

Sabine Lutz-Bonengel · Ulrike Schmidt · Tanja Schmitt
Stefan Pollak

Sequence polymorphisms within the human mitochondrial genes MTATP6, MTATP8 and MTND4

Received: 15 July 2002 / Accepted: 19 November 2002 / Published online: 6 May 2003

© Springer-Verlag 2003

Abstract By sequencing the control region of mitochondrial DNA, the majority of human DNA samples can be differentiated. A further increase in differentiation probability may be possible, e.g. by extending the sequenced region to coding regions of the mitochondrial genome. Restriction to those positions that do not result in a change of the amino acids guarantees that the information thus obtained does not refer to phenotypically relevant information. In the present study the sequence data of the mitochondrial genes MTATP6, MTATP8 and MTND4 were collected from 109 subjects and analyzed in order to define variable positions suitable for identification purposes. There were 32 variable base positions among 850 bases studied from MTATPase genes and 1,200 bases of the MTND4 gene showed 28 variable positions. “Hot spots” for base exchanges were found in both regions and one position (position 11719 in the MTND4 gene) seems to be suitable for SNP investigation for forensic purposes.

Keywords Mitochondrial DNA · Coding region · MTATP · MTND4 · Single nucleotide polymorphism

Introduction

Apart from the nuclear genome, human cells contain an additional genome of 16,569 base pairs (bp) within the mitochondria (Anderson et al. 1981). Due to the high copy number per cell (Robin and Wong 1988), the tissue concentration of mitochondrial DNA (mtDNA) is 10^3 to 10^4 -fold elevated in comparison to genomic DNA (Bogenghagen and Clayton 1974; Veltri et al. 1990). Therefore, the forensic value of mtDNA analysis lies in the possibility to obtain DNA from even small traces where only

highly degraded or very small quantities of nuclear DNA are available (Wilson et al. 1995; Lutz et al. 1996). A major restriction for mtDNA typing, the small size of the mitochondrial compared to the nuclear genome, is partially counteracted by the higher mutation rate within mitochondria. Accumulation of free radicals, lack of protective histones and limited DNA repair cause a 10-fold higher mutation frequency of mtDNA in comparison to nuclear DNA (Richter et al. 1988). This provides the basis for a high sequence variability that manifests itself especially in the non-coding control region where selective pressure is lowered (Greenberg et al. 1983; Parsons et al. 1997).

Another restriction for mtDNA typing is maternal inheritance which makes it impossible to perform any determination of paternity or to distinguish between maternally related individuals. On the other hand the lack of recombination with a paternal counterpart makes spontaneous mutation the only source of change in the mitochondrial genome (Hutchison et al. 1974; De Francesco et al. 1980; Giles et al. 1980), permitting investigation of the maternal relationships even in those cases where representatives of several generations are missing (Gill et al. 1994; Lutz et al. 1999).

By sequencing the hypervariable regions HVI and HVII within the mitochondrial control region, the majority of human DNA samples can be differentiated (Budowle et al. 1999, 2002; Lutz et al. 1998). Another 20% of those samples showing identical HVI and HVII sequence patterns can be differentiated by the hypervariable region HVIII (Lutz et al. 1997). In spite of this high individualization potential of the mitochondrial control region (e.g. the haplotype diversity for 109 unrelated individuals from northern Germany is 0.9891, Pfeiffer et al. 1999), some mtDNA sequences cannot be distinguished within the hypervariable regions I–III. As an example Table 1 shows three frequent haplotypes with the frequency of occurrence in the german-speaking Caucasian population. There are 12 additional haplotypes found at frequencies of 0.5% or greater (this ignores length variation in the HVII C-stretch region, Parsons and Coble 2001).

As mtDNA is a single locus, haplotype frequencies cannot be obtained by simply multiplying the individual fre-

S. Lutz-Bonengel (✉) · U. Schmidt · T. Schmitt · S. Pollak
Institut für Rechtsmedizin, Klinikum der Universität Freiburg,
Albertstrasse 9, 79104 Freiburg, Germany
Tel.: +49-761-2036854, Fax: +49-761-2036858,
e-mail: lutz@sun11.ukl.uni-freiburg.de

Table 1 The three common haplotypes and their frequency in German-speaking countries (calculated from D-loop base, <http://www.d-loop-base.de>, Wittig et al. 2000)

Haplotype	Frequency
Anderson sequence	0.01
16519 C; 263G; 315.1 C	0.027
16519 C; 263G; 309.1 C; 315.1 C	0.039

quencies and the sequenced region must be compared with an entire catalogue of sequenced regions. To reach an acceptable probability of individualization with this method, the availability of sequence data from numerous unrelated subjects is extremely important. Several groups have already established such databases (Wittig et al. 2000; Miller and Budowle 2001; Röhl et al. 2001; Forster et al. 2002; Walther Parson, personal communication).

Apart from extensive databasing it is also necessary to try to increase the possibility of differentiation within the mtDNA molecule. For example, the region investigated for polymorphisms could be extended to the coding regions of the mitochondrial genome as well.

Until recently, almost all studies on human identification based on mtDNA sequencing have been confined to the control region. It is believed that the DNA sequences in the coding region of mtDNA are highly conserved in healthy individuals (Johns 1995), therefore polymorphisms in the mtDNA coding regions have only rarely been explored for human identity testing or human population genetics.

The fact that coding regions are more polymorphic than assumed can also be seen by analysis of the cytochrome b gene. Variations in this gene have been studied regarding species identification (Zehner et al. 1998; Parson et al. 2000), population dispersal (Andreu et al. 1999; Lee et al. 2002) and in terms of the role that mtDNA mutations play in human diseases (Andreu et al. 1998). By sequencing 32 Caucasian individuals Andreu et al. (1999) found 27 variants and could define 4 different genotypes for the cytochrome b gene.

Further variations in the mtDNA coding region have been published on the basis of the presence or absence of restriction enzyme recognition sites (Torrioni et al. 1996; Wallace et al. 1999) or of complete mtDNA sequences (Ingman et al. 2000; Finnillä et al. 2001; Herrnstadt et al. 2002).

Ingman et al. (2000) determined the substitution rate of the different parts of the human mitochondrial genomes. In this study MTND4 showed the second largest substitution rate of all mitochondrial genes (1.78×10^{-8} substitutions per site per year). The substitution rates of the genes MTATP 6 and MTATP 8 are also relatively high (1.52×10^{-8} and 1.33×10^{-8} , respectively), therefore we decided to search for polymorphism within these three genes.

As it is impractical to sequence the entire coding region when investigating a forensic case, it seems reasonable to determine particularly variable positions and to develop a typing method using SNPs (single nucleotide polymorphisms) (Holyoake et al. 2001; Jobling 2001; Parsons and

Coble 2001). SNPs are defined by the presence of two or more alternative bases in a certain position of DNA sequence (Brookes 1999). Different methods exist for the detection of the single nucleotide variations such as RFLP (restriction fragment length polymorphism) (Yoneda et al. 1996; Butler et al. 1998), ASO (allele-specific oligonucleotide) (Wong and Senadheera 1997), primer elongation reaction on oligonucleotide microarrays (Erdogan et al. 2001), homogeneous fluorescent PCR assays (Parsons and Coble 2001) or light cycler detection assays (Hiratsuka et al. 2002).

Restriction to synonymous changes guarantees that the information thus obtained does not refer to certain individual properties (e.g., hereditary diseases), but can be used for individualization only. This has to be emphasized because for several hereditary diseases, for example LHON (Leber's hereditary optic neuropathy) and Leigh syndrome, mutations of mitochondrial DNA have been described some of which are located within the MTATP6 gene (Akagi et al. 2001; Carrozzo et al. 2001; Schon et al. 2001). In addition, prediction of phenotypic information from DNA analyzed for forensic purposes is not allowed by German law (§81e StPO).

In the present study the sequences of the mitochondrial genes MTATP6, MTATP8 and MTND4 (see map under <http://www.mitomap.org> and Mitomap 2001) were collected and analyzed for each of 109 unrelated individuals in order to define polymorphisms suitable for forensic identification purposes.

Material and methods

Samples

A total of 109 DNA samples from maternally unrelated persons were analyzed. Samples stemmed partly from peripheral blood from corpses ($n=48$), and partly from buccal swabs from living volunteers ($n=61$).

Nucleotides in the mtDNA were numbered according to the Cambridge reference sequence (CRS) reported by Anderson et al. (1981).

DNA extraction

DNA was extracted using standard techniques as previously described (Walsh et al. 1991; Lutz et al. 1996).

Amplification of mtDNA

The PCR primers used to amplify the whole control region and parts of the genes MTATP6, MTATP8 and MTND4 are listed in Table 2. Amplification was carried out in 25 μ l reaction volumes using a Biozym PTC-200 thermal cycler (MJ Research, Waltham MA). Each reaction contained 1.25 U of Taq DNA polymerase (Applied Biosystems, Foster City, CA), 5 pmol each primer, 50 μ M each dNTP, 10 mM Tris/HCl (pH 8.3), 50 mM KCl and 1.5 mM MgCl₂. Negative and blank reagent controls were used throughout. After amplification the products were confirmed using 2% agarose gel electrophoresis and separated from residual primers, dNTPs and buffer using the QIAquick PCR purification kit (Qiagen, Hilden, Germany).

Table 2 Sequences of primers used for mitochondrial DNA amplification and sequencing

PCR Primer	Region	Nucleotide sequence
^a L15995	d-loop	5' GTAAAACGACGGCCAGTGAAGTCCACCATTAGCACCCAAAG 3'
^a H16488	d-loop	5' GGAAACAGCTATGACCATGAGGAACCAGATGTCGGATACAG 3'
^a L16221	d-loop	5' GTAAAACGACGGCCAGTGAACAAGCAAGTACAGCAATCAAC 3'
^a H259	d-loop	5' GGAAACAGCTATGACCATGGATGTCTGTGTGGAAAGTGGCT 3'
^a L182	d-loop	5' GTAAAACGACGGCCAGTGACGCACCTACGTTCAATATTAC 3'
^a H619	d-loop	5' GGAAACAGCTATGACCATGGGTGATGTGAGCCCGTCTAA 3'
^a L8353	ATPase 6,8	5' GTAAAACGACGGCCAGTGAGATTAAGAGAACCAACACCT 3'
^a H8805	ATPase 6,8	5' GGAAACAGCTATGACCATGTAGATAGTTGGGTGGTTGGT3'
L8734	ATPase 6	5' CTAAAGGACGAACCTGATCTC 3'
^a L8839	ATPase 6	5' GTAAAACGACGGCCAGTGATATCTATAAACCTAGCCATGG 3'
^a H9203	ATPase 6	5' GGAAACAGCTATGACCATGGTGATTGGTGGGTCATTATG 3'
L10708	ND4	5' GTGGGCCTAGCCCTACTAGTCTC 3'
H11071	ND4	5' TGGCTGTGAATGTTATAATTAAGGAG 3'
L11062	ND4	5' CTCTACCTCTCTATACTAATCTCC 3'
H11325	ND4	5' GTCATATTAAGTTGTTGGCTCAGG 3'
L11421	ND4	5' TGACTCCCTAAAGCCCATGT 3'
H11479	ND4	5' GTGTGAGGCGTATTATACCATAGC 3'
H11835	ND4	5' GCGAGGCTTGCTAGAAGTC 3'
L11782	ND4	5' ACGCACTCACAGTCGCATC 3'
H12152	ND4	5' GATTCACAATCTGATGTTTTGGTT 3'
Sequencing primers		
M13		5' GTAAAACGACGGCCAGTGA 3'
Rev		5' GGAAACAGCTATGACCATG 3'

Primers are numbered according to the location of the 3'-ends in the reference sequence. *L* and *H* designate the light and the heavy strands of the mtDNA molecule, respectively.

^aL strand primers show the sequence of the universal primer at the 5'-end, *H* strand primers carry the sequence of the reverse primer.

Cycle sequencing, electrophoresis, detection and analysis

The purified PCR products were sequenced by the dideoxy chain-termination procedure of Sanger et al. (1977) using the cycle sequencing method. Electrophoresis and detection of the fluorescence-labeled chain termination products were performed with an Applied Biosystems DNA Sequencer Model 310.

Analysis was performed using ABI 310 analysis software version 3.4.1. Sequences were aligned and compared with the CRS using the Sequence Navigator software version 1.0.1 (Applied Biosystems).

Results and discussion

Within the genes MTATP6 and MTATP8 a stretch of approximately 850 bases was studied. The 109 samples analyzed showed a total of 33 positions that varied from the Anderson sequence (see Table 3), however, all sequenced individuals showed an G in position 8860, confirming the revision of the CRS by Andrews et al (1999). This study demonstrated that at this position the original CRS represents a rare polymorphism. Thus 32 positions were found to differ from the revised CRS within the genes MTATP6 and MTATP8.

We could not detect any hypervariable regions within the genes MTATP6 and MTATP8. However, there were positions with frequent substitutions: 8697A was found in 10 individuals and 9055A was found in 8 individuals, each showing a transition from G to A.

Within the genes MTATP6 and MTATP8 silent mutations without an effect on the amino acid sequence of the coded protein were found in approximately 50% (15 out

of 32) of the variable positions analyzed. The selection pressure on MTATP6 and MTATP8 thus seems to be surprisingly low.

Mutations causing an exchange of amino acids (17 out of 32) (see Table 3 positions in bold type) were excluded from further investigation as genetic information collected for merely forensic purposes (individualization) must never contain any phenotypically relevant information (e.g., hereditary diseases).

The 15 silent polymorphisms form a total number of 15 haplotypes (Table 3). The most common haplotype is haplotype 1, which corresponds to the revised CRS and accounts for 83 of the 109 samples, 9 samples showed 1 single exchange at position 8697 (haplotype 8), 4 at position 8592 (haplotype 4) and 2 at position 8512 (haplotype 3). The other 11 haplotypes, which showed up to 2 sequence polymorphisms, were observed only once.

In the region of the MTND4 gene approximately 1,200 bases have been analyzed (see Table 4) and among 109 unrelated individuals, 28 variable positions were found. Although greater in length, this region shows a higher sequence conservation than the MTATP6 and MTATP8 genes. In the genes MTATP6 and MTATP8, 17 out of 32 base exchanges resulted in an exchange of amino acids, whereas the sequence variations found in the MTND4 gene were almost exclusively silent mutations. Only 2 out of 28 sequence polymorphisms (position 10760 and 11025) would lead to a change of the amino acid sequence (see Table 4 positions in bold type). This suggests that compared to the MTATP6 and MTATP8 genes, the MTND4 gene is less polymorphic possibly due to a higher selection pressure.

Table 3 Sequence polymorphisms of 109 subjects between positions 8366–9202 (MTATP6 and MTATP8 genes; five bases of the MTATP6

Position	8400	8410	8446	8472	8512	8521	8531	8542	8592	8642	8645	8676	8684	8697	8701	8705	8708
Reference (Anderson)	T	C	A	C	A	A	A	T	G	A	A	C	C	G	A	T	A
Reference aa	Met	Pro	Lys	Pro	Lys	Glu	Thr	Cys	Leu	Asn	Asn	Ille	Thr	Met	Thr	Met	His
Substitution aa	Thr			Leu			Ala			Ser	Ser		Ille		Ala	Thr	Arg
Individual																	
1–42
43–47
48–51	A	.	.	.
52–55
56–58
59–61
62–64	A
65–66	G
67–68
69–70
71	G
72	G
73
74
75
76
77
78
79
80
81	.	T
82	A
83
84	C	.
85
86	.	.	.	T
87
88	T
89
90
91
92	G	.	.	G
93
94
95
96	.	.	G
97	T
98	A	.	.	.
99	A	.	.	.
100	A	.	.	.
101	A	.	.	.
102	A	.	.	.
103	A	.	.	.
104	A	.	.	.
105	C	A	.	.	.
106	A	G	.	.
107	A	.	.	.
108	C	G	A	.	.	.
109	A	.	.	.
Exchange frequency	1	1	1	1	2	1	1	1	4	1	1	1	1	10	1	1	1

Numbers on top indicate positions in the mitochondrial genome.

The reference nucleotides are according to the CRS.

Dots show matches with the reference nucleotides.

Letters show the bases which differ from the CRS.

Polymorphisms leading to an exchange of amino acids are in bold type.

gene are missing) with respect to CRS (Anderson et al. 1981)

8787	8790	8818	8836	8860	8901	8902	8950	8989	8994	9041	9055	9091	9093	9132	9142	MTATP6/ MTATP8 haplotypes (all)	MTATP6/ MTATP8 haplotypes (silent)
C	G	C	A	A	A	G	G	G	G	A	G	A	A	A	G		
Leu	Leu	Leu	Met Val	Thr Ala	Leu	Ala Thr	Val Ile	Ala Thr	Leu	His Arg	Ala Thr	Thr Ala	Thr	Leu	Val Ile		
.	.	.	.	G	1	1
.	.	.	.	G	1	1
.	.	.	.	G	8	8
.	.	.	.	G	1	1
.	.	.	.	G	1	1
.	.	.	.	G	1	1
.	.	.	.	G	4	4
.	.	.	.	G	3	3
.	.	.	.	G	1	1
.	.	.	.	G	1	1
.	.	.	.	G	5	1
.	.	.	.	G	10	1
.	.	.	.	G	.	.	A	14	1
.	.	.	.	G	G	17	1
.	.	.	.	G	G	.	.	.	19	1
.	.	.	.	G	A	21	1
.	.	.	.	G	.	.	.	A	15	1
.	.	.	.	G	1	1
.	.	.	.	G	1	1
.	.	.	.	G	1	1
.	.	.	.	G	2	2
.	.	.	.	G	4	4
.	.	.	.	G	1	1
.	.	.	.	G	9	1
.	.	.	.	G	1	1
.	.	.	G	G	23	1
.	.	.	.	G	A	18	1
.	.	.	.	G	7	1
.	.	.	.	G	A	18	1
.	.	.	.	G	A	18	1
.	.	.	.	G	A	18	1
.	.	.	.	G	A	11	1
.	.	.	.	G	A	18	1
.	.	.	.	G	1	1
A	.	.	.	G	G	13	13
.	.	.	.	G	22	22
.	.	.	.	G	6	6
.	.	.	.	G	8	8
.	.	.	.	G	8	8
.	.	.	.	G	8	8
.	.	T	.	G	8	8
.	.	.	.	G	A	12	12
.	.	.	.	G	16	16
.	.	.	.	G	G	.	20	20
.	.	.	.	G	24	24
.	A	.	.	G	26a	26
.	.	.	.	G	A	.	G	.	.	27a	27
.	.	.	.	G	A	28a	28
.	.	.	.	G	.	A	25	8
1	1	1	1	109	1	1	1	1	1	1	8	1	1	1	1		

Reference aa amino acid after Anderson.

Substitution aa amino acid substitution resulting from point mutation in the individual sequence.

Exchange frequency number of individuals showing base exchange at a certain position.

MTATP6/MTATP8 haplotypes all haplotypes within the 8366–9202 region considering both non-silent and silent mutations.

MTATP6/MTATP8 haplotypes silent haplotypes within 8366–9202 region considering only silent mutations.

Table 4 Sequence polymorphisms of 109 subjects between positions 10760–12137 (MTND4 gene) with respect to CRS (Anderson et al. 1981)

Position	10760	10790	10873	11025	11053	11147	11161	11167	11182	11251	11272	11299	11326	11332	11353
Reference (Anderson)	A	T	T	T	A	T	C	A	A	A	A	T	C	C	T
Reference aa	Met	Leu	Pro	Leu	Met	Leu	Thr	Trp	Glu	Leu	Leu	Thr	Ser	Ala	Ala
Substitution aa	Val			Pro											
Individual															
1–42
43–47
48–51	G
52–55	G
56–58
59–61	G
62–64
65–66
67–68
69–70	C
71
72
73
74
75
76
77	G
78	G
79
80
81
82	G
83	G	C
84
85	T	.
86
87	C	.	.	.
88	C
89	C	.	.	.
90	.	.	.	C	C	.	.	.
91	G	C	.	.	.
92	G	.	.	.	C	.	.	.
93	C	.	.	.
94	T	T	.
95
96
97
98	T	.	.	G
99	G
100	.	C	G
101	.	.	C	G	G
102	T	.
103
104	G
105	G
106	.	.	C
107	C	.	.	.
108	C	.	.	.
109	G
Exchange frequency	1	1	2	1	1	2	1	2	1	19	1	8	1	3	2

Numbers at the top indicate positions in the mitochondrial genome.

The reference nucleotides are according to the CRS.

Dots show matches with the reference nucleotides.

Letters show the bases which differ from the CRS.

Polymorphisms leading to an exchange of amino acids are in bold type.

Reference aa amino acid after Anderson.

11362	11377	11467	11485	11518	11560	11674	11719	11812	11818	11840	11864	11866	ND4 haplotypes (all)	ND4 haplotypes (silent)	MTATP6/ MTATP8 ND4 haplotypes (silent)
A	G	A	T	G	A	C	G	A	A	C	T	A			
Met	Lys	Leu	Gly	Leu	Trp	Thr	Gly	Leu	Pro	Leu	Leu	Leu			
.	1	1	1-1
.	.	G	A	16	16	1-16
.	A	G	18	18	8-18
.	A	13	13	1-13
.	A	7	7	1-7
.	A	A	22	22	1-22
.	1	1	4-1
.	1	1	3-1
.	.	.	.	G	6	6	1-6
.	2	2	1-2
.	1	1	1-1
.	1	1	1-1
.	1	1	1-1
.	1	1	1-1
.	1	1	1-1
.	1	1	1-1
.	3	3	1-3
.	5	5	1-5
.	C	.	8	8	1-8
.	G	9	9	1-9
.	1	1	2-1
.	4	4	4-4
.	10	10	1-10
.	A	7	7	1-7
.	.	G	14	14	1-14
G	A	15	15	1-15
.	.	G	A	19	19	1-19
.	.	G	A	21	21	1-21
.	.	G	A	.	G	.	.	.	23	23	1-23
.	.	G	A	26	19	1-19
.	.	G	A	27	27	1-27
.	.	G	C	.	.	.	A	28	28	1-28
.	.	G	C	.	.	.	A	.	.	T	.	.	30	30	1-30
.	.	G	.	A	.	.	A	12	12	1-12
.	A	7	7	13-7
.	A	7	7	22-7
.	.	G	A	16	16	6-16
.	2	2	8-2
.	A	13	13	8-13
.	A	G	24	24	8-24
.	A	G	29	29	8-29
.	.	G	A	20	20	12-20
.	T	A	17	17	16-17
.	A	13	13	20-13
.	A	G	18	18	24-18
.	A	11	11	26-11
.	A	G	A	25	25	27-25
.	.	G	A	19	19	28-19
.	A	G	18	18	8-18
1	4	18	2	1	2	1	43	8	1	1	1	1			

Substitution aa amino acid substitution resulting from point mutation in the individual sequence.

Exchange frequency number of individuals showing base exchange at a certain position.

ND4 haplotypes all haplotypes within 10760–12137 region considering both non-silent and silent base exchanges.

ND4 haplotypes silent haplotypes within 10760–12137 region considering only silent mutations.

MTATP6/MTATP8/ND4 haplotypes silent haplotypes including the three genes considering only silent base exchange.

As for the genes MTATP6 and MTATP8, for further investigation of the MTND4 gene we focused on the 26 positions with synonymous changes. In contrast to the relatively even distribution of the sequence polymorphisms in the genes MTATP6 and MTATP8, the MTND4 gene showed an accumulation of polymorphic sites between positions 11147 and 11377. In this sequence segment comprising 230 bp, 45 out of 130 base exchanges were observed at 12 out of 26 silent positions.

MTND4 base exchanges were found most frequently at positions 11719, 11251, 11467 and 11812, in 43, 19, 18 and 8 individuals, respectively.

The 26 silent mutation sites resulted in a total number of 29 haplotypes (see Table 4). The most common haplotype was again haplotype 1, which corresponds to the revised CRS and accounts for 54 out of 109 samples. Haplotypes 7, 13, 16 and 18 were found in 6 individuals, haplotypes 2, 19 and 22 in 3 individuals, haplotype 6 in 2 individuals each and 20 haplotypes were observed only once.

Considering all 3 analyzed genes combined, they could be split into 40 different haplotypes and the haplotype observed most frequently was that corresponding to the revised CRS sequence (48 out of 109 samples). Nearly all the sequence deviations observed were transitions, only one transversion was found within the MTATP6 gene at position 8787 (C to A, Table 3). Deletions or insertions were observed neither in the genes MTATP6 and MTATP8 nor in the MTND4 gene, although e.g., the MTND4 gene showed a sequence (CCCCCTCC) between positions 10947 and 10955 suggesting an increased slippage risk. The dominant role of the mutation type "transition" reflects the fact that mispair during replication is the major source of spontaneous mutations in mitochondrial DNA (Thomas and Beckenbach 1989).

Our findings agree with the results of Tzen et al. (2001) who identified clustered polymorphic sites in the mitochondrial coding region encompassing positions 8389–8865 by determination of mtDNA sequences of 119 unrelated Chinese. Within this 477 bp region they found 39 sites with a transition:transversion ratio of 37:2. Tzen et al. (2001) found two types of transversion (T to G and C to G) and one nucleotide deletion, but no nucleotide insertions. In our study only one C to A transversion was found and nucleotide deletions and insertions were not observed. Also, concordant with our examination the most common haplotype in the study of Tzen et al. (2001) was identical to the Anderson sequence except for position 8860G.

Ingman et al. (2000) showed that the number of polymorphic sites increases in proportion to the number of sequences taken into account and reported polymorphic sites of 646 coding region sequences (nps 577–16023 CRS) (Ingman et al. 2000; <http://www.genpat.uu.se/mtDB/>). Between the positions 8366–9207 (MTATP6 and MTATP8) they found 134 sequence changes and between positions 10760–12137 (MTND4) 147 sequence changes (surveyed on 10/10/2002).

Another study of complete mtDNA sequencing has been published by Herrnstadt et al. (2002) who analyzed

560 complete mtDNA coding region sequences from unrelated individuals of European, Asian, and African origin, presenting 636 sequence changes which occurred in the coding region of at least 2 individuals and 497 haplogroup-associated polymorphisms. For example, between the positions 8366–9207 (MTATP6 and MTATP8) and 10760–12137 (MTND4) they found the following haplogroup-specific positions for sequences of European origin: 8994 (haplogroup W), 9055 (haplogroup K, U2), and 11812 (haplogroup T2).

In our study we found 1 individual with a G to A transition at position 8994, 8 individuals with the same sequence change at position 9055 and 8 individuals with an A to G transition at position 11812. For all of these individuals a European origin has been confirmed.

A common HVI/HVII haplotype in the German population is 16519C, 263G, 315.1C (Lutz et al. 1998) and there are 12 additional haplotypes that occur at levels of

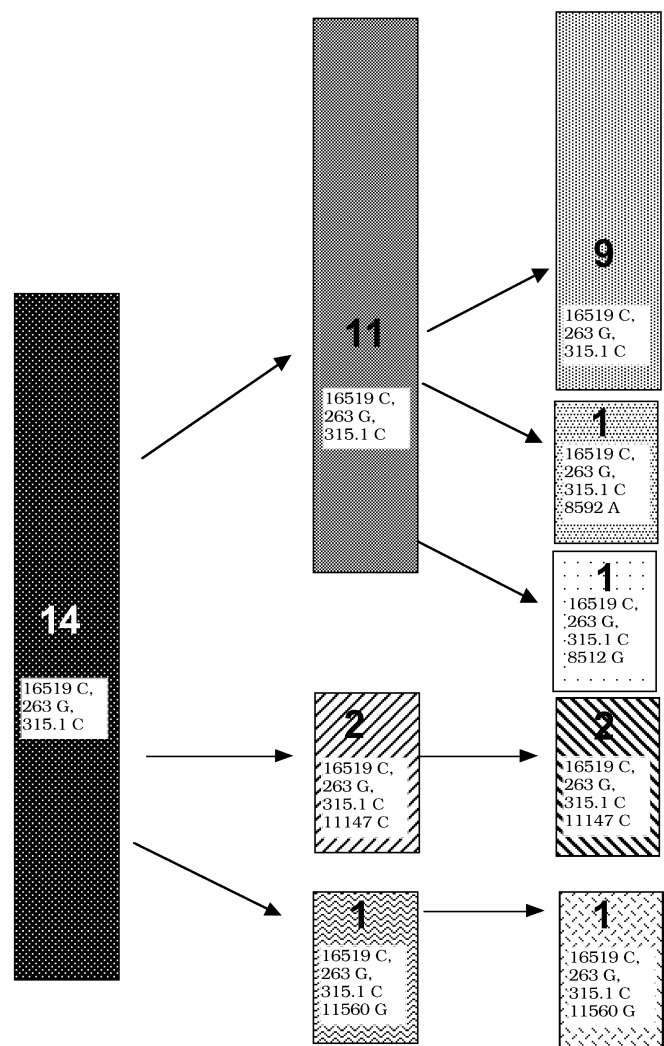


Fig. 1 Diagramme of the differentiation of the haplotype variant 1 (16519C, 263G, 315.1C) into three haplotypes after sequencing of the MTND4 gene and into five sequence types after additional sequencing of the genes MTATP6 and MTATP8

0.5% or greater (ignoring HVII C-stretch polymorphisms, Parsons and Coble 2001). To distinguish between individuals with the same control region haplogroup, it is necessary to include the coding region in the examination.

In order to test whether polymorphic positions within the mitochondrial genes MTATP6, MTATP8, and MTND4 allow a discrimination of samples with identical d-loop sequences, two particularly frequent d-loop sequences were chosen as examples (Table 1). The 14 sequences of haplotype 16519C, 263G, 315.1C could be subdivided into 3 different haplotypes after sequencing of the MTND4 gene and into 5 sequence haplotypes after additional sequencing of the genes MTATP6 and MTATP8 (see Fig. 1). The 16 sequences of haplotype 16519C, 263G, 309.1C, 315.1C could be distinguished into 2 haplotypes after sequencing of the genes MTATP6 and MTATP8 and into 4 haplotypes after additional sequencing of the MTND4 gene.

A total of 15 haplotypes have been observed in the ATPase genes, 29 in the MTND4 gene and by combining the 2 gene regions 40 lineages could be distinguished.

However, for the investigation of a forensic case sequencing of the entire coding region is too time-consuming and expensive. Analysis of a suitable panel of mtDNA SNPs should yield the necessary information at less cost and in less time.

For this reason SNP differentiation methods are increasingly being used. In order to also apply such methods in the field of mitochondrial DNA it will be necessary to identify as many SNP positions as possible but polymorphisms which occur rarely are not good SNP target sites. An important requirement for a polymorphism to serve as a SNP marker should be a high variability resulting in a high discrimination potential. However, for the three genes analyzed in this study, only position 11719 in the MTND4 gene fulfills this demand (11719A was found in 43 out of 109 individuals) and thus seems to be a potential candidate for SNP investigation. With the advance of more powerful gene chip-based methods allowing simultaneous screening for a high number of SNPs, less frequent polymorphisms will also be suitable for individualization.

Acknowledgements This research was supported by grants from the Deutsche Forschungsgemeinschaft (Po 579/2–4).

References

- Akagi M, Inui K, Tsukamoto H, Sakai N, Muramatsu T, Yamada M, Matsuzaki K, Goto Y, Nonaka I, Okada S (2001) A point mutation of mitochondrial ATPase 6 gene in Leigh syndrome. *Neuromuscul Disord* 12:53–55
- Anderson S, Bankier AT, Barrell BG, et al. (1981) Sequence and organization of the human mitochondrial genome. *Nature* 290:457–464
- Andreu AL, Bruno C, Shanske S et al. (1998) Missense mutation in the mtDNA cytochrome b gene in a patient with myopathy. *Neurology* 51:1444–1447
- Andreu AL, Bruno C, Hadjigeorgiou GM, Shanske S, DiMauro S (1999) Polymorphic variants in the human mitochondrial cytochrome b gene. *Mol Genet Metab* 67:49–52
- Andrews RM, Kubacka I, Chinnery PF, Lightowlers RN, Turnbull DM, Howell N (1999) Reanalysis and revision of the Cambridge reference sequence for human mitochondrial DNA. *Nat Genet* 23:147
- Bogenhagen D, Clayton DA (1974) The number of mitochondrial deoxyribonucleic acid molecules in mouse L and human HeLa cells. *J Biol Chem* 249:7991–7995
- Brookes AJ (1999) The essence of SNPs (Review). *Gene* 234:177–186
- Budowle B, Wilson MR, DiZinno JA, Stauffer C, Fasano MA, Holland MM, Monson KL (1999) Mitochondrial DNA regions HVI and HVII population data. *Forensic Sci Int* 103:23–35
- Budowle B, Allard MW, Fisher CL, Isenberg AR, Monson KL, Stewart JE, Wilson MR, Miller KW (2002) HVI and HVII mitochondrial DNA data in Apaches and Navajos. *Int J Legal Med* 116:212–215
- Butler JM, Wilson MR, Reeder DJ (1998) Rapid mitochondrial DNA typing using restriction enzyme digestion of polymerase chain reaction amplicons followed by capillary electrophoresis separation with laser-induced fluorescence detection. *Electrophoresis* 19:119–124
- Carozzo R, Tessa A, Vazquez-Memije ME et al. (2001) The T9176G mtDNA mutation severely affects ATP production and results in Leigh syndrome. *Neurology* 56:687–690
- De Francesco L, Attardi G, Croce CM (1980) Uniparental propagation of mitochondrial DNA in mouse-human cell hybrids. *Proc Natl Acad Sci U S A* 77:4079–4083
- Erdogan F, Kirchner R, Mann W, Ropers HH, Nuber UA (2001) Detection of mitochondrial single nucleotide polymorphisms using a primer elongation reaction on oligonucleotide microarrays. *Nucleic Acids Res* 29:E36
- Finnilä S, Lehtonen MS, 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, De Leo G, Flugy A, Salerno A, Ayala G, Kouvatsi A, Villems R, Romano V (2002) Continental and subcontinental distributions of mtDNA control region types. *Int J Legal Med* 116:99–108
- Giles RE, Blanc H, Cann HM, Wallace DC (1980) Maternal inheritance of human mitochondrial DNA. *Proc Natl Acad Sci U S A* 77:6715–6719
- Gill P, Ivanov PL, Kimpton C, Piercy R, Benson N, Tully G, Evett I, Hagelberg E, Sullivan KM (1994) Identification of the remains of the Romanov family by DNA analysis. *Nat Genet* 6:130–135
- Greenberg BD, Newbold JE, Sugino A (1983) Intraspecific nucleotide sequence variability surrounding the origin of replication in human mitochondrial DNA. *Gene* 21:33–49
- Herrnstadt C, Elson JL, Fahy E, Preston G, Turnbull DM, Anderson C, Ghosh SS, Olefsky JM, Beal MF, Davis RE, Howell N (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
- Hiratsuka M, Narahara K, Kishikawa Y, Ismail Hamdy S, Endo N, Agatsuma Y, Matsuura M, Inoue T, Tomioka Y, Mizugaki M (2002) A simultaneous LightCycler detection assay for five genetic polymorphisms influencing drug sensitivity. *Clin Biochem* 35:35–40
- Holyoake AJ, McHugh P, Wu M, O'Carroll S, Benny P, Sin IL, Sin FY (2001) High incidence of single nucleotide substitutions in the mitochondrial genome is associated with poor semen parameters in men. *Int J Androl* 24:175–182
- Hutchison CA III, Newbold JE, Potter SS, Edgell MH (1974) Maternal inheritance of mammalian mitochondrial DNA. *Nature* 251:536–538
- Ingman M, Kaessmann H, Pääbo S, Gyllensten U (2000) Mitochondrial genome variation and the origin of modern humans. *Nature* 408:708–713
- Jobling MA (2001) Y-chromosomal SNP haplotype diversity in forensic analysis. *Forensic Sci Int* 118:158–162

- Johns DR (1995) Seminars in medicine of the Beth Israel Hospital, Boston. Mitochondrial DNA and disease (Review). *N Engl J Med* 333:638–644
- Lee SD, Lee YS, Lee JB (2002) Polymorphism in the mitochondrial cytochrome B gene in Koreans an additional marker for individual identification. *Int J Legal Med* 116:74–78
- Lutz S, Weisser HJ, Heizmann J, Pollak S (1996) mtDNA as a tool for identification of human remains. *Int J Legal Med* 109:205–209
- Lutz S, Weisser HJ, Heizmann J, Pollak S (1997) A third hyper-variable region in the human mitochondrial D-loop. *Hum Genet* 101:384
- Lutz S, Weisser HJ, Heizmann J, Pollak S (1998) Location and frequency of polymorphic positions in the human mtDNA control region of individuals from Germany. *Int J Legal Med* 111:67–77
- Lutz S, Meier V, Amberg R (1999) Determination of relationship using mtDNA. *Rom J Legal Med* 7:158–164
- Miller KW, Budowle B (2001) A compendium of human mitochondrial DNA control region: development of an international standard forensic database. *Croat Med J* 42:315–327
- MITOMAP (2001) A human mitochondrial genome database. Center for Molecular Medicine, Emory University, Atlanta, GA, USA. <http://www.gen.emory.edu/mitomap.html>
- Parson W, Pegoraro K, Niederstatter H, Fogar M, Steinlechner M (2000) Species identification by means of the cytochrome b gene. *Int J Legal Med* 114:23–28
- Parsons TJ, Coble MD (2001) Increasing the forensic discrimination of mitochondrial DNA testing through analysis of the entire mitochondrial DNA genome (Review). *Croat Med J* 42:304–309
- Parsons TJ, Muniec DS, Sullivan K, Woodyatt N, Alliston-Greiner R, Wilson BDL, Holland KA, Weedn VW, Gill P, Holland MM (1997) A high observed substitution rate in the human mitochondrial DNA control region. *Nat Genet* 15:363–368
- 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
- Richter C, Park J-W, Ames BN (1988) Normal oxidative damage to mitochondrial and nuclear DNA is extensive. *Proc Natl Acad Sci U S A* 85:6465–6467
- Robin ED, Wong R (1988) Mitochondrial DNA molecules and virtual number of mitochondria per cell in mammalian cells. *J Cell Physiol* 13:507–513
- Röhl A, Brinkmann B, Forster L, Forster P (2001) An annotated mtDNA database. *Int J Legal Med* 115:29–39
- Sanger F, Nicklen S, Coulson AR (1977) DNA sequencing with chain terminating inhibitors. *Proc Natl Acad Sci U S A* 74:5463–5467
- Schon EA, Santra S, Pallotti F, Girvin ME (2001) Pathogenesis of primary defects in mitochondrial ATP synthesis. *Semin Cell Dev Biol* 12:441–448
- Thomas WK, Beckenbach AT (1989) Variation on salmonid mitochondrial DNA: evolutionary constraints and mechanisms of substitution. *J Mol Evol* 29:233–245
- Torroni A, Huoponen K, Francalacci P, Petrozzi M, Morelli L, Scozzari R, Obinu D, Savontaus ML, Wallace DC (1996) Classification of European mtDNAs from an analysis of three European populations. *Genetics* 144:1835–1850
- Tzen CY, Wu TY, Liu HF (2001) Sequence polymorphism in the coding region of mitochondrial genome encompassing position 8389–8865. *Forensic Sci Int* 120:204–209
- Veltri KL, Espiritu M, Singh G (1990) Distinct genomic copy number in mitochondria of different mammalian organs. *J Cell Physiol* 143:160–164
- Wallace DC, Brown MD, Lott MT (1999) Mitochondrial DNA variation in human evolution and disease (Review). *Gene* 238:211–230
- Walsh PS, Metzger DA, Higuchi R (1991) Chelex-100 as a medium for simple extraction of DNA for PCR-based typing for forensic material. *Biotechniques* 10:506–513
- Wilson MR, DiZinno J, Polansky D, Replogle J, Budowle B (1995) Validation of mitochondrial DNA sequencing for forensic casework analysis. *Int J Legal Med* 108:68–74
- Wittig H, Augustin C, Baasner A et al. (2000) Mitochondrial DNA in the Central European population. Human identification with the help of the forensic mt-DNA D-loop-base database. *Forensic Sci Int* 113:113–118
- Wong LJ, Senadheera D (1997) Direct detection of multiple point mutations in mitochondrial DNA. *Clin Chem* 43:1857–1861
- Yoneda M, Tanno Y, Tsuji S, Attardi G (1996) Detection and quantification of point mutations in mitochondrial DNA by PCR. *Methods Enzymol* 264:432–441
- Zehner R, Zimmermann S, Mebs D (1998) RFLP and sequence analysis of the cytochrome b gene of selected animals and man: methodology and forensic application. *Int J Legal Med* 111:323–327