

# Mitochondrial DNA Variability in Bosnians and Slovenians

B. A. Malyarchuk<sup>1,\*</sup>, T. Grzybowski<sup>2</sup>, M. V. Derenko<sup>1</sup>, J. Czarny<sup>2</sup>, K. Drobnič<sup>3</sup> and D. Mišćicka-Šliwka<sup>2</sup>

<sup>1</sup>*Institute of Biological Problems of the North, Russian Academy of Sciences, Portovaya str. 18, 685000 Magadan, Russia*

<sup>2</sup>*The Ludwik Rydygier Medical University, Forensic Medicine Institute, ul. Skłodowskiej-Curie 9, 85-094 Bydgoszcz, Poland*

<sup>3</sup>*Forensic Laboratory and Research Center, Ministry of the Interior, Vodovodna 95, 1000 Ljubljana, Slovenia*

---

## Summary

Mitochondrial DNA variability in two Slavonic-speaking populations of the northwestern Balkan peninsula, Bosnians (N = 144) and Slovenians (N = 104), was studied by hypervariable segments I and II (HVS I and II) sequencing and restriction fragment-length polymorphism (RFLP) analysis of the mtDNA coding region. The majority of the mtDNA detected in Southern Slavonic populations falls into the common West Eurasian mitochondrial haplogroups (e.g., H, pre-V, J, T, U, K, I, W, and X). About 2% of the Bosnian mtDNAs encompass East Eurasian and African lineages (e.g., M and L1b, respectively). The distribution of mtDNA subclusters in Bosnians, Slovenians and the neighbouring European populations reveals that the common genetic substratum characteristic for Central and Eastern European populations (such as Germans, Poles, Russians and Finns) penetrates also South European territories as far as the Western Balkans. However, the observed differentiation between Bosnian and Slovenian mtDNAs suggests that at least two different migration waves of the Slavs may have reached the Balkans in the early Middle Ages.

---

## Introduction

Human history of the Balkan Peninsula is very complex because of multiple migration processes that occurred there since the initial occupation by anatomically modern humans in Palaeolithic times. According to archaeological studies, the Balkans were settled during the Mesolithic and Neolithic periods (Pinhasi *et al.* 2000). Moreover, the Neolithic expansion from Anatolia into Europe spread through the Balkans and led to increasing population density in the Balkan area, although its Mesolithic populations seem to have been very scarce (Lahr *et al.* 2000; Pinhasi *et al.* 2000). Important episodes of the immediate impact on the cultural and linguistic nature of Europe were the Urnfield culture migrations that occurred in the late Bronze Age. It

has been suggested that the bearers of this culture colonized the entire area from the Baltic southwards, over the Alps to the Adriatic Sea and Apennines (Piggott, 1965). From about 1000 B.C., the Balkan Peninsula was already inhabited by the Illyrians in the west and by the Thracians in the south-east between the Danube and the Aegean Sea. According to archaeological and historical data, the Slavs invaded the Balkan Peninsula as early as the 2<sup>nd</sup> century A.D. and advanced as far as Greece and even settled on the Peloponnese and several Aegean islands (Sedov, 1995; Šavli *et al.* 1996). As a result of the settlement movements of the Slavs in the Balkans, beginning in the 5<sup>th</sup>-6<sup>th</sup> centuries A.D., the various South Slavonic groups gradually evolved (Šavli *et al.* 1996). It has been suggested that the Illyrians were assimilated by Slavonic tribes and became the ancestors of modern Slovenians, Croatians and Bosnians, whereas the Serbians may be the descendants of the Thracians assimilated by the Slavs. However, the origin and genetic history of the Southern Slavonic populations is still poorly understood.

\*Correspondence: Dr. Boris A. Malyarchuk, Genetics Laboratory, Institute of Biological Problems of the North, Portovaya str., 18, 685000 Magadan, Russia. Fax/Phone: 7 41322 34463. E-mail: malyar@ibpn.kolyma.ru

Genetic studies of the Southern Slavs by the use of molecular markers of mitochondrial DNA (mtDNA) and the Y chromosome have attracted the interest of population geneticists (Calafell *et al.* 1996; Comas *et al.* 1999; Richards *et al.* 2000; Rosser *et al.* 2000; Tolk *et al.* 2000; Zaharova *et al.* 2000; Sterlinko *et al.* 2001; Torroni *et al.* 2001). However, mtDNA population data are so far available only for Bulgarians (Calafell *et al.* 1996; Richards *et al.* 2000) and a Croatian sample from the Adriatic islands (Tolk *et al.* 2000). These studies have failed to demonstrate significant differences between Southern Slavonic populations and most other Europeans. It was found that Bulgarians are close to Western Europeans in terms of genetic distances (Calafell *et al.* 1996). The study of Croatians revealed that their mtDNA cluster composition, as well as frequency pattern, is generally similar to other European and Near Eastern populations, with some deviations toward Asian (A, F) and African (L2a) mtDNA haplogroups (Tolk *et al.* 2000). These results suggest that the genetic history of the South Slavonic populations inhabiting the Balkan Peninsula is very complex and needs further investigation. In order to characterize the mtDNA structure and diversity in Southern Slavs, we have performed sequence analysis of the two hypervariable segments of the mtDNA control region and typed the mtDNA for coding region RFLP markers in samples from Bosnian and Slovenian individuals. To investigate the human history of the Balkan Peninsula, these results are compared with the data from other European populations.

## Materials and Methods

### Population Samples

Population samples of 104 Slovenians and 144 Bosnians living in different regions of Slovenia and Bosnia-Herzegovina, respectively, were studied. All individuals were maternally unrelated and it was stated that their maternal grandmothers had been born in the area considered for this study. The Slovenian samples were collected randomly over the whole country. The sampled Bosnian individuals were of Serbian and Croatian origin. They speak a Bosnian dialect of the Serbo-Croatian language of the Slavonic linguistic group, to which the

Slovenian language also belongs. Nevertheless, it is accepted by linguists that the Slovenian language distinguishes itself from Southern Slavonic languages, to which the Serbo-Croatian language belongs, and may belong to Western Slavonic languages, thus preserving some proto-Slavonic linguistic characteristics and even a lexical relationship with Baltic languages (Šavli *et al.* 1996).

### mtDNA Analysis

Total genomic DNA was prepared from whole blood (Bosnian samples) and buccal swabs (Slovenian samples) by means of cell lysis in the presence of proteinase K and 1% SDS, followed by phenol/chloroform extractions. Hypervariable segments I and II (HVS I and II) of the mtDNA noncoding control region (CR) were amplified and sequenced as described elsewhere (Malyarchuk *et al.* 2002). The nucleotide sequences of HVS I from position 15999 to 16400 and HVS II from position 30 to 407 were determined and compared with the Cambridge reference sequence (CRS: Anderson *et al.* 1981; Andrews *et al.* 1999).

To determine the haplogroup status of the CR 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), Macaulay *et al.* (1999), and Finnilä *et al.* (2000). To screen the restriction site polymorphisms, 10–12  $\mu$ l of the respective unpurified PCR products were digested for two hours at 37°C with 5U of the appropriate endonucleases and electrophoresed in 8% polyacrylamide gels followed by ethidium bromide staining. The samples were typed for a restricted set of RFLPs (Table 1) that were diagnostic of all major West Eurasian and East Asian clusters, on the basis of the hierarchical mtDNA RFLP scheme (Macaulay *et al.* 1999; Torroni *et al.* 2001; Yao *et al.* 2002). Control region sequences, belonging to haplogroups pre-HV, N1b, L1b and Z, were identified on the basis of the HVS I and II motifs classification (Chen *et al.* 2000; Richards *et al.* 2000; Yao *et al.* 2002). Sequence classification into mtDNA subclusters was based on the nomenclatures of Richards *et al.* (1998; 2000) and Macaulay *et al.* (1999).

**Table 1** RFLP polymorphisms used to identify major Eurasian mtDNA haplogroups

Haplogroups	Characteristic restriction site(s)
West Eurasian:	
HV	– 14766 <i>MseI</i>
H	– 14766 <i>MseI</i> , – 7025 <i>AluI</i>
pre*V1	– 14766 <i>MseI</i> , – 15904 <i>MseI</i> , +4577 <i>NlaIII</i>
pre*V2	– 14766 <i>MseI</i> , +15904 <i>MseI</i> , +4577 <i>NlaIII</i>
V	– 14766 <i>MseI</i> , +15904 <i>MseI</i> , – 4577 <i>NlaIII</i>
U	+12308 <i>HinfI</i>
K	+10394 <i>DdeI</i> , +12308 <i>HinfI</i> , – 9052 <i>HaeII</i>
J	+10394 <i>DdeI</i> , – 13704 <i>BstNI</i>
T	+13366 <i>BamHI</i> , +15606 <i>AluI</i>
T1	+13366 <i>BamHI</i> , +15606 <i>AluI</i> , – 12629 <i>AvaII</i>
I	– 4529 <i>HaeII</i> , +8249 <i>AvaII</i> , +10032 <i>AluI</i> , +10394 <i>DdeI</i>
W	+8249 <i>AvaII</i> , – 8994 <i>HaeIII</i>
X	– 1715 <i>DdeI</i> , +14465 <i>AclI</i>
East Eurasian:	
M:	+10394 <i>DdeI</i> , +10397 <i>AluI</i>
C	+10394 <i>DdeI</i> , +10397 <i>AluI</i> , – 13259 <i>HincII</i> /+13262 <i>AluI</i>
D	+10394 <i>DdeI</i> , +10397 <i>AluI</i> , – 5176 <i>AluI</i>
E	+10394 <i>DdeI</i> , +10397 <i>AluI</i> , – 7598 <i>HhaI</i>
G	+10394 <i>DdeI</i> , +10397 <i>AluI</i> , +4830 <i>HaeII</i> / +4831 <i>HhaI</i>
A	+663 <i>HaeIII</i>
B	9-bp deletion
F	– 12406 <i>HpaI</i> / <i>HincII</i>

### Phylogenetic and Statistical Analysis

For phylogenetic analysis, all available published data on the HVS I –RFLP mtDNA variability in West Eurasian populations was used (Richards *et al.* 2000; Meinilä *et al.* 2001; Mogentale-Profizi *et al.* 2001; Forster *et al.* 2002). To classify the Slavonic mtDNA haplotypes, a phylogeographic approach, based on the phylogenetic analysis of the spatial distribution of mitochondrial haplotypes and haplogroups determined as a monophyletic clade, was performed (Richards *et al.* 1998).

The population genetic structure 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 1000 permutations. Intrapopulation diversities ( $h$ ) were calculated using the formulas (Nei & Tajima, 1981) as implemented in Arlequin 2.0.

For the CR sequence sharing analysis, HVS I and HVS II haplotypes of Bosnians and Slovenians, as well as other European populations, were compared.

Data from the following populations were used: 436 Poles and 201 Russians (Malyarchuk *et al.* 2002); 200 Southern Germans (Lutz *et al.* 1998); 101 Austrians (Parson *et al.* 1998); 150 Western Germans (Baasner *et al.* 1998; Baasner & Madea, 2000); 109 North-Western Germans (Pfeiffer *et al.* 1999); 192 Finns (Finnilä *et al.* 2001); 30 Bulgarians (Calafell *et al.* 1996); 83 Italians (Tagliabracci *et al.* 2001); 204 French-speaking individuals from France and Switzerland (Rousset & Mangin, 1998; Dimo-Simonin *et al.* 2000); and 241 Portuguese (Pereira *et al.* 2000).

### Results and Discussion

The analysis of HVS I and II variability in combination with RFLP typing of the coding region haplogroup-specific sites in a total sample of 248 Bosnian and Slovenian individuals allowed detection of 200 different mitochondrial haplotypes (Table 2). 106 different haplotypes were found among the 144 subjects from Bosnia-Herzegovina and 85 different haplotypes were identified in the sample of 104 Slovenians. This high mtDNA haplotype diversity ensures that only 10 shared HVS I and II haplotypes were found between the populations studied.

It has been found that the CR sequences are clustered, according to the Eurasian mtDNA classification (Macaulay *et al.* 1999; Richards *et al.* 2000; Finnilä *et al.* 2001; Yao *et al.* 2002), into 20 haplogroups and their subgroups: H, HV2, pre-V, pre-HV, U1, U2, U3, U4, U5, K, J\*, J1a, T\*, T1, I, N1b, X, W, M\*, and Z. In addition, a single African-specific sequence type, belonging to haplogroup L1b, was revealed in a Bosnian sample.

The haplogroup pre-HV, which is defined by HVS I motif 16126–16362 associated with +14766 *MseI* site and 73A variant in HVS II (Macaulay *et al.* 1999; Richards *et al.* 2000), was found in the Bosnian sample with a low frequency of 1.4%. A single haplotype belonging to haplogroup HV2, which is characterized by HVS I motif 16217 in combination with the site loss for 14766 *MseI* and 73G variant in HVS II, was revealed in Bosnians. Note that the latter haplogroup, designated previously as P, was described by Tambets *et al.* (2000). Both of these haplogroups appear to be typical for southern populations of West Eurasia,

**Table 2** mtDNA haplotypes and their distribution in Bosnians and Slovenians

HVS I (minus 16000)	HVS II	HG	BOS	SLO
CRS	CRS	H	1	
CRS	315.1	H		1
CRS	143 228 263 309.1 315.1	H	1	
CRS	146 195 263 309.1 315.1	H	1	
CRS	151 263 309.1 309.2 315.1	H	1	
CRS	152 263 315.1	H	2	
CRS	152 263 309.1 315.1	H	1	
CRS	263 309.1 315.1	H	5	1
CRS	263 309.1 309.2 315.1	H	2	2
CRS	263 315.1	H	4	2
CRS	56.1 263 309.1 309.2 315.1	H		3
CRS	73 263 309.1 315.1	H		1
CRS	73 263 315.1	H	1	
CRS	73 185 263 309.1 315.1	H		1
CRS	195 257 263 309.1 315.1	H		1
311	257 263 315.1	H	1	
92	263 315.1	H	1	
93 212 222	263 315.1	H	1	
93 335	237 263 315.1	H		1
124	195 263 309.1 315.1	H		1
129	263 309.1 309.2 315.1	H	2	
131 325	152 263 309.1 309.2 315.1	H		1
136	263 315.1	H	1	
142 325	146 263 309.1 309.2 315.1	H	1	
162	73 263 309.1 315.1	H		2
51 162 213 266	73 263 309.1 315.1	H		2
51 162 213 266	73 263 315.1	H		1
168	152 263 315.1	H	1	
184	44.1 55 57 146 152 263 309.1 309.2 315.1	H	1	
CRS	55 57 263 309.1 315.1	H	1	
CRS	150 153 263 315.1	H		1
189 193.1	146 153 204 263 309.1 315.1	H	1	
189	153 204 263 309.1 315.1	H	1	
189 193.1 390	153 204 263 309.1 309.2 315.1	H	1	
189 193.1	153 204 207 263 309.1 309.2 315.1	H	1	
189 193.1	152 228 263 309.1 315.1	H		1
183C 189 193.1 356	263 315.1	H	1	
189 356 362	263 315.1	H	1	
189 356 362	263 315.1	H	1	
185 189 356 362	263 315.1	H	1	
354	146 263 315.1	H	2	
193 354	263 309.1 315.1	H	1	
209	263 315.1	H		1
221	146 263 309.1 315.1	H	1	
93 221	150 263 309.1 315.1	H		1
223	263 315.1	H		1
256 311	263 309.1 315.1	H	1	
256 352	152 263 309.1 315.1	H	1	
261	263 309.1 315.1	H		1
261 291	73 200 263 309.1 315.1	H	1	
263	263 309.1 315.1	H		1
265C	152 263 308.1T 315.1	H		1
274	263 309.1 315.1	H		1
291 390	263 309.1 315.1	H	1	

**Table 2** (continued)

HVS I (minus 16000)	HVS II	HG	BOS	SLO
291 390	263 309.1 309.2 315.1	H	1	
293	262 263 309.1 315.1	H	1	
293 311	195 263 309.1 309.2 315.1	H	2	
278 293 311	195 198 263 315.1	H	1	
92 140 265 293 311	195 263 315.1	H		1
294	263 309.1 309.2 315.1	H	1	
304	263 309.1 315.1	H	2	2
304	263 315.1	H		2
304	189 263 315.1	H	1	
304	228 263 315.1	H		2
304	263 309.1 309.2 315.1	H	1	1
304	146 263 309.1 309.2 315.1	H	1	
304 368	263 309.1 309.2 315.1	H		1
129 304	263 315.1	H		1
166C 304	263 309.1 315.1	H	1	
172 304 311	263 315.1	H		1
261 304	93 263 315.1	H	1	
294 304	263 309.1 309.2 315.1	H	1	
294 304	146 263 315.1	H		1
311	146 263 309.1 315.1	H	2	
311	263 309.1 309.2 315.1	H	1	
311	93 249D 263 309.1 315.1	H	1	
311	263 309.1 315.1	H	3	
311	152 263 309.1 315.1	H	1	
311	93 152 195 263 309.1 315.1	H	1	
311	263 315.1	H	1	1
311	195 263 315.1	H		1
258 311	146 263 315.1	H		1
189 193.1 311	N/A	H		1
362	195 239 263 309.1 315.1	H		1
362	239 263 309.1 309.2 315.1	H		1
362	42.1 239 263 309.1 315.1	H		1
390A/G	263 309.1 309.2 315.1	H		1
298	72 263 309.1 309.2 315.1	V	1	1
298	72 263 309.1 315.1	V	3	
298	72 200 263 309.1 315.1	V	1	
153 298	72 93 263 309.1 309.2 315.1	V		1
169 298	72 263 309.1 315.1	V		2
209 298	72 188 204 207 263 309.1 315.1	V	1	
234 298	73 263 309.1 315.1	V	1	
298	195 263 309.1 309.2 315.1	pre*V1	1	
298	263 309.1 309.2 315.1	pre*V1		1
298 311	195 263 309.1 315.1	pre*V1		1
298 311	72 195 263 315.1	pre*V1		1
140 222 298	185 235 263 309.1 315.1	pre*V1	1	
126 168 266C/T 304 362	60.1 60.2 64 71D 263 309.1 309.2 315.1	pre-HV	1	
126 168 266 304 362	60.1 60.2 64 71D 263 309.1 309.2 315.1	pre-HV	1	
217 243 261	72 73 152 195 263 309.1 315.1	HV2	1	
69 126	73 146 185 189 228 263 295 309.1 315.1	J*	1	
69 126	73 146 185 188 222 228 263 295 315.1	J*		1
69 126	73 185 188 228 263 295 315.1	J*		2
69 126	73 185 210 228 263 295 309.1 315.1	J*		1
69 126	73 185 228 263 295 315.1	J*		1
69 126	73 185 263 295 315.1	J*		1

**Table 2** (continued)

HVS I (minus 16000)	HVS II	HG	BOS	SLO
69 126	73 152 228 263 295 309.1 315.1	J*		1
69 126	73 228 263 295 309.1 315.1	J*		1
69 126 172 189	73 185 228 263 295 315.1	J*	1	
69 126 220C	73 185 228 263 295 309.1 309.2 315.1	J*	1	
69 126 366	73 185 188 228 263 295 309.1 315.1	J*	1	
69 126 290 311	73 185 188 228 263 295 309.1 315.1	J*	1	
69 93 126	73 185 228 263 295 315.1	J*	1	
69 126 366	73 185 188 228 263 295 309.1 315.1	J*	1	
69 126 261	73 185 189 204 228 263 295 315.1	J*	1	
69 92 126 261	73 146 185 189 228 263 295 315.1	J*		1
69 126 304	73 185 263 295 315.1	J*		1
69 126 145 231 261	73 150 152 195 215 263 295 309.1 315.1 319	J1a	1	
69 126 145 222	73 195 263 315.1	J1a	1	
126 294 296	73 151 263 309.1 315.1	T*		1
126 294 296 324	73 263 315.1	T*	2	
126 257D 294 296 304	73 152 263 315.1	T*	1	1
126 294 296 304	73 263 309.1 315.1	T*	1	
126 294 296 304	73 146 151 263 315.1	T*		1
126 294 296 304	73 152 263 315.1	T*		1
126 270 294 296 304	73 263 315.1	T*		1
126 192 290 294	73 263 315.1	T*	1	
126 163 186 189 294	73 195 263 309.1 315.1	T1	1	
126 163 186 189 292 294	73 195 263 309.1 315.1	T1	1	
126 163 186 189 207 261 270 294	73 152 263 315.1 384	T1		1
224 311	73 94 263 315.1	K	1	
224 311	73 263 315.1	K	1	
224 311	73 146 195 198 263 309.1 315.1	K	2	
224 311	73 263 309.1 315.1	K	1	
224 311	73 263 264 315.1	K		1
93 224 311	73 114 263 315.1	K		1
224 311 320	73 146 152 263 315.1	K		1
224 261 311	73 150 263 279 315.1	K		1
93 224 265A/C 311 362	73 263 309.1 315.1	K	1	
183C 189 249 304	73 263 285 315.1 385	U1	1	
126 182C 183C 189 249 353 360 362	73 183 263 285 309.1 309.2 315.1	U1	1	
51 129C 182C 183C 189 234 260 356 362	73 152 217 263 309D 315.1 394	U2		1
343	73 150 263 315.1	U3	1	
343	73 150 195 263 315.1	U3		1
343 390	73 150 263 315.1	U3		1
356	73 263 309.1 315.1	U4	1	
356	73 195 263 309.1 315.1	U4		1
356	73 195 263 315.1	U4	1	
356	73 195 217 263 309.1 309.2 315.1 316	U4	2	
356	73 146 152 185 195 263 315.1	U4	1	
356 362	73 195 263 315.1	U4	1	
145 356 362	73 195 263 309.1 315.1	U4		1
134 356	73 152 194 195 263 315.1	U4	1	
179 356	73 195 263 309.1 315.1	U4	1	
179 356	73 195 263 315.1	U4		1
311 356	73 146 152 195 263 309.1 315.1	U4		1
356 362	73 195 263 310	U4		2
192 256 270	73 263 315.1	U5a	2	1
192 256 270 399	73 263 315.1	U5a	1	
192 256 270 399	73 263 309.1 315.1	U5a	1	

**Table 2** (*continued*)

HVS I (minus 16000)	HVS II	HG	BOS	SLO
192 256 270 399	73 146 263 315.1	U5a	2	
192 256 270 291 399	73 263 315.1	U5a	1	
192 256 270 291 311 355 399	73 263 309.1 309.2 315.1	U5a		1
192 222 256 270 399	73 263 309.1 315.1	U5a	1	
192 249 256 270 357 399	73 263 315.1	U5a	1	
256 270 399	73 152 263 315.1	U5a	1	
189 192 234 242 256 270 311 362	73 151 152 263 309.1 315.1	U5a		1
256 270 286 320 399	73 183 263 309.1 315.1	U5a		1
114A 256 270 294	73 263 309.1 309.2 315.1	U5a		1
114A 192 256 270 294	73 263 309.1 315.1	U5a		3
189 270	73 150 152 263 309.1 309.2 315.1	U5b	1	
183C 189 270	73 150 152 263 315.1 315.2	U5b	1	1
186 189 270 293 294	73 150 263 309.1 315.1	U5b	1	
189 213 270 311	73 150 228 263 315.1	U5b	1	
93 189 193.1 270 301	73 150 207 263 315.1	U5b	1	
144 189 193.1 270	73 150 263 309.1 315.1	U5b	1	
144 189 270	73 150 263 309.1 315.1	U5b	1	
192 311	73 150 263 315.1	U5b		1
239 311	73 150 263 315.1	U5b		1
145 176G 244 390	73 146 263 309.1 315.1	N1b	1	
129 148 223	73 199 204 250 263 309.1 315.1	I	2	
129 148 223 391	73 199 250 263 309.1 315.1	I		1
129 223 391	73 199 204 250 263 315.1	I		1
93 129 223 391	73 199 204 207 250 263 315.1	I	1	
129 172 223 311 391	73 199 204 250 263 315.1	I	1	
292	73 119 189 195 204 207 263 315.1	W		1
223 292	73 189 194 195 204 207 263 309.1 315.1	W	1	
223 292	73 152 189 194 195 204 207 214 263 309.1 315.1	W		1
223 292 311	73 143 189 195 204 207 263 315.1	W	1	
223 292 295	73 119 189 195 204 207 263 315.1	W		1
192 223 292 325	73 152 189 194 195 204 207 263 309.1 315.1	W		1
172 223 231 292	73 189 194 195 199 204 207 263 315.1	W		1
189 193.1 223 278	73 153 195 225 226 263 309.1 315.1	X	1	
183C 189 193.1 223 278	73 152 153 195 263 295 315.1	X	1	
182C 183C 189 223 278	73 153 189 195 225 226 263 315.1	X		1
126 187 189 223 264 270 274 278 293 311 362 400	73 152 182 185T 189 195 247 263 315.1 357	L1b	1	
129 223 291 298	73 263 309.1 315.1	M*	1	
185 223 260 298	73 152 185 249D 263 309.1 315.1	Z	1	

Sample codes: BOS – Bosnians; SLO – Slovenians. Mutations are shown indicating positions relative to the CRS (Anderson *et al.* 1981). The nucleotide positions in HVS I and II sequences correspond to transitions; transversions are further specified. Haplogroup names (HG) are given in capital letters according to the mtDNA classification (Macaulay *et al.* 1999; Richards *et al.* 2000). Heteroplasmic nucleotides are indicated by a slash. The presence of insertions or deletions is indicated by .1, .2 and .3 or D, respectively, following the nucleotide position.

because pre-HV reaches its highest frequencies in the Near East, and HV2 extends from the West Mediterranean area to India (Tambets *et al.* 2000). It is noteworthy that the Bosnian HV2 sample is characterized by the nucleotide motif 16217–16243 found previously in Indians (Tambets *et al.* 2000).

Haplogroup pre-V sequences, defined by the CR motif 16298–72 (Torroni *et al.* 2001), were found in

Bosnians and Slovenians at relatively high frequencies (>6%). These samples were further identified as belonging to haplogroup V or paraphyletic group pre\*V on the basis of RFLP analysis (Table 1). Those of the pre\*V-samples defined by loss or gain of the 15904 *Mse*I site were designated as pre\*V1 and pre\*V2, respectively. The present study showed that Bosnians and Slovenians possess pre\*V1 cluster sequences at frequencies

**Table 3** Haplogroup distributions (no. of individuals and % values in parentheses) in Bosnians and Slovenians in comparison with Poles and Russians

Haplogroup	Bosnians (144)	Slovenians (104)	Poles (436)	Russians (201)
H	69 (47.92)	49 (47.12)	197 (45.18)	85 (42.29)
HV*	1 (0.69)	0	4 (0.92)	4 (1.99)
pre-V	9 (6.25)	7 (6.73)	21 (4.82)	11 (5.47)
pre-HV	2 (1.39)	0	0	1 (0.50)
J	10 (6.94)	10 (9.62)	34 (7.80)	16 (7.96)
T*	5 (3.47)	5 (4.81)	41 (9.40)	18 (8.96)
T1	2 (1.39)	1 (0.96)	9 (2.06)	4 (1.99)
K	6 (4.17)	4 (3.85)	15 (3.44)	6 (2.99)
U1	2 (1.39)	0	0	2 (1.00)
U2	0	1 (0.96)	4 (0.92)	3 (1.49)
U3	1 (0.69)	2 (1.92)	2 (0.46)	2 (1.00)
U4	8 (5.56)	6 (5.77)	22 (5.05)	7 (3.48)
U5a	10 (6.94)	8 (7.69)	23 (5.28)	15 (7.46)
U5b	7 (4.86)	3 (2.88)	15 (3.44)	6 (2.99)
U7	0	0	1 (0.23)	1 (0.50)
U8	0	0	2 (0.46)	0
U*	0	0	1 (0.23)	0
I	4 (2.78)	2 (1.92)	8 (1.83)	5 (2.49)
W	2 (1.39)	5 (4.81)	16 (3.67)	4 (1.99)
X	2 (1.39)	1 (0.96)	8 (1.83)	7 (3.48)
N1b	1 (0.69)	0	1 (0.23)	0
N1c	0	0	1 (0.23)	0
R*	0	0	2 (0.46)	1 (0.50)
L1b	1 (0.69)	0	0	0
L3	0	0	1 (0.23)	0
M	2 (1.39)	0	8 (1.83)	3 (1.49)
h ( $\pm$ s.e.)	0.75 $\pm$ 0.04	0.75 $\pm$ 0.04	0.77 $\pm$ 0.02	0.80 $\pm$ 0.03

Data for Poles and Russians are from Malyarchuk *et al.* (2002).

of 1.4% and 2.9%, respectively, comparable with those observed in the Mediterranean area (Torroni *et al.* 2001). Results of HVS II analysis showed that Bosnian and Slovenian V mtDNAs generally harboured the motif 72C-73A, with a single exception of 72T-73G found in the Bosnian sample. Meanwhile, the majority of pre\*V1 HVS II sequences (4 out of 5) had the combination 72T-73A.

Table 3 summarizes frequencies of the mtDNA haplogroups and subgroups found in Bosnians and Slovenians presented here, as well as in Polish and Russian samples studied previously (Malyarchuk *et al.* 2002). All of these samples are characterized by a similar pattern of mtDNA haplogroup distributions and comparison between populations did not reveal statistical differences ( $F_{ST} = -0.0018$ ,  $p = 0.899$ ). However, taking into consideration the neighbouring Southern European populations, such as Albanians and Greeks (designated, according to Richards *et al.* (2000), as East Mediterraneans [EMT]), Bulgarians and Romanians (as South East Europeans [SEE]), and populations of Italy

(designated as Central Mediterraneans [CMT]), it has been found that the populations compared produced a very low, but significant, differentiation level ( $F_{ST} = 0.0058$ ,  $p = 0$ ). AMOVA results suggest that the patterns of mtDNA diversity reflect some genetic differences between South European populations (Table 4). Significant differences (with a  $p$  value  $< 0.01$ ) were observed between populations of Italy and all Slavonic-speaking populations. In addition, significant  $F_{ST}$  values were revealed for Bosnians and East Mediterraneans. The Polish sample, representing the most northern population of Slavs, was found to be significantly different from East Mediterraneans as well as from South East Europeans.

The main mitochondrial haplogroup of the Bosnian and Slovenian sequences is H, which is also the most frequent haplogroup in Europe (Richards *et al.* 2000). This haplogroup comprises the majority of the Bosnian (48%) and Slovenian (47%) samples. It is known that haplogroup H, according to the complete mtDNA coding region sequence, is characterized by a considerable branching substructure and forms at least four large

	BOS	SLO	POL	RUS	SEE	EMT
SLO	-0.0057					
POL	-0.0002	-0.0028				
RUS	-0.0003	-0.0025	-0.0020			
SEE	0.0044	0.0035	0.0050*	0.0018		
EMT	0.0098*	0.0070	0.0087*	0.0022	0.0021	
CMT	0.0217*	0.0191*	0.0164*	0.0086*	0.0068*	0.0018

**Table 4** Matrix of  $F_{ST}$  values from mitochondrial haplogroup frequencies in European populations

Populations coded as BOS (Bosnians), SLO (Slovenians), POL (Poles), RUS (Russians), SEE (South East Europeans), EMT (East Mediterraneans), CMT (Central Mediterraneans). \* - significant differences ( $p < 0.01$ ).

subclusters (Finnilä *et al.* 2001; Herrnstadt *et al.* 2002). Moreover, based on mtDNA CR sequence data in Europeans, a great number of clearly defined nucleotide motifs was found (Richards *et al.* 2000; Tambets *et al.* 2000; Finnilä *et al.* 2001; Malyarchuk *et al.* 2002). Among them, several subclusters, such as 16362-239, 16293-16311-195, 16162-73, 16189-16356, 16304-16311, and 16294-16304 were found to be characteristic for Germans, Poles and Russians, therefore indicating the common genetic substratum for these Central and Eastern European populations (Malyarchuk *et al.* 2002). Table 5 shows examples of HVS I subclusters found in Bosnians and Slovenians, in comparison with several European populations. These subclusters are represented by

mtDNA haplotypes, which differ by the fewest number of base substitutions.

Importantly, almost all of the HVS I subclusters observed in Bosnians and Slovenians, (both or separately) can be accounted for in many European populations. Bosnians and Slovenians share, albeit at low frequency, an H-subcluster determined by HVS I motif 16221, which is also characteristic for other South European populations, with the highest frequency of 2% found in Albanians and Greeks (Richards *et al.* 2000). However, Bosnians and Slovenians show considerable differences in frequencies of subclusters 16162, 16189-16356, 16354 and 16362. Subcluster 16362, relatively frequent in European populations, was not found in Bosnians

**Table 5** Frequency of the mtDNA subclusters (no. of individuals and % values in parentheses) found in Bosnians and Slovenians in comparison with other European populations

HVS I subclusters	HG	BOS 144	SLO 104	RUS 201	POL 436	SEE 233	EMT 149	WG 259	SG 301	FIN 403
16221	H	1 (0.7)	1 (1.0)	0	0	1 (0.4)	3 (2.0)	0	2 (0.7)	0
16294 16304	H	1 (0.7)	1 (1.0)	0	3 (0.7)	2 (0.9)	0	2 (0.7)	4 (1.3)	0
16189 16356	H	4 (2.8)	0	5 (2.5)	16 (3.7)	2 (0.9)	2 (1.3)	5 (1.9)	8 (2.7)	3 (0.7)
16278 16293 16311	H	1 (0.7)	0	4 (2.0)	3 (0.7)	0	0	0	1 (0.4)	11 (2.7)
16293 16311	H	2 (1.4)	0	0	4 (0.9)	0	2 (1.3)	2 (0.7)	3 (1.0)	0
16354	H	3 (2.1)	0	8 (4.0)	3 (0.7)	4 (1.7)	1 (0.7)	0	0	4 (1.0)
16263	H	0	1 (1.0)	0	0	0	0	5 (1.9)	2 (0.7)	0
16304 16311	H	0	1 (1.0)	2 (1.0)	1 (0.2)	0	0	0	3 (1.0)	0
16162	H	0	5 (4.8)	2 (1.0)	5 (1.1)	0	0	10 (3.9)	6 (2.0)	18 (4.5)
16092 16293 16311	H	0	1 (1.0)	1 (0.5)	11 (2.5)	4 (1.7)	4 (2.7)	0	1 (0.4)	2 (0.5)
16362	H	0	3 (2.9)	6 (3.0)	10 (2.3)	5 (2.1)	3 (2.0)	8 (3.1)	3 (1.0)	1 (0.3)
16069 16126 16261	J*	1 (0.7)	1 (1.0)	1 (0.5)	0	6 (2.6)	8 (5.4)	1 (0.4)	2 (0.7)	0
16069 16126 16366	J*	2 (1.4)	0	1 (0.5)	2 (0.5)	0	0	0	4 (1.3)	0
16114A 16192 16256 16270 16294	U5a	0	4 (3.8)	0	2 (0.5)	0	1 (0.7)	1 (0.4)	1 (0.3)	16 (4.0)
16144 16189 16270	U5b	2 (1.4)	0	3 (1.5)	2 (0.5)	0	0	0	0	35 (8.7)
16192 16311 <sup>a</sup>	U5b	0	1 (1.0)	0	0	1 (0.4)	0	0	0	10 (2.5)

Populations coded as BOS (Bosnians), SLO (Slovenians), POL (Poles), RUS (Russians), SEE (South East Europeans), EMT (East Mediterraneans), WG (West Germans), SG (South Germans and Austrians), and FIN (Finns).

<sup>a</sup>This sequence type, as well as Slovenian sequence 16239-16311, belongs to subcluster U5b being found in association with +4732RsaI variant characteristic for the 7768G/14182C-branch of the U5 cluster (Finnilä *et al.* 2001).

and vice versa – subcluster 16354, frequent in Central-Eastern European populations (especially in Russians), was not observed in Slovenians. Similarly, subcluster 16189–16356, which is present in all European populations analyzed, was not found in Slovenians. On the contrary, Slovenians have a high frequency (4.8%) of subcluster 16162, which is characteristic for Central and Eastern European populations (especially for Germans and Finns), but seems not to be typical for Southern Europeans, including Bosnians. However, the western neighbors of Slovenians, Veneto-speaking Italians from Barco and Posina, possess this H-subcluster at a high frequency of about 6% (Mogentale-Profizi *et al.* 2001). It should be noted, however, that Italian Veneti differ greatly from Slovenians by an increased frequency of the haplogroup TJ (35% on average; Mogentale-Profizi *et al.* 2001).

The H-subcluster, determined by motif 16293–16311 that is characterized by a pan-European distribution (Richards & Macaulay, 2000), was found both in Bosnians and Slovenians. However, its Eastern European branch 16278–16293–16311, frequent in Russians and Finns, occurred only in Bosnians. The member of another branch, 16092–16293–16311, was observed only in Slovenians – this sequence type (16092–16140–16265–16293–16311) is identical to those characteristic for Poles (Malyarchuk *et al.* 2002). As sequence type belonging to subgroup 16304–16311, which is present in Russians, Poles and South Germans, was found in the Slovenian sample, whereas another rare H-branch defined by motif 16294–16304 was found in both Bosnians and Slovenians.

It is well known that European populations contain a large number of closely related mtDNA lineages, and geographic patterns of mtDNA variations may exist at the level of individual lineages (Helgason *et al.* 2000; Richards *et al.* 2000). Therefore, analysis of the distribution of rare mtDNA lineages and their combinations in populations may be an informative tool for studying ethnic history. For example, H-haplotype 16223 revealed in the mtDNA pool of Slovenians is present in European populations being found in South Germans and Ukrainians (Lutz *et al.* 1998; Malyarchuk & Derenko, 2001). Rare H-haplotype 16263, found in Slovenians, has so far been observed only in Germans (Table 5) and in French-speaking populations (Rousselet

& Mangin, 1998; Dimo-Simonin *et al.* 2000). Besides the H-sequences, there are also several geographically informative J- and U5-sequence types. Sequence type 16069–16126–16366 and its derivatives occurred at noticeable frequencies in Bosnians and Germans as well as in Poles and Russians. Another J\*-subcluster, defined by HVS I motif 16069–16126–16261 (with J\*-HVS II motif 185–228), was found at low frequency both in Bosnians and Slovenians, but is present at surprisingly high frequency in Southern European populations, with a peak (>5%) in Greeks and Albanians (Richards *et al.* 2000). Similarly, the J\*-HVS I sequence type 16069–16126 was found in more than 7% of Slovenians, as well as Albanians and Greeks. This frequency is 10 times higher than the corresponding value in Bosnians and essentially exceeds frequencies observed in different European populations, according to the database of Richards *et al.* (2000). Interestingly, this 16069–16126 HVS I sequence type, revealed in 8 out of 104 Slovenians, appears to be very diverse, being found in association with 7 different HVS II sequences. Since, in Europe, haplogroup J appears to have arrived from the Near East in the Neolithic period (Richards *et al.* 1998, 2000), the high levels of the root type 16069–16126 found in Slovenians possibly may represent a traces of the Neolithic migrations from the Near East.

Distribution of other mtDNA lineages seems to reflect the historical contacts between populations of the Northwestern Balkans and Northern/Eastern Europe. The U5b1-lineage with motif 16144–16189–16270 occurs at a frequency of 1.4% in Bosnians, but it is well known that this U5-subcluster has a restricted Northern/Eastern European distribution, being found frequently (8%–52%) in Finns and Saami as well as rarely (0.5%–1.5%) in Slavonic (Russians and Poles) and Baltic (Lithuanians) populations (Sajantila *et al.* 1995; Orekhov *et al.* 1999; Meinilä *et al.* 2001; Kasperaviciute & Kucinskas, 2002; Malyarchuk *et al.* 2002). In addition, Bosnians are characterized by the presence of the Asian-specific haplogroup Z (0.7%), which was previously revealed in Europe in Saami, Finns and Russians (Sajantila *et al.* 1995; Delghandi *et al.* 1998; Orekhov *et al.* 1999; Malyarchuk & Derenko, 2001; Meinilä *et al.* 2001). In Slovenians, the U5a-lineage defined by the substitution 16114A was found at a relatively high frequency of 3.8%. To date, this lineage that has been

found with a similarly high frequency only in Finns (Meinilä *et al.* 2001). In addition, Slovenians are characterized by the presence of another lineage frequently occurring in Finns – the U5b-haplotype 16192-16311 (Table 5).

The mitochondrial gene pool of Bosnians has some peculiarities, clearly distinguishing this population from Slovenians. Among them, there is a high frequency of an H-sequence type that differs from the CRS-HVS I only at the position 16311. Although 11 Bosnians (7.6%) carry this HVS I sequence type (which is, nevertheless, characterized by 6 different HVS II sequences), it has, to date, been found with a similar frequency (6%) only in the Alpine region, according to the database of Richards *et al.* (2000). Another mtDNA haplotype frequent in Bosnians (2.8%) is a sequence type with CR motif 16189-153-204. The results of screening for this H-subgroup in published HVS I and II data sets from different European populations demonstrate that only one such haplotype has been found in the Austrian population (according to data of Parson *et al.* 1998). Meanwhile, H-16189 sequences are widespread in Europe, but almost all of them do not carry the specific nucleotide combination 153-204 in HVS II. For instance, 6.5% of Finns have HVS I sequence type 16093-16189 or simply 16189, but all of these sequences do not have any specific HVS II motifs (Finnilä *et al.* 2001; Meinilä *et al.* 2001). However, one cannot exclude the possibility that the “Bosnian-specific” H-16189 subcluster is present in the Eastern Mediterranean region, as well as in Southeastern Europe (among Bulgarians and Romanians), since these populations are characterized by 4% of H-16189 sequences according to the HVS I database of Richards *et al.* (2000). Therefore, additional studies on HVS I and II variability are required to clarify this question. In addition, in Bosnians CR sequences have been found (at a frequency of 1.4%) belonging to haplogroup H and having HVS II motif 055-057, which is very rare in Europe, being previously observed only in Icelanders (Helgason *et al.* 2000). There is also a considerable difference between subcluster U5b sequences present in Bosnians and Slovenians, since all Bosnian U5b members (4.9%) are defined by the 16189-16270 motif, whereas in Slovenians, two out of three U5b-sequences belong to the specific branch determined by a possible back mutation from

T to C at the U5-diagnostic position 16270 (Meinilä *et al.* 2001).

In summary, this study reveals that the genetic pool of Bosnians has more similarities with other Southern European populations – the presence of typically South European mtDNA haplogroups, such as pre-HV, HV2 and U1 (Richards *et al.* 2000), and even one African-specific L1b sequence, which is observed, nevertheless, at low frequencies in Southern Italy (Rickards *et al.* 2000). Meanwhile, other examples presented here show that Bosnians and Slovenians possess interesting combinations of mtDNA subclusters and separate CR sequence types, characteristic mostly of Central and Eastern Europeans. This suggests that the common genetic substratum observed in modern German, Slavonic and western Finno-Ugric populations also penetrates also South East European populations, reaching territory as far as the Western Balkans. Meanwhile, genetic evidence concerning observed differentiation of Bosnians and Slovenians at the level of mtDNA subclusters may suggest that different groups of the Slavs penetrated the Balkans during their move to the south. According to archaeological data, the Slavs invaded the Balkan Peninsula as early as the 2<sup>nd</sup> century A.D., but after the 5<sup>th</sup> century A.D. large Slavonic movements came from two directions into the Balkans: from the east over the Carpathian Mountains and also from the north across Pannonia into the Western Balkans (Sedov, 1979; Šavli *et al.* 1996). The observed similarity between Bosnian and Russian/Finnish mtDNAs (such as U5b1, Z, H-16354, H-16278-16293-16311), and Slovenian and German/Polish/Finnish mtDNAs (such as H-16162, H-16263, U5b-16192-16311, U5a-16114A) allows us to speculate that the differences between the Slovenian and Bosnian mtDNA pools may have been partially due to the two different Slavonic migration waves that reached the Balkans in the early Middle Ages from different regions of Europe. It is noteworthy that Slavonic-speaking populations (Poles, Russians, Belorussians, Ukrainians, Czechs, Slovaks) are characterized by a high frequency (30%–50%) of Y chromosome haplogroups, defined by the M17 marker (Malaspina *et al.* 2000; Rosser *et al.* 2000; Semino *et al.* 2000). This haplogroup, R1a according to the Y Chromosome Consortium (2002), is rare in Western and Southern Europe, including Greeks (8%–21%),

Albanians (10%–12%), Italians (2%–4%), but notably it is present in Slavonic-speaking populations of the Balkan Peninsula, being found at a relatively high frequency in Croats (29.3%), Slovenians (37.0%) and Macedonians (35.0%), with the possible exception of Bulgarians (12%) (Malaspina *et al.* 2000; Rosser *et al.* 2000; Semino *et al.* 2000). This high level of the haplogroup R1a in the majority of the Southern Slavs appears to be a good indicator of the genetic impact of the expanding Slavonic waves in past times. A more extended Y chromosome microsatellite analysis of R1a haplotypes in the Slavonic populations is currently underway, to clarify the complex pattern of population expansions in the Balkan area.

## Acknowledgements

We are very grateful to Ewa Lewandowska for her excellent technical assistance. The authors would like to thank two anonymous reviewers for the useful comments. This work was supported by the Russian Fund for Basic Research (grant 00-06-80448), and the grant from the Ludwik Rydygier Medical University in Bydgoszcz, Poland (BW66/02).

## References

- 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. & 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.
- 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**, 35–49.
- Chen, Y.-C., Olckers, A., Schurr, T. G., Kogelnik, A. M., Huoponen, K. & Wallace, D. (2000) mtDNA variation in the South African Kung and Khwe - and their genetic relationships to other African populations. *Am J Hum Genet* **66**, 1362–1383.
- Comas, D., Reynolds, R. & Sajantila, A. (1999) Analysis of mtDNA HVRII in several human populations using an immobilised SSO probe hybridisation assay. *Eur J Hum Genet* **7**, 459–468.
- Delghandi, M., Utsi, E. & Krauss, S. (1998) Saami mitochondrial DNA reveals deep maternal lineage clusters. *Hum Hered* **48**, 108–114.
- Dimo-Simonin, N., Grange, F., Taroni, F., Brandt-Casadevall, C. & Mangin, P. (2000) Forensic evaluation of mtDNA in a population from west Switzerland. *Int J Legal Med* **113**, 89–97.
- 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.
- Finnilä, S., Lehtonen, M. S. & Majamaa, K. (2001) Phylogenetic network for European mtDNA. *Am J Hum Genet* **68**, 1475–1484.
- Forster, P., Cali, F., Röhl, A., Metspalu, E., D'Anna, R. & Mirisola, M., *et al.* (2002) Continental and subcontinental distributions of mtDNA control region types. *Int J Legal Med* **116**, 99–108.
- Helgason, A., Sigurdardóttir, 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.
- Herrnstadt, C., Elson, J. L., Fahy, E., Preston, G., Turnbull, D. M. & Anderson, C., *et al.* (2002) Reduced-median-network analysis of complete mitochondrial DNA coding-region sequences for the major African, Asian, and European haplogroups. *Am J Hum Genet* **70**, 1152–1171.
- Kasperaviciute, D. & Kucinskas, V. (2002) Variability of the human mitochondrial DNA control region sequences in the Lithuanian population. *J Appl Genet* **43**, 255–260.
- Lahr, M. M., Foley, R. A. & Pinhasi, R. (2000) Expected regional patterns of Mesolithic-Neolithic human population admixture in Europe based on archaeological evidence. In: *Archaeogenetics: DNA and the population prehistory of Europe* (eds C. Renfrew & K. Boyle), pp. 81–88. Cambridge: McDonald Institute for Archaeological Research.
- 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.
- 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.
- Malaspina, P., Cruciani, F., Santolamazza, P., Torroni, A., Pangrazio, A. & Akar, N., *et al.* (2000) Patterns of male-specific inter-population divergence in Europe, West Asia and North Africa. *Ann Hum Genet* **64**, 395–412.

- 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., 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.
- Meinilä, M., Finnilä, S. & Majamaa, K. (2001) Evidence for mtDNA admixture between the Finns and the Saami. *Hum Hered* **52**, 160–170.
- Mogentale-Profizi, N., Chollet, L., Stepanovich, A., Dubut, V., Poggi, C. & Pradie, M. P., *et al.* (2001) Mitochondrial DNA sequence diversity in two groups of Italian Veneto speakers from Veneto. *Ann Hum Genet* **65**, 153–166.
- Nei, M. & Tajima, F. (1981) DNA polymorphism detectable by restriction endonucleases. *Genetics* **97**, 145–163.
- 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.
- Pereira, L., Prata, M. J. & Amorim, A. (2000) Diversity of mtDNA lineages in Portugal: not a genetic edge of European variation. *Ann Hum Genet* **64**, 491–506.
- Pfeiffer, H., Brinkmann, B., Hühne, J., Rolf, B., Morris, A. A. & Steighner, R., *et al.* (1999) Expanding the forensic German mitochondrial DNA control region database: genetic diversity as a function of sample size and microgeography. *Int J Legal Med* **112**, 291–298.
- Piggott, S. (1965) *Ancient Europe from the beginnings of agriculture to classical antiquity*. Edinburgh: Edinburgh University Press.
- Pinhasi, R., Foley, R. A. & Lahr, M. M. (2000) Spatial and temporal patterns in the Mesolithic–Neolithic archaeological record of Europe. In: *Archaeogenetics: DNA and the population prehistory of Europe* (eds. C. Renfrew & K. Boyle), pp. 45–56. Cambridge: McDonald Institute for Archaeological Research.
- Richards, M. & Macaulay, V. (2000) Genetic data and the colonization of Europe: genealogies and founders. In: *Archaeogenetics: DNA and the population prehistory of Europe* (eds. C. Renfrew & K. Boyle), pp. 139–151. Cambridge: McDonald Institute for Archaeological Research.
- 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., Sykes, B. & Guida, V., *et al.* (2000) Tracing European founder lineages in the Near Eastern mtDNA pool. *Am J Hum Genet* **67**, 1251–1276.
- Rickards, O., Martinez-Labarga, C., Casalotti, R., Castellana, G., Tunzi Sisto, A. M. & Mallegni, F. (2000) MtDNA variability in extinct and extant populations of Sicily and Southern Italy. In: *Archaeogenetics: DNA and the population prehistory of Europe* (eds. C. Renfrew & K. Boyle), pp. 175–183. Cambridge: McDonald Institute for Archaeological Research.
- Rosser, Z. H., Zerjal, T., Hurler, M. E., Adojaan, M., Alavantic D. & Amorim, A., *et al.* (2000) Y-chromosomal diversity in Europe is clinal and influenced primarily by geography, rather than by language. *Am J Hum Genet* **67**, 1526–1543.
- Rousselet, F. & Mangin, P. (1998) Mitochondrial DNA polymorphisms: a study of 50 French Caucasian individuals and application to forensic casework. *Int J Legal Med* **111**, 292–298.
- 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.
- Šavli, 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).
- Sedov, V. V. (1995) *The Slavs in the early Middle Ages*. Moscow: Archaeological Fund (in Russian).
- Semino, O., Passarino, G., Oefner, P. J., Lin, A. A., Arbuzova, S. & Beckman, L. E., *et al.* (2000) The genetic legacy of Paleolithic Homo sapiens sapiens in extant Europeans: a Y chromosome perspective. *Science* **290**, 1155–1159.
- Sterlinko, H., Pajnic, I. Z., Balazic, J. & Komel, P. (2001) Human Y-specific STR haplotypes in a Slovenian population sample. *Forensic Sci Int* **120**, 226–228.
- Tagliabracci, A., Turchi, C., Buscemi, L. & Sassaroli, C. (2001) Polymorphism of the mitochondrial DNA control region in Italians. *Int J Legal Med* **114**, 224–228.
- Tambets, K., Kivisild, T., Metspalu, E., Parik, J., Kaldma, K. & Laos, S., *et al.* (2000) The topology of the maternal lineages of the Anatolian and Trans-Caucasus populations and the peopling of Europe: some preliminary considerations. In: *Archaeogenetics: DNA and the population prehistory of Europe* (eds. C. Renfrew & K. Boyle), pp. 219–235. Cambridge: McDonald Institute for Archaeological Research.
- Tolk, H. V., Pericic, M., Barac, L., Martinovic Klaric, I., Janijevic, B. & Rudan, I., *et al.* (2000) MtDNA haplogroups in the populations of Croatian Adriatic Islands. *Coll Anthropol* **2**, 267–279.

- Torroni, A., Bandelt, H.-J., Macaulay, V., Richards, M., Cruciani, F. & Rengo, C., *et al.* (2001) A signal, from human mtDNA, of postglacial recolonization in Europe. *Am J Hum Genet* **69**, 844–852.
- 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., 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.
- Y Chromosome Consortium (2002) A nomenclature system for the tree of human Y-chromosomal binary haplogroups. *Genome Res* **12**, 339–348.
- Yao, Y.-G., Kong, Q.-P., Bandelt, H.-J., Kivisild, T. & Zhang, Y.-P. (2002) Phylogeographic differentiation of mitochondrial DNA in Han Chinese. *Am J Hum Genet* **70**, 635–651.
- Zaharova, B., Gilissen, A., Andonova-Baklova, S., Cassiman, J.-J., Decorte, R. & Kremensky, I. (2000) Network analysis of Y-chromosome compound haplotypes in three Bulgarian population groups. In: *Archaeogenetics: DNA and the population prehistory of Europe* (eds. C. Renfrew & K. Boyle), pp. 237–243. Cambridge: McDonald Institute for Archaeological Research.

Received: 10 October 2002

Accepted: 26 February 2003