Ann. Hum. Genet. (2002), **66**, 261–283 ' *University College London*
DOI: 10.1017}S0003480002001161 *Printed in the United Kingdom*

Mitochondrial DNA variability in Poles and Russians

B. A. MALYARCHUK", T. GRZYBOWSKI#, M. V. DERENKO", J. CZARNY#, M. WOZ;NIAK#
and D. MIS,CICKA-S;LIWKA#

" *Institute of Biological Problems of the North, Russian Academy of Sciences, Portovaya str. 18, 685000 Magadan, Russia*

The Ludwik Rydygier University School of Medical Sciences, Forensic Medicine Institute, ul. Sklodowskiej-Curie 9, 85-094 Bydgoszcz, Poland

(Received 11.2.02. Accepted 11.4.02)

summary

Mitochondrial DNA (mtDNA) sequence variation was examined in Poles (from the Pomerania-Kujawy region; $n = 436$) and Russians (from three di erent regions of the European part of Russia; $n = 201$), for which the two hypervariable segments (HVS I and HVS II) and haplogroup-specific coding region sites were analyzed. The use of mtDNA coding region RFLP analysis made it possible to distinguish parallel mutations that occurred at particular sites in the HVS I and II regions during mtDNA evolution. In total, parallel mutations were identified at 73 nucleotide sites in HVS I (17.8 %) and 31 sites in HVS II (7.

ern and Southern (Gavli *et al*. 1996). In the north, the Lusatian culture was succeeded by the Pomeranian culture, extending over the coastal region from the mouth of the Oder to the mouth of the Vistula. The Przeworsk group encompassed the southern parts of present-day Poland. In the 2nd and 3rd centuries A.D., this group spread northward into the swampy Pripet and united there with the Zarubincy culture. It has been suggested that out of this culture the Eastern Slavonic language group developed (Rybakov, 1981; Gavli *et al*. 1996). Archaeologists report that the Slavs invaded the Balkan peninsula as early as the 2nd century A.D., and since this settlement movement of the Southern Slavs gradually evolved (Sedov, 1979). All of these 'migration' hypotheses claim that modern Slavonic groups are the result of an admixture between pre-Slavonic European populations and Slavonic tribes, whose homeland was probably in central Europe (Sedov, 1979; Alekseeva & Alekseev, 1989). This theory also predicts that diverse modern Slavonic populations may have certain combinations of genetic markers derived from the gene pool of the assumed ancestral Proto-Slavonic population.

A high-resolution analysis of maternal mtDNA lineages appears to be a highly informative approach in the reconstruction of the past demographic events, when large enough samples are available (Helgason *et al*. 2000). MtDNA sequences can be used to create a detailed pattern of the spatially resolved distribution of maternal lineages in Slavonic populations, and to trace a number of shared maternal lineages unique for Slavonic groups, connecting them among themselves and to other neighbours such as present-day German and Finno-Ugric populations. However, the mtDNA data sets for Slavonic populations living in Southern, Central, and Eastern Europe are either incomplete or virtually non-existent for many regional groups of Slavs, especially for populations inhabiting the East European Plain. Population samples of Slavs have been analyzed in di erent ways: some covering only HVS I sequences, others also including coding-region RFLPs (Malyarchuk *et al*. 1995; Calafell *et al*. 1996; Orekhov *et al*. 1999; Richards *et al*. 2000; Tolk *et al*. 2000; Malyarchuk & Derenko, 2001). In addition, almost all of these mtDNA studies have not addressed specific questions about the origin and early dispersal of Slavs in Europe. To date, it is known that Slavonic populations sharing the same language group (such as Russians, Ukrainians, Bulgarians) display a large amount of interpopulation genetic variation (Malyarchuk & Derenko, 2001). Moreover, we have not found any specific combinations of unique mtDNA types that clearly distinguish Russians from Germans and the neighboring Eastern European populations.

To obtain a better characterization of Slavonic mtDNA variability, we present here mtDNA diversity data in Poles and Russians, based on the HVS I and HVS II sequences typed for the presence of major West Eurasian haplogroup-specific markers.

materials and methods

Population samples

A population sample of 436 Poles from the Pomerania-Kujawy region of the northern part of Poland was studied. In addition, three population samples of Russians from the European region of Russia were analysed: 62 unrelated individuals from the south (Stavropol region), 76 from the centre (Orel region) and 63 from the east (Saratov region).

MtDNA analysis

DNA samples from the blood of individuals studied were used for mtDNA amplification and sequencing. PCR amplification of the entire noncoding region was performed using the primers L15926 and H00580. The temperature profile for 30 cycles of amplification was 94 °C for 20 sec, 50 °C for 30 sec, and 72 °C for 2.5 min (Thermal Cycler 9700; Perkin Elmer, USA). The resulting amplification product was diluted 1000-fold and 4 μl aliquots were added to an array of second-round, nested PCR reactions (32 cycles) to generate DNA templates for sequencing. The primer sets L15997}M13(−21)H16401 and M13(−21)L15997}H16401 were used to generate

both strands of the hypervariable segment I (HVS I). Similarly, the primer sets L00029}M13-(−21)H00408 and H00408}M13(−21)L00029 were used for hypervariable segment II (HVS II). Both primer sequences, and nomenclature, were used according to Sullivan *et al.* (1992). Negative controls were prepared for both the DNA extraction and the amplification process.

timate the diversity of mtDNA haplotypes, the average number of transitions on the reconstructed phylogeny from ancestral type to each sample (ρ) was used, according to the methods of Forster *et al.* (1996).

For the CR sequence sharing analysis, HVS I and HVS II haplotypes of Poles and Russians, as well as other European populations, were compared. 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); 109 North-Western Germans (Pfeiffer *et al.* 1999); and 192 Finns (Finnilä *et al.* 2001*b*).

results and discussion

Sequence variability in Poles and Russians

In the present study, the nucleotide sequences of HVS I from position 15991 to 16400 and HVS II from position 20 to 420 have been determined in 436 Poles and 201 Russians. Comparison to the Cambridge reference sequence (Anderson *et al.* 1981) showed that 140 nucleotide sites were polymorphic in HVS I (34.

Table 1. *Parallel mutations detected in the mtDNA HVS I in Poles and Russians*

Position	Nucleotide change	Poles	*n*	Russians	*n*	Total
16051	A ! G	H, U2, U5	3	H, U2	2	3
16069	C ! T	J	1	K, J	2	2
16071	C ! T	W	1	R*	1	2
16086	T ! C	U*, I, X	3		0	3
16092	T ! C	H, J1b, D	3	H, K, J*	3	5
16093	T ! C	H, K, U4, U5, J1b, C, G	7	H, K, U2, U4, U5, J*, X	7	10
16126	T ! C	(J*, T*)	1	U7, (J*, T*), pre-HV, pre-V, X, D	6	6
16129	G ! A	H, U5, T*, T1, I, W, M*	7	H, U4, U1, I	4	9
16140	T ! C	H, U5, T*	3		0	3
16145	G ! A	J1b, J1a, N1b	3	U5, J*, J1b	3	5
16146	A ! G	U8	1	T*	1	2
16148	C ! T	H, M*	2	H, U7	2	3
16150	C ! T	pre-V	1	U4	1	2
16153	G ! A	H, pre-V	2	pre-V	1	2
16168	C ! T	H, U3	2		0	2
16169	C ! T	H, pre-V	2		0	2
16170	A ! G	H, T1	2		0	2
16172	T ! C	H, K, U5, J1b, T*, I	6	J1b, I	2	6
16179	C ! T	H, U4, U5, U8, R*	5		0	5
16186	C ! T	T1	1	J*, T1	2	2
16189	T ! C	H, K, U4, U2, U5, J*, J1a, T*, T1, pre-V, X, C, D	13	H, U1, U2, U5, J*, T1, I, W, X, D	10	16
16192	C ! T	H, K, U5, T*, W, M*	6	H, U5, J1b, W	4	7
16193	C ! T	H, J2	2	H	1	2
16213	G ! A		0	H, J*	2	2
16218	C ! T	H, pre-V	2		0	2
16221	C ! T	H, U4	2		0	2
16222	C ! T	H, U5, J*, J1b	4	U5, J1b, T*	3	5
16223	C ! T	U4, (I, N1b,N1c, W, X, L3, C, D, G, M*, E)	2	(I, W, X, D, G, M*)	1	2
16224	T ! C	K	1	K, U1	2	2
16227	A ! G	T*, R*, G	3	G	1	3
16231	T ! C	H, J1a	2	H	1	2
16234	C ! T	K, U4, U5, T*, C, G	6	HV*, G, M*	3	8
16239	C ! T	H, U*	2		0	2
16241	A ! G	J1a	1	X	1	2
16243	T ! C	H	1	T1	1	2
16249	T ! C	H, J*	2	H, U1	2	3
16256	C ! T	H, K, U2, U5	4	H, U2, U5	3	4
16260	C ! T	T	1	U5	1	2
16261	C ! T	H, J1a, J1b	3	H, J*, J1b	3	4
16265	A ! G	H, N1c	2	H	1	2
16266	C ! T	H, U5, X, R*, D	5		0	5
16270	C ! T	H, U5, pre-V	3	H, U5	2	3
16271	T ! C	H, T*	2	T*	1	2
16274	G ! A	H, K, X, R*	4	H	1	4
16278	C ! T	H, J2, X, R*, G, E	6	H, X, G	3	6
16286	C ! T	H, U5	2		0	2
16288	T ! C	U5, C	2	H, U1	2	4
16290	C ! T	H	1	H, J*	2	2
16291	C ! T	H, K, U5, pre-V, M*	5	H, U5	2	5
16292	C ! T	T*, W	2	U5, T*, W	3	3
16293	A ! G	H, K, I	3	H, T1	2	4
16294	C ! T	H, U4, U5, T	4	T, I	2	5
16295	C ! T	W	1	HV*, W	2	2
16298	T ! C	T*, pre-V, C, M*	4	pre-V	1	4
16300	A ! G	X	1	H, X, M*	3	3
16304	T ! C	H, T*	2	H, U5, T*, I	4	4
16309	A ! G	U7, G	2	U7	1	2

Table 1. (*Cont.*)

Position	Nucleotide change	Poles	*n*	Russians	*n*	Total
16311	T ! C	H, K, U*, U5, J*, T*, R*, HV*, pre-V, I, W	11	H, K, J*, J1a, HV*, I	6	12
16316	A !					

Table 2. *Parallel mutations detected in the mtDNA HVS II in Poles and Russians*

Position	Nucleotide change	Poles	*n*	Russians	*n*	Total
64	C ! T	H	1	pre-HV	1	2
73	G ! A	H, HV*, U5, I, N1c	5	H, HV*, pre-HV	3	6
93	A ! G	H, HV*, pre-V	3	H, pre-V	2	3
143	G ! A	H, U4, W	3		0	3
146	T ! C	H, K, U4, U5, J*, J1b, T*, W, X	9	H, K, U7, T*, HV*, R*	6	12
150	C ! T	H, K, U*, U3, U5, J1a, J2, T*, HV*, W, D	11	H, U1, U3, U5, J1a, R*	6	13
151	C ! T	H, K, pre-V, L3	4	H, U7, pre-HV	3	6
152	T ! C	H, K, U2, U3, U4, U5, U7, J*, J1a, J2, T*, T1, I, N1b, W, X, L3, C, G	19	H, K, U4, U2, U3, U5, U7, T*, T1, pre-HV, R*	11	21
153	A ! G	X	1	X, M*	2	2
182	C ! T	H, X	2		0	2
189	A ! G	K, pre-V, N1c, W	4	H, W	2	5
194	C ! T	pre-V, W, R*	3	W	1	3
195	T ! C	H, K, U4, U3, J1a, T*, T1, pre-V, N1c, W, X, R*, L3	13	H, U4, U1, U5, U7, J1a, T1, pre-V, W, X	10	16
198	C ! T	H, X	2	U4, U5, pre-HV	3	5
199	T ! C	H, T, I, W	4	T, I, X	3	5
200	A ! G	U5, HV*	2	H	1	3
204	T ! C	H, U4, U5, (I, N1c, W)	4	H, K, U5, J1a, (I, W), X	6	7
207	G ! A	H, T*, (I, N1c, W)	3	H, (I, W)	2	3
210	A ! G	J*, N1c	2		0	2
215	A ! G	H, K, U4, J1a	4	U4, W	2	5
228	G ! A	H, U2, U4, J*, pre-V	5	J*, pre-V	2	5
236	T ! C	H	1	U4	1	2
239	T ! C	H, I	2	H	1	2
240	A ! G	T	1	H	1	2
250	T ! C	K, I	2	I	1	2
263	G ! A		0	H, T*	2	2
279	T ! C	H, T*	2		0	2
295	C ! T	H, J, R*	3	J	1	3
310	T ! C	U4, T*, C, M*	4	H, U4	2	5
319	T ! C	H, J1a, T*	3		0	3
385	A ! G		0	U1, T1	2	2

Mutations are shown indicating positions relative to the HVS II sequence that di ers from the revised CRS at np

Table 4. *Haplogroup distributions* (*no. of individuals and % values in parentheses*) *in Poles and Russians*

Haplogroup	Poles (436)	Russians (201)
H	197 (45.18)	85 (42.29)
HV*	4 (0.92)	4 (1.99)
pre-V	21 (4.82)	11 (5.47)
pre-HV	0	1 (0.50)
J	34 (7.80)	16 (7.96)
T*	41 (9.40)	18 (8.96)
T1	9 (2.06)	4 (1.99)
K	15 (3.44)	6 (2.99)
U1	0	2 (1.00)
U2	4 (0.92)	3 (1.49)
U3	2 (0.46)	2 (1.00)
U4	22 (5.05)	7 (3.48)
U5	38 (8.72)	21 (10.45)
U7	1 (0.23)	1 (0.50)
U8	2 (0.46)	0
U*	1 (0.23)	0
I	8 (1.83)	5 (2.49)
W	16 (3.67)	4 (1.99)
X	8 (1.83)	7 (3.48)
N1b	1 (0.23)	0
N1c	1 (0.23)	0
R*	2 (0.46)	1 (0.50)
L3	1 (0.23)	0
M	8 (1.83)	3 (1.49)

allowed detection of 455 di erent mitochondrial haplotypes (see Appendix): 329 haplotypes among 436 Poles and 158 haplotypes among 201 Russians. This high resolution ensured that only 32 shared HVS I and II haplotypes were found between Poles and Russians.

+15904*Mse*I. In addition, the HVS I haplogroup pre-V sequences lacking 16298C variant were found in one Polish and one Russian individual. It may be noted that haplogroup pre-V frequencies in Poles and Russians correspond to those observed in other Western, Central and Northern European populations (Table 1 in Torroni *et al*. 2001).

Phylogenetic studies have shown that haplogroups J and T stem from a common node which is distinguished from the ancestral node R* by polymorphisms at nps 4216, 11251, 15452 and 16126 (Macaulay *et al*. 1999; Finnilä *et al*. 2001*b*). Both haplogroups are widely distributed in European populations as well as in the Polish and Russian samples presented here. Haplogroup T represents 11.

Table 5. *U4a sequence types distribution in di erent populations*

HVS I sequence	HVS II sequence	Coding region markers	Sample origin
CRS	73 195 263 310	4646 12308	Poles"
CRS	73 152 195 263 310	4646 12308	Russians"
CRS	66 73 195 263 310 315D	4646 12308	Poles"
CRS	73 195 263 310	ND	Austrians#
16129 16362	73 195 263 310	4646 12308	Russians"
16189	73 195 263 310	4646 12308	Poles"
16294	73 195 263 310	4646 12308	Poles"
16294	ND	12308	Nenets$
16263	73 195 263 310	4646 12308	Russians"
16356	73 143 195 263 310	4646 12308	Poles"
16356	73 195 263 310 315D	ND	Germans%
16223 16356	73 195 263 310	4646 12308	Poles", Finns&

Data from the following studies were analyzed: " Present study, # Parson *et al*. 1998, $ Saillard *et al*. 2000, % Baasner & Madea, 2000, & Finnila$ *et al*. 2001*a*. ND, not determined.

U5b. The distribution of the subgroup U5a and U5b frequencies in Poles and Russians is approximately equal, with the U5a subgroup prevailing over U5b – 5.3 % and 3.4 % in Poles, and 7.5 % and 3 % in Russians.

U4 (with CR motif 16356-195) is the next relatively frequent subgroup in the populations studied, being found at a frequency of 5 % in Poles and 3.5 % in Russians. Phylogeographic studies revealed that two major founder clusters characterize U4, determined by HVS I motifs 16356 and 16134-16356 (Richards *et al*.1998, 2000); it was also suggested that the latter subgroup appears to be specific for Central and Eastern European populations. In this study, 16134-16356 sequences with low frequencies of 1.4 % in Poles and 0.5 % in Russians were observed. Perhaps more importantly, among Poles and Russians 14 HVS I sequences which belong to haplogroup U (+12308*Hin*fI) have been identified, but they do not share any mutations with subgroup-specific polymorphisms within haplogroup U (Table 5). All of these sequence types, as well as some members of subgroup U4, are characterized by HVS II motif 73-310. These samples were tested for the presence of a U4-diagnostic site +4643*Rsa*I and it was found that all of them belong to the U4-subgroup. In accordance with the style of established mtDNA nomenclature (Richards *et al*. 1998; Macaulay *et al*. 1999) we designated U4-sequences with the 310C variant in HVS II as belonging to clade U4a. Analysis of the published HVS I and II data allowed us to reveal U4a sequence types, although at a low frequency, in populations of Finno-Ugric-speaking Finns and Nenets, and German-speaking populations of Austrians and Germans (Table 5). Nevertheless, the current data on population distribution of U4a sequences led us to assume that the majority of them are characteristic for Poles and Russians, where this U4-subcluster was found with a frequency of 2.3 % and 2.0 %, respectively.

The geographic picture of the U4a sequence distribution remains unclear, since many published population data on the HVS I and II variability appear to be insu cient to determine an exact phylogenetic status of the CR sequences (such as CRS-73, for instance) without the support of coding-region sites. This study has observed CRS-73 sequences belonging to haplogroups H and HV*. Therefore, additional detailed studies are required to elucidate the origin and diversification of the U4a subcluster in Europe. In addition, phylogenetic relationships between control region sequences belonging to the U4 subgroup remain ambiguous, and therefore, the branching order of these sequence types cannot be resolved (Figure 1). The median network demonstrates that two possible phylogenetic directions are possible. The first scenario suggests that mutation at np 310 appeared later than marker mutation at np 16356, and further diversification of the U4a occurred after back-mutation at np 16356. On the contrary, the second scenario suggests that mutation at np 310 outstripped change at np 16356, and hence, the U4a subcluster may be

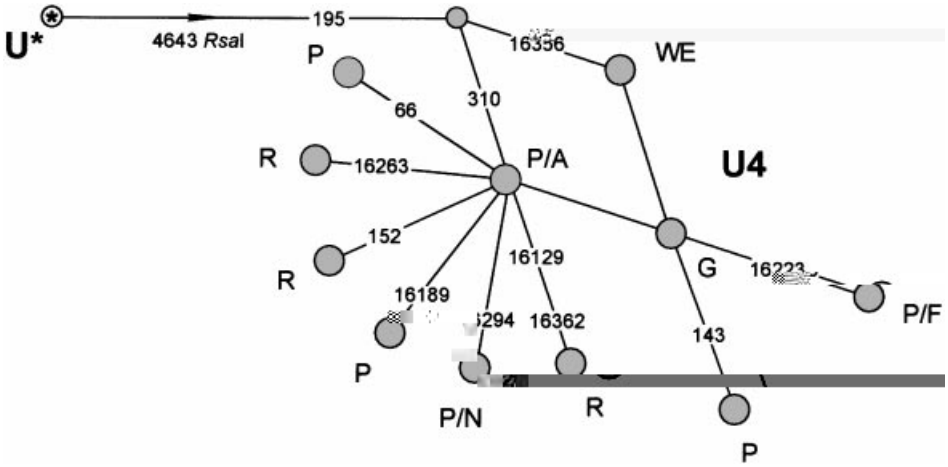

Fig. 1. Schematic phylogenetic network of the subhaplogroup U4 sequence types. The node U*, labelled by an asterisk (*), is defined by the RFLP variant +12308*Hin*fI and 73G variant in the HVS II, in comparison with the revised CRS (Andrews *et al.* 1999). The deletion event at np 315 was not considered. Any diversity within the node defined by 16356 variant alone is not shown. Reticulation in the network indicates ambiguity in the topology. RFLP variant is shown with the arrow pointing in the direction of a site gain. The nodes in the network represent the haplotypes found in populations (Table 5) as well as hypothetical intermediate haplotypes (empty nodes). Labelled nodes are U4a haplotypes observed in Poles (P), Russians (R), Germans (G), Austrians (A), Finns (F), Nenets (N), or in West Eurasians (WE).

considered as an ancestral state for the U4-phylogeny. Unfortunately, with the data on nucleotide stability in HVS I and II regions, we were unable to resolve this inconsistency due to an almost equal instability of nps 16356 and 310 (Tables 1 and 2). According to our data, the variant 16356C appeared twice and independently in the background of haplogroups U and H, but the variant 310C occurred independently in the background of haplogroup U, and rarely in association with H, T*, C, M* lineages. In addition, diversity estimates calculated for the two subsets of U4, with and without the 310C variant, gave similar values of $\rho = 0.929$ for U4a and $\rho = 1.083$ for the remaining U4-HVS I sequences found in Poles and Russians.

Besides subgroups U5 and U4, several minor U-subclusters were found in Polish and Russian mtDNA pools. Subgroup U1 with HVS I motif 16189-16249 (Macaulay *et al.* 1999), accompanied by variant 285T in HVS II, was present at a frequency of 1.0 % in Russians. Subgroup U2 sequences, characterized by HVS I and II motif 16051-16129C-152-217-340, were observed at low frequencies, both in Poles (0.9 %) and in Russians (1.5 %). U3-sequences with CR motif 16343-150 appear to be rare in Poles and Russians, being found at frequencies of 0.5 % and 1.0 %, respectively. Similarly, U7 sequences with CR motif 16309-16318T}C-152 were present in the populations studied at low frequency (less than 1.0 %). The U8-subgroup (Finnilä *et al.* 2001*b*), which is defined by motif 16342-282, was observed only in Poles at a frequency of 0.5 %.

The remaining haplogroups I, W, X, N1b, N1c observed in Poles and Russians belong to the macro-haplogroup N, which also encompasses all aforementioned clusters of haplogroups (HV, JT, UK) as members of the macro-haplogroup R (Macaulay *et al.* 1999; Richards *et al.* 2000). Haplogroup I, characterized by CR motif 16129-16223-16391-199-204-250, occurred in Poles and Russians at a frequency of 1.8 % and 2.5 %, respectively. N1b and N1c sequences are defined by tentative HVS I motifs 16145-16176G-16223 and 16223-16265, correspondingly (Richards *et al.* 2000), and were found as individual haplotypes in Poles. Haplogroup W sequences (CR motif 16223-16292-189-204-207) were observed in Poles and Russians at frequencies of 3.7 % and 2.0 %. Topology of the phylogenetic network of haplogroups I and W was resolved based on mtDNA variability in the coding region, with the exception of reticulation composed of polymorphisms at nps 1719 and 8251 (Finnilä *et al.* 2001*b*). According to the phylogeny suggested by

Finnilä *et al*. (2001*b*

Table 6. *The frequency of shared haplotypes found in Poles* (*POL*) *and in Russians* (*RUS*) *in comparison with Germans* (*GER*) *and Finns* (*FIN*)

HG	HVS I sequence	HVS II sequence	POL (436)	RUS (201)	GER (560)	FIN (192)
H	CRS	263	9.4	8.0	8.9	0
H	CRS	146 195 263	0.2	0.5	0.2	0
H	CRS	152 263	1.4	1.0	2.7	0
H	CRS	195 263	0.2	0.5	0.5	0
H	16093	263	0.5	2.5	0.2	0
H	16129	263	0.7	1.0	0.2	0
H	16274	146 263	0.2	0.5	0	0
H	16304	263	1.6	1.5	1.1	1.0
H	16051 16162 16259	73 263	0.2	0.5	0	0
H	16189 16356	263	0.5	0.5	0.4	0
H	16080 16189 16356	263	0.2	0.5	0	0
H	16189 16356 16362	263	0.7	0.5	0.4	0
H	16311	263	0.7	0.5	0.9	1.0
H	16278 16293 16311	195 263	0.2	2.0	0	0
H	16354	263	0.5	2.5	0	0
H	16362	239 263	1.4	2.0	0.5	0
HV*	CRS	73 263	0.5	0.5	?	0
K	16224 16311	73 146 152 263	0.7	1.5	0.5	2.6
U4	16093 16356	73 195 215 263	0.2	0.5	0	0
U5	16192 16256 16270	73 263	1.0	0.5	0.7	0
U5	16192 16256 16270 16399	73 152 263	0.2	0.5	0	0
U5	16192 16222 16256 16270 16399	73 263	0.2	0.5	0.2	0
U5	16256 16270 16399	73 263	0.2	1.5	0.5	0
U5	16256 16270 16399	73 152 263	0.2	1.0	0	0
J*	16069 16126	73 185 263 295	0.7	0.5	0.2	0
J*	16069 16126	73 185 228 263 295	0.2	0.5	1.1	0
J*	16069 16126 16311	73 185 263 295	0.2	0.5	0	0
J1b	16069 16126 16145 16172 16222 16261	73 242 263 295	0.7	1.5	0.2	0
T*	16126 16294 16296	73 263	0.5	1.0	1.3	0
T*	16126 16294 16296 16304	73 263	1.6	3.0	1.4	0
T1	16126 16163 16186 16189 16294	73 152 195 263	1.2	1.0	0.7	1.6
pre-V	16298	72 263	2.1	1.5	1.3	2.1
I	16129 16172 16223 16311 16391	73 199 203 204 250 263	0.2	0.5	0.4	3.1
X	16189 16223 16255 16278	73 153 195 225 227 263	0.2	0.5	0.2	0.5

HG denotes mitochondrial haplogroup. A question mark (?) denotes that haplogroup a liation of the CR sequence type cannot be determined without additional coding-region markers.

Table 7. *The frequency of shared HVS I subclusters found in Poles* (*POL*) *and in Russians* (*RUS*) *in comparison with Germans* (*GER*) *and Finns* (*FIN*)

HG	HVS I subclusters	POL (436)	RUS (201)	GER (560)	FIN (192)
H	16129, 16129-16316	0.7	2.5	0.7	0.5
H	16256, 16256-16352, 16256-16319	0.2	2.5	0.4	0
H	16291	0.2	0.5	0.4	0
H	16278-16293-16311	0.7	1.5	0.2	0
H	16092-16140-16265-16293-16311	2.8	0.5	0.2	0
H	16192-16304-16311	0.2	1.0	0.5	0
H	16189-16356	1.6	0.5	0.4	0
H	16080-16189-16356	0.5	1.0	0.2	0
H	16189-16356-16362	1.2	1.0	1.1	0
H	16354	0.7	4.5	0	1.0
pre-V	16153-16298	0.5	1.0	0.2	2.6
J*	16069-16126-16311	0.2	0.5	0	1.0
U2	16051-16129C-16189-16256	0.5	0.5	0.2	0
U4	16093-16356	0.2	0.5	0	0
U4a	CRS, 16356	2.3	2.0	0.4	0.5
U5	16192-16222-16256-16270-16399	0.2	1.0	0.2	0
U5	16192-16256-16270-16291-16399	0.9	0.5	0.5	0
W	16223-16292-16325	0.7	0.5	0	0

HG denotes mitochondrial haplogroup.

investigate genetic similarity between them at the level of shared mtDNA haplotypes and their subclusters, the total number of CR sequence types has been reduced by means of removing polymorphic variants at unstable sites (such as A ! C transversions at nps 16182 and 16183) and insertions of additional C residues in the HVS I and II poly-C tracts. As a result, out of 297 and 142 CR sequence types observed in Poles and Russians, respectively, 34 are shared between these population samples. Table 6 shows the distribution of shared haplotypes between Poles and Russians in comparison with Germans and Finns. The latter populations were selected in accordance with their geographic proximity and the historical evidence concerning their participation in the formation of modern Poles and Russians. The results of this analysis indicate that only a small fraction of the CR haplotypes (10 out of 34 haplotypes) appear to be actually shared between Poles and Russians, not being found in German and Finnish gene pools. These haplotypes belong to five di erent haplogroups – H, HV*, U4, U5, and J*. It should be noted, however, that the majority of these haplotypes belong to subclusters which can be found in common among many West Eurasian populations.

In order to identify subclusters of CR haplotypes which are specific mainly for Poles and Russians, the distribution of haplotypes that di er by the fewest number of base substitutions in Poles and Russians and their neighbours, Germans and Finns, were analyzed. Although almost all of the mtDNA subclusters observed in Poles and Russians can be accounted for in many European populations, this analysis allowed us to reveal at least 16 subclusters of relatively rare haplotypes which have a preferential distribution among Poles and Russians (Table

16189-16356 and its branches 16080-16189-16356 and 16189-16356-16362, was found frequently in Poles (3.3%) as well as in Russians (2.5%) and Germans (1.7%). A relatively high occurrence of H1-sequences determined by motif 16192-16304-16311 is characteristic for Russians in comparison to Poles and Germans, but another H1-branch (with motif 16294-16304) is clearly common between Germans and Poles.

Taking into account the data presented in Tables 6 and 7, one can conclude that we were not able to find any specific combinations of unique mtDNA haplotypes and their subclusters clearly distinguishing Poles and Russians, as Slavonic-speaking populations, from the neighboring European populations such as Germans and Finns. This trend was also noted in a previous study on the HVS I-RFLP variation in Russians in comparison with Western and Eastern European populations (Malyarchuk & Derenko, 2001). One possible exception is subgroup U4a. This subgroup comprises 10 (2.3%) out of 436 Poles, 4 (2.0%) out of 201 Russians, 2 (0.4%) out of 560 Germans (Parson *et al*. 1998; Baasner & Madea, 2000) and 1 (0.25%) out of 403 Finns (Finnilä *et al*. 2001*a*). Given the relatively high frequency and diversity of U4a among Poles and Russians and its low frequency in the neighbouring German and Finnish populations, one can suggest a central-eastern European origin of U4a. It is possible that the subsequent dispersal of this mtDNA subgroup in Eastern European populations was due to Slavonic migrations. Undoubtedly, to elucidate the origin of the U4a subgroup, additional analysis is required followed by a much more extensive sampling of Slavonic and other European populations.

conclusion

Analysis of mtDNA variation, performed by means of sequencing two hypervariable segments and assaying of haplogroup-diagnostic polymorphisms in the coding region, appears to be an effective genetic tool for inferring the genetic history of populations (Macaulay *et al*. 1999). In the present study, we have found that Poles and Russians are characterized by the same West Eurasian mtDNA haplogroups which describe at least 95% of mtDNA variations in Europe and the Near East (Torroni *et al*. 1996; Richards *et al*. 1998, 2000). Although there is a good correspondence between CR sequences and RFLPs grouping into monophyletic mtDNA clusters, and this looks to be strongest when only HVS I sequences are used in comparison (Bandelt *et al*. 2000); the addition of the HVS II sequences may be extremely useful in the resolution of the phylogenetic relationships among some uncertain mitochondrial lineages. A good example of this is subgroup U4a which does not have a certain HVS I motif, but can be recognized on the basis of HVS II information.

Despite the high level of mtDNA variation in Poles and Russians, both populations exhibit a similar pattern of mtDNA haplogroup distribution, which is also typical for many European populations studied. Moreover, the analysis of distribution of CR haplotypes and subclusters shared between Poles and Russians has shown that both Slavonic populations share them mainly with Germans and Finns. These data allow us to suggest that Europeans, despite their linguistic differences, originated in the common genetic substratum which predates the formation of the most modern European populations. It seems that considerable genetic similarity between European populations, which has been revealed by mtDNA variation studies, was

appendix

mtDNA haplotypes and their distribution in Polish and Russian populations

HVS I (minus 16000)	HVS II	HG	POL	RUS
CRS	263 309.1 315.1	H	12	8
CRS	263 315.1	H	22	5
CRS	263 309.1 309.2 315.1	H	7	3
CRS	263 309.1 315.1 337	H	1	
CRS	143 228 263 309.1 315.1	H	1	
CRS	146 263 309.1 315.1	H	2	
CRS	146 152 263 309.1 315.1	H		1
CRS	146 263 309.1 309.2 315.1	H	1	
CRS	146 195 263 309.1 315.1	H	1	1
CRS	150 263 315.1	H	1	
CRS	93 152 263 309.1 315.1	H	1	

appendix (*cont.*)

HVS I (minus 16000)	HVS II	HG	POL	RUS

appendix (*cont.*)

HVS I (minus 16000)	HVS II	HG	POL	RUS
298	72 263 315.1	pre-V		1
298	72C}T 195 263 309.1 315.1	pre-V		1
298	72 195 263 309.1 309.2 315.1	pre-V		1
298	72 263 309.1 315.1	pre-V	6	2
298	72 263 309.1 309.2 315.1	pre-V	3	
298	72 195 228 263 309.1 315.1	pre-V		1
104CA 298	72 151 228 263 309.1 315.1	pre-V	1	
126 298	72 263 309.1 315.1	pre-V		2
150 298	42.1 72 263 309.1 309.2 315.1	pre-V	1	
169 298	72 263 309.1 315.1	pre-V	1	
216 298	263 309.1 309.2 315.1	pre-V	2	
218 298	72 189 194 263 309.1 309.2 315.1	pre-V	1	
270 298	72 263 315.1	pre-V	1	
298 311	195 263 309.1 309.2 315.1	pre-V	1	
153 298	72 93 263 309.1 315.1	pre-V	1	
153 298	72 93 195 263 309.1 315.1	pre-V		1
153 298	72 93 195 263 315.1	pre-V		1
153 189 298	72 93 195 263 309.1 315.1	pre-V	1	
291 298	72 93 195 263 309.1 309.2 315.1	pre-V	1	
CRS	73 263 315.1	HV*		1
CRS	73 263 309.1 315.1	HV*	2	
067 355	150 200 263 309.1 315.1	HV*	1	

appendix (*cont.*)

HVS I (minus 16000)	HVS II	HG	POL	RUS
126 294 296 304	73 207 263 309.1 309.2 315.1	T*	1	
294 296 304	73 263 315.1	T*	2	
126 129 294 296 304	73 263 315.1	T*	1	

appendix (*cont.*)

HVS I (minus 16000)	HVS II	HG	POL	RUS
192 256 270 286 320 399	73 183 263 315.1	U5a	2	
076 192 256 270 399	73 263 309.1 315.1	U5a		1
145 189 192 256 270 399	73 195 263 309.1 315.1	U5a		1
256 270	73 263 309.1 315.1	U5a	1	
256 270	73 263 315.1	U5a	1	
256 270 399	73 263 315.1	U5a	1	
256 270 399	73 263 309.1 315.1	U5a		3
256 270 399	73 152 263 309.1 315.1	U5a	1	2
051 256 270 399	73 263 309.1 315.1	U5a	1	
256 270 362 399	73 152 204 263 309.1 315.1	U5a	1	
256 260 270 291 399	73 263 309.1 315.1	U5a		1
256 270 291 294 399	73 263 309.1 315.1	U5a	1	
114CA 192 256 270 294	73 263 309.1 315.1	U5a	1	
114CA 192 256 270 294	263 309.1 315.1	U5a	1	
189 270	73 150 263 315.1	U5b		1
189 270 291	73 150 152 263 315.1	U5b	1	
093 189 270	73 150 152 263 315.1	U5b	2	
093 189 270	73 150 152 263 270.1 315.1	U5b		1
093 189 193.1 270	73 150 263 309.1 315.1	U5b	1	
179 189 193 193.1 270	73 150 263 315.1	U5b	1	
183C 189 193.1 270 286	73 150 152 263 315.1	U5b	1	
093 182C 183C 189 193.1 270	73 150 152 263 315.1 315.2	U5b	1	
140 174 183C 189 193.1 270 288 311	73 150 263 309.1 315.1	U5b	1	
093 258 270 292 362	73 150 263 309.1 315.1	U5b		1
189 325	73 150 152 263 315.1	U5b	1	
189 217 234 270 398	73 150 263 315.1	U5b	1	
189 192 270 398	73 150 263 315.1	U5b	2	
189 270 398	73 150 263 315.1	U5b	1	
144 189 270	73 150 263 315.1	U5b		2
144 189 266 270	73 150 263 292 315.1	U5b	1	
144 183C 189 193.1 241AT 270	73 150 263 315.1	U5b	1	
144 189 193.1 270	73 150 152 243 263 315.1	U5b		1
309 318AT 362	60 73 152 263 315.1	U7	1	
073D 126 148 309 318AC	73 146 151 152 195 263 315.1	U7		1
146 342	73 263 282 309.1 315.1	U8	1	
179 342	73 263 282 309.1 315.1	U8	1	
179 187 227 245 266 274 278 362	73 194 195 246 315.1	R*	1	
071 355 357	73 81 146 150 152 263 283 309.1 315.1	R*		1
311	73 263 295 315.1	R*	1	
129 223 391	73 152 199 204 207 250 263 315.1	I	1	
129 223 391	73 152 199 204 207 239 250 263 309.			

appendix (*cont.*)

HVS I (minus 16000)	HVS II	HG	POL	RUS
092 102 164 182C 183C 189 193.1 223 266 362	42.1 73 150 263 309.1 309.2 315.1	D	1	
223 227 234 278 362	73 263 309.1 315.1	G		1
093 209 223 227 234 278 309 362	73 152 263 315.1	G	1	
129 148 192 223 291 298	73 263 310	M*	1	
223 234 300 316 362	73 153 263 315.1	M*		1
223 278 362	73 260 263 309.1 315.1	E	1	

Sample codes: POL, Poles; RUS, Russians. Mutations are shown indicating positions relative to the CRS (Anderson *et al*

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

Ingman, M., Kaessmann, H., Pääbo, S. & Gyllensten, U. (2000). Mitochondrial genome variation and the origin of modern humans. *Nature* **408**, 708–713.

Kivisild, T., Bandelt, H.-J., Wang, J., Derenko, M., Malyarchuk, B., Golubenko, M., *et al.* (2001). Mitochondrial DNA tree for Eastern Asian populations. In: *Abstracts of the First workshop on information technologies application to problems of biodiversity and dynamics of ecosystems in North Eurasia* (*WITA'2001*), p. 302. Novosibirsk: Institute of Cytology and Genetics.

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

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

Maca-Meyer, N., Gonzalez, A. M., Larruga, J. M., Flores, C. & Cabrera, V. M. (2001). Major genomic mitochondrial lineages delineate early human expansions. *BMC Genetics* **2**, 13.

Macaulay, V., Richards, M., Hickey, E., Vega, E., Cruciani, F., Guida, V., *et al.* (1999). The emerging tree

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

Torroni, A., Huoponen, K., Francalacci, P., Petrozzi, M., Morelli, L., Scozzari, R., *et al.* (1996). Classification of European mtDNAs from an analysis of three European populations. *Genetics* **144**, 1835–1850.

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

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

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