Ann. Hum. Genet. (2001), **65**, 63–78
Printed in Great Britain

Mitochondrial DNA variability in Russians and Ukrainians: Implication to the origin of the Eastern Slavs

B. A. MALYARCHUK AND M. V. DERENKO
Institute of Biological Problems of the North, 685000 Magadan, Russia

(Received 13.6.00. Accepted 9.11.00)

SUMMARY

In order to investigate the origin of the Eastern Slavs, mitochondrial DNA (mtDNA) sequence variation was examined in Russians and Ukrainians by hypervariable segment I (HVS I) sequencing and restriction analysis of the haplogroup-specific sites. No significant differences were found for Russians and Ukrainians when compared to other Europeans – in fact, they fall within the range of gene diversity seen throughout Europe and exhibit the unimodal pattern of pairwise sequence differences. Moreover, HVS I sequences in the Russians and Ukrainians are similar or identical to those found in eastern and western European populations. Despite the small genetic distances between Europeans, phylogenetic analysis reveals a considerable heterogeneity of Eastern Slavonic populations – they do not cluster together onto a phylogenetic tree. Analysis of distribution of rare HVS I types shared between populations of Eastern Slavs and other West Eurasians has shown that Russians share rare haplotypes mainly with Germans and Finno–Ugric populations. Of these, subhaplogroup H1 sequence types, which are defined by different combinations of nucleotides 16192T, 16294T, 16304C, 16311C and 16320T, are found predominantly in common between Russians and German-speaking populations. The data obtained allow us to conclude that the Slavonic migrations in early Middle Ages from their putative homeland in central Europe to the east of Europe were accompanied mostly by the same mtDNA types characteristic for the pre-Slavonic populations of eastern Europe.

INTRODUCTION

Analysis of mitochondrial DNA (mtDNA) polymorphism has become a useful tool for human population and molecular evolution studies (Wallace, 1995). The maternal mode of inheritance of the mitochondrial genome and the high rate of base substitutions allow the use of mtDNA polymorphisms for inferring the pattern of prehistoric female migrations and peopling of different regions of the world.

Archaeological and anthropological studies indicate that the colonization of the Eastern European plain by Slavs from the central part of Europe began in the early Middle Ages (Niederle, 1896; Alekseeva, 1973; Sedov, 1979). However, anthropologically the Russians, Ukrainians and Belorussians – the main modern ethnic groups of Eastern Slavs – are heterogeneous (Alekseeva & Alekseev, 1989). There are several theories on the origin of Eastern Slavs. These theories are based on the morphological, cultural and genetic variation represented in modern and ancient Eastern Slavonic populations. Currently, hypotheses can be classified into two groups: hybridization and transformation theories. Hybridization theories claim that modern Russians, Ukrainians and Belorussians are the result of an admixture between Slavonic tribes, whose homeland was probably in central Europe, and pre-Slavonic populations of eastern Europe, such as Finno–Ugric (on the north-west and east of eastern

Correspondence: Dr Boris A. Malyarchuk, Genetics Laboratory, Institute of Biological Problems of the North, Portovaya str., 18, 685000 Magadan, Russia. Tel/Fax: 741322 34463.
E-mail: ibpn@online.magadan.su

Europe), Baltic (on the west) and Iranic tribes (on the south) (Alekseeva, 1973). This theory predicts that Eastern Slavs have genes deriving from both the Slavonic and pre-Slavonic people.

On the other hand, transformation theory posits that Russians, as well as Ukrainians, gradually evolved from ancient populations of eastern Europe, at least from the Late Bronze Age (Alekseev, 1989). In this theory, in the early Middle Ages Slavonic people might have contributed culturally but not genetically to the formation of Eastern Slavonic tribes, imposing a language of the Slavonic group. In this scenario, a model of linguistic replacement, and most likely an elite dominance process (Renfrew, 1994), is assumed.

However these theories of the Eastern Slavs origin were primarily based on the morphological evidence. Therefore, it is necessary to study the origin and formation of the modern Russians, Ukrainians and Belorussians on the basis of molecular genetic evidence. The first studies of the mtDNA non-coding hypervariable segment I (HVS I) sequences, as well as restriction enzyme analysis in the coding region, have allowed us to conclude that the Eastern Slavs are closely related to the western European populations (Malyarchuk *et al.* 1995; Malyarchuk, 1997). Recent analysis of the mtDNA HVS I of Russian populations from the European part of Russia (Orekhov *et al.* 1999) is completely in agreement with this conclusion. To study the origin of the Eastern Slavs in detail, we present here mtDNA diversity data in Russians and Ukrainians, based on the HVS I sequences typed for the presence of major West Eurasian haplogroup-specific markers, described by Macaulay *et al.* (1999).

subjects and methods

Subjects and mtDNA analysis

DNA samples from 50 Russians and 18 Ukrainians were obtained from maternally-unrelated individuals residing in Magadan, north-eastern Siberia, and originating from different regions of the former Soviet Union.

PCR fragments, encompassing the entire mtDNA first hypervariable segment, were amplified using amplification primers L15926 and H16498 (Di Rienzo & Wilson, 1991). Nucleotide sequences between nucleotide positions (nps) 16000 and 16400 were determined by use of the Sanger dideoxy chain-termination method and Sequenase enzyme with amplification primers as sequencing primers. The HVS I sequences obtained were compared with the Cambridge reference sequence – CRS (Anderson *et al.* 1981).

To determine the RFLP haplogroup status of the Eastern Slavonic sequences, RFLP typing was performed by restriction endonuclease analysis of PCR amplified mtDNA fragments using the same primer pairs and amplification conditions as described by Torroni *et al.* (1996, 1997). The samples were typed for a restricted set of RFLPs that were diagnostic of all major west Eurasian clusters, on the basis of the hierarchical mtDNA RFLP scheme (Macaulay *et al.* 1999).

To determine haplogroup H sequences, all samples were tested for 14766*Mse*I, 10394*Dde*I, and 7025*Alu*I. Samples lacking these three sites were assigned to cluster H. Haplogroup H samples were tested for 4332*Ava*II (Torroni *et al.* 1994) and those harbouring +4332*Ava*II and 16304C in HVS I were assigned to subhaplogroup H1. All non-H samples harbouring −14766*Mse*I and −10394*Dde*I were tested for 4577*Nla*III, and samples lacking the *Nla*III site were classified as cluster V. All non-H and non-V samples (−14766*Mse*I and −10394*Dde*I) were determined as HV*.

All non-HV samples were tested for 12308*Hin*fI. Those with +12308*Hin*fI were assigned to clusters U and K, and were further determined as belonging to haplogroup K or to subgroups of the haplogroup U on the basis of the HVS I motif information (Richards *et al.* 1998; Macaulay *et al.* 1999).

The remaining samples were tested for 13366*Bam*HI, 15606*Alu*I, 15925*Msp*I, and 12629*Ava*II. Those with +13366*Bam*HI, +15606*Alu*I, and −15925*Msp*I were assigned to cluster T. The haplogroup T sequences lacking the 12629*Ava*II site were classified as T1, whereas those with +12629*Ava*II were determined as T*.

The remaining samples were tested for 13704*Bst*OI, and those with −13704*Bst*OI and +10394*Dde*I were classified as J.

The remaining mtDNAs were further classified as follows: +14465*Acc*I to cluster X; +10394*Dde*I and +10397*Alu*I to cluster M. M-sequences were classified as belonging to haplogroup Z on the basis of the HVS I motif information (Schurr *et al*. 1999). Haplogroups I and W sequences were also determined on the basis of the HVS I motifs classification (Richards *et al*. 1998; Macaulay *et al*. 1999).

Phylogenetic analyses

The phylogenetic relationships between mitochondrial haplotypes comprising various combinations of the HVS I sequences and RFLPs were analysed by the median-network method (Bandelt *et al*. 1995).

To infer the phylogenetic relationships between populations, HVS I sequence data from 19 populations from Europe and Asia Minor (total sample of 1045 individuals) were used. Previously published HVS I sequences were edited to a standard length (nps 16051–16362), deletion and insertion polymorphisms were removed. Populations coded as: TUR1 – 29 Turks (Calafell *et al*. 1996); TUR2 – 45 Turks (Comas *et al*. 1996); ADY – 50 Adygei (Macaulay *et al*. 1999); BUL – 30 Bulgarians (Calafell *et al*. 1996); TUS – 49 Tuscans (Francalacci *et al*. 1996); GER1 – 200 Southern Germans (Lutz *et al*. 1998); GER2 – 50 Western Germans (Baasner *et al*. 1998); SWI – 74 Swiss (Pult *et al*. 1994); AUS – 101 Austrians (Parson *et al*. 1998); FIN – 50 Finns, VFI – 34 Volga Finnic, EST – 47 Estonians, KAR – 83 Karelians, SWE – 32 Swedes (Sajantila *et al*. 1995); RUS1 – 50 Russians from Magadan (present data); UKR – 18 Ukrainians (present data); RUS2 – 55 Russians from Kostroma, RUS3 – 14 Russians from Rjazan, RUS4 – 34 Russians from Kursk (Orekhov *et al*. 1999).

The relationships between these populations were studied through intermatch-mismatch diversity values using the Sendbs program (provided by N. Takezaki, National Institute of Genetics, Tokyo). Several DNA distances between populations were calculated after a 1000 haplotype permutation between populations, and the neighbour-joining (NJ) tree was built.

Comparative analysis of mtDNA data

For a sequence sharing analysis, published HVS I haplotypes from Eurasian populations (in a combined set of 2799 individuals) were compared with the HVS I sequences of Russians and Ukrainians presented here. In addition to the populations mentioned above, data from the following populations were used: 42 Middle Easterners and 69 Sardinians (Di Rienzo & Wilson, 1991); 45 Israeli Druze (Macaulay *et al*. 1999); 85 Algerians and 54 Portuguese (Côrte-Real *et al*. 1996); 18 Berbers and 56 Canary Islanders (Pinto *et al*. 1996); 139 Spanish (Côrte-Real *et al*. 1996; Pinto *et al*. 1996; Torroni *et al*. 1999); 105 Basques (Bertranpetit *et al*. 1995; Côrte-Real *et al*. 1996); 22 Turks, 29 Finns, 33 Danes, 107 Northern Germans, 49 Bavarians (Richards *et al*. 1996); 32 Russian Germans from a German community in western Siberia (Baranov *et al*. 1999); 73 Finns and 28 Swedes (Kittles *et al*. 1999); 261 British (Piercy *et al*. 1993; Richards *et al*. 1996); 70 individuals from the eastern Italian Alps (Stenico *et al*. 1996); 205 subjects from central Asian admixed populations of East Asian and European ancestry: 55 Kazakh, 95 Kirghiz, 55 Uighur (Comas *et al*. 1998); 300 individuals from India (Kivisild *et al*. 1999*a*). HVS I sequences of the Ossetians, Georgians, and Armenians from the Caucasus, represented in the form of median networks (Metspalu *et al*. 1999), were also analysed.

In addition, HVS I sequence data on 572 East Asians and native Siberians (Shields *et al*. 1993; Horai *et al*. 1996; Kolman *et al*. 1996; Derenko & Shields, 1997; Kazakovtseva, 1998) were incorporated into the analysis.

RESULTS AND DISCUSSION

Sequence variability in Russians and Ukrainians

In the present study, we have sequenced the mtDNA region between nps 16000 and 16400

encompassing hypervariable segment I (HVS I) in 50 Russians and 18 Ukrainians. Overall, 48 nucleotide sites were polymorphic (Table 1). Only two sites presented transversions, both C to A, at nps 16114 and 16242. Among transitions, there is a high proportion of pyrimidine substitutions (37 positions) compared to purine transitions (9 positions).

Sequence comparison detected 54 different mitochondrial haplotypes (Table 1), 9 of them were found more than once in the total sample, 6 were shared between Russians and Ukrainians. Forty-four HVS I sequence types were found in 50 Russians. The most frequently occurring sequence types in the Russians were the CRS (Anderson *et al*. 1981) – one of the most frequent HVS I sequences in West Eurasia – and the 16069T–16126C haplotype belonging to haplogroup J (in 4 and 3 individuals, respectively). The results confirm the high variability of the mtDNA control region in the Ukrainians, with 16 different haplotypes in 18 individuals studied. The most frequent sequences in this small sample of Ukrainians were haplotypes 16223T (in two individuals) and 16126C–16294T–16296T–16304C (in two individuals).

The pairwise nucleotide difference distributions for the Russians and Ukrainians, as well as for most European populations, are clearly bell-shaped, unimodal being consistent with exponentially growing populations with different periods of expansion (Rogers & Harpending, 1992). Russians show peaks at four differences, with a mean of 4.56 ± 0.52 differences (standard error based on 1000 bootstraps), whereas Ukrainians exhibit peaks at five differences and a higher mean (4.76 ± 0.64). These means of pairwise differences are in accordance with previous data for European populations (Comas *et al*. 1997).

Population phylogenetic analysis

By means of phylogenetic analysis it has been shown recently (Francalacci *et al*. 1996; Comas *et al*. 1997) that European populations demonstrate limited genetic differentiation. Limited geographic patterning of mtDNA diversity in Europe was observed also when autocorrelation statistics were used (Simoni *et al*. 2000). Low differentiation suggests that most mitochondrial sequences in eastern and western Europe have a local and common ancestry in the Upper Paleolithic and Neolithic ages.

Although all European populations have very small genetic distances between them (not shown), the results of phylogenetic analysis demonstrate a considerable heterogeneity of Russian populations. Figure 1 shows a NJ tree for 19 populations, which was constructed on the basis of the HVS I sequences. This population tree demonstrates a major division between populations from Asia Minor and southern Europe (Turks and Tuscans) and the remaining populations, which can be divided into two subclusters: one that includes the Ukrainians, Bulgarians and Adygei, and another subcluster that includes the majority of European populations analysed. We should note, however, that the majority of clusters in the NJ tree had very low bootstrap probabilities ($< 50\%$). The cluster that had more support (bootstrap value is 54%) included the Bulgarians and Adygei.

It is interesting that Russian populations do not cluster together and that Russians from the southern European part of Russia (Kursk – RUS4) are closely related to 'southern' populations, such as Turks and Tuscans, whereas the Russians from the central-northern European part of Russia (Kostroma – RUS2) and Russians from Magadan (RUS1) are clustered together with the Volga Finnic sample, Swedes and Austrians. The sample from the central European region of Russia (Rjazan – RUS3) is closely related to the Southern Germans and Swiss. The Slavonic-speaking Bulgarians and Ukrainians occupy the third position on the mitochondrial phylogenetic tree together with the Adygei from Northern Caucasus. It seems that these populations are grouped together on the basis of an anthropological criterion because the Pontic anthropological type characterizes them (Alekseeva & Alekseev, 1989).

In general, the results of phylogenetic analysis indicate that Slavonic populations sharing the

Table 1. *HVS I sequences and RFLP status of Russian and Ukrainian samples*

Sample number[a]	HVS I sequence[b]	RFLP status[c]
RUS 22	CRS	H
RUS 58	CRS	H
RUS 67	CRS	H
RUS 76	CRS	H
UKR 60	CRS	H
RUS 75	168	H
RUS 54	218	H
RUS 74	93	H
RUS 21	129	H
RUS 80	278	H
RUS 69	278 390	H
RUS 19	362	H
UKR 83	362	H
UKR 84	223	H
UKR 86	223	H
RUS 36	325	H
UKR 15	93 325	H
RUS 4	189 311	H
UKR 88	189 362	H
RUS 31	304 311	H1
RUS 29	93 192 304 311	H1
RUS 64	189 356	H
RUS 65	189 274	H
UKR 85	93 271	H
UKR 41	278 293 311	H
RUS 61	147 169 189	H
RUS 11	189 243	H
RUS 78	93 189 298 311	V
RUS 14	129 223 291 298	V
RUS 50	311	HV*
RUS 81	311	HV*
RUS 26	266 311	HV*
RUS 49	153	HV*
RUS 62	286 304	U*
UKR 7	242 C-A 274 288 356 362	U4
RUS 30	356	U4
UKR 17	356	U4
RUS 12	114 C-A 192 256 270 294	U5
RUS 5	192 256 270 362 399	U5
RUS 20	192 256 270	U5
UKR 82	192 256 270 311	U5
UKR 57	256 270 399	U5[d]
RUS 48	256 278 399	U5
RUS 35	224 311	K
UKR 6	224 311	K
UKR 89	69 126	J*
RUS 27	69 126	J*
RUS 72	69 126	J*
RUS 73	69 126	J*
RUS 2	126 260 355 362	pre-HV*
RUS 13	126 163 186 189 294	T1
UKR 39	126 163 186 189 294	T1
RUS 17′	126 163 186 189 271 294	T1
RUS 51	126 294	T*
RUS 77	126 294 296	T*
UKR 56	126 294 296 304	T*
UKR 79	126 294 296 304	T*
RUS 33	93 126 294 296 304	T*
RUS 0	111 126 294 296 304 311 327	T*
RUS 8	126 153 294 362	T*
RUS 23′	126 189 292 294	T*
RUS 3	294 304	T*

Table 1. (*cont.*)

Sample number[a]	HVS I sequence[b]	RFLP status[c]
RUS 1	189 223 278	X
UKR 87	86 189 223 278	X
RUS 23	172 189 223 255 278	X
RUS 46	223 292	W
RUS 71	129 294 311 391	I
RUS 9	129 185 223 224 260 298	Z

[a] RUS, Russians; UKR, Ukrainians.
[b] Nucleotide positions (−16000), between 16000 and 16400, that are different from the Cambridge reference sequence (CRS; Anderson *et al.* 1981). Mutations are transitions, transversions are further specified.
[c] Haplogroup membership was determined as described in Materials and Methods.
[d] Despite being −7025*Alu*I, this sample has a clear U5 HVS I and RFLP motif.

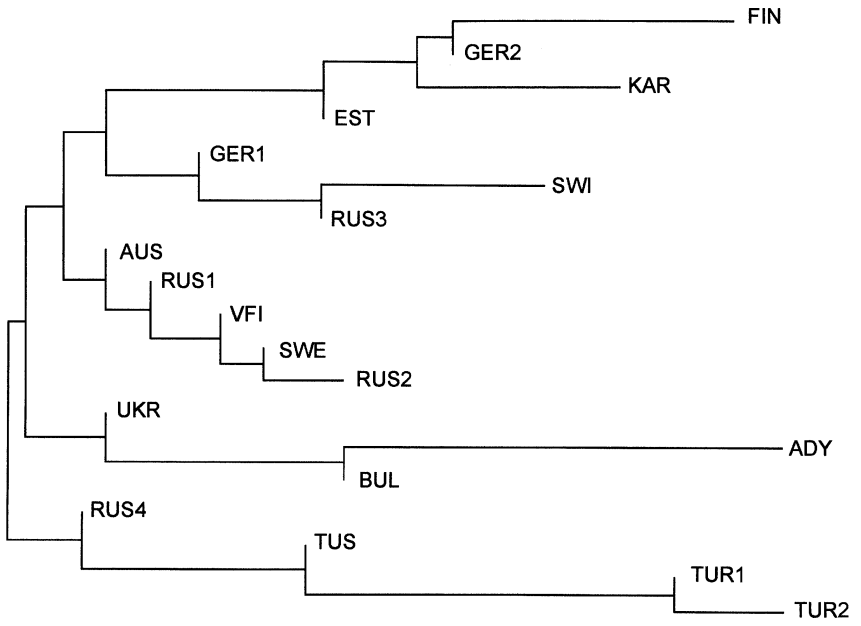

Fig. 1. Neighbour-joining tree for 19 West Eurasian populations, based on mtDNA HVS I sequences. Populations coded as in Materials and Methods.

same language group display a large amount of interpopulation genetic variation.

Shared HVS I sequence types between Eastern Slavs and other European populations

Together with the published sequence data for 103 Russians from European regions of Russia (Orekhov *et al.* 1999), a total of 153 sequences were analysed in the Russians. There were 99 distinct types of HVS I sequences in total. We identified twelve sequence types shared among Russian populations. However, only two sequence types were found in common between all four Russian populations, as well as between Russians and Ukrainians – CRS and sequence type 16069T–16126C, belonging to haplogroup J. Two HVS I sequences (16304C and 16224C–16311C) were shared between three Russian populations, and the remaining eight sequence types were found in common among two populations – 16356C, 16189C–16356C, 16093C, 16311C, 16362C, 16325C, 16162G, 16126C–16163G–16186T–16189C–16294T.

In order to create the most appropriate model for the origins of the Russians and Ukrainians,

representing the main large ethnic groups of Eastern Slavs, we examined the HVS I sequence variation in a population of Ukrainians and in four populations of Russians (including data from Orekhov *et al.* 1999) in comparison with published sequence data for 41 populations (2696 individuals in total) from West Eurasia, North Africa, Central Asia, and India. In addition, HVS I sequence data on 572 Asians, including native Siberians, were incorporated into the analysis.

This analysis revealed that among 99 HVS I sequence types that were found in Russians, 43 were unique to Russians. Among 16 HVS I sequence types in Ukrainians studied, 4 were unique to this population. We should note that all unique types observed in Russians and Ukrainians appear to be derived from the HVS I sequence types previously described in Eurasian populations. Therefore, we have not found any specific combinations of unique mtDNA types clearly distinguishing Russians from the neighbouring eastern European populations, such as Estonians, Karelians, Volga Finnic and Adygei.

Among HVS I sequence types shared between Eastern Slavs and other populations (56 in Russians and 12 in Ukrainians), the majority of these HVS I sequences were found in common among many western Eurasian populations. The maximum numbers of identical sequence types between Russians and other populations in pairwise comparisons were observed between Russians and Southern Germans (20 shared types), Northern Germans (18 types), Austrians (16 types), Estonians (15 types), and Karelians (15 types). A relatively large number of common types were also found between Ukrainians and Northern Germans (10 types), and Russians (9 types).

To identify the closest populations to Russians, we have excluded from analysis the most frequent HVS I sequence types, which were in common between more than six populations. As a result, we found 27 relatively rare mtDNA types. Of these, 14 sequence types were shared between Russians and German populations (Southern, Northern and Western Germans [6, 5 and 1 shared types, respectively], and Austrians [5 shared types]), and 13 sequence types were found in common between Russians and eastern European Finnic populations (Volga Finnic, Estonians, and Karelians [3, 6 and 4 types, respectively]), whereas 8 sequence types were shared among these German and Finnic populations. Therefore, we identified two groups of populations – central European (Germans and Austrians) and eastern European (Volga Finnic, Karelians, and Estonians) – that shared maximum numbers of rare HVS I sequence types with Russians. Though these results have to be treated with caution due to sampling errors, this kind of analysis has enabled us to detect two weak signals in the mitochondrial gene pool of Russians suggesting their dual (central and eastern European) origin and supporting the hybridization model for the origin of modern Russians. Undoubtedly, to elucidate the origin of the Eastern Slavs, an additional analysis of lineage sharing is required after a much more extensive sampling of Slavonic and other European populations has been performed.

RFLP and HVS I combined data from the Russians and Ukrainians

To identify the HVS I sequences of the Eastern Slavs, RFLPs typing of the haplogroup-diagnostic sites was performed. The HVS I sequences and RFLP status of these mtDNAs are reported in Table 1. The phylogenetic relationships of the RFLP-typed HVS I sequences are shown in Fig. 2.

The HVS I sequences of Russians and Ukrainians are clustered into eight haplogroups, according to a nomenclature which is now commonly accepted (Macaulay *et al.* 1999): H, V, HV*, U, K, J, T, X, though the haplogroups I, W and Z only harbour single individuals.

The main mitochondrial haplogroup of the Eastern Slavonic sequences is group H, which is the most frequent haplogroup in Europe and also common in the Near East (Richards *et al.* 1998). Haplogroup H comprises the majority of the Russian (38.0%) and Ukrainian (44.0%) samples. Although most of these sequences have

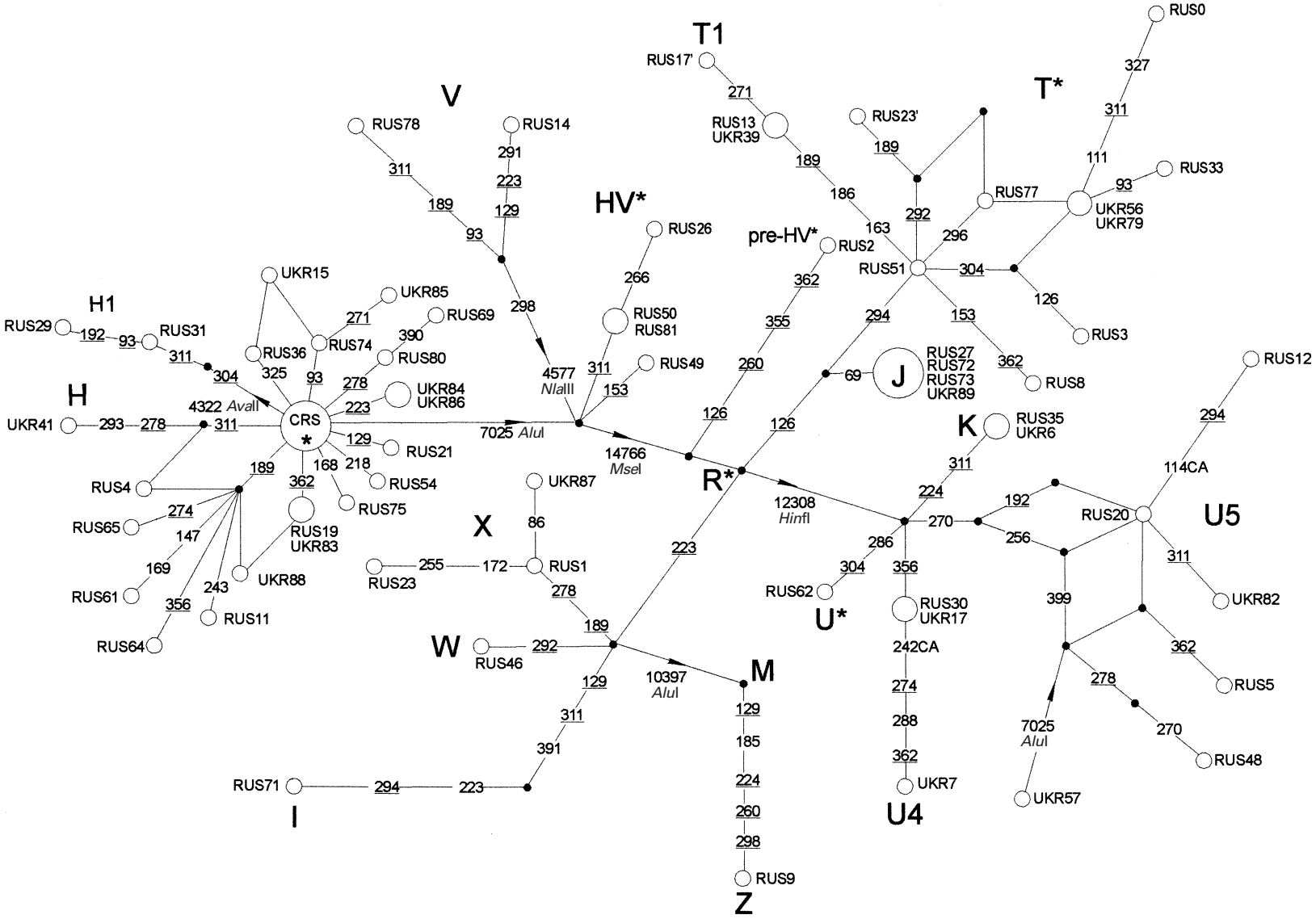

Fig. 2. Network (Bandelt *et al.* 1995) of the Russian and Ukrainian HVS I sequences. RFLP variants required to distinguish haplogroups HV*, H, V, U+K and M members are used in the construction of the network. The arrows on the branches indicate the direction of a site gain at each of the restriction sites. The node marked * matches the revised Cambridge reference sequence (CRS; Andrews *et al.* 1999), which was found in individuals RUS22, RUS58, RUS67, RUS76 and UKR60. Links are labelled by the nucleotide positions (nps) in HVS I (minus 16000) to designate transitions; transversions are further specified. HVS I mutations and RFLP variants are shown indicating nps relative to the CRS. Underlining indicates nucleotide positions which have mutated more than once. Haplogroup names are given in letters according to mtDNA classification (Richards *et al.* 1998; Macaulay *et al.* 1999). Circle size is proportional to the haplotype frequency in populations. Samples from other populations (marked by black dots) are used to refine the network. Reticulations indicate ambiguity in the topology.

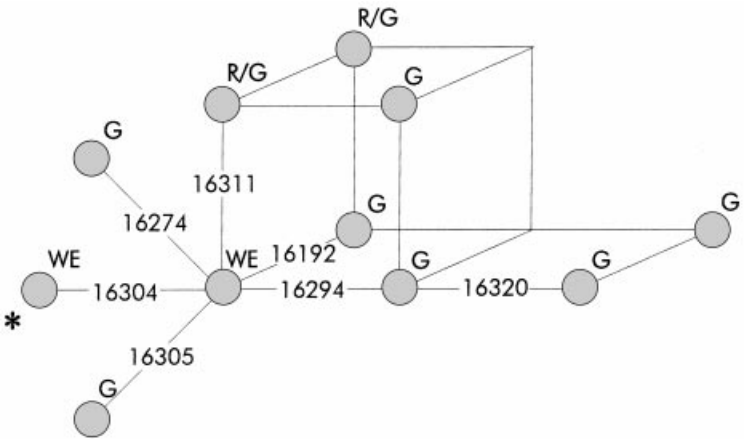

Fig. 3. Schematic phylogenetic network of the subhaplogroup H1 sequence types. The node labelled by an asterisk (*) corresponds to the CRS. HVS I mutations diagnostic for the main lineages within H1 (with motif 16304C) are displayed on the branches. Any diversity within the nodes is not shown. Reticulations in the network indicate ambiguity in the topology and are represented by cycles. The nodes in the network represent the combinations of mutations 16192T, 16294T, 16304C, 16311C and 16320T determining the haplotypes found in populations (Table 2) as well as hypothetical intermediate haplotypes (empty nodes). Labelled nodes are H1 haplotypes observed in Germanic-speaking populations (G), in Russians (R), in both populations (G/R), or in West Eurasians (WE).

been observed in other European populations, some peculiarities of the H-sequences in Eastern Slavs were observed. Thus, in the Ukrainians, we have found an HVS I sequence type differing from the CRS at np 16223 (individuals UKR84 and UKR86). Note that 16223T HVS I sequences have been previously described at low frequencies in different West Eurasian populations, but most of these sequences were assigned to the cluster IWX* (Richards *et al.* 1998). Screening of the published HVS I and II data allowed us to reveal the only 16223T-sequence which was observed in a Southern German population (GER161 in Lutz *et al.* 1998) and can be determined as belonging to the haplogroup H, being found in association with 00073A variant.

Another interesting feature of the Eastern Slavonic haplogroup H diversity is the presence of the lineage with motif 16304C–16311C (individuals RUS29 and RUS31). Recently, Helgason *et al.* (2000), in their study of the mtDNA lineages in the Icelanders, described separate subhaplogroup H1 within haplogroup H HVS I sequences. The candidate ancestral HVS I sequence type of H1 is 16304C in association with the 16519T. This haplotype represents the majority of subhaplogroup H1 mtDNAs, being found in different populations of West Eurasia, and is central to the phylogeny of the remaining H1 haplotypes (Fig. 3). Subhaplogroup H1 shows some degree of genetic substructure – Helgason *et al.* (2000) noted that the only lineages with motifs 16304C–16274A and 16304C–16305G show restricted north-western European geographic distribution, with prevalence of the lineage 16304C–16274A in Germany and Iceland and lineage 16304C–16305G in Iceland and the Faroe Islands. However, additional H1 lineages, which are characterized by different combinations of nucleotides at nps 16192, 16294, 16304, 16311 and 16320, are observed in the mitochondrial gene pool of Europeans. Analysis of the published data allowed us to find at least 14 HVS I sequence types comprising different nucleotide combinations at positions mentioned above (Table 2). The members of this set of subhaplogroup H1 sequences have been observed in different Germanic-speaking people and in the Russians (Table 2), although at a low frequency, but the majority of them are characteristic for populations of Germany.

It is difficult to estimate distribution and

Table 2. *Subhaplogroup H1 HVS I sequence types, which are defined by different combinations of nucleotides 16192T, 16294T, 16304C, 16311C and 16320T, in different populations*

HVS I sequence	Haplogroup H status confirmation	Sample origin
16093 16192 16304 16311	−14766*Mse*I, −7025*Alu*I	Russians[1]
16192 16304 16311	ND	Russian Germans[2]
16192 16304	00073A	Austrians[3]
16192 16304	ND	Bavarians[4]
16192 16304	ND	Orcadians[5]
16192 16214 16304	00073A	Southern Germans[6]
16304 16311	−14766*Mse*I, −7025*Alu*I	Russians[1]
16304 16311 16391	−7025*Alu*I, 00073A	Icelanders[7]
16124 16304 16311	00073A	Southern Germans[6]
16172 16304 16311	ND	Swedes[8]
16294 16304 16311	ND	Icelanders[8]
16222A 16294 16304 16311	ND	Icelanders[8]
16294 16304	00073A	Southern Germans[6]
16294 16304	ND	Orcadians[5]
16189 16294 16304	00073A	Southern Germans[6]
16294 16304 16320	00073A	Southern Germans[6]
16294 16304 16320	00073A	Western Germans[9]
16294 16304 16320	00073A	British[10]
16294 16304 16320	00073A	Austrians[3]
16192 16294 16304 16320	ND	Danes[4]

HVS I sequences, identified as belonging to haplogroup H by using of RFLP analysis or HVS II sequencing, from the following studies were analysed: [1] present study, [2] Baranov *et al.* 1999, [3] Parson *et al.* 1998, [4] Richards *et al.* 1996, [5] Miller *et al.* 1996, [6] Lutz *et al.* 1998, [7] Helgason *et al.* 2000, [8] Sajantila *et al.* 1995, [9] Baasner *et al.* 1998, [10] Piercy *et al.* 1993.

ND, not determined.

diversity of the subhaplogroup H1 sequences in European populations now because of the absence of any additional markers in mtDNA coding regions required for any subhaplogroup H1 status confirmation. We should note only that the variant +4332*Ava*II (4336C) seems to be diagnostic for subhaplogroup H1, being found in association with HVS I sequence types with motif 16304C–16311C in Russians (individuals RUS29 and RUS31) and in combination with the −16303*Rsa*I variant (16304C) in HVS I, as well as with motif −16303*Rsa*I/−16310*Rsa*I (or 16304C–16311C) (Torroni *et al.* 1994, 1996, 1998). However, additional detailed studies are required to elucidate the origin and diversification of the subhaplogroup H1 mtDNAs.

In addition, phylogenetic relationships between the majority of HVS I sequences belonging to the H1 remain ambiguous and, therefore, the branching order of these H1-sequences cannot be resolved. For instance, a set of sequence types, which are characterized by different combinations of nucleotides at nps 16192, 16294, 16304, 16311 and 16320, cannot be differentiated phylogenetically into separate lineages of the subhaplogroup H1 because of the high degree of reticulations in the topology of phylogenetic networks (Fig. 3).

Among HVS I sequence types shared between Russians and Ukrainians, there is a small lineage within haplogroup H determined by nucleotide 16325C – haplotypes 16325C, 16093C–16325C in the present study and 16325C, 16304C–16325C in Orekhov *et al.* (1999). This lineage is curiously rare in other European populations, being probably found in Germans (16189C–16325C, 16256T–16325C), Algerians (16320T–16325C) and Basques (16320T–16323C–16325C). The origin and distribution of this rare 16325C lineage in Eurasian populations remains unclear.

In addition, the analysis of geographic distribution of the HVS I sequences has revealed that similar or identical HVS I sequence types determined by combinations of the nucleotides at nps 16304, 16311, and 16325 are found not only in European populations, but also in Indian

ones. However, it was shown recently (Kivisild *et al.* 1999*b*) that the Indian mtDNAs with motifs 16304C–16311C, 16266T–16304C–16311C, 16266T–16304C–16325C belong to the haplogroups R* and R1, although some HVS I sequences (motif 16309G–16325C) are members of haplogroup HV*. It is not clear whether the similarity of the HVS I molecular structures in Europeans and Indians is caused by independent mutations at nps 16304, 16311, and 16325 on the RFLP-backgrounds of haplogroup H (in Europe) and haplogroups R* and HV* (in India), or that this demonstrates an ancient genetic link between Europeans and Indians, supporting the common origin of Indian and European mtDNA lineages (Kivisild *et al.* 1999*a*). Although the positions 16304, 16311, and 16325 have been identified as fast sites in a European data (Richards *et al.* 1998; Macaulay *et al.* 1999), one cannot exclude the possibility, as it was proposed by Richards *et al.* (1998), that cluster H includes several of the oldest subclusters, including ones based on haplotypes 16304C, 16311C, and 16325C. Additional studies are required to clarify this.

The node designated as HV* (Richards *et al.* 1998) is highly important in mtDNA phylogeny because two haplogroups, H and V, the most frequent in Europe, descend from it. It was suggested (Metspalu *et al.* 1999) that the expansion (and possibly origin) of haplogroup HV* began from the Caucasus area – around 8.0% of Ossetians, Armenians, and Georgians are ascribed to haplogroup HV*, whereas it is rare among western Europeans. Among Eastern Slavs studied, three HVS I sequence types in Russians – 16311C (RUS50, RUS81), 16266T–16311C (RUS26), and 16153A (RUS49) – are identified as belonging to haplogroup HV* (RFLP motif: −14766*Mse*I, +7025*Alu*I). HVS I sequences 16311C are frequently found in West Eurasian populations (Richards *et al.* 1998) and belong mostly to haplogroup H. Though it has been shown recently (Metspalu *et al.* 1999) that sequence type 16311C, as one of the possible subfounders of haplogroup HV*, appears to be widespread among populations of the Caucasus, such as Ossetians and Georgians. In addition, on the basis of mtDNA variation in the Icelanders, Helgason *et al.* (2000) have suggested that 16311C lineages belonging to HV* form a separate cluster HV2.

In Eastern Slavs, 16126C-sequences are represented by three haplogroups – pre-HV*, J and T. Sequence type RUS2 is member of the lineage that is defined by HVS I motif 16126C–16362C and belongs to cluster pre-HV* (Richards *et al.* 2000). Haplogroup J sequences in Russians and Ukrainians presented here, as well as in other Russian populations (Orekhov *et al.* 1999), are characterized by reduced diversity – almost all of them (83.0%) have sequence type 16069T–16126C and belong to subgroup J*. A similar pattern of the haplogroup J subgroups' distribution is characteristic of other eastern European populations, such as Estonians and Karelians.

Haplogroup T (HVS I motif: 16126C–16294T) represents 18.0% and 16.7% of the Russian and Ukrainian mtDNAs, respectively, and includes two distinct subgroups, T* and T1. This haplogroup is also one of the most frequent in other Russian populations studied, with a frequency of 11.6%, on average (Orekhov *et al.* 1999). Some description of the haplogroup T variation in Russians and Ukrainians was presented recently (Malyarchuk & Derenko, 1999), although in the context of a study of the molecular instability of the haplogroup T HVS I sequences. We should note only that sequence type RUS3 (16294T–16304C) shares identical HVS I sequence with subhaplogroup H1 sequences (Table 2). Nevertheless, this sample is characterized by T-specific RFLP motif +13366*Alu*I/−13367*Ava*II, +15606*Alu*I, −15925*Msp*I and is identified as belonging to subgroup T*.

Haplogroups U and K sequences, which are defined by a variant +12308*Hin*fI, were found in 14.0% of the Russian mtDNAs and in 27.8% of the Ukrainian mtDNAs. Of these, haplogroup K sequences are relatively rare in the Russian sample studied (2.0%). On the contrary, haplogroup U itself is widely distributed in Eastern Slavonic populations and is represented by subgroups U*, U4 and U5. Some striking

Table 3. *Molecular instability in subgroup U5 HVS I sequence types in several populations*

HVS I sequence	Unstable nucleotide position	Sample origin
16192 16256 16270 **16291**		Germans[1], Russians[2], Selkups[3], Adygei[4]
16093 16192 16256 16270 **16291**		Russians[2], Finns[5]
16129 16192 16256 16270 **16291**		Russians[2],Germans[6]
16093 16148 16192 16256 16270 **16291**		British[7]
16093 16256 16270 **16291**	16192	Finns[5]
16156 16192 16270 **16291**	16256	Kets[3]
16156 16192 16270	16256	Kets[3]
16156 16192 16268 16270 **16291**	16256	Kets[3]
16076 16192 16256 16270		Volga Finnic[8]
16076 16192 16256	16270	Russians[2]
16256 16270 **16278** 16399		Germans[1]
16256 **16278** 16399	16270	Ukrainians[9]

Marker mutations are shown in bold. HVS I sequences from the following studies were analysed: [1] Lutz *et al.* 1998, [2] Orekhov *et al.* 1999, [3] Kazakovtseva 1998, [4] Macaulay *et al.* 1999, [5] Kittles *et al.* 1999, [6] Richards *et al.* 1996, [7] Piercy *et al.* 1993, [8] Sajantila *et al.* 1995, [9] present study.

peculiarities were found in Russian and Ukrainian U-sequences.

U5, the most frequent and ancient subgroup of haplogroup U in Europe (Torroni *et al.* 1996; Richards *et al.* 1998), is represented in the Russians and Ukrainians by subgroups U5a1 and U5a1a. In addition, U5b-sequences (including Saami-specific U5b1 mtDNAs) have been found at a low frequency in the Russians (Orekhov *et al.* 1999). Subgroup U5a1 has the whole HVS I motif 16192T–16256T–16270T, whereas other U5-subgroups are defined by motifs with the loss of 16192T, 16256T or both (Table 1 in Richards *et al.* 1998). As suggested by Richards *et al.* (1998), one cannot exclude the possibility that some subgroups of U5-sequences were generated by back-mutations at HVS I diagnostic nucleotide variants (from U5a1 to U5a1a, for instance). Here, we provide additional evidence of this, based on the compatibility analysis of the rare nucleotide variants (marker mutations) with various combinations of mutations at nps 16192, 16256, and 16270 (Table 3). The transition C to T at np 16291 is rare in subgroup U5 sequences and is usually associated with U5a1-diagnostic motif. These sequences are spread evenly in different populations of West Eurasia (Table 3). A search in the mtDNA sequence database reveals, however, a specific set of U5-sequences with variant 16291T developed, most likely, locally in two different regions of Eurasia – in northern Europe (among the Finns) and in western Siberia (among the Kets). It is interesting that divergence of these haplotypes was accompanied by two directions of evolutionary changes – by the loss of variant 16192T in the Finns and by the loss of variant 16256T, as well as arising of population-specific variant 16156A, in the Kets.

Recently, Finnilä *et al.* (2000) have shown that np 16192 seems to be prone to recurrent mutations and that np 16270 has experienced a back-mutation. On the basis of mtDNA coding and non-coding regions variation in the Finns, Finnilä *et al.* (2000) found mutations 5656A→G and 12618G→A shared by subhaplogroups U5a (with motif 16189C–16192T–16270T) and U5b (with motif 16189C–16270T), implying that the 16189C–16270T haplotype evolved by back-mutation at np 16192 from the 16189C–16192T–16270T haplotype. Moreover, recently Howell & Smejkal (2000) found an extremely high rate of reversion (hypermutation) from T to C at np 16192 in subhaplogroup U5 mtDNAs that harboured the 16189C–16192T combination of HVS I polymorphisms.

Two last cases in Table 3 illustrate instability of the U5-sequnces at np 16270, the major motif of the subhaplogroup U5 as a whole. HVS I sequences marked by variant 16278T (rare on

the subgroup U5 background) are found in common only between Ukrainians and Southern Germans. Haplotypes marked by variant 16076T (rare in the West Eurasian data set) are shared only between Russians and Finns. These examples, demonstrating kinship between Eastern Slavs and Germans, as well as Finno–Ugric populations, at the level of rare mtDNA sequences, give additional support to the hybridization model for the origin of Eastern Slavs.

Haplogroups X, W, and I sequences, characterized by the ancestral state at position 16223 (Richards *et al.* 1998), occurred at a low frequency (4% and less) in the data set presented here, as well as in other populations of Russians (Orekhov *et al.* 1999). The remaining HVS I sequence RUS9 belongs to haplogroup Z. Note that haplogroup Z sequences are found at a high frequency (26.2%) in the Tungusic-speaking Evens of eastern Siberia (Derenko & Shields, 1997), as well as at frequencies 5.8% and 6.4%, respectively, in the Koryaks and Itel'men of the Kamchatka peninsula (Schurr *et al.* 1999), though the origin of haplogroup Z seems to be central Asian/eastern Siberian (Denisova *et al.* 1999).

Taking into account the data of Orekhov *et al.* (1999), one should conclude that the Mongoloid admixture in Russians appears to be insignificant (less than 3.0%) and represented by haplogroups C and Z sequences. The presence of the Saami-specific haplogroup Z, as well as subgroup U5b1 sequences in the mitochondrial gene pool of Russians, we consider as a consequence of local Finno–Ugric tribes assimilation by Slavs during their movement to the North of eastern Europe. The presence of Asian-specific components, such as haplogroup C sequences, in the mitochondrial gene pool of Russians may be explained by their complicated ethnic history, including long-lasting interactions with Asians. However, in the study of mtDNA sequences only the female lineages are taken into account, whereas Mongoloid morphological traits in the Russians, revealed by anthropologists (Alekseeva, 1973), might have been derived from male migrants.

CONCLUSION

Analysis of mtDNA variation in Eastern Slavonic populations allows us to suggest a few conclusions. First, the Russians and Ukrainians are characterized by the same West Eurasian mtDNA haplogroups that describe at least 95% of mtDNA variations in Europe and the Near East (Torroni *et al.* 1996; Richards *et al.* 1996, 1998). Moreover, mtDNA haplotypes of the Russians and Ukrainians are profoundly close to those found in western and eastern Europe. We have observed a large amount of identical HVS I sequence types shared between Eastern Slavs and other Europeans. Second, phylogenetic analysis suggests that Eastern Slavs are heterogeneous at the population level. Their heterogeneity is also confirmed by the low amount of HVS I sequence types shared between populations of Russians. Third, analysis of distribution of rare HVS I types shared between populations of Eastern Slavs and other West Eurasians has shown that Russians share these mainly with Germans and Finno–Ugric populations, such as Volga Finnic, Karelians, and Estonians. This supports the hybridization model for the origin of Eastern Slavs. The most intriguing finding of this study, is a specific lineage 16304C–16311C within the subhaplogroup H1, representing 4% of the Russian mtDNAs studied. These mtDNAs are found predominantly in common between Russians and Germanic-speaking populations and were most likely introduced to the East of Europe from central Europe during the Slavonic migrations.

Overall, the data obtained allow us to conclude that the Slavonic migrations in early Middle Ages from their putative homeland in central Europe to the East of Europe were mostly accompanied by the same mtDNA types characteristic for the pre-Slavonic populations of eastern Europe. The minor female contribution into Slavonic migrations may also explain the picture observed.

In addition, these data highlight the difficulty in the identification of clear coevolutionary

patterns between linguistic and genetic relationships in particular human populations – in Slavonic populations, for instance. The evidence from the present study suggests that the assumption of a common central European origin of Slavs should be tested with studies that include mtDNA and Y chromosome loci and Slavonic populations from southern and western Europe. An appropriate linguistic model explaining distribution of Slavonic languages and culture on the background of genetic heterogeneity of modern Slavonic populations is also required.

We thank to G. Denisova (IBPN, Magadan) and I. Dmitrenko (PILC 'Living Arctic', Magadan) for technical assistance; to V. Orekhov (VIGG, Moscow) for help in this study. The authors would like to thank two anonymous reviewers for their useful comments. This work was supported by the Russian Fund for Basic Research (grant 00–06–80448) and the State Program 'Frontiers in Genetics' (grant 99–4–30).

REFERENCES

Alekseev, V. P. (1989). *Historical anthropology and ethnogenesis*. Moscow: Nauka (in Russian).

Alekseeva, T. I. (1973). *Ethnogenesis of Eastern Slavs*. Moscow: Moscow State University (in Russian).

Alekseeva, T. I. & Alekseev V. P. (1989). Anthropological view of the origin of Slavs. *Priroda* **881**, 60–69 (in Russian).

Anderson, S., Bankier, A. T., Barrell, B. G., de Bruijn, M. H. L., Coulson, A. R., Drouin, J., *et al.* (1981). Sequence and organization of the human mitochondrial genome. *Nature* **290**, 457–465.

Andrews, R. M., Kubacka, I., Chinnery, P. F., Lightowlers, R. N., Turnbull, D. M. & Howell, N. (1999). Reanalysis and revision of the Cambridge reference sequence for human mitochondrial DNA. *Nature Genet.* **23**, 147.

Baasner, A., Schäfer, C., Junge, A. & Burkhard, M. (1998). Polymorphic sites in human mitochondrial DNA control region sequences: population data and maternal inheritance. *Forensic Sci. Internatl.* **98**, 169–178.

Bandelt, H.-J., Forster, P., Sykes, B. C. & Richards, M. B. (1995). Mitochondrial portraits of human populations using median networks. *Genetics* **141**, 743–753.

Baranov, P. O., Babenko, V. N., Ivanova, A. V., Kobzev, V. F., Romashenko, A. G. & Voevoda, M. I. (1999). The peculiarities of the mitochondrial gene pool of the Russian Germans. *Genetika* **35**, 249–254 (in Russian).

Bertranpetit, J., Sala, J., Calafell, F., Underhill, P. A., Moral, P. & Comas, D. (1995). Human mitochondrial DNA variation and the origin of Basques. *Ann. Hum. Genet.* **59**, 63–81.

Calafell, F., Underhill, P., Tolun, A., Angelicheva, D. & Kalaydjieva, L. (1996). From Asia to Europe: mitochondrial DNA sequence variability in Bulgarians and Turks. *Ann. Hum. Genet.* **65**, 35–49.

Comas, D., Calafell, F., Mateu, E., Perez-Lezaun, A. & Bertranpetit, J. (1996). Geographic variation in human mitochondrial DNA control region sequence: the population history of Turkey and its relationship to the European populations. *Mol. Biol. Evol.* **13**, 1067–1077.

Comas, D., Calafell, F., Mateu, E., Perez-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., Perez-Lezaun, A., Bosch, E., Martinez-Arias, R., Clarimon, J., *et al.* (1998). Trading genes along the Silk Road: mtDNA sequences and the origin of Central Asian populations. *Am. J. Hum. Genet.* **63**, 1824–1838.

Côrte-Real, H. B. S. M., Macaulay, V. A., Richards, M. B., Hariti, G., Issad, M. S., Cambon-Thomsen, A., *et al.* (1996). Genetic diversity in the Iberian peninsula determined from mitochondrial sequence analysis. *Ann. Hum. Genet.* **60**, 331–350.

Denisova, G. A., Derenko, M. V. & Malyarchuk, B. A. (1999). A partial central Asian/eastern Siberian origin of the Saami mtDNAs. *Am. J. Hum. Genet.* **65** (Suppl.), A200: 1101.

Derenko, M. V. & Shields, G. F. (1997). Mitochondrial DNA sequence diversity in three North Asian aboriginal population groups. *Mol. Biol.* (Moscow) **31**, 665–669.

Di Rienzo, A. & Wilson, A. C. (1991). Branching pattern in the evolutionary tree for human mitochondrial DNA. *Proc. Natl. Acad. Sci. USA* **88**, 1597–1601.

Finnilä, S., Hassinen, I. E., Ala-Kokko, L. & Majamaa, K. (2000). Phylogenetic network of the mtDNA haplogroup U in Northern Finland based on sequence analysis of the complete coding region by conformation-sensitive gel electrophoresis. *Am. J. Hum. Genet.* **66**, 1017–1026.

Francalacci, P., Bertranpetit, J., Calafell, F. & Underhill, P. A. (1996). Sequence diversity of the control region of mitochondrial DNA in Tuscany and its implications for the peopling of Europe. *Am. J. Phys. Anthropol.* **100**, 443–460.

Helgason, A., Sigurardóttir, S., Gulcher, J. R., Ward, R. & Stefánsson, K. (2000). mtDNA and the origin of the Icelanders: deciphering signals of recent population history. *Am. J. Hum. Genet.* **66**, 999–1016.

Horai, S., Murayama, K., Hayasaka, K., Matsubayashi, S., Hattori, Y., Fucharoen, G., et al. (1996). mtDNA polymorphism in east Asian populations, with special reference to the peopling of Japan. *Am. J. Hum. Genet.* **59**, 579–590.

Howell, N., Kubacka, I. & Mackey, D. A. (1996). How rapidly does the human mitochondrial genome evolve? *Am. J. Hum. Genet.* **59**, 501–509.

Howell, N. & Smejkal, C. B. (2000). Persistent heteroplasmy of a mutation in the human mtDNA control region: hypermutation as an apparent consequence of simple-repeat expansion/contraction. *Am. J. Hum. Genet.* **66**, 1589–1598.

Kazakovtseva, M. A. (1998). Mitochondrial DNA polymorphism in native and Eastern Slavonic populations

of the West Siberia. Ph.D. Thesis. Novosibirsk: Institute of Cytology and Genetics (in Russian).

Kittles, R. A., Bergen, A. W., Urbanek, M., Virkkunen, M., Linnoila, M., Goldman, D. & Long, J. C. (1999). Autosomal, mitochondrial, and Y chromosome DNA variation in Finland: evidence for a male-specific bottleneck. *Am. J. Phys. Anthropol.* **108**, 381–399.

Kivisild, T., Bamshad, M. J., Kaldma, K., Metspalu, M., Metspalu, E., Reidla, M., et al. (1999*a*). Deep common ancestry of Indian and western-Eurasian mitochondrial DNA lineages. *Current Biology* **9**, 1331–1334.

Kivisild, T., Kaldma, K., Metspalu, M., Parik, J., Papiha, S. & Villems, R. (1999*b*). The place of the Indian mitochondrial DNA variants in the global network of maternal lineages and the peopling of the Old World. In *Genomic diversity: applications in human population genetics* (eds. S. Papiha, R. Deka & R. Chakraborty), pp. 135–152. New York : Kluwer Academic/Plenum Publishers.

Kolman, C. J., Sambuughin, N. & Bermingham, E. (1996). Mitochondrial DNA analysis of Mongolian populations and implications for the origin of New World founders. *Genetics* **142**, 1321–1334.

Lutz, S., Weisser, H.-J., Heizmann, J. & Pollak, S. (1998). Location and frequency of polymorphic positions in the mtDNA control region of individuals from Germany. *Int. J. Legal Med.* **111**, 67–77.

Malyarchuk, B. A. (1997). Distribution of mitochondrial DNA markers in Caucasoid populations of Eurasia. *Russ. J. Genet.* **33**, 836–840.

Malyarchuk, B. A. & Derenko, M. V. (1999). Molecular instability of the mitochondrial haplogroup T sequences at nucleotide positions 16292 and 16296. *Ann. Hum. Genet.* **63**, 489–497.

Malyarchuk, B. A., Derenko, M. V. & Solovenchuk, L. L. (1995). Types of mitochondrial DNA control region in Eastern Slavs. *Russ. J. Genet.* **31**, 723–727.

Macaulay, V., Richards, M., Hickey, E., Vega, E., Cruciani, F., Guida, V., *et al.* (1999). The emerging tree of West Eurasian mtDNAs: a synthesis of control-region sequences and RFLPs. *Am. J. Hum. Genet.* **64**, 232–249.

Metspalu, E., Kivisild, T., Kaldma, K., Parik, J., Reidla, M., Tambets, K. & Villems, R. (1999). The Trans-Caucasus and the expansion of the Caucasoid-specific: human mitochondrial DNA. In *Genomic diversity: applications in human population genetics* (eds. S. Papiha, R. Deka & R. Chakraborty), pp. 121–133. New York: Kluwer Academic/Plenum Publishers.

Miller, K. W. P., Dawson, J. L. & Hagelberg, E. (1996). A concordance of nucleotide substitutions in the first and second hypervariable segments of the human mtDNA control region. *Int. J. Legal Med.* **109**, 107–113.

Niederle, L. (1896). O puvodu slovanu v Praze. Praha (in Czech).

Parson, W., Parsons, T. J., Scheithauer, R. & Holland, M. M. (1998). Population data for 101 Austrian Caucasian mitochondrial DNA d-loop sequences: Application of mtDNA sequence analysis to a forensic case. *Int. J. Legal Med.* **111**, 124–132.

Orekhov, V., Poltoraus, A., Zhivotovsky, L. A., Spitsyn, V., Ivanov P. & Yankovsky, N. (1999). Mitochondrial DNA sequence diversity in Russians. *FEBS Letters* **445**, 197–201.

Pinto, F., Gonzalez, A. M., Hernandez, M., Larruga, J. M. & Cabrera, V. M. (1996). Genetic relationship between the Canary Islanders and their African and Spanish ancestors inferred from mitochondrial DNA sequences. *Ann. Hum. Genet.* **60**, 321–330.

Piercy, R., Sullivan, K. M., Benson, N. & Gill, P. (1993). The application of mitochondrial DNA typing to the study of white Caucasian genetic identification. *Int. J. Leg. Med.* **106**, 85–90.

Pult, I., Sajantila, A., Simanainen, J., Georgiev, O., Schaffner, W. & Pääbo, S. (1994). Mitochondrial DNA sequences from Switzerland reveal striking homogeneity of European populations. *Biol. Chem. Hoppe Seyler* **375**, 837–840.

Renfrew, C. (1994). World linguistic diversity. *Sci. Am.* **270**, 104–110.

Richards, M., Côrte-Real, H., Forster, P., Macaulay, V., Wilkinson-Herbots, H., Demaine, A., *et al.* (1996). Paleolithic and neolithic lineages in the European mitochondrial gene pool. *Am. J. Hum. Genet.* **59**, 185–203.

Richards, M. B., Macaulay, V. A., Bandelt, H.-J. & Sykes, B. C. (1998). Phylogeography of mitochondrial DNA in western Europe. *Ann. Hum. Genet.* **62**, 241–260.

Richards, M. B., Macaulay, V. A., Hickey, E., Vega, E., Forster, P., Bandelt, H.-J., *et al.* (2000). Tracing European founder lineages in the Near Eastern mtDNA pool. *Am. J. Hum. Genet.* **67**, 1251–1276.

Rogers, A. R. & Harpending, H. (1992). Population growth makes waves in the distribution of pairwise genetic distances. *Mol. Biol. Evol.* **9**, 552–569.

Sajantila, A., Lahermo, P., Anttinen, T., Lukka, M., Sistonen, P., Savontaus, M. L., *et al.* (1995). Genes and languages in Europe – an analysis of mitochondrial lineages. *Genome Res.* **5**, 42–52.

Schurr, T. G., Sukernik, R. I., Starikovskaya, Y. B. & Wallace, D. C. (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.

Sedov, V. V. (1979). *Origin and early history of Slavs.* Moscow: Nauka (in Russian).

Shields, G. F., Schmiechen, A. M., Frazier, B. L., Redd, A., Voevoda, M. I., Reed, J. K. & Ward, R. H. (1993). mtDNA sequences suggest a recent evolutionary divergence for Beringian and northern North American populations. *Am. J. Hum. Genet.* **53**, 549–562.

Simoni, L., Calafell, F., Pettener, D., Bertranpetit, J. & Barbujani, G. (2000). Geographic patterns of mtDNA diversity in Europe. *Am. J. Hum. Genet.* **66**, 262–278.

Stenico, M., Nigro, L., Bertorelle, G., Calafell, F., Capitanio, M., Corrain, C., *et al.* (1996). High mitochondrial sequence diversity in linguistic isolates of the Alps. *Am. J. Hum. Genet.* **59**, 1363–1375.

Torroni, A., Bandelt, H.-J., D'Urbano, L., Lahermo, P., Morral, P., Sellito, D., *et al.* (1998). MtDNA analysis reveals a major late Paleolithic population expansion from southwestern to northeastern Europe. *Am. J. Hum. Genet.* **62**, 1137–1152.

Torroni, A., Cruciani, F., Rengo, C., Sellitto, D., López-Bigas, N., Rabionet, R., *et al.* (1999). The A1555G mutation in the 12S rRNA gene of human mtDNA: recurrent origins and founder events in families affected by sensorineural deafness. *Am. J. Hum. Genet.* **65**, 1349–1358.

Torroni, A., Huoponen, K., Francalacci, P., Petrozzi, M., Morelli, L., Scozzari, R., *et al.* (1996). Classification of

European mtDNAs from an analysis of three European populations. *Genetics* **144**, 1835–1850.

Torroni, A., Lott, M. T., Cabell, M. F., Chen, Y. S., Lavergne, L. & Wallace, D. C. (1994). mtDNA and the origin of Caucasians: identification of ancient Caucasian-specific haplogroups, one of which is prone to a recurrent somatic duplication in the D-loop region. *Am. J. Hum. Genet.* **55**, 760–776.

Torroni, A., Petrozzi, M., D'Urbano, L., Sellitto, D., Zeviani, M., Carrara, F., *et al.* (1997). Haplotype and phylogenetic analyses suggest that one European-specific mtDNA background plays a role in the expression of Leber hereditary optic neuropathy by increasing the penetrance of primary mutations 11778 and 14484. *Am. J. Hum. Genet.* **60**, 1107–1121.

Wallace, D. C. (1995). Mitochondrial DNA variation in human evolution, degenerative disease and aging. *Am. J. Hum. Genet.* **57**, 201–223.