doi: 10.1111/j.1469-1809.2007.00410.x

Mitochondrial DNA Variability in Slovaks, with Application to the Roma Origin

B. A. Malyarchuk[1,*], M. A. Perkova[1], M. V. Derenko[1], T. Vanecek[2], J. Lazur[3], P. Gomolcak[4]

[1]*Institute of Biological Problems of the North, Russian Academy of Sciences, Portovaya str. 18, 685000 Magadan, Russia*
[2]*Department of Pathology, Medical Faculty Hospital, Charles University, Pilsen, Czech Republic*
[3]*Department of Laboratory Medicine, LABMED, Kosice, Slovakia*
[4]*Institute of Pathology, Slovak Medical University, Bratislava, Slovakia*

Summary

To gain insight into the mitochondrial gene pool diversity of European populations, we studied mitochondrial DNA (mtDNA) variability in 207 subjects from western and eastern areas of Slovakia. Sequencing of two hypervariable segments, HVS I and HVS II, in combination with screening of coding region haplogroup-specific RFLP-markers, revealed that the majority of Slovak mtDNAs belong to the common West Eurasian mitochondrial haplogroups (HV, J, T, U, N1, W, and X). However, a few sub-Saharan African (L2a) mtDNAs were detected in a population from eastern part of Slovakia. In addition, about 3% of mtDNAs from eastern Slovakia encompass Roma-specific lineages. By means of complete mtDNA sequencing we demonstrate here that the Roma-specific M-lineages observed in gene pools of different Slavonic populations (Slovaks, Poles and Russians), belong to Indian-specific haplogroups M5a1 and M35. Moreover, we show that haplogroup J lineages found in gene pools of the Roma and some Slavonic populations (Czechs and Slovaks) belong to new subhaplogroup J1a, which is defined by coding region mutation at position 8460.

Keywords: mitochondrial DNA, Slovak population, Roma, molecular phylogeography

Introduction

Despite the divergent views on the ethnic history and origin of Slavs, it is generally accepted that their prehistory begins in 2000 B.C. from the Central European community of Urnfield cultures (Sedov, 1979; Šavli et al. 1996). Later, in the middle of the first millennium B.C. the Proto-Slavs were formed on the basis of Lusatian (Lausitz) culture. This emerged in Central Europe and spread over a region that reached from the central basin of the Oder River and the Bohemian mountain ridge, as far east as the Ukraine, and as far north as the shores of the Baltic Sea (Gimbutas, 1971; Sedov, 1979). In Roman times, the territory of Przeworsk culture, which is identified as Slavonic, expands to the south-east (Upper Dniester, Ukraine) and to the south (northern-eastern Slovakia) (Sedov, 1979). During the 5^{th}-6^{th} centuries A.D., the formation of the early medieval Slavonic cultures took place, and among them the Prague-Korchak culture is thought to have developed on the basis of Przeworsk culture (Gimbutas 1971; Sedov, 1979). According to anthropological and archaeological data, the homeland of the Slavs is located on the broad territories encompassing the eastern part of the Czech Republic, Slovakia, southern Poland and western Ukraine (Sedov, 1979; Alekseeva & Alekseev, 1989). At present these territories are mainly inhabited by Western (Czechs, Slovaks and Poles) and Eastern Slavs (Ukrainians).

Modern molecular genetic approaches are widely used for reconstruction of ethnic history of different peoples. Among these approaches, analysis of the maternally inherited mitochondrial DNA (mtDNA) seems to be the most appropriate tool for characterization of gene pools and for tracking maternal gene flow. Previous mtDNA studies of European populations have shown that Slavonic populations share a common genetic substratum characteristic of Central and Eastern European populations, such as German, Baltic and western Finno-Ugric populations (Malyarchuk et al. 2002; Pliss et al. 2006). It has been found that this genetic substratum also penetrates south-eastern European populations (such as the Bosnians and

* Corresponding author: Dr. Boris A. Malyarchuk, Genetics Laboratory, Institute of Biological Problems of the North, Portovaya str., 18, 685000 Magadan, Russia. Fax/Phone: +7 4132 634463. E-mail: malyarchuk@ibpn.ru

Slovenians), thus reaching territory as far as the Western Balkans (Malyarchuk et al. 2003). Meanwhile, previous genetic data sets for Slavonic populations were incomplete because some ethnic groups of Slavs remain unexplored. Little is known about the mtDNA structure and diversity in Slovak populations inhabiting the region that is very important in respect of understanding the ethnic history of the Slavs as a whole. Recent mtDNA variability study in Czechs, the neighbors of Slovaks, has shown that they are genetically similar with adjacent European populations, but characterized by a small frequency of East Eurasian (2.8%) and Roma-specific (2.8%) mtDNA lineages (Malyarchuk et al. 2006b). Therefore, the aim of the present study was to characterize the mtDNA variation in Slovaks from western and eastern areas of Slovakia, based on variation of the HVS I and HVS II sequences, followed by a hierarchical survey of mtDNA haplogroup-specific restriction fragments length polymorphism (RFLP) markers. In addition, to determine the phylogenetic status of some unclassified mtDNA lineages (mostly Roma-specific) found in the gene pool of Slovaks, complete sequencing of mitochondrial genomes was performed.

Materials and Methods

Population Samples

A population sample of 207 individuals from two areas of Slovakia, western (n = 70) (area of Bratislava) and eastern (n = 137) (area of Kosice), was studied. These samples were represented by randomly chosen unrelated Caucasians from Slovakia. All individuals studied were maternally unrelated and originated from the area considered for this study. Appropriate informed consent was obtained from all participants in this study.

mtDNA Sequencing and RFLP Analysis

The total DNA was isolated from blood samples with a QIAamp DNA Blood Mini kit (QIAGEN, Hilden, Germany) according to manufacturer's protocol. To generate control region mtDNA templates for sequencing, the temperature profile for 32 cycles of amplification was: 94°C for 20 sec, 50°C for 30 sec, and 72°C for 1 min 30 sec, with initial denaturation at 94°C for 2 min (GeneAmp PCR System 9700; Applied Biosystems, Foster City, CA). The primer sets L15997/H16401 and L29/H408 were used to amplify the hypervariable segments I and II (HVS I and II), respectively. PCR products were purified by ultrafiltration (Microcon 100; Amicon) and sequenced directly from both strands using the BigDye Terminator v. 3.1 sequencing kit (Applied Biosystems, Foster City, CA). The nucleotide sequences of HVS I from position 15999 to 16400 and HVS II from position 30 to 407 were determined on ABI 3130 Genetic Analyzer (Applied Biosystems, Foster City, CA) and compared with the revised Cambridge reference sequence (rCRS) (Andrews et al. 1999). DNA sequence data were analyzed using SeqScape v. 2.5 software (Applied Biosystems, Foster City, CA). Nucleotide positions showing point indels and transversions located between positions 16180–16193 and 303–315 were excluded from further analysis.

Complete sequencing of mitochondrial genomes representing M5a and M35 lineages in one Slovak and one Russian individual and J1 lineages in three Slovak individuals was performed using the methodology described in details by Torroni et al. (2001).

All Slovak samples were subjected to RFLP analysis of coding region sites that were diagnostic of all major Eurasian clusters (haplogroups and subhaplogroups), on the basis of the hierarchical mtDNA RFLP scheme as described elsewhere (Malyarchuk et al. 2002; Malyarchuk et al. 2006b). 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) and Finnilä et al. (2000).

Phylogenetic and Statistical Analysis

All available published data on HVS I mtDNA variability or RFLPs in Slavonic-speaking populations were used for comparative purposes: Bulgarians (Calafell et al. 1996; Richards et al. 2000), Russians (Malyarchuk et al. 1995; Orekhov et al. 1999; Richards et al. 2000; Malyarchuk et al. 2002; Belyaeva et al. 2003; Malyarchuk et al. 2004); Ukrainians (Malyarchuk & Derenko, 2001), Belorussians (Belyaeva et al. 2003), Poles (Richards et al. 2000; Malyarchuk et al. 2002), Czechs (Richards et al. 2000; Vanecek et al. 2004; Malyarchuk et al. 2006b), Bosnians (Malyarchuk et al. 2003; Cvjetan et al. 2004), Slovenians (Malyarchuk et al. 2003; Zupanic Pajnic et al. 2004), Croatians (Tolk et al. 2000; Cvjetan et al. 2004), Serbians, Herzegovinians and Macedonians (Cvjetan et al. 2004).

Genetic variation was analyzed using methods implemented in the Arlequin 2.0 software (Schneider et al. 2000). The statistical significance of F_{ST}-values was estimated by permutation analysis using 10000 iterations. Chi-square analysis of haplogroup frequencies in populations was performed by means of the program CHIRXC, which estimates the probability of homogeneity using Monte Carlo simulation (1000 runs) (Zaykin & Pudovkin, 1993).

Multidimensional scaling (MDS) analysis of pairwise interpopulation F_{ST} values was performed by means of the software package STATISTICA (StatSoft, Inc., Tulsa, OK, USA). For MDS analysis, data from the following populations were used: 200 Southern Germans (Lutz et al. 1998), 101 Austrians (Parson et al. 1998), 150 Western Germans (Baasner et al. 1998; Baasner & Madea, 2000), 179 Czechs from western Bohemia (Malyarchuk et al. 2006b), 104 Slovenians and 144 Bosnians (Malyarchuk et al. 2003), 436 Poles and 201 Russians (Malyarchuk et al. 2002), 50 Finns, 47 Estonians and 83 Karelians (Sajantila et al. 1995).

Most parsimonious trees of the mtDNA haplogroups were reconstructed manually, following a parsimony approach, and confirmed by use of the program Network 4.2. Sequence

classification into mtDNA subclusters was based on the nomenclatures of Palanichamy et al. (2004) and Carelli et al. (2006) for subhaplogroup J1 and Sun et al. (2006) for haplogroups M5a and M35. Five completely sequenced mitochondrial genomes were submitted to the GenBank under accession numbers EF583175-EF583179.

Results

The analysis of HVS I and II variation, in combination with RFLP typing of the coding region haplogroup-diagnostic sites allowed detection of 187 different mtDNA haplotypes in a total sample of 207 Slovak individuals (Table 1). Out of these 187 haplotypes, 126 were revealed in eastern Slovakia (for n = 137) and 69 haplotypes were found in western Slovakia (for n = 70). Only 8 haplotypes were shared by two Slovak samples studied (Table 1). The overwhelming majority of Slovak mtDNA haplotypes were classified into West Eurasian haplogroups (such as H, HV0, HV2, HV3, J, T, U, I, W, X, N1b) (Table 2). Only two distinct haplotypes belonging to African-specific haplogroup L2a were detected in eastern Slovakia, this finding was not unexpected because low frequencies of African mtDNAs were previously found in different European populations (Salas et al. 2002; Malyarchuk et al. 2004; Malyarchuk & Czarny, 2005; Pereira et al. 2005). In eastern Slovaks, two haplotypes belonging to macrogroup M were also detected (Table 1). However, in contrast to the previously studied Czech population from western Bohemia (Malyarchuk et al. 2006b), samples from Slovakia do not display any East Eurasian mtDNAs. One of the Slovak M-haplotype belongs to subhaplogroup M1b and is identical to M1b1a-haplotypes revealed in Italians and Bedouins from southern Israel (Olivieri et al. 2006) as well as in Saudi Arabs (Abu-Amero et al. 2007). A second M-lineage detected in Slovaks is defined by variants at positions 16129–16223-16230–16233-16304–16344. This lineage is identical to those revealed previously in gene pools of the Bulgarian Roma at frequency of 3.6% (Gresham et al. 2001). Based on the presence of the 16129 variant, Gresham et al. (2001) suggested that this lineage belongs to Indian-specific haplogroup M5. Nevertheless, to determine its exact phylogenetic status we completely sequenced our Slovak sample (Slv227) and compared it with Indian M-haplotypes published by Sun et al. (2006) (Fig. 1). As a result, we have found that our sample belongs to haplogroup M35 due to mutations at positions 199 and 12561. Moreover, it shared transition at 15928 with the South Indian sample T17 (from Andhra Pradesh) that allowed us to define a new Indian/Roma branch called as M35b.

Previously, using complete mtDNA sequencing approach we found that another frequent Roma lineage with motif: 16129–16223-16291–16298 belongs to Indian-specific haplogroup M5 (Malyarchuk et al. 2006a). However, due to the progress of mitochondrial genome sequencing in Indian populations (Kivisild et al. 2006; Sun et al. 2006), it is possible to define the exact phylogenetic position of these Romani mtDNAs more precisely. We present here the completely sequenced Roma-specific M5 haplotype revealed in the Russian gene pool (individual Rus44 from study by Malyarchuk & Derenko, 2001) and compare it with the published M5-mtDNAs of Roma (Polish individual PL173; Malyarchuk et al. 2006a) and Indian/Asian origin (individuals R61 and As32 from studies by Sun et al. (2006) and Kivisild et al. (2006), respectively). As a result, we are able to considerably improve the M5 phylogenetic tree reconstructed by Sun et al. (2006). The results presented in Fig. 1 show that M5a1 trunk is defined by transitions at 4916 and 15287, and the Roma-specific branch of this subgroup designated as M5a1b is determined by transitions at four positions (1303, 3954, 6461 and 9833). It is worthwhile to note that one of the M5a1a samples (R61) was taken from South India (Andhra Pradesh), while another sample (As32) was of unknown Asian ancestry. Thus, our findings point to a possible Indian origin of both Roma-specific haplogroups, M5a1b and M35b.

Previously, we have found that the Polish Roma population is characterized by high incidence (18.8%) of haplogroup J1* lineage, defined by HVS I motif 16069–16126-16145–16222-16235–16261-16271 (Malyarchuk et al. 2006a). This and a similar haplotype, lacking only the 16271 transition, are very rare in European Roma populations, being found only in the Spanish, Bulgarian and Hungarian Roma (Gresham et al. 2001; Egyed et al. 2007). Among Europeans, such haplotypes have been revealed only in French (0.5%; Dubut et al. 2004), Hungarian (0.5%; Egyed et al. 2007) and Czech (about 3%; Vanecek et al. 2004; Malyarchuk et al. 2006b) populations. In the present study, we have found that 2.9% of individuals from eastern Slovakia are characterized by exactly the same J1*-haplotype. Taking into account its similarity with J1-haplotypes revealed in Southwestern Pakistani populations (Quintana-Murci et al. 2004) and thus assuming that this haplotype might have been characteristic of the ancestral Romani population (Malyarchuk et al. 2006a), we completely sequenced two J1*-samples revealed in Slovakia (Fig. 2). Comparison with published data (pooled in MitoMap mtDNA tree (Ruiz-Pesini et al. 2007)) demonstrated that these J1*-haplotypes belong to a new subhaplogroup J1a defined by transition at 8460. This subhaplogroup appears to be a sister clade to subhaplogroup J1b. Therefore, the J1a contribution to the Roma, and through them to some European gene pools, can be caused

Table 1 mtDNA haplotypes in Slovak populations

HVS I (minus 16000)	HVS II	HG	N
85 129	152 263 309iC 315iC	H*	E(1)
93 129 291 316	263 315iC	H*	W(1)
104 183AC 189 193iC 317AT	263 309iCC 315iC	H*	E(1)
129	263 309iCC 315iC	H*	E(1)
129	263 315iC	H*	E(1)
129 184	146 263 315iC	H*	E(1)
153 346 355 399	146 195 263 309iC 315iC	H*	E(1)
174 362	152 263 309iC 315iC	H*	E(1)
189 193dC 256	152 195 263 315iC	H*	E(1)
218	143 195 263 315iC	H*	E(1)
234	152 263 309iC	H*	E(1)
234	152 263 309iC 315iC	H*	E(1)
311	152 200 263 315iC	H*	E(1)
311	263 315iC	H*	E(1)
311 355	263 315iC	H*	W(1)
324	204 263 309iC 315iC	H*	W(1)
362	152 263 309iCC	H*	E(1)
366	263 309iC 315iC	H*	W(1)
CRS	146 152 195 263 309iC 315iC	H*	E(1)
CRS	183 263 309iC 315iC	H*	W(1)
CRS	207 263 309iC 315iC	H*	W(1)
CRS	262 263 309iC 315iCC	H*	W(1)
CRS	263 309iC 315iC	H*	E(1)
CRS	263 315iC	H*	E(2)
51 270	263 315iC	H1	W(1)
93	152 263 309iC 315iC	H1	E(1)
179 311	263 309dC 315iC	H1	E(1)
184	93 263 315iC	H1	W(1)
189	152 263 309iC 315iC	H1	E(1)
189	263 315iC	H1	E(1)
189 295	263 315iC	H1	E(1)
189 295	263 309iC 315iC	H1	E(1)
209	263 315iC	H1	W(1)
261	263 309iC 315iC	H1	E(1)
263	263 315iC	H1	E(1)
311	257 263 315iC	H1	W(1)
183AC 189 362	257 263 309iCCC 315iC	H1	W(1)
183AC 189 193iC 235 260	263 315iC	H1	W(1)
183AC 189 288	263 309iC	H1	E(1)
CRS	73 263 315iC	H1	W(1)
CRS	143 263 309iC 315iC	H1	W(1)
CRS	195 257 263 309iCC 315iC	H1	E(1)
CRS	263 315iC	H1	E(2)
CRS	263 309iC 315iC	H1	E(1)
134 162	73 263 309iC 315iC	H1a	E(1)
162	73 263 309iC 315iC	H1a	E(1) W(1)
162	73 263 315iC	H1a	E(1)
162	73 263 309iCC 315iCC	H1a	E(1)
80 189 193iC 356	151 263 309iC 315iC	H1b	E(1)
188 189 356 362	263 315iC	H1b	E(1)
189 356	263 315iC	H1b	W(1)
189 356	151 263 309iCC 315iC	H1b	E(1)
189 356 362	263 315iC	H1b	E(1)
183AC 189 356 362	263 309iC	H1b	E(1)

Table 1 Continued.

HVS I (minus 16000)	HVS II	HG	N
183AC 189 193iC 356	263 309iC 315iC	H1b	E(1)
CRS	263 309iC 315iC	H2	W(1)
CRS	249dA 263 309iCC 315iC	H2	E(1)
CRS	315iC	H2	E(1) W(1)
354	73 263 309iC 315iC	H2a	E(1)
157 192	263 315iC	H4	W(1)
311	146 200 263 309iCC 315iC	H4	E(1)
CRS	73 263 315iC	H4	W(1)
CRS	263 315iC	H4	E(1)
304	146 150 195 198 263 309iC 315iC	H5	W(1)
304	146 263 315iC	H5	W(1)
304	146 195 263 315iC	H5	E(1)
114 172 304 311	146 263 315iC	H5a	W(1)
218 304	263 309iC 315iC	H5a	E(1)
304	152 263 309iC 315iC	H5a	E(1)
304	263 309iC 315iC	H5a	E(2) W(1)
304	263 315iC	H5a	E(1) W(1)
304 343	263 309iC 315iC	H5a	E(1)
362	152 239 263 309iC 315iC	H6	E(1)
362	239 263 309iC 315iC	H6	E(2) W(2)
362	239 263 309iCC 315iC	H6	E(1)
362 400	239 263 315iC	H6	E(1)
362	152 239 263 315iC	H6	E(1)
265	263 309iCC 315iC	H7	E(1)
CRS	263 315iC	H7	E(1)
CRS	263 309iC 315iCC	H7	E(1)
CRS	263 309iC 315iC	H7	E(1)
278 293 311	195 263 309iC 315iC	H11a	E(1)
293 311	195 263 315iC	H11a	E(1) W(1)
214 217 335	73 151 152 195 246 263 309iC 315iC	HV2	E(1)
214 217 335	73 151 152 195 246 263 279 309iC 315iC	HV2	E(1)
172 300 311	146 263 309iC 315iCC	HV3	W(1)
311	263 309iC 315iC	HV3	E(1)
311 354	263 315iC	HV3	E(1)
113AC 172 311	263 309iCC 315iCC	HV3	W(1)
129 145 223 391	73 146 152 199 204 207 250 263 309iC	I	E(1)
129 172 223 311 391	73 199 203 204 250 263 315iC	I1	E(2) W(1)
129 172 223 311 391	73 199 203 204 250 263 309iC 315iC	I1	E(1)
129 172 223 311 391	73 189 199 203 204 250 263 315iC	I1	W(1)
129 172 223 311 391	73 199 204 250 263 315iC	I1	E(1)
129 223 311 368 391	73 152 199 204 250 263 309iC 315iC	I1	E(1)
129 223 311 391	73 189 193 199 204 250 263 309iC	I1	E(1)
86 129 223 391	73 152 199 204 207 239 250 263 315iC	I3	E(1)
69 126 145 222 235 261 271	73 263 295 309iC	J1a	E(2)
69 126 145 222 235 261 271	73 263 295 309iC 315iC	J1a	E(1)
69 126 145 222 235 261 271 290	73 263 295 309iC 315iC	J1a	E(1)
69 126 145 172 222 261	73 242 263 295 309iC 315iC	J1b1	W(1)
69 126 261	73 263 295 309iC	J1c	E(1)
69 92 126	73 185 228 263 295 309iC 315iC	J1c	W(1)
69 126 172 189	73 143 185 228 263 295 315iC	J1c	W(1)
69 126 153 390	73 150 185 195 228 263 295 315iC	J1c	W(1)
69 126 147 209 242 311	73 185 188 228 263 295 309iC 315iCC	J1c	W(1)
69 126 189 193iC 245	73 185 228 263 295 309iC 315iCC	J1c	E(1)
69 126	73 185 263 295 309iC 315iC 356iC	J1c	E(1)

Table 1 Continued.

HVS I (minus 16000)	HVS II	HG	N
69 126	73 185 195 228 263 295 309iC 315iC	J1c	E(1)
69 126	73 185 188 228 263 295 309iC 315iC	J1c	E(1)
69 126	73 146 185 188 222 228 263 295 315iC	J1c	E(1)
69 126	73 146 185 188 228 263 295 309iC 315iC	J1c	W(1)
69 126	73 185 188 228 263 295 315iC	J1c	E(1)
69 126 145 231 261	73 150 152 195 215 263 295 315iC	J2a	W(1)
69 126 145 231 261	73 150 152 195 215 263 295 311 315iC 319	J2a	W(1)
224 311	73 263 315iC	K	E(1)
224 311	73 146 195 263 309iC 315iC	K	E(1)
224 311	73 263 309iC 315iC	K	E(1)
51 223 278 294 309 390	73 143 146 152 195 263 309iC 315iC	L2a	E(1)
172 189 192 218 223 278 294 309 390	73 143 146 152 195 263 309iC 315iC	L2a	E(1)
129 185 189 193dC 223 249 311	73 195 200 263 309iC 315iC	M1b1a	E(1)
129 223 230 233 304 344	73 199 263 309iC 315iC	M35b	E(1)
145 176CG 223 390	73 152 263 315iC	N1b	W(1)
126 294 296	73 263 309iC 315iCC	T	W(1)
126 294 296	73 263 315iC	T	W(1)
126 294 296	73 263 309iC 315iC	T	W(1)
126 172 186 189 294 298 399	73 263 315.C	T	W(1)
126 274 294 304	73 263 315.C	T	W(1)
126 181 189 193iC 294 296	73 152 263 309iC 315iC	T	W(1)
126 294 296 362	73 263 309iC 315iC	T	W(1)
126 182AC 183AC 189 193iC 294 296 298	73 150 195 263 315iC	T	W(1)
126 182AC 183AC 189 294 296	73 195 263 315iC	T	W(1)
126 294 296 304	73 263 315iC	T	E(1)
126 140 189 294 296 311	73 195 263 309iC 315iC	T	E(1)
126 172 294 304	73 195 263 309iC 315iC	T	E(1)
126 294 304 318AT	73 152 263 309iCC 315iC	T	E(1)
126 234 294 296 304	73 146 263 309iC 315iC	T	E(1)
126 292 294 296 304	73 146 263 315iC	T	E(1)
126 153 207AC 294	73 150 263 309iC 315iC	T	E(1)
126 294 296 304	73 263 315iC	T	E(1)
126 294 304	73 152 263 309iC 315iC	T	E(1)
126 163 189 221 243 294 311	73 263 315iC	T1	W(1)
93 126 163 186 189 294	73 263 309iC 315iC	T1	E(1)
126 163 186 189 294	73 152 195 263 309iC 315iC	T1	E(2)
129 183AC 189 247 249 288	73 146 150 152 195 263 285 309iC 315iC 385	U1	W(1)
51 129GC 182AC 183AC 189 291300 362	73 152 217 263 309dC 315iC 340	U2	W(1)
51 129GC 183AC 189 193iC 362	73 152 204 263 309dC 315iC 340	U2	E(1)
356	73 263 309iC 315iC	U4	W(1)
261 356	73 195 263 315iC	U4	E(1)
356	73 146 152 195 263 309iCC 315iC	U4	E(1)
179 356	73 195 263 309iC 315iC	U4	W(1)
179 356	73 195 263 315iC	U4	E(1)
134 356	73 152 195 263 296 315iC	U4a	E(1)
356	73 146 195 263 310	U4a	W(1)
189 356	73 195 263 310	U4a	E(1)
356	73 195 263 310	U4a	E(1)
CRS	73 195 263 310	U4a	E(2)
256 270 320 399	73 195 263 309iC 315iC	U5a	W(1)
256 270 399	73 263 309iC	U5a	W(1)
192 234 256 270	73 263 309iC 315iC	U5a	E(1)
256 270 399	73 263 309iC 315iC	U5a	E(1)
93 183AC 189 193iC 270	73 150 152 263 315iC	U5b	W(1)

Table 1 Continued.

HVS I (minus 16000)	HVS II	HG	N
270	73 150 263 315iC	U5b	W(1)
174 192 311	73 150 263 315iC	U5b	W(1)
192 270	73 150 204 207 235 263 309iCC 315iC	U5b	W(1)
144 189 270	73 150 263 309iCC 315iC	U5b	E(1)
189 270	73 150 152 263 315iC	U5b	E(1)
147 183AC 189 270	73 150 263 309iCC 315iC	U5b	E(1)
189 193iC 270	73 150 152 263 309iC 315iC	U5b	E(1)
183AC 189 193iC 270	73 150 152 263 315iC	U5b	E(1)
93 189 270	73 150 263 315iC	U5b	E(1)
144 189 270	73 150 263 315iC	U5b	E(1)
129 189 194iC 270	73 150 152 263 285 309iC 315iC	U5b	E(1)
318AT	73 152 263 309iC	U7	E(1)
298	72 263 309iC 315iC	HV0	E(2) W(1)
298	72 152 263 315iC	HV0	E(1)
298	72 195 228 263 309iC 315iC	HV0	E(1)
298	72 263 309iCC 315iC	HV0	E(3)
223 292	73 189 194 195 199 204 207 263 315iC	W	W(1)
193 223 292	73 119 152 189 195 204 207 263 315iC	W	W(1)
192 223 292 325	73 189 194 195 204 207 263 309iC 315iC	W	W(1)
292	73 143 189 195 204 207 263 315iC	W	E(1)
223 292 295 324	73 189 195 204 207 263 315iC	W	W(1)
223 292	73 189 195 204 207 263 309iC 315iC	W	W(1)
189 223 278	73 150 153 195 225 226 263 309iC 315iC	X	W(1)
93 184CA 189 223 278	73 153 195 225 226 263 315iC	X	E(1)
189 192 223 278 399	73 143 195 225 226 235 263 309iC 315iC	X	E(1)

Number of subjects (N) is shown in parentheses; region belonging to eastern and western Slovakia is indicated by letters "E" and "W", respectively. Mutations are shown indicating positions relative to the revised mtDNA CRS (Andrews et al. 1999). The nucleotide positions in HVS I and II sequences correspond to transitions; transversions are further specified. Haplogroup names (HG) are given in large letter according to the mtDNA classification (Richards et al. 2000; Palanichamy et al. 2004; Achilli et al. 2005; Olivieri et al. 2006; Sun et al. 2006; Torroni et al. 2006). The presence of insertions or deletions is referred by "i" and "d", respectively, following the nucleotide position.

by gene flow from Indo-Pakistani region. One of the haplogroup J Slovak sample (Slv174; 16069–16126–16261–73–263–295) was characterized by uncertain HVS II motif, due to the lack of any subhaplogroup-diagnostic mutation according to classification developed in Palanichamy et al. (2004) and Carelli et al. (2006). To determine its phylogenetic status, we completely sequenced this sample and found that it belongs to subhaplogroup J1c being accompanied by transition at 14798, despite the lack of diagnostic mutation at position 228. It is unclear however, whether this sample should occupy the ancestral node for J1c-phylogeny or it is the result of back-mutation at position 228 (Fig. 2).

As in other European populations, the most frequent haplogroup in Slovaks is haplogroup H that encompasses 45.4%. Similar frequencies of this haplogroup (44%) have been previously revealed in Czech population (Malyarchuk et al. 2006b). Among the Slovaks, twelve subhaplogroups (H*, H1*, H1a, H1b, H2*, H2a1, H4, H5*, H5a, H6, H7, and H11a) were found using an RFLP approach described by Loogväli et al. (2004). Table 2 shows haplogroups frequency distribution in western and eastern Slovaks, in comparison with the Czechs. Some differences in between-population distribution of H-subgroups can be revealed (for instance, the lower frequency of H1b, H6 and H7 or higher frequency of H5* in western Slovaks), however, they are statistically insignificant. Between-population comparisons for distribution of all mtDNA haplogroups and subhaplogroups (Table 2) demonstrate the absence of statistical differences between Slovak and Czech populations ($P > 0.35$ in all cases). When comparisons on the level of individual haplogroups were performed, the only haplogroup W was revealed more frequently in western Slovaks than in eastern Slovaks and Czechs ($P = 0.018$ and 0.004, respectively). Chi-square analysis of mtDNA haplogroup frequencies in Slovaks in comparison with Slavonic-speaking populations, such as Czechs, Poles, Bosnians, Slovenians and Russians (according

HG	Slovaks			Czechs
	West region (n = 70)	East region (n = 137)	In total (n = 207)	(n = 179)
H*	7 (10.0)	18 (13.1)	25 (12.1)	26 (14.5)
H1*	8 (11.4)	13 (9.5)	21 (10.1)	16 (8.9)
H1a	1 (1.4)	4 (2.9)	5 (2.4)	5 (2.8)
H1b	1 (1.4)	6 (4.4)	7 (3.4)	8 (4.5)
H2*	2 (2.9)	2 (1.5)	4 (1.9)	6 (3.4)
H2a1	0	1 (0.7)	1 (0.5)	3 (1.7)
H4	2 (2.9)	2 (1.5)	4 (1.9)	1 (0.6)
H5*	2 (2.9)	1 (0.7)	3 (1.4)	4 (2.2)
H5a	3 (4.3)	6 (4.4)	9 (4.3)	2 (1.1)
H6	2 (2.9)	6 (4.4)	8 (3.9)	3 (1.7)
H7	0	4 (2.9)	4 (1.9)	2 (1.1)
H11a	1 (1.4)	2 (1.5)	3 (1.4)	3 (1.7)
HV0	1 (1.4)	7 (5.1)	8 (3.9)	4 (2.2)
HV2	0	2 (1.5)	2 (1.0)	0
HV3	2 (2.9)	2 (1.5)	4 (1.9)	3 (1.7)
I	2 (2.9)	8 (5.8)	10 (4.8)	5 (2.8)
J1a	0	4 (2.9)	4 (1.9)	5 (2.8)
J1b1	1 (1.4)	0	1 (0.5)	2 (1.1)
J1c	5 (7.1)	7 (5.1)	12 (5.8)	12 (6.7)
J2a	2 (2.9)	0	2 (1.0)	2 (1.1)
N1a	0	0	0	1 (0.6)
N1b	1 (1.4)	0	1 (0.5)	2 (1.1)
N9a	0	0	0	1 (0.6)
K	0	3 (2.2)	3 (1.4)	7 (3.9)
M1b	0	1 (0.7)	1 (0.5)	0
M35b	0	1 (0.7)	1 (0.5)	0
T*	9 (12.9)	9 (6.6)	18 (8.7)	17 (9.5)
T1	1 (1.4)	3 (2.2)	4 (1.9)	5 (2.8)
U1	1 (1.4)	0	1 (0.5)	0
U2	1 (1.4)	1 (0.7)	2 (1.0)	1 (0.6)
U3	0	0	0	4 (2.2)
U4*	2 (2.0)	3 (2.2)	5 (2.4)	2 (1.1)
U4a	1 (1.4)	5 (3.6)	6 (2.9)	3 (1.7)
U5a	2 (2.9)	2 (1.5)	4 (1.9)	9 (5.0)
U5b	4 (5.7)	8 (5.8)	12 (5.8)	6 (3.4)
U7	0	1 (0.7)	1 (0.5)	0
U8a	0	0	0	1 (0.6)
W	5 (7.1)	1 (0.7)	6 (2.9)	1 (0.6)
X	1 (1.4)	2 (1.5)	3 (1.4)	3 (1.7)
A	0	0	0	1 (0.6)
D	0	0	0	1 (0.6)
M7	0	0	0	2 (1.1)
L2a	0	2 (1.5)	2 (1.0)	0

Data for Czechs are from Malyarchuk et al. (2006b).

Table 2 Haplogroup distributions (no. of individuals and % values in parentheses) in Slovaks in comparison with Czech population

to data by Malyarchuk et al. (2006b)) has also revealed the population homogeneity (with $P > 0.15$ in all comparisons).

A similar conclusion (namely a lack of between-population differences in Slovaks and Czechs) also follows from an analysis of molecular variance (AMOVA) performed on the level of HVS I sequences. For such analysis, different European populations – Slavonic-speaking (Slovaks, Czechs, Slovenians, Bosnians, Poles, Russians), German-speaking (West Germans, South Germans, Austrians) and Finno-Ugric-speaking (Finns, Karelians, Estonians) – were used. Analysis of between-population

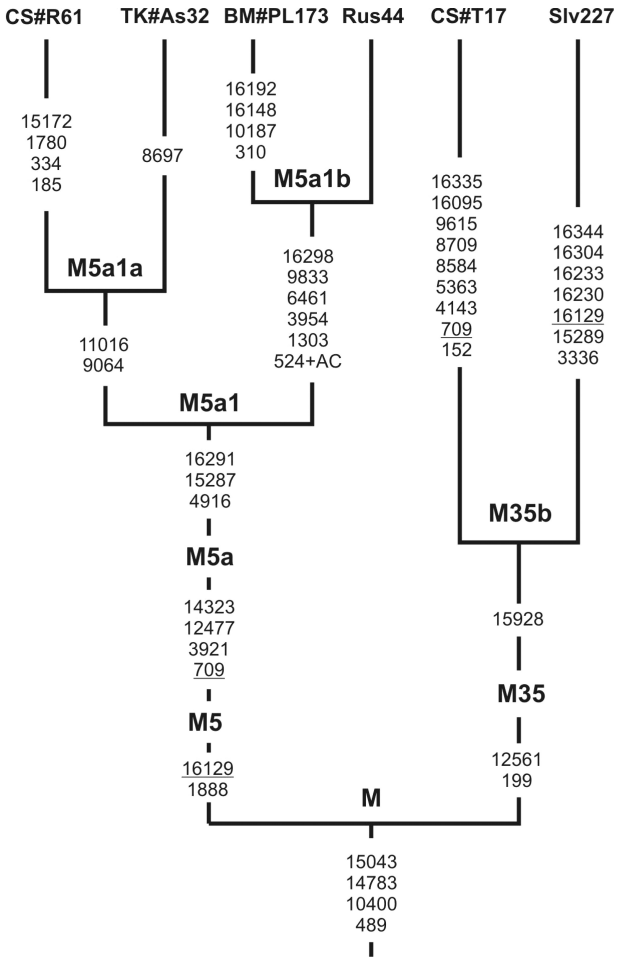

Figure 1 The phylogenetic tree of haplogroups M5a and M35 based on complete mitochondrial genome sequences. The tree is rooted in macrohaplogroup M. Numbers along links refer to substitutions scored relative to the rCRS (Andrews et al. 1999). Transversions are further specified, recurrent mutations are underlined. A plus sign (+) denotes an insertion. Four additional complete sequences were taken from the literature (Kivisild et al. 2006; Malyarchuk et al. 2006a; Sun et al. 2006) and designated by TK, BM, and CS, respectively, followed by "#" and the original sample code. For sample TK#As32 only coding region information was available.

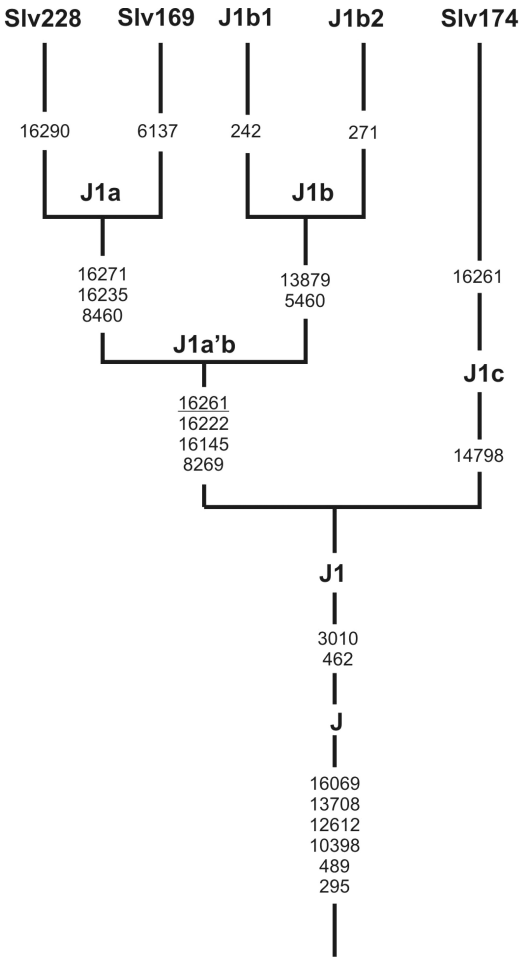

Figure 2 The phylogenetic tree of subhaplogroup J1 based on complete mitochondrial genome sequences. The tree is rooted in haplogroup J. Numbers along links refer to substitutions scored relative to the rCRS (Andrews et al. 1999). Recurrent mutations are underlined. For subhaplogroups J1b1 and J1b2 only diagnostic mutations are shown according to classification (Palanichamy et al. 2004).

differentiation based on F_{ST}-distances revealed that only 0.16% of variation was due to differences among populations ($P = 0.013$). Significant pairwise F_{ST}-differences ($P < 0.05$) were found mainly between some Finno-Ugric populations (Finns and Karelians) and Slavonic and German populations (Table 3). The MDS analysis performed on the basis of pairwise F_{ST} values revealed that Slovak populations do not cluster together. Western Slovaks are located together with the Czechs and Austrians (in accordance with their geographic proximity), whereas eastern Slovaks are placed close to Slovenians (Fig. 3).

Discussion

Although the mtDNA analysis appears to be an effective tool in the reconstruction of population history, such studies of European populations do not allow researchers to identify any specific features clearly distinguishing the Slavs from neighboring populations (Malyarchuk & Derenko 2001; Malyarchuk et al. 2002). This means that questions such as the reconstruction of proto-Slavonic migrations and

Table 3 Between-population differences based on pairwise F_{ST}-distances inferred from mtDNA HVS I variation data

Population	CZE	POL	SLO	BOS	RUS	AUS	SGE	WGE	EST	FIN	KAR	SLV (W)
POL	0.0013											
SLO	0.0031	0.0009										
BOS	0.0032	0.0006	0.0003									
RUS	−0.0028	−0.0008	0.0019	−0.0011								
AUS	−0.0031	0.0010	0.0015	−0.0007	−0.0046							
SGE	0.0007	0.0001	0.0004	−0.0012	−0.0014	−0.0008						
WGE	0.0010	−0.0024	−0.0010	0.0021	−0.0026	0.0003	0.0005					
EST	−0.0007	−0.0031	0.0017	−0.0016	−0.0006	−0.001	−0.0007	−0.0082				
FIN	0.0085*	0.0038	0.0076*	0.0017	0.0087*	0.0082*	0.0043	0.0026	−0.0018			
KAR	0.0078*	0.0052*	0.0089*	0.0034	0.0056	0.0087*	0.0062*	−0.0009	−0.0063	0.0010		
SLV (W)	−0.0038	−0.0006	0.0041	0.0044	−0.0024	−0.0023	0.0012	0.0009	−0.0030	0.0058	0.0094*	
SLV (E)	0.0008	0.0023	0.0067*	0.0021	0.0005	0.0025	0.0017	0.0036	−0.0024	0.0050	0.0040	0.0007

Abbreviations for the populations are as follows: SLV (W) – western Slovaks, SLV (E) – eastern Slovaks, CZE – Czechs, POL – Poles, SLO – Slovenians, BOS – Bosnians, RUS – Russians, AUS – Austrians, SGE – South Germans, WGE – West Germans, EST – Estonians, FIN – Finns, KAR – Karelians. *significant differences ($p < 0.05$).

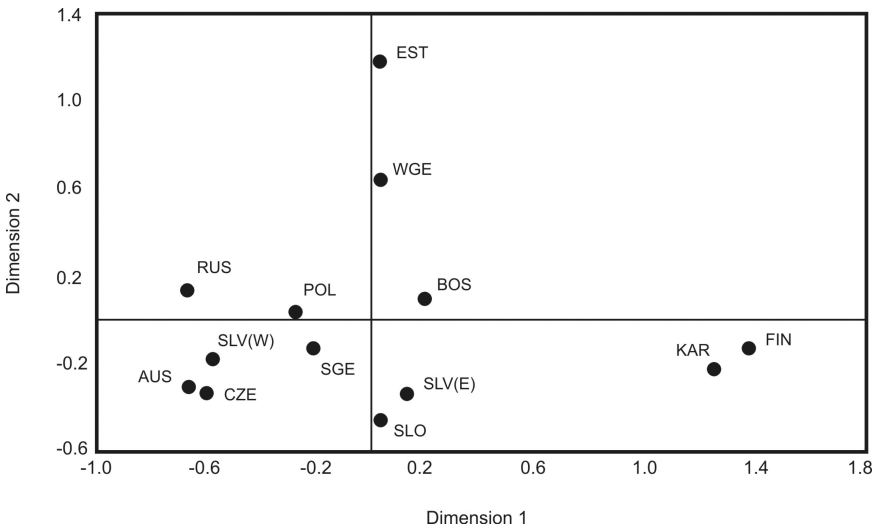

Figure 3 Multidimensional scaling plot of F_{ST} distances between Slovaks and the surrounding European populations based on mtDNA HVS I variation data (stress value 0.001). Populations designated as in Table 3.

the search for the Slav homeland may not be answered by use of mtDNA markers. Meanwhile, Y-chromosome studies have shown that two genetically distant groups of the Slavs – Southern and Western/Eastern ones – can be recognized by means of SNP and/or STR analyses (Pericic et al. 2005; Rebala et al. 2007). Results of recent Y-STR analysis in different Slavonic populations, including Slovaks (Rebala et al. 2007), suggest that the expansion of medieval Slavs started from their putative homeland in the basin of the middle Dnieper (Ukraine). However, this conclusion seems to be uncertain because the distribution pattern of insignificant P-values for population pairwise F_{ST}-distances (Table 1 of Rebala et al. (2007)) indicates that Y-STR genetic homogeneity in Slavs extends from Slovakia, Ukraine and central Belarus to western parts of Russia, thus suggesting that Slavonic expansion might also have started in Central Europe. In any case, additional studies based on STR-SNP diversity analysis of Y-chromosome and complete mtDNA sequencing, in combination with autosomal loci variation, are required to solve the key problems of the genetic history of the Slavs.

One of the important questions is the nature and extent of non-European admixture in gene pools of Slavs. It is worth noting that gene pools of Slovaks and Czechs are characterized by a common component represented by the

Roma-specific mtDNA lineages. The Roma (also known as Gypsies), who are believed to be of Indian origin, nowadays represent a large population spread over all of Europe, with their highest concentrations in southeastern Europe and the Iberian Peninsula (Kalaydjieva et al. 2001). It is suggested that by the 13th century the Roma had entered the Balkans and some groups moved slowly through the territories inhabited by the Slavs (Demeter et al. 2000). The first migration of small groups of the Roma from Hungary throughout Bohemia to the north and west of Europe took place at the beginning of the 15th century (Demeter et al. 2000). According to the latest Census in 2001 the Roma form the minority group, representing about 3.5% of the total Slovak population. However, not all Roma people may have claimed themselves as Roma, so the real number is unclear. Genetic studies have shown that the Roma populations share a common genetic history, as evidenced by mtDNA and Y-chromosomal markers as well as by several disease loci (Gresham et al. 2001; Kalaydjieva et al. 2001; Morar et al. 2004; Malyarchuk et al. 2006a). However, despite the obvious evidence that some Roma founder mtDNA lineages belong to haplogroups widespread across South Asia, there is no detailed phylogenetic characterization of such mtDNAs. In the present study, we extend considerably our knowledge about the phylogenetic position of Roma-specific lineages revealed in gene pools of different Slavonic groups (such as Slovaks, Czechs, Poles, and Russians), by using complete mtDNA sequencing approach. As a result, we have found that some Roma-specific M-lineages belong to the Indian haplogroups M5a1 and M35. Moreover, a novel subhaplogroup J1a, which is present with a marked frequency in gene pools of the Roma as well as Slovaks and Czechs, has been described. Further reconstruction of the phylogeographic structure of subhaplogroup J1a, requires new complete mtDNA data from India and Southwest Asia.

Acknowledgements

This research received support from the Program of Basic Research of Russian Academy of Sciences "Biodiversity and gene pools dynamics". B.A.M. was supported in part by the Russian Science Support Foundation. We are grateful to Dr. Tomasz Grzybowski for help in this study.

References

Abu-Amero, K. K., Gonzalez, A. M., Larruga, J. M., Bosley, T. M. & Cabrera, V. M. (2007). Eurasian and African mitochondrial DNA influences in the Saudi Arabian population. *BMC Evol. Biol* **7**, 32.

Achilli, A., Rengo, C., Battaglia, V., Pala, M., Olivieri, A., Fornarino, S., Magri, C., Scozzari, R., Babudri, N., Santachiara-Benerecetti, A. S., Bandelt, H. -J., Semino, O. & Torroni, A. (2005). Saami and Berbers – an unexpected mitochondrial DNA link. *Am J Hum Genet* **76**, 883–886.

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

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. *Nat Genet* **23**, 147.

Baasner, A. & Madea, B. (2000). Sequence polymorphisms of the mitochondrial DNA control region in 100 German Caucasians. *J Forensic Sci* **45**, 1343–1348.

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 Int* **98**, 169–178.

Belyaeva, O., Bermisheva, M., Khrunin, A., Slominsky, P., Bebyakova, N., Khusnutdinova, E., Mikulich, A. & Limborska, S. (2003). Mitochondrial DNA variations in Russian and Belorussian populations. *Hum Biol* **75**, 647–660.

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* **60**, 35–049.

Carelli, V., Achilli, A., Valentino, M. L., Rengo, C., Semino, O., Pala, M., Olivieri, A., Mattiazzi, M., Pallotti, F., Carrara, F., Zeviani, M., Leuzzi, V., Carducci, C., Valle, G., Simionati, B., Mendieta, L., Salomao, S., Belfort, R. Jr, Sadun, A.A. & Torroni, A. (2006). Haplogroup effects and recombination of mitochondrial DNA: novel clues from the analysis of Leber hereditary optic neuropathy pedigrees. *Am J Hum Genet* **78**, 564–574.

Cvjetan, S., Tolk, H. V., Barac Lauc, L., Colak, I., Dordevic, D., Efremovska, L., Janicijevich, B., Kvesic, A., Martinovic Klaric, I., Metspalu, E., Pericic, M., Parik, J., Popovic, D., Sijacki, A., Terzic, R., Villems, R. & Rudan, P. (2004). Frequencies of mtDNA haplogroups in Southeastern Europe – Croatians, Bosnians and Herzegovinians, Serbians, Macedonians and Macedonian Romani. *Coll Anthropol* **28**, 193–198.

Demeter, N., Bessonov, N. & Kutenkov, V. (2000). *The Roma history. New view. Voronezh: N. Mikluho-Maklay Institute of Ethnology and Anthropology*, Russian Academy of Sciences (in Russian).

Dubut, V., Chollet, L., Murail, P., Cartault, F., Beraud-Colomb, E., Serre M. & Mogentale-Profizi, N. (2004). MtDNA polymorphisms in five French groups: importance of regional sampling. *Eur J Hum Genet* **12**, 293–300.

Egyed, B., Brandstätter, A., Irwin, J. A., Pádár, Z., Parsons, T. J. & Parson, W. (2007). Mitochondrial control region sequence variations in the Hungarian population: Analysis of population samples from Hungary and from Transylvania (Romania). *Forensic Sci Internatl: Genetics*. **1**, 158–162.

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.

Gimbutas, M. (1971). *The Slavs*. New York, NY: Praeger Publishing.

Gresham, D., Morar, B., Underhill, P. A., Passarino, G., Lin, A. A., Wise, C., Angelicheva, D., Calafell, F., Oefner, P. J., Shen, P., Tournev, I., de Pablo, R., Kicinskas, V., Perez-Lezaun, A., Marushiakova, E., Popov, V. & Kalaydjieva, L. (2001). Origins and divergence of the Roma (Gypsies). *Am J Hum Genet* **69**, 1314–1331.

Kalaydjieva, L., Gresham, D. & Calafell, F. (2001). Genetic studies of the Roma (Gypsies): a review. *BMC Med Genet* **2**, 5.

Kivisild, T., Shen, P., Wall, D. P., Do, B., Sung, R., Davis, K., Passarino, G., Underhill, P. A., Scharfe, C., Torroni, A., Scozzari, R., Modiano, D., Coppa, A., de Knijff, P., Feldman, M., Cavalli-Sforza, L. L. & Oefner, P. J. (2006). The role of selection in the evolution of human mitochondrial genomes. *Genetics* **172**, 373–387.

Loogväli, E. -L., Roostalu, U., Malyarchuk, B. A., Derenko, M. V., Kivisild, T., Metspalu, E., Tambets, K., Reidla, M., Tolk, H. -V., Parik, J., Pennarun, E., Laos, S., Lunkina, A., Golubenko, M., Barac, L., Pericic, M., Balanovsky, O. P., Gusar, V., Khusnutdinova, E. K., Stepanov, V., Puzyrev, V., Rudan, P., Balanovska, E. V., Grechanina, E., Richard, C., Moisan, J. P., Chaventre, A., Anagnou, N. P., Pappa, K. I., Michalodimitrakis, E. N., Claustres, M., Golge, M., Mikerezi, I., Usanga, E. & Villems, R. (2004). Disuniting uniformity: a pied cladistic canvas of mtDNA haplogroup H in Eurasia. *Mol Biol Evol* **21**, 2012–2021.

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. & Czarny, J. (2005). African DNA lineages in the mitochondrial gene pool of Europeans. *Mol Biol (Mosk)* **39**, 703–709.

Malyarchuk, B. A. & Derenko, M. V. (2001). Mitochondrial DNA variability in Russians and Ukrainians: Implication to the origin of the Eastern Slavs. *Ann Hum Genet* **65**, 63–78.

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

Malyarchuk, B. A., Grzybowski, T., Derenko, M. V., Czarny, J., Wozniak, M. & Miscicka-sliwka, D. (2002). Mitochondrial DNA variability in Poles and Russians. *Ann Hum Genet* **66**, 261–283.

Malyarchuk, B. A., Grzybowski, T., Derenko, M. V., Czarny, J., Drobni(, K. & Miscicka-Sliwka, D. (2003). Mitochondrial DNA variability in Bosnians and Slovenians. *Ann Hum Genet* **67**, 412–425.

Malyarchuk, B., Derenko, M., Grzybowski, T., Lunkina, A., Czarny, J., Rychkov, S., Morozova, I., Denisova, G. & Miscicka-sliwka, D. (2004). Differentiation of mitochondrial DNA and Y chromosome in Russian populations. *Hum Biol* **76**, 877–900.

Malyarchuk, B. A., Grzybowski, T., Derenko, M. V., Czarny, J. & Miscicka-sliwka, D. (2006a). Mitochondrial DNA diversity in the Polish Roma. *Ann Hum Genet* **70**, 195–206.

Malyarchuk, B. A., Vanecek, T., Perkova, M. A., Derenko, M. V. & Sip, M. (2006b). Mitochondrial DNA variability in the Czech population, with application to the ethnic history of Slavs. *Hum Biol* **78**, 681–696.

Morar, B., Gresham, D., Angelicheva, D., Tournev, I., Gooding, R., Guergueltcheva, V., Schmidt, C., Abicht, A., Lochmuller, H., Tordai, A., Kalmar, L., Nagy, M., Karcagi, V., Jeanpierre, M., Herczegfalvi, A., Beeson, D., Venkataraman, V., Carter, K. W., Reeve, J., de Pablo, R., Kucinskas, V. & Kalaydjieva, L. (2004). Mutation history of the Roma/Gypsies. *Am J Hum Genet* **75**, 596–609.

Olivieri, A., Achilli, A., Pala, M., Battaglia, V., Fornarino, S., Al-Zahery, N., Scozzari, R., Cruciani, F., Behar, D. M., Dugoujon, J. M., Coudray, C., Santachiara-Benerecetti, A. S., Semino, O., Bandelt, H.-J. & Torroni, A. (2006). The mtDNA legacy of the Levantine early Upper Palaeolithic in Africa. *Science* **314**, 1767–1770.

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.

Palanichamy, M. G., Sun, C., Agrawal, S., Bandelt, H. -J., Kong, Q. -P., Khan, F., Wang, C. -E., Chaudhuri, T. K., Palla, V. & Zhang, Y. -P. (2004). Phylogeny of mitochondrial DNA macrohaplogroup N in India, based on complete sequencing: Implications for the peopling of South Asia. *Am J Hum Genet* **75**, 966–978.

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.

Pereira, L., Cunha, C., Alves, C. & Amorim, A. (2005). African female heritage in Iberia: a reassessment of mtDNA lineage distribution in present times. *Hum Biol* **77**, 213–229.

Pericic, M., Lauc, L. B., Klaric, I. M., Rootsi, S., Janicijevic, B., Rudan, I., Terzic, R., Colak, I., Kvesic, A., Popovic, D., Sijacki, A., Behluli, I., Dordevic, D., Efremovska, L., Bajec, D. D., Stefanovic, B. D., Villems, R. & Rudan, P. (2005). High-resolution phylogenetic analysis of southeastern Europe traces major episodes of paternal gene flow among Slavic populations. *Mol Biol Evol* **22**, 1964–1975.

Pliss, L., Tambets, K., Loogvali, E. -L., Pronina, N., Lazdins, M., Krumina, A., Baumanis, V. & Villems, R. (2006). Mitochondrial DNA portrait of Latvians: Towards the understanding of the genetic structure of Baltic-speaking populations. *Ann Hum Genet* **70**, 439–458.

Quintana-Murci, L., Chaix, R., Wells, R. S., Behar, D. M., Sayar, H., Scozzari, R., Rengo, C., Al-Zahery, N., Semino, O., Santachiara-Benerecetti, A. S., Coppa, A., Ayub, Q., Mohyuddin, A., Tyler-Smith, C., Mehdi, S. Q., Torroni, A. & McElreavey, K. (2004). Where West meets East: The complex mtDNA landscape of the southwest and central Asian corridor. *Am J Hum Genet* **74**, 827–845.

Rebala, K., Mikulich, A. I., Tsybovsky, I. S., Sivakova, D., Dzupinkova, Z., Szczerkowska-Dobosz, A. & Szczerkowska, Z. (2007). Y-STR variation among Slavs: evidence for the Slavic homeland in the middle Dnieper basin. *J Hum Genet* **52**, 406–414.

Richards, M. B., Macaulay, V. A., Hickey, E., Vega, E., Sykes, B., Guida, V., Rengo, C., Sellito, D., Cruciani, F., Kivisild, T., Villems, R., Thomas, M., Rychkov, S., Rychkov, O., Rychkov, Yu., Golge, M., Dimitrov, D., Hill, E., Bradley, D., Romano, V., Cali, F., Vona, G., Demaine, A., Papiha, S., Triantaphyllidis, C., Stefanescu, G., Hatina, J., Belledi, M., DiRienzo, A., Novelletto, A., Oppenheim, A., Norby, S., Al-Zaheri, N., Santachiara-Benerecetti, S., Scozzari, R., Torroni, A. & Bandelt, H. -J. (2000). Tracing European founder lineages in the Near Eastern mtDNA pool. *Am J Hum Genet* **67**, 1251–1276.

Ruiz-Pesini, E., Lott, M.T., Procaccio, V., Poole, J. C., Brandon, M. C., Mishmar, D., Yi, C., Kreuziger, J., Baldi, P. & Wallace, D. C. (2007). An enhanced MITOMAP with a global mtDNA mutational phylogeny. *Nucleic Acids Res* **35** (Database issue), D823–828.

Sajantila, A., Lahermo, P., Anttinen, T., Lukka, M., Sistonen, P., Savontaus, M. L., Aula, P., Beckman, L., Tranebjaerg, L., Gedde-Dahl, T., Issel-Tarver, L., Di Rienzo, A. & Paabo, S. (1995). Genes and languages in Europe – an analysis of mitochondrial lineages. *Genome Res* **5**, 42–52.

Salas, A., Richards, M., De la Fe, T., Lareu, M. V., Sobrino, B., Sanchez-Diz, P., Macaulay, V. & Carracedo, A. (2002). The making of the African mtDNA landscape. *Am J Hum Genet* **71**, 1082–1111.

Savli, J., Bor, M. & Tomažic, I. (1996). *Veneti. First builders of European community. Tracing the history and language of early ancestors of Slovenes*. Wien, Boswell: Editiones Veneti.

Schneider, S., Roessli, D. & Excoffier, L. (2000). *Arlequin ver.2.0: A software for population genetics data analysis*. Genetics and Biometry laboratory, University of Geneva, Switzerland.

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

Sun, C., Kong, Q. P., Palanichamy, M. G., Agrawal, S., Bandelt, H. -J., Yao, Y. G., Khan, F., Zhu, C. L., Chaudhuri, T. K. & Zhang, Y. P. (2006). The dazzling array of basal branches in the mtDNA macrohaplogroup M from India as inferred from complete genomes. *Mol Biol Evol* **23**, 683–690.

Tolk, H. V., Pericic, M., Barac, L., Martinovic Klaric, I., Janicijevic, B., Rudan, I., Parik, J., Villems, R. & Rudan, P. (2000). MtDNA haplogroups in the populations of Croatian Adriatic Islands. *Coll Anthropol* **24**, 267–279.

Torroni, A., Huoponen, K., Francalacci, P., Petrozzi, M., Morelli, L., Scozzari, R., Obinu, D., Savontaus, M. -L. & Wallace, D. C. (1996). Classification of European mtDNAs from an analysis of three European populations. *Genetics* **144**, 1835–1850.

Torroni, A., Rengo, C., Guida, V., Cruciani, F., Sellitto, D., Coppa, A., Calderon, F. L., Simionati, B., Valle, G., Richards, M., Macaulay, V. & Scozzari, R. (2001). Do the four clades of the mtDNA haplogroup L2 evolve at different rates? *Am J Hum Genet* **69**, 1348–1356.

Torroni, A., Achilli, A., Macaulay, V., Richards, M. & Bandelt, H.-J. (2006). Harvesting the fruit of the human mtDNA tree. *Trends Genet* **22**, 339–345.

Vanecek, T., Vorel, F. & Sip, M. (2004). Mitochondrial DNA D-loop hypervariable regions: Czech population data. *Int J Legal Med* **118**, 14–18.

Zaykin, D. V. & Pudovkin, A. J. (1993). Two programs to estimate significance of Chi-square values using pseudo-probability test. *J Heredity* **84**, 152.

Zupanic Pajnic, I., Balazic, J. & Komel, R. (2004). Sequence polymorphism of the mitochondrial DNA control region in the Slovenian population. *Int J Legal Med* **118**, 1–4.

Received: 5 September 2007
Accepted: 27 September 2007