Int J Legal Med (2001) 114 : 169–172

ORIGINAL ARTICLE

H. Pfeiffer · P. Forster · C. Ortmann · B. Brinkmann

The results of an mtDNA study of 1200 inhabitants of a German village in comparison to other Caucasian databases and its relevance for forensic casework

Received: 29 January 2000 / Accepted: 24 April 2000

Abstract Mitochondrial DNA control region sequences were determined in 1200 male volunteers from one village area of Lower Saxony for the hypervariable region 1 (HV1). The 154 variable positions found resulted in 460 different haplotypes with a haplotype diversity value of 0.98165. The number of different haplotypes showed a nearly linear increase with the number of individuals typed. The haplotype diversity approached saturation level at a value of approximately 0.981 after typing 400 individuals. Furthermore, the number of different haplotypes and the haplotype diversity were calculated for four short amplicons of HV1 in order to establish the most variable section with a high efficiency for forensic casework.

Key words Mitochondrial DNA · Hypervariable region 1 · Forensic science

Introduction

The high polymorphism of the hypervariable region 1 (HV1) of the mitochondrial DNA D-loop and its use for forensic casework has been demonstrated by many authors (Piercy et al. 1993; Lee et al. 1997; Lutz et al. 1998, 1999; Parson et al. 1998; Rousselet and Mangin 1998; Pfeiffer et al. 1998, 1999). Due to high regional polymorphism the application of mtDNA analysis in forensic casework benefits from large local databases for estimating the probability of identity by chance (Allen et al. 1998). The first part of our study presents the collection of 1200 mtDNA HV1 sequences from male volunteers living in a village in Lower Saxony. The correlation between the number of different haplotypes and the population sample size was explored in order to demonstrate the high variability of the HV1 region even in a local context. The data set was compared to other existing databases from Caucasian population groups (Lutz et al. 1998, 1999; Parson et al. 1998; Pfeiffer et al. 1999).

For poor quality DNA, for instance from ancient bone material, the amplification of short mtDNA fragments is more promising than amplification of the whole hypervariable region since DNA undergoes damage and degradation influenced by environmental factors (Bär et al. 1988). The molecular genetic analyses of the Tyrolean Ice Man (Handt et al. 1994) and of the Neandertal-type specimen found in western Germany (Krings et al. 1997) demonstrated that mtDNA sequences can be retrieved from human remains after a long time interval since death. However, old skeletal material was demonstrated generally to allow amplification of DNA sequences no more than 200 base pairs (bp) long. For this reason, when working with ancient bone fragments, primers spanning short mtDNA segments are required for amplification. The short amplicons should be as informative as possible and should provide a high polymorphic content. The aim of the second part of our study was to calculate the number of different haplotypes and the haplotype diversity (Nei 1987) from the 1200 samples for four short amplicons of HV1 in order to find out the most informative segment for application in forensic casework when stain mtDNA is limited and assumed to be degraded.

Supplementary material: Data on the mtDNA sequences in HVI from 1200 male volunteers from one rural area in Lower Saxony, Germany (Table S1), are available in electronic form on Springer-Verlag's server at http://link.springer.de/link/service/journals/00414/index.htm.

H. Pfeiffer (✉) · C. Ortmann · B. Brinkmann
Institute of Legal Medicine, University of Münster,
Von Esmarch-Strasse 62, 48149 Münster, Germany
e-mail: pfeiffh@uni-muenster.de;
Fax: +49-251-8355158

P. Forster
McDonald Institute for Archaeological Research,
University of Cambridge, Downing Street,
Cambridge, CB2 3ER, UK

Materials and methods

DNA sampling and extraction

Saliva samples were collected from 1200 male volunteers between 16 and 55 years old living in one village area 20 km from Braun-

Table 1 Mean number of different haplotypes and mean haplotype diversity for random selected sub-populations from the database (n = 1200)

Population sample size	Number of random selections	Mean number of different haplotypes	Mean standard deviation	Mean haplotype diversity	Mean standard deviation
50	10	38	3	0.9619	0.0063
100	10	69	3	0.9718	0.0059
150	5	95	4	0.9762	0.0036
200	5	121	3	0.9766	0.0038
250	5	144	2	0.9777	0.0019
300	5	167	4	0.9785	0.0022
350	5	187	3	0.9792	0.0018

schweig, a city in Lower Saxony with 240,000 inhabitants and a modern industry structure. The village has approximately 5300 inhabitants, and many of the residents have moved out from the city. Approximately 5% of the typed volunteers were not German Caucasians but came, in most cases, from Turkey or Poland. The samples were collected in the course of a voluntary mass-screening after a homicide, where the only evidential material was a hair shaft that did not belong to the victim. From this hair shaft, for purposes of comparison, only data for the hypervariable region 1 of mtDNA (HV1) existed.

The samples were preselected neither for ethnicity nor for maternal unrelatedness, and therefore represent the current genetic composition of the village. DNA extraction was performed with Chelex 100 (Walsh et al. 1991).

MtDNA amplification and sequencing

The amplification and sequencing for the hypervariable region 1 (HV1) and the short fragments of HV1 were performed according to the instructions of the Armed Forces DNA Identification Laboratory (1413 Research Blvd., Rockville, MD USA 20850D3125) using the following primer pairs (Holland et al. 1995):

– F15971 (5′ TTA ACT CCA CCA TTA GCA CC)
– R16410 (5′ GAG GAT GGT GGT CAA GGG AC)

– F15971 (5′ TTA ACT CCA CCA TTA GCA CC)
– R16175 (5′ TGG ATT GGG TTT TTA TGT A)

– F15971 (5′ TTA ACT CCA CCA TTA GCA CC)
– R16251 (5′ GGA GTT GCA GTT GAT GT)

– F16144 (5′ TGA CCA CCT GTA GTA CAT AA)
– R16410 (5′ GAG GAT GGT GGT CAA GGG AC)

– F16190 (5′ CCC CAT GCT TAC AAG CAA GT)
– R16410 (5′ GAG GAT GGT GGT CAA GGG AC)

DNA sequencing was carried out on an ABI Prism 310 automated sequencer with BigDye Terminator sequencing reagents (ABI-Perkin Elmer, Weiterstadt, Germany). The analysis of the sequences and the calculation of the haplotype diversity was performed as described previously (Pfeiffer et al. 1999), and all sequences were checked by two investigators.

All polymorphisms compared with the Anderson sequence (Anderson 1981) were combined in one Microsoft Excel (Version 5.0) table. The sequencing reaction was carried out in both directions and only confirmed deviations from the reference sequence (Anderson 1981) were included in the database. In cases of length heteroplasmy in the poly-C strand, the polymorphisms before the C-stretch in the reverse reaction and behind the C-stretch in the forward sequencing reaction were confirmed by repeating amplification and sequencing reactions.

The number of different haplotypes and the haplotype diversity of the total HV1 was calculated for randomly selected sample groups between 50 and 1200 using the Microsoft Excel sorting function. With the same function, the number of different haplotypes and the haplotype diversity of the four short mtDNA amplicons were calculated for all 1200 samples. The number of haplotypes and the haplotype diversity (Nei 1987) for the sample size 50 and 100 is the mean value of 10 and for the sample size from 150 to 350 of 5 randomly selected population sub-samples (Table 1). The mean standard deviation was calculated for all these groups, and the values varied between 2 and 4 for the number of different haplotypes and from 0.0018 to 0.0063 for the haplotype diversity (Table 1).

Results and discussion

A total of 1200 males from one village area of Lower Saxony were typed in HV1 for the nucleotide positions (nps) 16024–16365 according to the reference sequence (Anderson et al. 1981) (Table S1). The 154 variable positions

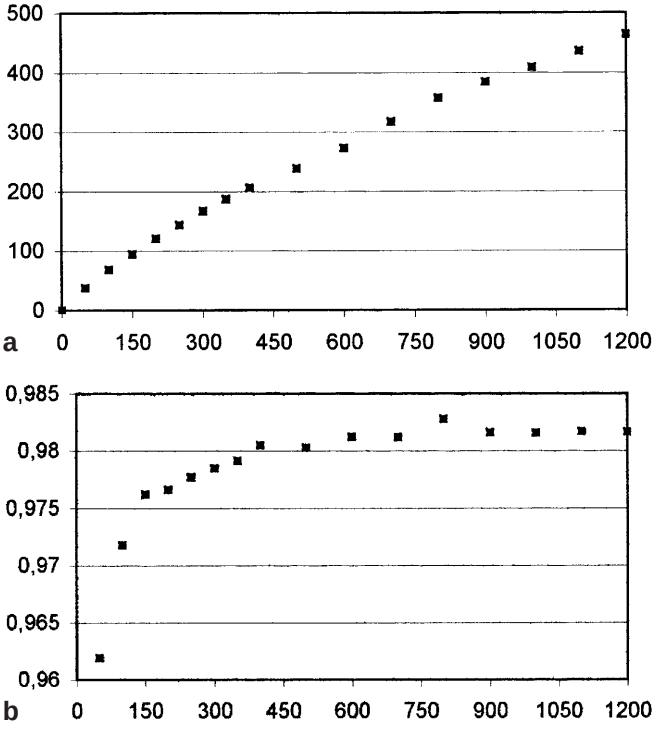

Fig. 1 Correlation between population sample size and **a** number of different mtDNA haplotypes and **b** haplotype diversity (Nei 1987) in HV1. The samples were randomly selected from the database; the value for the sample sizes 50 and 100 is the mean value of 10 and for the sample sizes from 150 to 350 of 5 randomly selected population sub-samples, respectively

Table 2 Number of different haplotypes and haplotype diversity (Nei 1987) of four short mtDNA amplicons in HV1 obtained from 1200 individuals

Primer set	Analysed segment (nps)	Number of different haplotypes	Haplotype diversity (Nei 1987)
F15971 R16175	16024–16142	45	0.608077
F15971 R16251	16024–16222	174	0.828584
F16144 R16410	16166–16365	356	0.962861
F16190 R16410	16212–16365	256	0.937146

Table 3 Number of polymorphic positions in HV1 and number of different haplotypes in three Caucasian populations

Reference	Size of database	Number of polymorphic positions in HV1	Number of different haplotypes
Lutz et al. (1998)	200	88	125
Parson et al. (1998)	101	74	68
Pfeiffer et al. (1999)	109	63	71

Table 4 Frequencies of the most commonly occurring haplotypes in this study compared to three other Caucasian populations. The frequencies were estimated from the number of haplotypes found in relation to the sample size of the data base (The most frequent haplotypes in this study with exception of the Anderson sequence were *1:* 16224 C, 16311 C; *2:* 16126 C, 16163 G, 16186 T, 16189 C, 16294 T; *3:* 16304 C; *4:* 16069 T, 16126 C; *5:* 16126 C, 16294 T, 16296 T, 16304 C; *6:* 16298 C; *7:* 16362 C; *8:* 16311 C; *9:* 16189 C; *10:* 16069 T, 16126 C, 16145 A, 16231 C, 16261 T; *11:* 16129 A, 16223 T; *12:* 16189 C, 16192 T, 16270 T; *13:* 16189 C, 16356 C)

Individual haplotype and frequency (%) in this study	Frequency (%) of individual haplotypes in literature		
	Lutz et al. (1998)	Parson et al. (1998)	Pfeiffer et al. (1999)
Anderson/10.3	15	21.8	11.9
1/3.3	2.5	1	2.8
2/2.9	1	0	0
3/2.7	2	0	1.8
4/2.6	2.5	2	7.3
5/2.6	3.5	1	0
6/2.4	2	1	2.8
7/2.2	1	0	0.9
8/1.8	1.5	1	3.7
9/1.3	1	2	1.8
10/1.2	1	3	0
11/1.1	1.5	1	0
12/0.9	0	0	0
13/0.9	1	0	0

resulted in 460 different haplotypes. With increasing number of individuals, the number of different haplotypes increases nearly linearly without approaching saturation level (Fig. 1 a). The haplotype diversity (Nei 1987) approached saturation level at a value of approximately 0.981 after typing 400 individuals (Fig. 1 b), and therefore from a database size of 400 HV1 sequences the possibility of 2 randomly selected individuals having identical mtDNA types is 2%.

Of the 1200 individuals, 305 showed unique mtDNA HV1 sequences. The most frequent haplotype was the reference sequence (Anderson et al. 1981), which occurred in 124 samples (10.3%), and 12 haplotypes showed up more than 10 times . An intact poly-cytosine tract between nps 16184 and 16193 without a T at position 16189 resulting in length heteroplasmy (Bendall and Sykes 1995) with the characteristic out-of-reading frame sequence beyond the C-stretch, was found in 144 cases (12%). In such cases for confirmation of the base call the application of primer sets F15971/R16175 and F16190/16410 can be recommended for forensic use (Table 2).

We conclude from these data that there is a very high variability in HV1 even in a local area where the number of maternally related individuals is expected to be higher than in big cities. The high variability of the data set can be explained by the fact that people from the nearby big city moved into the village area and that about 5% of the volunteers did not belong to the German Caucasian population group. However, the variability of this data set does not seem to be higher than in other Caucasian population groups (Table 3). The number of different haplotypes in HV1 in other Caucasian population groups (Lutz et al. 1998; Parson et al. 1998; Pfeiffer et al. 1999) correlates with the sample size of the database in the same manner as in our data set (Table 1; Fig. 1). Most of the frequent haplotypes in our data also occur with a high frequency in other Caucasian databases (Table 4). Our data do not suggest at which population sample size the saturation of the number of variable sites will occur. Although in our data set a total of 154 variable positions were found, in other Caucasian population groups (Lutz et al. 1998; Parson et al. 1998; Pfeiffer et al. 1999) variable positions and haplotypes occur which did not show up in our data. As the most commonly occurring haplotypes found in this study are also essentially the most common in other Caucasian population samples, it would seem that even those population groups from different regions are basically similar (Table 4). However, there also occur exceptional differences in frequencies of certain haplotypes, for instance the Anderson sequence or haplotype 4 in this study (Table 4). These differences may be due to small sample size of the data bases compared to the 1200 samples investigated in this study. It seems from this comparison that with the increase of the sample size of a mtDNA data base of one population group, the frequencies of certain commonly occurring haplotypes remain relatively stable while new unique haplotypes show up.

In forensic case work the amplification of short DNA segments is generally more sensitive and efficient (Mannucci et al. 1994). In these cases the amplification of multiple overlapping small fragments is a very common prac-

tice (Holland et al.1995). The calculations from our database showed that the primer set F16144/R16410 is the most informative short amplicon concerning the individual variability (Table 2). However, this primer set cannot be applied for sequencing in both directions in cases with a length heteroplasmy in the poly-C stretch. Therefore, in forensic and archaeological case work, when only a small amount of mtDNA sequences can be obtained, we recommend the use of the primer pair F16190/R16410 among other small mtDNA fragments (Table 2).

Acknowledgements We would like to thank Kirsten Waterkamp and Karin Möller for their excellent technical assistance in the laboratory.

References

Allen M, Engström A-S, Meyers S, Handt O, Saldeen T, von Haeseler A, Pääbo S, Gyllensten U (1998) Mitochondrial DNA sequencing of shed hairs and saliva on robbery caps: sensitivity and matching probabilities. J Forensic Sci 43: 453–464

Anderson S, Bankier AT, Barrell BG, de Bruijn MHL, Coulson AR, Drouin J, Eperon IC, Nierlich DP, Roe BA, Sanger F, Schreier PH, Smith AJH, Staden R, Young IG (1981) Sequence and organisation of the human mitochondrial genome. Nature 290: 457–464

Bär W, Kratzer A, Mächler M, Schmidt W (1988) Postmortem stability of DNA. Forensic Sci Int 39: 39–70

Bendall KE, Sykes BC (1995) Length heteroplasmy in the first hypervariable segment of the human mtDNA control region. Am J Hum Genet 57: 248–256

Handt O, Richards M, Trommsdorff M, Kilger C, Simanainen J, Georgiev O, Bauer K, Stone A, Hedges R, Schaffner W, Utermann G, Sykes B, Pääbo S (1994) Molecular genetic analysis of the Tyrolean Ice Man. Science 264: 1775–1778

Holland MM, Fisher DL, Roby RK, Ruderman J, Bryson C, Weedn VW (1995) Mitochondrial DNA sequence analysis of human remains. Crime Lab Digest 22: 109–115

Krings M, Stone A, Schmitz RW, Krainitzki H, Stoneking M, Pääbo S (1997) Neandertal DNA sequences and the origin of modern humans. Cell 90: 19–30

Lee SD, Shin CH, Kim KB, Lee YS, Lee JB (1997) Sequence variation of mitochondrial DNA control region in Koreans. Forensic Sci Int 87: 99–116

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

Lutz S, Weisser H-J, Heizmann J, Pollak S (1999) Location and frequency of polymorphic positions in the mtDNA control region of individuals from Germany (erratum). Int J Legal Med 112: 205–209

Mannucci A, Sullivan KM, Ivanov PL, Gill P (1994) Forensic application of a rapid and quantitative DNA sex test by amplification of the X-Y homologous gene amelogenin. Int J Legal Med 106: 190–193

Nei M (1987) Molecular evolutionary genetics. Columbia University Press, New York Chichester, West Sussex, pp 177–179

Parson W, Parsons TJ, Scheithauer R, Holland MM (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

Pfeiffer H, Steighner R, Fisher R, Yoon C-L, Mörnstad H, Holland MM (1998) Mitochondrial DNA extraction and typing from isolated dentin – experimental evaluation in a Korean population. Int J Legal Med 111: 309–313

Pfeiffer H, Brinkmann B, Hühne J, Rolf B, Morris AA, Steighner R, Holland MM, Forster P (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

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

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

Walsh PS, Metzger DA, Higuchi R (1991) Chelex-100 as a medium for simple extraction of DNA for PCR-based typing from forensic material. Biotechniques 10: 506–513