Acta Biochimica Polonica

Vol. 51 No. 4/2004

883–895

QUARTERLY

Comparison between the Polish population and European populations on the basis of mitochondrial morphs and haplogroups✪

Janusz Piechota[1], Katarzyna Tońska[1], Magda Nowak[1], Dagmara Kabzińska[1], Anna Lorenc[2] and Ewa Bartnik[1,3]✉

[1]*Department of Genetics, University of Warsaw, Warszawa, Poland;* [2]*Postgraduate School of Molecular Medicine, Warszawa, Poland;* [3]*Institute of Biochemistry and Biophysics, Warszawa, Poland*

Received: 14 September, 2004; revised: 02 November, 2004; accepted: 03 November, 2004

Key words: morph, haplogroup, haplotype, homoplasy, mtDNA evolution, Polish population

Polymorphisms in mitochondrial DNA (mtDNA) were analyzed in 152 samples from the Polish population using restriction enzymes *Ava*I, *Bam*HI, *Hae*II, *Hpa*I and *Pst*I. Additionally, each sample was classified into the appropriate haplogroup. When required, appropriate fragments were sequenced to establish the exact polymorphic sites. We found one new morph for *Pst*I and six new morphs for *Ava*II. Some detected morphs have previously been described as population specific morphs in different regions of the world. All polymorphisms were classified into 31 different haplotypes. 21 of them were detected in single individuals. The Polish population was compared with other populations from different regions. Moreover, we have obtained evidence for mutation hot spots in the mtDNA coding region. Our results indicate that *Ava*II morph and haplogroup composition of the Polish population is similar to other European populations and has a distribution typical for this part of the world. However, statistically significant differences in haplogroup composition were found between the Polish population and Italian and Finnish populations.

Analysis of mitochondrial DNA (mtDNA) variation has very often been used to analyze human evolution. In the 1980's and early 1990's restriction fragment length polymor-

✪This work was supported by: the State Committee for Scientific Research Grants PBZ-KBN-022/P05/199, 4P05E08220, and the Center of Excellence in Molecular Biotechnology ICA-CT-2000-70010.

✉Correspondence to: Ewa Bartnik, Department of Genetics, University of Warsaw, A. Pawińskiego 5a, 02-106 Warszawa, Poland; phone: (48 22) 592 2247; fax: (48 22) 592 2244; e-mail: ebartnik@ibb.waw.pl

Abbreviations: mtDNA, mitochondrial DNA; RFLPs, restriction fragment length polymorphism.

phisms (RFLPs) were frequently used for mtDNA analysis. This is based on the assumption that a mutation in mtDNA can cause the disappearance of a restriction enzyme site or the appearance of a new one. This is in turn linked to the changes in length of fragments obtained after mtDNA cleavage with a restriction enzyme. Restriction patterns were described as morphs characteristic for a given restriction enzyme. Thus, instead of comparing whole mtDNA sequences the comparison would be limited to morphs and haplotypes found in the population. The advantage of this method was its relative speed, simplicity and limitation of the processed data. However, this latter trait is currently the most serious disadvantage of this method. The use of this technique considerably restricts the length of the analyzed sequence — depending on the restriction enzyme used to about ten to several dozen nucleotides. Some RFLPs are still used in the classification of the mtDNA genomes into appropriate haplogroups. Haplogroups have been shown to somehow reflect the evolution of the mtDNA.

Currently RFLP analysis of mtDNA has been performed for many populations from various parts of the world. Also, haplogroups have been characterized for many populations. We decided to perform such investigations for the Polish population and to compare the results with those obtained for other populations.

MATERIALS AND METHODS

The analyzed sample comprised 152 persons. All of them were patients of the Children's Memorial Health Institute or the National Research Institute of Mother and Child in Warsaw. They were not suffering from mitochondrial diseases and originated from different parts of Poland. All of them consented to molecular analysis of their DNA.

DNA was extracted from PBL by standard methods (Sambrook & Russell, 2001). In most cases, the Southern technique was used for analysis, but regions with a high density of restriction sites were analyzed using PCR. RFLP analysis was performed for five enzymes: *Ava*I, *Bam*HI, *Hae*II, *Hpa*I and *Pst*I. The polymorphic sites found in RFLP analysis were sequenced in order to establish the type of mutation and its exact location. All samples were checked for affiliation to European haplogroups: H, I, J, K, T, V, X, W and U. In this case mtDNA fragments were ampli-

Table 1. Haplogroup classification

Haplogroup	Restriction site
H	7025(–) *Alu*I
J	13704(–) *Bst*NI, 16065(–) *Hin*fI
I	1715 D*de*I(–), 8249(+) *Ava*II, 10028(+) *Alu*I, 16389(+) *Bam*HI
K	9052(–) *Hae*II, 12308(+) *Hin*fI
T	13367(+) *Bam*HI, 15606(+) *Alu*I
U	12308(+) *Hin*fI
V	4577(–) *Nle*III
W	8249(+) *Ava*II, 8994(–) *Hae*III

Classification of the samples to the appropriate haplogroup was made on the basis of haplogroup-specific polymorphisms described by Torroni *et al.* (1996). Polymorphisms are coded as follows: the number indicates the position in the mtDNA, +/– in the brackets indicate appearance/disappearance of the restriction site and the name of the restriction enzyme.

fied by PCR and analyzed with the appropriate restriction enzyme. Analyzed restriction sites are listed in Table 1.

For PCR analysis of the samples and sequencing of the detected polymorphisms primers described by Torroni *et al.* (1992) and Taylor *et al.* (2001) were used.

DNA probes (p1 and p2) for Southern analysis were produced by random primer extension in the presence of [α-^{32}P]dATP. The probes covered almost all mtDNA. Template DNA was generated in a PCR reaction using Bio-X-Act polymerase (BIOLINE) and total human DNA as a template. The conditions of the reaction were according to the manufacturer's instructions. Two-step PCR was used. In the first step (10 cycles) the annealing temperature 58.5°C (p1) or 57.5°C (p2) was applied. In the second step (25 cycles) the annealing temperature was 59°C in both cases. The primers corresponded to mtDNA residues: p1F (16453–16472), p1R (9199–9180), p2F (8282–8301), p2R (16220–16193).

Confidence intervals for haplotype and morph frequencies in the population were calculated from the binomial distribution. χ^2 test was used to calculate the statistical significance of differences found in interpopulation comparisons.

RESULTS AND DISCUSSION

Morphs for *Ava*II restriction enzyme

Among the 152 mtDNA samples one new morph for *Pst*I and six new morphs for *Ava*II (Table 2) were found. All of them appear in the population at very low levels. The estimated frequencies are between 0.02 and 3.6 percent (for confidence interval = 0.95). *Ava*II-39 is similar to A*va*II-34$^{\text{Italy}}$ morph described as a population specific morph in North Italy (Brega *et al.*, 1994). *Ava*II-37 and *Ava*II-38 have not been found in other populations probably because they cannot be de-

Table 2. New morphs for *Ava*II and *Pst*I enzymes

Enzyme	Morph	Description	Polymorphism
*Ava*II	35	8249(+)	G8251A
		13367(–)	G13368A
	36	16390(–)	G16390A
		13367(–)	G13368A
		12629(–)	C12633A
	37	545(+)	A546G
	38	2646(+)	A2649G
	39	13367(–)	G13368A
		16569(+)	T3C
	40	16390(–)	G16391A
		13367(–)	G13368A
		4249(+)	G8251A
*Pst*I	4	9020(–)	G9025A

Six new morphs for the *Ava*II and one new morph for the *Pst*I restriction enzyme were found in the Polish population. Some polymorphisms for *Ava*II (8249(+), 13367(–), 16390(–)) were previously described, but they have never been found in such combinations with other polymorphisms.

tected by Southern blot analysis, which often is the only technique used in RFLP analysis. Because of its inaccuracy the Southern blot analysis precludes detecting morphs which differ slightly from previously described morphs. Restriction sites for this enzyme are very polymorphic, so large number of various morphs was described for *Ava*II restriction enzyme (Table 3). *Ava*II restriction enzyme has commonly been used for analysis of different populations, so that it was possible to compare results obtained for the Polish population with frequencies of *Ava*II morph obtained for populations from quite different parts of the world (Table 4).

Morph composition of the Polish population is very similar to other European populations and differs significantly from Asian and African populations. *Ava*II morph composition and its level are comparable between the Polish population and other European populations, e.g. Caucasians, from Finland, Czechoslovakia, and different parts of Italy. A high level of the *Ava*II-1 morph (70–85%) is typical for this region of the world. *Ava*II-9 (about 10%), *Ava*II-5, *Ava*II-3, *Ava*II-2 (about 3–10% of each) also have relatively high representation. These morphs are also observed in the Polish population.

*Ava*II-6 morph can be found in various parts of the world, but its level is very low. It has been found in African, European and Oriental populations. The *Ava*II-8 morph has been found for Caucasians (Johnson *et al.*, 1983). On the contrary, *Ava*II-23^{Wolof} was previously described for the Wolof tribe living in the Senegal region of the West Africa (Scozarri *et al.*, 1988).

Asian populations seem to be more homogenous, with *Ava*II-1 being the most frequent morph. *Ava*II-3, *Ava*II-5, *Ava*II-9 are less frequent than in the European populations. Instead of *Ava*II-13 morph typical for Europe, *Ava*II-12 morph is observed. Moreover, some specific morphs (e.g. 25^{Aeta}, 26^{Vedda}, 30^{Hindu}) are observed at a relatively high

Table 3. Description of *Ava*II morphs found in various populations

Morph	Description
1	Standard
2	8249(+)
3/3$^{\bullet}$	16390(–)
4	~8275(+)/~15890(+)/16390(–)
5	8249(+)/16390(–)
6	15829(+)
7/28	13367(–)/16390(–)
8	~4776(+)
9	13367(–)
10	3876(+)
11	~8229(+)/~15870(+)
12	12629(–)
13	12629(–)/13367(–)
13^{Jap}	16503(+)
14	8249(+)/~12130(+)/16390(–)
15	4311(+) or 4336(+)
17	~5229(+)
18^{Hindu}	15487(+)
19^{Italy*}	2621(+)/8249(+)/16390(–)
20^{Italy*}	4332(+)/15829(+)
21^{Sard*}	4332(+)/8249(+)
22^{Sard*}	4332(+)/8249(+)/12629(+)/13367(–)
23^{Wolof}	8342(+)
24^{Cal}	6460(+)
24^{Korean}	~14900(+)
25	8393(+)
25^{Aeta}	~6000(+)
26^{Vedda}	~6700(+)
29^{Hindu}	5260(+)/15882(+)
30^{Hindu}	14258(+)
34^{Italy}	13367(–)/8(+)

*Restriction sites were reconstructed on the basis of hybridization data in Brega *et al.* (1986) and detected polymorphisms published in the MITOMAP database (http://www. mitomap.org).

level of 5–10%. They are absent in other parts of the world.

In Africa morphs similar to European ones are observed, but the levels at which they occur are quite different. For example, in Senegal (West Africa) the most frequent morphs are *Ava*II-1 (55%) and *Ava*II-3 (41%). However, the *Ava*II-2 morph (66%) is the most frequent in South Africa, whereas *Ava*II-1 morph (22%) occurs at a three times lower level. Also for Bushmen the *Ava*II-2 morph (59%) is typical and more frequent than the *Ava*II-1 morph (12%). A relatively high level of the *Ava*II-5 morph (about 10%) is observed in Africa, but in Senegal this morph is totally absent. Also the *Ava*II-9 morph is almost absent in Africa and was found only in Senegal.

Haplogroup composition of Polish population

147 samples were classified into 8 European haplogroups (Table 5). Haplogroup H is the most frequent one and constitutes 38% of the sample, which is a typical European percentage. No samples with haplogroup X were

Table 4. Comparison of *Ava*II morphs found in the Polish population with other world populations

Population	Ref.	n	Morph											
			1	2	3	4	5	6	7	9	12	13	15	Other
Poland		152	113	5	6		3	2		15				8(1) 35(1) 36(1) 37(1) 38(1) 39(1) 40(1) 23^{Wolof}(1)
Finland	1	110	89	8	2		5			3		1	2	
Czechoslovakia	2	64	50	5			4	1		4				
Apulia (Italy)	3	87	53	5	5		8	2	2	8		3	1	
Calabaria (Italy)	4	60	45	3	3		1			6		1		23^{Cal}(1)
Roma (Italy)	5	95	70	3	1		2	1		13		1	2	19^{Italy}(1) 20^{Italy}(1)
Northern Italy	2	99	83	3			1			7		3	1	34^{Italy}(1)
Sardinia	5	136	116	2	1	1				10		4		21^{Sard}(1) 22^{Sard}(1)
Desulo (Sardinia)	6	85	57							28				
Galtelli (Sardinia)	6	83	80							3				
Tonara (Sardinia)	6	101	87							14				
Orosei (Sardinia)	6	108	101				1	1		3				25(2)
Israel (Ashkenazi)	7	21	14	2	4					1				
Israel (Sephardic)	7	18	14	1	2					1				
Israel (Yews)	8	38	29				5			1		1	1	
Israel (Arabs)	8	39	27		2		8				1	1		
Near East	9	46	44				2							
Caucasians	9	50	37	2	1		4		1	3				8(1) 10(1)
Hindu	10	79	64	1	2		2			2	1			
Japan	11	116	102	6	5						2			13^{Jap}(1)
Japan	12	74	66	1							1			10(1)
Ainu (Asia)	12	48	48											
Aeta (Asia)	12	37	33											25^{Aeta}(4)
Vedda (Asia)	12	20	13		1		1							26^{Vedda}(5)
Korea (Asia)	12	64	62	1										24^{Korean}(1)
Africa	9	74	20	25	16	2	9		1					11(1)
Southern Africa	13	139	30	92	3		11	1						
Senegal	14	186	102	2	76			2		1				17(2) 23^{Wolof}(1)
Bushmen	9	34	4	20	1	2	7							
Bantu	9	40	16	5	15		2		1					11(1)

The *Ava*II morphs composition of the Polish population shows typical European distribution, which is quite different from the Asian and African populations. References: [1] Vilkki *et al.* (1988); [2] Brega *et al.* (1994); [3] de Benedictis *et al.* (1989b); [4] de Benedictis *et al.* (1989a); [5] Brega *et al.* (1986); [6] Sartoris *et al.* (1988); [7] Tikochinski *et al.* (1991); [8] Bonne-Tamir *et al.* (1986); [9] Johnson *et al.* (1983); [10] Semino *et al.* (1991); [11] Horai & Matsunaga (1986); [12] Harihara *et al.* (1988); [13] Soodyall & Jenkins (1992); [14] Scozzari *et al.* (1988).

found. Five samples remained unclassified. Four of them do not posses any polymorphism characteristic for European haplogroups. In one sample an ambiguity between haplogroup U and H occurred.

Haplogroup composition of Polish population had been studied by Malyarchuk *et al.* (2002). Some differences are observed between the results obtained here and described by Malyarchuk *et al.* (2002) (Table 5). A lower level of the H haplogroup and a higher level of the U haplogroup were found in our sample. The observed differences may be a result of different origins of the samples. Our sample covers the whole region of Poland. Malyarchuk *et al.* (2002) studied Poles only from the Pomerania–Kujawy region of the northern part of Poland. However, comparison of these two sets of results shows that the observed differences are not statistically significant (χ^2 = 5.90, D.F. = 7, P = 0.55). Also, no statistical differences were found between the Polish population and the population of Russians from the European part of Russia and mixed population from Finland, Sweden and Tuscany region (χ^2 = 11.40, D.F. = 14, P = 0.65). These results are consistent with the conclusions of Malyarchuk *et al.* (2002).

However, statistically significant differences were found between our sample and the samples of Finnish populations studied by Finnilä *et al.* (2001) and Niemi *et al.* (2003) (χ^2 = 64.66, D.F. = 14, $P << 0.001$). Finnish populations seem to be very differentiated. Relatively high levels of the U and W haplogroups and a low level of the T haplogroup is characteristic for the Finnish population studied by Finnilä *et al.* (2001). The sample studied by Niemi *et al.* (2003) is characterized by very high level of the H haplogroup and relatively low levels of the T and V haplogroups. Differences between Polish and Italian populations also appeared to be statistically significant (χ^2 = 20.45, D.F. = 7, P = 0.005). This is the result of a low percentage of the V haplogroup and a high percentage of other than European haplogroups in the Italian population. Significant differences in the

Table 5. Characterization of haplogroups in the Polish population

Haplogroup	Poland	Poland	Russia	Finland+ Sweden+ Tuscany	Finland	Finland	Italy
H	57 (38%)	45%	42%	41%	39%	51%	39%
U	32 (21%)	16%	18%	16%	28%	21%	17%
T	18 (12%)	11%	11%	12%	3%	6%	11%
J	12 (8%)	8%	8%	11%	5%	6%	5%
V	11 (7%)	6%	8%	3%	6%	4%	2%
K	9 (6%)	3%	3%	7%	3%	6%	8%
W	3 (2%)	4%	2%	2%	10%	1%	4%
I	5 (3%)	2%	2%	2%	3%	3%	3%
Other	5 (3%)	5%	6%	7%	4%	3%	13%
n	152	436	201	134	480	400	275
Reference	1	2	2	3	4	5	6

Haplogroup composition was established for the Polish population and compared with results for the other populations and results obtained for the Polish population by Malyarchuk *et al.* (2002); n, number of individuals in the analyzed sample; references: [1] this work; [2] Malyarchuk *et al.* (2002); [3] Torroni *et al.* (1996); [4] Finnilä *et al.* (2001); [5] Niemi *et al.* (2003); de Benedictis *et al.* (1999).

Table 6. Haplotype classification of the detected polymorphisms

No.	Polymorphism	Haplogroup	Frequency in the sample
1		H	52
2	15829(+) *Ava*II	H	2
3	8342(+) *Ava*II	H	1
4	8249(+) *Ava*II (**I**,**W**), 9020(–) *Pst*I	H	1
5	[13367(–) *Ava*II, 13366(+) *Bam*HI(**T**)]	H	1
6		U	28
7	4529(–) *Hae*II (**I**)	U	1
8	~4776(+) *Ava*II	U	1
9	8249(+) *Ava*II (**I**,**W**)	U	1
10	2646(+) *Ava*II	U	1
11		J	11
12	16390(–) *Ava*II	J	1
13		V	6
14	16390(–) *Ava*II	V	1
15	4830(+) *Hae*II	V	1
16	16390(–) *Ava*II, 12406(–) *Hpa*I	V	1
17	16390(–) *Ava*II, 4830(+) *Hae*II, 12406(–) *Hpa*I	V	1
18	8994(–) *Hae*III (**W**), 16390(–) *Ava*II	V	1
19		T	15
20	8249(+) *Ava*II **(I**,**W)**	T	1
21	16390(–) *Ava*II, 12629(–) *Ava*II	T	1
22	16569(+) *Ava*II	T	1
23		K	9
24		W	2
25	7025(–) *Alu*I (**H**)	W	1
26		I	3
27	[13367(–) *Ava*II, 13366(+) *Bam*HI (**T**)]	I	1
28	4577(–) *Nla*III (**V**)	I	1
29		Unknown	3
30	545(+) *Ava*II	Unknown	1
31	7025(–) *Alu*I (**H**), 12308(+) *Hin*fI (**U**, **K**)	H or U	1

Polymorphisms and haplogroups found in the sample were classified into 31 different haplotypes. For simpler analysis haplogroup characteristic polymorphisms, if they were in standard configuration, were omitted. Bold letters in brackets after polymorphisms indicate the haplogroup, for which given polymorphisms is characteristic. Polymorphisms resulting from the same single mutation are in the square brackets. Parsimony rule was applied for samples with presence of polymorphisms from different haplogroups.

haplogroup composition of the Finnish population are observed between the results obtained by Finnilä *et al.* (2001) and Niemi *et al.* (2003). This phenomenon may be possibly explained by different selection criteria of the subjects in each analyzed sample. The other possible explanation is that the samples were not equally representative for the whole region of Finland.

Despite described differences, all analyzed populations possess typical European distribution of the studied haplogroups. A high level of the H haplogroup (about 40%) is characteristic for the European population (Torroni *et al.*, 1996). Levels of the U, T and J haplogroups are also relatively frequent (usually up to 20%). The frequencies of the V, K, W and I haplogroups are relatively low and usually do not exceed 10%. Four subjects unclassified into any of the European haplogroups may be immigrants or descendants of immigrants from outside of Europe. However, we do not have an access to the genealogy of these subjects, so we cannot confirm this hypothesis.

Diversity of mitochondrial genomes

All detected polymorphisms were classified into 31 different haplotypes (Table 6). The most frequent one occurred in 34% of the analyzed persons. 21 haplotypes were found in single individuals, thus in only 14% of the analyzed persons. This indicates that in the population most haplotypes occur with a very low frequency. The frequency in population of haplotypes found in single individuals was estimated to be 0.02–3.6%.

A high level of mtDNA diversity is observed in every general population. After investigating 112 persons Cann *et al.* (1984) found 149 polymorphic restriction sites but 69 of them were only found in single persons. Only 7 polymorphisms were found in more than 20% of the analyzed subjects. The real diversity of the mtDNA emerged after sequencing the whole mitochondrial genomes. Ingman *et al.* (2000) sequenced 53 entire mitochondrial genomes derived from various parts of the world. They found 657 polymorphic changes, but only 283 of these were present in more than one person. Similar data were presented for 192 persons from the Finnish population (Finnilä *et al.,* 2001). In the entire mtDNA 413 polymorphisms have been found, which were classified to 134 haplotypes. The most frequent one has been found in only 15 persons (8% of studied sample).

From the data presented here several conclusions can be drawn. First, using restriction enzymes to analyze the diversity of mitochondrial genomes, we lose a great deal of the information contained in the analyzed mtDNA sequences. Second, the question should be asked whether there is a sequence characteristic for a given population. It was shown by Finnilä *et al.* (2001), that an enormous diversity is already present in a single population. Can a sequence be considered to be characteristic for a population if even though it is indeed the most common sequence it is found in only 8% of that population? If so, when choosing one sequence at random from the population the probability is 92% to find an uncharacteristic one.

Parallel mutations in the mtDNA

A fast mutation rate is characteristic for mitochondrial DNA. The published estimations of mutation rates for the control region, from phylogenetic studies, range between 0.025–0.26 mutation events per site/1 million years (Myr) (Pearson *et al.*, 1997). The mutation rate for hypervariable regions of mtDNA, HVR1 and HVR2 is much higher and ranges between 0.32 and 2.1/site/Myr (Pearson *et al.*, 1997; Sigurðadóttir *et al.*, 2000). The mutation rate for the coding region is lower (0.02–0.04/site/Myr), but is 10 times higher in comparison with the mutation rate in the nucleus. Such a high mutation rate increases the risk of parallel identical mutations in different branches of the evolutionary tree. The

presence of the same mutation in different haplogroups is referred to a homoplasy state. The homoplasy creates serious problems with the analysis of human evolution. The problem is complicated even further by the hot spots of mutations – points in mitochondrial DNA in which the mutation events are very frequent and occur many times in the mitochondrial DNA.

Due to high rate of D loop mtDNA evolution this region is not suitable for analysis of human evolution (Ingman *et al.*, 2000). The evolution of the D-loop sequences is impossible to determine because of the high level of homoplasy. Mutation hot spots are another problem, as mutations and reversions are particularly frequent at these region (Paabo, 1996). For example Malyarchuk *et al.* (2002) identified 73 hypervariable sites in HVR I (17.8%) and 31 hypervariable sited in HVR II (7.73%) at which more than one independent mutation was observed.

So far the coding region was believed not to be affected by the problem of homoplasy. However, the presented data indicate that the situation may be different. Finnilä *et al.* (2001) have found 21 parallel mutations in the coding region. Herrnstadt *et al.* (2002) found 174 polymorphisms, which were present in more than one haplogroup.

In our studies we found 6 polymorphisms present in more than one haplogroup (Table 7). Five of them are polymorphisms that were classified as haplogroup specific polymorphisms. The presence of these highly polymorphic sites creates serious problem with the classification the samples into appropriate haplogroups. We found 11 samples possessing polymorphisms characteristic for two different haplogroups. Ten of them were classified into haplogroups using parsimony rules, but in one case (haplotype No. 31) we were not able to classify the sample into the proper haplogroup because of the simultaneous presence of polymorphism 7025(–) *Alu*I (haplogroup H) and 12308(+) *Hin*fI (U and K haplogroups). These two polymorphisms were the only ones which were used in classification into H and U haplogroups.

Two mutations, G16390A and G16391A, have been found in our sample, both causing the disappearance of the same restriction site for *Ava*II in position 16390. A second mutation also causes the appearance of a new restriction site for *Bam*HI in the position 16389 and is a polymorphism characteristic for haplogroup I. In the Polish population, we have found mutation G16390A in morphs *Ava*II-3 and *Ava*II-36 (haplotypes 12, 14, 18, 21) whereas mutation G16391A was found in morphs *Ava*II-5 and *Ava*II-40 (haplotypes 26, 28, 27). Moreover, mutation G16390A was found in three different haplogroups (J, V and T). Mutation G16391A appeared, as expected, only in haplogroup I.

In the polymorphic site 12629 for *Ava*II present in the *Ava*II-36 morph (haplotype No. 21) we found the mutation C12633A, but a mutation C12633T (Tanaka *et al.*, 1994) and mutations in positions 12630 (Finnilä *et al.*, 2001) and 12631 (Marzuki *et al.*, 1991) have

Table 7. Parallel mutations observed in the sample

Polymorphism	Haplogroup
8249(+) *Ava*II	**I/W**, H, U, T
[13367(–) *Ava*II, 13366(+) *Bam*HI]	**T**, H, I
4529(–) *Hae*II	**I**, U
16360(–) *Ava*II	J, V, T
7025(–) *Alu*I	**H**, W, *U?*
8994(–) *Hae*III	**W**, V
12308(+) *Hin*fI	**U/K**, *H?*

We found several polymorphisms which are the result of parallel mutations in different branches of mitochondrial evolutionary tree. Bold letters indicate haplogroup for which given polymorphism is the diagnostic (characteristic) one. *U?*, *H?* - the ambiguity of haplotype No. 31 (see Table 6). It is unknown if 7025(–) *Alu*I polymorphism appeared in haplogroup U or 12308(+) *Hin*fI polymorphism appeared in haplogroup H. Polymorphisms resulting from the same single mutation are in the square brackets.

also been described. Mutation G8251A causing appearance of the *Ava*II restriction site in the position 8249 had been described as a polymorphism specific for I and W haplogroups. However, we found this polymorphism also in H, U and T haplogroups. Mutation G13368A removing *Ava*II restriction site in the position 13367 creates new restriction site for *Bam*HI in the same position. *Bam*HI(+) 13667 polymorphism is specific for T haplogroup, but we were able to find this polymorphism in H and I haplogroups as well. Homoplasy in sites 8251 and 12630 has been already observed (Herrnstadt *et al.*, 2002). We conclude that sites 16390, 13368, 8251 and 12630–33 are mutation hot spots in mitochondrial DNA. Mutations G8251A and G13368A seem to be the most important as they create new restriction sites for *Ava*II and *Bam*HI, which are diagnostic polymorphisms for W, I and T haplogroups.

Morphs as markers of population isolation

Morph composition of a small closed population may be significantly different from the surrounding populations. Table 4 shows that small populations are more homogeneous. In these populations specific morphs, absent in surrounding areas, may be observed. For example, eight *Ava*II morphs were found in the whole Sardinia region (Brega *et al.*, 1986). The same research was performed in four small villages (2000–4500 inhabitants) (Sartoris *et al.*, 1988). In three of them only 2 morphs (*Ava*II-1 and *Ava*II-9) were detected. The *Ava*II-1 morph appears in Galtelli at the level of 96%, which is a significantly higher value in comparison with other European populations. The frequency of the *Ava*II-9 morph in Desulo is 33%, and this value is about three times higher than in any other population. Three additional morphs (*Ava*II-5, *Ava*II-6 and *Ava*II-25) are observed in Orosei, but their frequency in the population is very low. The same situation is observed in four Asian villages by Harihara *et al.* (1988). In one of them, Ainu, only the *Ava*II-1 morph was found. In the remaining three villages totally new morphs were detected, which seem to be specific to each village. Some of the morphs (e.g. 25^{Aeta}, 26^{Vedda}) appear at a relatively high level of 10%.

The described situation is probably the result of genetic drift, a phenomenon well known for small enclosed populations. Genetic drift may lead to elimination of rare mtDNA variants or, on the contrary, to their fixation in the population. As a result, morph composition of small enclosed populations may be significantly different from the overall population. The fact that large populations are usually divided into smaller subpopulations is a serious problem in population comparisons. The differences in analyzed populations may be due to nonrepresentative groups used for analysis (e.g. too small groups or taken from only part of the territory covered by the populations).

Population specific morphs

Many morphs have been described as specific for the population in which they were found. Often these morphs are later found in different populations. We have for instance found the *Ava*II-23^{Wolof} morph in the Polish population; it had originally been described for the Wolof tribe in Senegal (Scozarri *et al.*, 1988). The *Ava*II-39 morph was probably described earlier as *Ava*II-34^{Italy} morph (Brega *et al.*, 1994). In general, morphs specific for a population have been found in one or two individuals in the analyzed group. Thus in the population they may occur at a level lower than 1%. It is easy to calculate that the probability of finding these morphs in a random sample composed of 100 persons is lower than 63%. In general the number of samples does not exceed 100 (Table 4). In the Polish population morphs and haplotypes have been found which are more frequent than 2% (confidence level = 0.95). If we wished to increase

the sensitivity of the method to 1% we would have to increase the number of analyzed samples to 300. Because of the large variation of mtDNA the analyzed sequences would have to encompass thousands of individuals, in order to obtain an accurate genetic profile of a given population. Otherwise the differences between populations may be due to the too small number of analyzed samples in which only some of the mtDNA haplotypes occurring in the population have been detected.

CONCLUSIONS

Our results show that Polish population has morph and haplogroup composition and distribution typical for European populations. Our results are in agreement with the results obtained by Malyarchuk *et al.* (2002). We found that haplogroup distribution in the Polish population is significantly different in comparison with data published for Italian and Finnish populations. The Italian population has relatively high levels of non-European haplogroups. Samples taken to the analysis of the Finnish population exhibit different haplogroup distributions.

In each population, there is an enormous differentiation of mitochondrial sequences, thus the analyzed samples should be as large as possible in order to precisely define the genetic profile of the population. Many mitochondrial haplotypes occur in the population at a very low level. Thus, at least some of the morphs described as specific for a population occur over a wide geographic area. However, their frequencies in the analyzed populations are too low, thus they have not been found in every analyzed sample.

Numerous mutations have appeared many times in different mtDNA lines. This creates a serious problem for mtDNA evolution analysis, and thus for human evolution analysis. Homoplasy may have led to the construction of phylogenetic trees with an inappropriate topology. Parallel mutations occurring in the restriction sites characteristic for the haplogroups may lead to classification of the analyzed DNA sample to an incorrect haplogroup. To avoid this problem as many restriction sites as possible should be taken to the RFLP analysis.

REFERENCES

Anderson SA, Bankier T, Barrell BG, de Bruijn MH, Coulson AR, Drouin J, Eperon IC, Nierlich DP, Roe BA, Sanger F, Schreier PH, Smith AJ, Stadeni R, Young IG. (1981) Sequence and organization of the human mitochondrial genome. *Nature.*; **290**: 457–65.

Bonne-Tamir B, Johnson MJ, Natali A, Wallace DC, Cavalli-Sforza LL. (1986) Human mitochondrial DNA types in two Israeli populations – a comparative study at the DNA level. *Am J Hum Genet.*; **38**: 341–51.

Brega A, Mura G, Caccio S, Semino O, Brdicka R, Santachiara-Benerecetti AS. (1994) MtDNA polymorphisms in a sample of Czechoslovaks and in two groups of Italians. *Gene Geogr.*; **8**: 45–54.

Brega A, Scozzari R, Maccioni L *et al.* (1986) Mitochondrial DNA polymorphisms in Italy. I. Population data from Sardinia and Rome. *Ann Hum Genet.*; **50**: 327–38.

Cann RL, Brown WM, Wilson AC. (1984) Polymorphic sites and the mechanism of evolution in human mitochondrial DNA. *Genetics.*; **106**: 479–99.

de Benedictis G, Rose G, Caccio S, Picardi P, Quagliariello C. (1989a) Mitochondrial DNA polymorphism in Calabria (southern Italy). *Gene Geogr.*; **3**: 33–40.

de Benedictis G, Rose G, Passarino G, Quagliariello C. (1989b) Restriction fragment length polymorphism of human mitochondrial DNA in a sample population from Apulia (southern Italy). *Ann Hum Genet.*; **53**: 311–8.

de Benedictis G, Rose G, Carrieri G, De Luca M, Falcone E, Passarino G, Bonafe M, Monti D, Baggio G, Bertolini S, Mari D,

Mattace R, Franceschi C. (1999) Mitochondrial DNA inherited variants are associated with successful aging and longevity in humans. *FASEB J.*; **13**: 1532–6.

Finnilä S, Lehtonen MS, Majamaa K. (2001) Phylogenetic network for European mtDNA. *Am J Hum Genet.*; **68**: 1475–84.

Harihara S, Saitou N, Hirai M et al. (1988) Mitochondrial DNA polymorphism among five Asian populations. *Am J Hum Genet.*; **43**: 134–43.

Herrnstadt C, Elson JL, Fahy E *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**: 1152–71.

Horai S, Matsunaga E. (1986) Mitochondrial DNA polymorphism in Japanese. II. Analysis with restriction enzymes of four or five base pair recognition. *Hum Genet.*; **72**: 105–17.

Ingman M, Kaessmann H, Paabo S, Gyllensten U. (2000) Mitochondrial genome variation and the origin of modern humans. *Nature.*; **408**: 708–13.

Johnson MJ, Wallace DC, Ferris SD, Rattazzi MC, Cavalli-Sforza LL. (1983) Radiation of human mitochondria DNA types analyzed by restriction endonuclease cleavage patterns. *J Mol Evol.*; **19**: 255–71.

Malyarchuk BA, Grzybowski T, Derenko MV, Czarny J, Wozniaki M, Miscicka-Sliwka D. (2002) Mitochondrial DNA variability in Poles and Russians. *Ann Hum Genet.*; **66**: 261–83.

Marzuki S, Noer AS, Lertrit P, Thyagarajan D, Kapsa R, Utthanaphol P, Byrne E. (1991) Normal variants of human mitochondrial DNA and translation products: the building of a reference data base. *Hum Genet.*; **88**: 139–45.

Niemi AK, Hervonen A, Hurme M, Karhunen PJ, Jylhai M, Majamaa K. (2003) Mitochondrial DNA polymorphisms associated with longevity in a Finnish population. *Hum Genet.*; **112**: 29–33.

Paabo S. (1996) Mutational hot spots in the mitochondrial microcosm. *Am J Hum Genet.*; **59**: 493–6.

Parsons TJ, Muniec DS, Sullivan K, Woodyatt N, Alliston-Greiner R, Wilson MR, Berry DL, *et al.* (1997) A high observed substitution rate in the human mitochondrial DNA control region. *Nat Genet.*; **15**: 363–8.

Sambrook J, Russell DW. (2001) *Molecular Cloning: A Laboratory Manual.* (ed. 3) pp 6.4–6.11 Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York.

Sartoris S, Varetto O, Migone N, Cappello N, Piazza A, Ferrara GB, Ceppellini R. (1988) Mitochondrial DNA polymorphism in four Sardinian villages. *Ann Hum Genet.*; **52**: 327–40.

Scozzari R, Torroni A, Semino O, Sirugo G, Brega A, Santachiara-Benerecetti AS. (1988) Genetic studies on the Senegal population. I. Mitochondrial DNA polymorphisms. *Am J Hum Genet.*; **43**: 534–44.

Semino O, Torroni A, Scozzari R, Brega A, Santachiara Benerecetti AS. (1991) Mitochondrial DNA polymorphisms among Hindus: a comparison with the Tharus of Nepal. *Ann Hum Genet.*; **55**: 123–36.

Sigurðadóttir S, Agnar H, Gulcher JR, Stefansson K, Donelly P. (2000) The mutation rate in the human mtDNA control region. *Am J Hum Gen.*; **66**: 1599–609.

Soodyall H, Jenkins T. (1992) Mitochondrial DNA polymorphisms in Khoisan populations from southern Africa. *Ann Hum Genet.*; **56**: 315–24.

Tanaka M, Ozawa T. (1994) Strand asymmetry in human mitochondrial DNA mutations. *Genomics.*; **22**: 327–35.

Taylor RW, Taylor GA, Durham SE, Turnbull DM. (2001) The determination of complete human mitochondrial DNA sequences in single cells: implications for the study of somatic mitochondrial DNA point mutations. *Nucleic Acids Res.*; **29**: E74–4.

Tikochinski Y, Ritte U, Gross SR, Prager EM, Wilson AC. (1991) mtDNA polymorphism in

two communities of Jews. *Am J Hum Genet.*; **48**: 129–36.

Torroni A, Huoponen K, Francalacci P. (1996) Classification of European mtDNAs from an analysis of three European populations. *Genetics.*; **144**: 1835–50.

Torroni A, Schurr TG, Yang CC *et al.* (1992) Native American mitochondrial DNA analysis indicates that the Amerind and the Nadene populations were founded by two independent migrations. *Genetics.*; **130**: 153–62.

Torroni A, Huoponen K, Francalacci P, Petrozzi M, Morelli L, Scozzari R, Obinu D, Savontausi ML, Wallace DC. (1996) Classification of European mtDNAs from an analysis of three European populations. *Genetics.*; **144**: 1835–50.

Vilkki J, Savontaus ML, Nikoskelainen EK. (1988) Human mitochondrial DNA types in Finland. *Hum Genet.*; **80**: 31.