% for both) in the Ket. In our earlier study the Ket population had a higher frequency of the DQA1\*0301/02/03 allele (24%) than of the DQA1\*0103 allele (13%). However, the frequencies of the DQA1\*0501/03/05 alleles in the two Ket populations were almost the same (30 and 29.4%).

At the DQB1 locus, seven and 10 alleles were obtained in the Ket and Nganasan, respectively. The most frequent allele in both populations was DQB1\*0301/09 (Table 3), which is frequent in Asia and America. The second most frequent allele was DQB1\*0602/06111 in the Ket and DQB1\*03032 in the Nganasan. These two alleles are frequent in central Siberian populations and in east Siberian populations, respectively.

Haplotypes were assigned on the basis of the haplotypes defined in a family study of a Tuvan population (Uinuk-ool *et al.*, 2002) and of homozygotes found in the present populations. In total, 19 HLA-DRB1-DQA1-DQB1 haplotypes were defined (Table 4), eight of which were identical to those obtained in the family study. Most of the haplotypes have a worldwide distribution and probably represent old haplotypes that originated before human expansion onto different continents. Six haplotypes are more specific for different populations: they are present in a few populations with low frequencies and probably arose later after differentiation of the populations.

There are two haplotypes, one (DRB1\*0403-DQA1\*0301/04/05-DQB1\*0302) present in the Nganasan (12.5%) and in most Siberian populations, and another (HLA-DRB1\*0404-DQA1\*0301/04/05-DQB1\*0302) found in the Ket of the present study (2.9%) and in the Mansi (2.9%), Buryats (4%), Negidals (5.7%), Ulchi (0.7%) and Koryaks (0.5%) of earlier studies (Grahovac *et al.*,



**Figure 2.** Neighbour-joining tree of human populations based on gene frequencies at two HLA class II loci, DRB1 and DQB1. The sources of population frequency data used (numbers in bracket) are as follows: (1) Fort *et al.* (1998); (2) Renquin *et al.* (2001); (3) Pimtanonthai *et al.* (2001); (4) Doherty *et al.* (1992); (5) Jungerman *et al.* (1997); (6) Sanchez-Velasco *et al.* (1999); (7) Lindblom & Svejgaard (1992); (8) Kapustin *et al.* (1999); (9) Muro *et al.* (2001); (10) Saruhan-Direskeneli *et al.* (2000); (11) Arnaiz-Villena *et al.* (1999); (12) Grubic *et al.* (1995); (13) Uinuk-ool *et al.* (2002); (14) Mizuki *et al.* (1998); (15) Munkhbat *et al.* (1997); (16) Amirzargar *et al.* (2001); (17) Mack *et al.* (2000); (18) Mizuki *et al.* (1997); (19) Chandanayingyong *et al.* (1997); (20) Geng *et al.* (1995); (21) Hashimoto *et al.* (1994); (22) Grahovac *et al.* (1998); (23) Monsalve *et al.* (1998); (24) Cerna *et al.* (1993); (25) Layrisse *et al.* (2001); (26) Hollenbach *et al.* (2001); (27) Krylov *et al.* (1995); (28) Leffell *et al.* (2002). In the notation of the original study (Mack *et al.*, 2000), Hawaii 1 and Hawaii 2 were Hawaii 38 and Hawaii 13, and PNG Lowlander 1 and PNG Lowlander 2 were PNG Lowland and PNG Lowland MV, respectively. PNG, Papua New Guinea.

1998; Uinuk-ool *et al.*, 2002). These haplotypes could also be observed on other continents, but they are found at the highest frequencies in some populations of Native Americans.

In the Ket we could define 13 haplotypes, of which the haplotypes HLA-DRB1\*1201/06-DQA1\*0501/03/05-DQB1\*0301/09 and HLA-DRB1\*1301-DQA1\*0103-DQB1\*0602/06111 were the most common, with frequencies of 17.7 and 23.5%, respectively. The first of these two haplotypes is common worldwide but occurs at the highest frequencies in Asia. It is present in Siberia

in all populations at high frequencies. The second haplotype, in contrast, has thus far only been found in the African Amhara population (Fort *et al.*, 1998), Iranians (Amirzargar *et al.*, 2001), Siberian Mansi (2.9%) and Tuvas (3.4%) (Uinuk-ool *et al.*, 2002). Both haplotypes have been confirmed by a family study in the Tuvan population.

The two most common haplotypes in the Nganasan were HLA-DRB1\*09012-DQA1\*0301/04/05-DQB1\*03032 (25%) and HLA-DRB1\*1101-DQA1\*0501/03/05-DQB1\*0301/09 (18.75%). The former is the most

**Table 4.** DRB1-DQA1-DQB1 haplotype frequencies in two Siberian populations (%)

DRB1	DQA1	DQB1	Ket (2n = 34)	Nganasan (2n = 48)
0101	0101/04/05 <sup>a</sup>	0501	–	2.1
15011/12 <sup>a</sup>	01021/22 <sup>a</sup>	0602/06111 <sup>a</sup>	2.9	2.1
15011	0103	0601	2.9	–
15011	0103	0602/06111 <sup>a</sup>	2.9	–
0401	0301/02/03 <sup>a</sup>	0301/09 <sup>a</sup>	–	2.1
0401	0301/02/03 <sup>a</sup>	0302	2.9	–
0403	0301/02/03 <sup>a</sup>	0302	–	12.5
0404	0301/02/03 <sup>a</sup>	0302	2.9	–
0408	0301/02/03 <sup>a</sup>	0302	2.9	–
1101	0501/03/05 <sup>a</sup>	0301/09 <sup>a</sup>	8.8	18.75
1201/06 <sup>a</sup>	0501/03/05 <sup>a</sup>	0301/09 <sup>a</sup>	17.7	12.5
1301	0103	0602	23.5	2.1
1301	0103	0603/14 <sup>a</sup>	–	2.1
1401	0101/04/05 <sup>a</sup>	0502	–	2.1
1401	0101/04/05 <sup>a</sup>	0503	–	2.1
1403	0501/03/05 <sup>a</sup>	0301/09 <sup>a</sup>	5.9	4.16
0701	0201	0201/02 <sup>a</sup>	11.8	10.4
0801	0401	0402	11.8	2.1
09012	0301/02/03 <sup>a</sup>	03032	2.9	25

<sup>a</sup> These alleles could not be distinguished by the methods used.

frequent haplotype in east Siberian populations (Uinuk-ool *et al.*, 2002) and is otherwise found elsewhere in Asia, but it is rare in Europe and in Native American populations. The latter is a common haplotype, with the highest frequency in European populations.

Using the frequency data from the three loci, three kinds of tree were constructed. The first was based on the frequencies of DRB1 alleles (not shown), the second on DRB1 and DQB1 frequencies (Fig. 2), and the third on DRB1, DQA1 and DQB1 frequencies (Fig. 3). The populations used for the trees included those from a previous study (Uinuk-ool *et al.*, 2002), but several new populations were added: Iranian, Murcian, Yucpa, Yupik Eskimo, Cameroonian, Aka Pigmy and Bantu Congolese (for references, see Fig. 2 legend). Almost all added populations cluster according to their geographical origin and language, with some exceptions probably caused by gene flow between the populations (Figs 2 and 3).

On all the trees, the Nganasan cluster together with the Tofalar and Todja populations from south Siberia, as well as with the Evenki and Ket from north-western Siberia (Grahovac *et al.*, 1998; Uinuk-ool *et al.*, 2002). This finding is in agreement with the historical and linguistic data available on these populations. The Nganasan language belongs to the Samoyedic language group, which the Tofalars and Todjas spoke originally. Only recently did the Tofalars and Todjas adopt the Turkish language of their neighbours, the Tuvans (Levin & Potapov, 1964). These observations support the notion that one of the northernmost populations of Siberia, the Nganasan, originated in south Siberia, but were later pushed to the north by invading new populations (Forsyth, 1992). The sharing of certain elements of material culture, such as reindeer

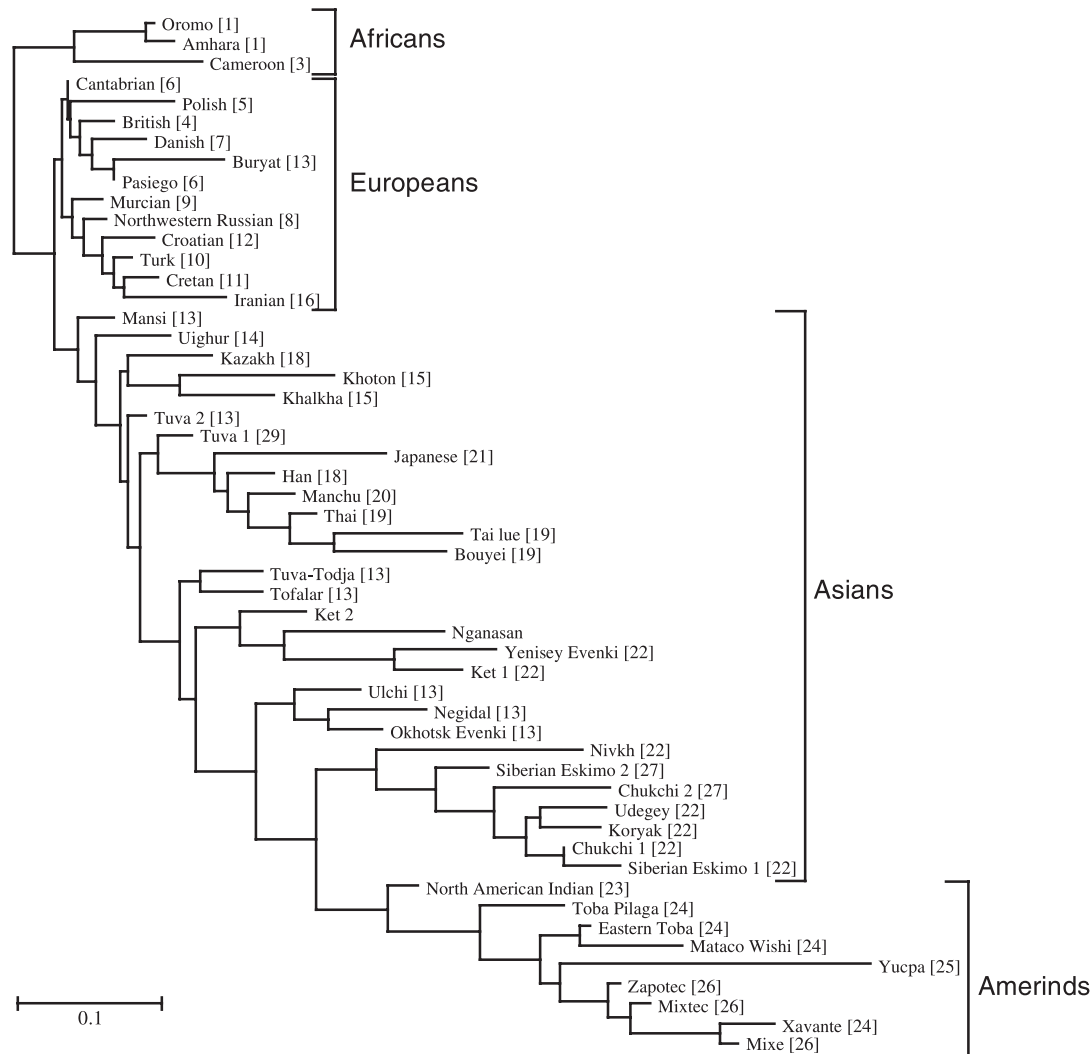
breeding, also indicates an affinity of the Nganasan to the Tofalars and Todjas. The clustering of these populations supports the southern origin of the Samoyed-speaking ethnic groups in the Sayan Mountains.

On all dendrograms, the Ket cluster with the populations described above. This clustering, too, is in agreement with ethnographical data. The Ket are the only people who speak the Ket language, which is unrelated to any other language. Several tribes speaking this language inhabited the valley of the Yenisey River from south to north. After the Russian invasion in the 17th century most of them mixed with Russians, Tungusic-speaking Evenki, and Turkic-speaking tribes and lost their culture and language. Only the more northern Ket tribes retained their original language. It appears that some of the extinct southern Ket-speaking tribes were involved in the ethnogenesis of the Tofalar and Todja people (Vanshtain, 1961; Levin & Potapov, 1964). This mixing probably explains why these populations cluster together on the HLA trees.

The mitochondrial DNA (mtDNA) diversity of the same samples, determined by Derbeneva *et al.* (2002), indicates the presence of both Asian (A, C, D and Z) and Caucasian (H, H2, H3, H8, U2, U4, U5, U7, J2 and W) haplogroups. The authors conclude that the high diversity and frequencies of H and U haplogroups represent traces of ancient, undifferentiated Upper Palaeolithic populations of modern *Homo sapiens*. In the present study, the Ket and Nganasan do indeed show a minor admixture of Caucasoid alleles. Similarly, the Mansi and Buryat populations tested in our previous study (Uinuk-ool *et al.*, 2002) showed high frequencies of Caucasoid HLA alleles. Unfortunately, the HLA polymorphism does not allow us to make inferences about the Neolithic origin of these alleles because the mutation rate in HLA genes is low compared to that of mtDNA and most of the HLA alleles now in existence originated in Africa or soon after the emigration of the ancestral population of *Homo sapiens* out of Africa. Indeed, in none of the studies of Siberian populations have new HLA alleles been found. The differences found among the populations in most cases are attributable to different distributions of the alleles and their different frequencies. It is therefore difficult to determine whether the European alleles found, for example, in the Mansi population (Uinuk-ool *et al.*, 2002) came from ancient Eurasians or whether they are the result of recent Russian admixture.

In summary, the geographic distribution and phylogeny of DRB1, DQA1 and DQB1 alleles suggest that the founding populations of the Nganasan and Ket originated in southern Siberia during the late Palaeolithic-early Neolithic period, and this suggestion is consistent with archaeological, anthropological and genetic (mitochondrial) evidence (Simchenko *et al.*, 1976; Khlobystin, 1998; Vasiliev, 2000; Derbeneva *et al.*, 2002).

The similarity in the composition of the HLA class II polymorphism between the two Ket populations sampled from different locations (present study and Grahovac *et al.*, 1998) and the grouping of both the Ket and the Nganasan populations with south and central Siberian



**Figure 3.** Neighbour-joining tree of human populations based on gene frequencies at three HLA class II loci: DRB1, DQA1 and DQB1. The sources of population frequency data used are given in brackets after the population names (see Fig. 2 legend); an additional reference is (29) Martinez-Laso *et al.* (2001).

populations, supported by historical and anthropological information, indicate that the obtained data are reliable despite the small sample size. However, to obtain more complete information about the genetic diversity of Siberian populations, and thus to resolve the discrepancies between the divergence time of Siberian populations from other Asians and the archaeological dating of the occupation of Siberia, more samples should be tested.

### Acknowledgements

We thank Dr Jan Klein for advice and discussion and Ms. Jane Kraushaar for editorial assistance.

### References

- Alekseenko, E.A. (1967) *Ket Ethnic History*. Nauka, Moscow.  
 Amirzargar, A., Mytilineos, J., Farjadian, S., Doroudchi, M., Scherer, S., Opelz, G. & Ghaderi, A. (2001) Human leukocyte

- antigen class II allele frequencies and haplotype association in Iranian normal population. *Human Immunology*, **62**, 1234.  
 Arnaiz-Villena, A., Iliakis, P., Gonzales-Hevilla, M., Longas, J., Gomez-Casado, E., Sfyridaki, K., Trapaga, J., Silvera-Redondo, C., Matsouka, C. & Martinez-Laso, J. (1999) The origin of Cretan populations as determined by characterizations of HLA alleles. *Tissue Antigens*, **53**, 213.  
 Cerna, M., Falco, M., Friedman, H., Raimondi, E., Maccagno, A., Fernandez-Vina, M. & Stastny, P. (1993) Differences in HLA class II alleles of isolated South American Indian populations from Brazil and Argentina. *Human Immunology*, **37**, 213.  
 Chandanayingyong, D., Stephens, H.A., Klaythong, R., Sirikong, M., Udee, S., Longta, P., Chantangpol, R., Bejrachandra, S. & Rungruang, E. (1997) HLA-A-B-DRB1-DQA1 and -DQB1 polymorphism in Thais. *Human Immunology*, **53**, 174.  
 Clayton, J. & Lonjou, C. (1997) Allele and haplotype frequencies for HLA loci in various ethnic groups. In: *HLA: Genetic Diversity of the Twelfth International Histocompatibility Workshop and Conference*, Vol. 1 (ed. by D. Charron), p. 665. Medical and Scientific International Publisher, Paris.

- Dawkins, R.L., Degli-Esposti, M.P., Abraham, L.J., Zhang, W. & Christiansen, F.T. (1991) Conservation versus polymorphism of the MHC in relation to transplantation, immune responses and autoimmune disease. In: *Molecular Evolution of the Major Histocompatibility Complex*, Vol. H59 (ed. by J. Klein & D. Klein), p. 391. Springer-Verlag, Heidelberg.
- Derbeneva, O.A., Starikovskaya, E.B., Volodko, N.V., Wallace, D.C. & Sukernik, R.I. (2002) [Mitochondrial DNA variation in the Ket and Nganasan and its implication for the initial peopling of Northern Eurasia.] *Genetika*, **38**, 1554 (in Russian).
- Doherty, D.G., Vaughan, R.W., Donaldson, P.T. & Mowat, A.P. (1992) HLA DQA, DQB, and DRB genotyping by oligonucleotide analysis. Distribution of alleles and haplotypes in British Caucasoids. *Human Immunology*, **34**, 53.
- Dolgikh, B.O. (1952) Origin of Nganasan. *Academy of Sciences USSR*, **18**, 5.
- Dolgikh, B.O. (1960) Tribal composition of native Siberian people in the 17th century. *Proceedings of the Institute of Ethnography of the Academy of Sciences USSR*, **55**, 3.
- Forsyth, J. (1992) *A History of the Peoples of Siberia*. Cambridge University Press, Cambridge.
- Fort, M., de Stefano, G.F., Cambon-Thomsen, A., Giraldo-Alvarez, P., Dugoujon, J.M., Ohayon, E., Scano, G. & Abbal, M. (1998) HLA class II allele and haplotype frequencies in Ethiopian Amhara and Oromo populations. *Tissue Antigens*, **51**, 327.
- Geng, L., Imanishi, T., Tokunaga, K., Zhu, D., Mizuki, N., Xu, S., Geng, Z., Gojobori, T., Tsuji, K. & Inoko, H. (1995) Determination of HLA class II alleles by genotyping in a Manchu population in the northern part of China and its relationship with Han and Japanese populations. *Tissue Antigens*, **46**, 111.
- Gilbert, D.G. (1995) *Seqpup Version 0.4j: A Biosequence Editor and Analysis Application*. <http://iubio.bio.indiana.edu/soft/molbio>.
- Goebel, T. (1999) Pleistocene human colonization of Siberia and peopling of the Americas: an ecological approach. *Evolutionary Anthropology*, **8**, 208.
- Goebel, T. & Aksenov, M. (1995) Accelerator radiocarbon dating of the initial Upper Paleolithic in southeast Siberia. *Antiquity*, **69**, 349.
- Grahovac, B., Sukernik, R.I., O'Huigin, C., Zaleska-Rutczynska, Z., Blagitko, N., Raldugina, O., Kosutic, T., Satta, Y., Figueroa, F., Takahata, N. *et al.* (1998) Polymorphism of the HLA class II locus in Siberian populations. *Human Genetics*, **102**, 27.
- Grubic, Z., Zunec, A., Naipal, A., Kastelan, A. & Giphart, M.J. (1995) Molecular analysis of HLA class II polymorphism in Croatsians. *Tissue Antigens*, **46**, 293.
- Hashimoto, M., Kinoshita, T., Yamasaki, M., Tanaka, H., Imanishi, T., Ihara, H., Ichikawa, Y. & Fukunishi, T. (1994) Gene frequencies and haplotypic associations within the HLA region in 916 unrelated Japanese individuals. *Tissue Antigens*, **44**, 166.
- Hollenbach, J.A., Thomson, G., Cao, K., Ferndndez-Vina, M., Erlich, H.A., Bugawan, T.L., Winkler, C., Winter, M. & Klitz, W. (2001) HLA diversity, differentiation, and haplotype evolution in Mesoamerican natives. *Human Immunology*, **62**, 378.
- Jungerman, M., Sanchez-Mazas, A., Fichna, P., Ivanova, R., Charron, D., Hors, J. & Djoulah, S. (1997) HLA class II DRB1, DQA1 and DQB1 polymorphisms in the Polish population from Wielkopolska. *Tissue Antigens*, **49**, 624.
- Kapustin, S., Lyschov, A., Alexandrova, J., Imyanitov, E. & Blinov, M. (1999) HLA class II molecular polymorphisms in healthy Slavic individuals from North-Western Russia. *Tissue Antigens*, **54**, 517.
- Khlobystin, L.P. (1998) *Ancient History of Taimyr and the Formation of the North Eurasian Cultures*. Dmitriy Bulanin, St. Petersburg.
- Kimura, A. & Sasazuki, T. (1992) Eleventh international histocompatibility workshop reference protocol for the HLA DNA-typing technique. In: *HLA 1991. Proceedings of the Eleventh International Histocompatibility Workshop and Conference*, Vol. 1 (ed. by K. Tsuji, M. Aizawa & T. Sasazuki), p. 397. Oxford University Press, Oxford.
- Kimura, A., Hoshino, S., Yoshida, M., Harada, H. & Sasazuki, T. (1992) Polymerase chain reaction-single strand conformation polymorphism (PCR-SSCP) analysis of HLA genes: a practical HLA-matching system. In: *HLA 1991. Proceedings of the Eleventh International Histocompatibility Workshop and Conference*, Vol. 1 (ed. by A. Tsuji, M. Aizawa & T. Sasazuki), p. 511. Oxford University Press, Oxford.
- Klein, J., O'Huigin, C., Kasahara, M., Vincek, V., Klein, D. & Figueroa, F. (1991) Frozen haplotypes in Mhc evolution. In: *Molecular Evolution of the Major Histocompatibility Complex*, Vol. H59 (ed. by J. Klein & D. Klein), p. 261. Springer-Verlag, Heidelberg.
- Krylov, M., Erdesz, S., Alexeeva, L., Benevolenskaya, L., Arnett, F.C. & Reveille, J.D. (1995) HLA class II and HLA-B27 oligotyping in two Siberian native population groups. *Tissue Antigens*, **46**, 382.
- Layrisse, Z., Guedez, Y., Dominguez, E., Paz, N., Montagnani, S., Matos, M., Herrera, F., Ogando, V., Balbas, O. & Rodriguez-Larralde, A. (2001) Extended HLA haplotypes in a Carib Amerindian population: the Yucpa of the Perija Range. *Human Immunology*, **62**, 992.
- Leffell, M.S., Fallin, M.D., Erlich, H.A., Fernandez-Vijna, M., Hildebrand, W.H., Mack, S.J. & Zachary, A.A. (2002) HLA antigens, alleles and haplotypes among the Yup'ik Alaska natives: report of the ASHI Minority Workshops, Part II. *Human Immunology*, **63**, 614.
- Levin, M.G. & Potapov, L.P. (eds) (1964) *The Peoples of Siberia*. University of Chicago Press, Chicago.
- Lindblom, B. & Sveigaard, A. (1992) HLA genes and haplotypes in the Scandinavian populations. In: *HLA 1991. Proceedings of the Eleventh International Histocompatibility Workshop and Conference* (ed. by K. Tsuji, M. Aizawa & T. Sasazuki), p. 651. Oxford University Press, Oxford.
- Mack, S.J., Bugawan, T.L., Moonsamy, P.B., Erlich, J.A., Trachtenberg, E.A., Paik, Y.K., Begovich, A.B., Saha, N., Beck, H.P., Stoneking, M. *et al.* (2000) Evolution of Pacific/Asian populations inferred from HLA class II allele frequency. *Tissue Antigens*, **55**, 383.
- Martinez-Laso, J., Sartakova, M.L., Allende, L., Konenkov, V.I., Moscoso, J., Silvera-Redondo, C., Pacho, A., Trapaga, J., Gomez-Casado, E. & Arnaiz-Villena, A. (2001) HLA markers in Tuvinians: a population with both Oriental and Caucasoid characteristics. *Annals of Human Genetics*, **65**, 245.
- Mizuki, N., Ohno, S., Ando, H., Sato, T., Imanishi, T., Gojobori, T., Ishihara, M., Ota, M., Geng, Z., Geng, L. *et al.* (1997) Major histocompatibility complex class II alleles in Kazak and Han populations in the Silk Route of northwestern China. *Tissue Antigens*, **50**, 527.
- Mizuki, N., Ohno, S., Ando, H., Sato, T., Imanishi, T., Gojobori, T., Ishihara, M., Goto, K., Ota, M., Geng, Z. *et al.* (1998) Major histocompatibility complex class II alleles in an Uyur population in the Silk Route of Northwest China. *Tissue Antigens*, **51**, 287.
- Mohabeer, A.J., Hiti, A.L. & Martin, W.J. (1991) Non-radioactive single strand conformation polymorphism (SSCP) using the Pharmacia 'PhastSystem'. *Nucleic Acids Research*, **19**, 3154.
- Monsalve, M.V., Edin, G. & Devine, D.V. (1998) Analysis of HLA class I and class II in Na-Dene and Amerindian populations from British Columbia, Canada. *Human Immunology*, **59**, 48.
- Munkhbat, B., Sato, T., Hagihara, M., Sato, K., Kimura, A., Munkhtuvshin, N. & Tsuji, K. (1997) Molecular analysis of HLA polymorphism in Khoton-Mongolians. *Tissue Antigens*, **50**, 124.
- Muro, M., Marin, L., Torio, A., Moya-Quiles, M.R., Minguela, A., Rosique-Roman, J., Sanchis, M.J., Garcia-Calatayud, M.C., Garcia-Alonso, A.M., Alvarez, L. *et al.* (2001) HLA polymorphism in the Murcia population (Spain): in the cradle of the archaologic Iberians. *Human Immunology*, **62**, 910.

- Nei, M. (1987) *Molecular Evolutionary Genetics*. Columbia University Press, New York.
- Nei, M., Tajima, F. & Tateno, Y. (1983) Accuracy of estimated phylogenetic trees from molecular data. II. Gene frequency data. *Journal of Molecular Evolution*, **19**, 153.
- Orita, M., Iwahana, H., Kanazawa, H., Hayashi, K. & Sekiya, T. (1989) Detection of polymorphisms of human DNA by gel electrophoresis as single-strand conformation polymorphisms. *Proceedings of the National Academy of Sciences USA*, **86**, 2766.
- Pimthanohai, N., Hurley, C.K., Leke, R., Klitz, W. & Johnson, A.H. (2001) HLA-DR and -DQ polymorphism in Cameroon. *Tissue Antigens*, **58**, 1.
- Renquin, J., Sanchez-Mazas, A., Halle, L., Rivalland, S., Jaeger, G., Mbayo, K., Bianchi, F. & Kaplan, C. (2001) HLA class II polymorphism in Aka Pygmies and Bantu Congolese and a reassessment of HLA-DRB1 African diversity. *Tissue Antigens*, **58**, 211.
- Ruhlen, M. (1988) *A Guide to the World's Languages. Vol. 1: Classification*, pp. 64; 127; 134. Stanford University Press, Stanford, CA.
- Saitou, N. & Nei, M. (1987) The neighbor-joining method: a new method for reconstructing phylogenetic trees. *Molecular Biology and Evolution*, **4**, 406.
- Sanchez-Velasco, P., Escribano de Diego, J., Paz-Miguel, J.E., Ocejio-Vinyals, G. & Leyva-Cobian, F. (1999) HLA-DR, DQ nucleotide sequence polymorphisms in the Pasiegos (Pas valleys, Northern Spain) and comparison of the allelic and haplotypic frequencies with those of other European populations. *Tissue Antigens*, **65**.
- Saruhan-Direskeneli, G., Uyar, F.A., Bakar, S. & Eraksoy, M. (2000) Molecular analysis of HLA-DRB1-DQA1 and -DQB1 polymorphism in Turkey. *Tissue Antigens*, **55**, 171.
- Shimkin, D. (1994) Ket. In: *Encyclopedia of World Cultures, Vol. 6: Russia and Euroasia/China* (ed. by P. Friedrich & N. Diamond), p. 183. G. K. Hall, New York.
- Simchenko, Y.B. (1976) *The Culture of Reindeer Hunters in Northern Siberia*. Nauka, Moscow.
- Takahata, N. (1991) Trans-species polymorphism of HLA molecules, founder principle, and human evolution. In: *Molecular Evolution of the Major Histocompatibility Complex* (ed. by J. Klein & D. Klein), p. 29. Springer-Verlag, Heidelberg.
- Takezaki, N. & Nei, M. (1996) Genetic distances and reconstruction of phylogenetic trees from microsatellite DNA. *Genetics*, **144**, 389.
- Uinuk-ool, T.S., Takezaki, N. & Klein, J. (2003) Ancestry and kinships of native Siberian populations: the HLA evidence. *Evolutionary Anthropology*, **12**, 231.
- Uinuk-ool, T.S., Takezaki, N., Sukernik, R.I., Nagl, S. & Klein, J. (2002) Origin and affinities of indigenous Siberian populations as revealed by HLA class II gene frequencies. *Human Genetics*, **110**, 209.
- Vanshtain, C.I. (1961) *Tuvans-Todjans*. Publisher of Eastern Literature, Moscow (in Russian).
- Vasiliev, S.A. (2000) The Siberian mosaic: upper Palaeolithic adaptations and change before the last glacial maximum. In: *Hunters of the Golden Age* (ed. by W. Roebroeks, M. Mussi, J. Svoboda & K. Fennema), p. 173. University of Leiden, Leiden.
- Vasiliev, S.A., Kuzmin, Y.V., Orlova, L.A. & Dementiev, V.N. (2002) Radiocarbon-based chronology of the Paleolithic in Siberia and its relevance to the peopling of the New World. *Radiocarbon*, **44**, 503.