Int J Legal Med (1998) 111 : 67–77

© Springer-Verlag 1998

ORIGINAL ARTICLE

S. Lutz · H.-J. Weisser · J. Heizmann · S. Pollak

Location and frequency of polymorphic positions in the mtDNA control region of individuals from Germany

Received: 21 May 1997 / Received in revised form: 9 September 1997

Abstract In order to identify polymorphic positions and to determine their frequency in the human mitochondrial D-loop containing region, the mitochondrial DNA (mtDNA) control region of 200 unrelated individuals from Germany were amplified and directly sequenced. Sequence comparison led to the identification of 190 mitochondrial lineages as defined by 202 variable positions. The most frequently occurring lineage comprised 5 individuals, whereas 186 types of D-loop sequences were observed in only one individual. Of the sequences studied 7% are not unique but show at least one counterpart with an identical haplotype. The majority (61%) of the control regions investigated showed between four and eight nucleotide positions deviating from the reference sequence. The maximum number of deviations observed in a single control region was 18. The majority of the variable positions in the D-loop region (88%) are located within three hypervariable regions. Sequence variations are caused by nucleotide substitutions, insertions or deletions. As compared to insertions and deletions, nucleotide substitutions make up the vast majority of the mutations (90%). We have predominantly found transitions (75%) and a significantly lower frequency of transversions (15%) whereas insertions (6%) as well as deletions (4%) are rather rare. Upon sequencing the mitochondrial control region from 200 German Caucasians the genetic diversity was estimated at 0.99. The probability of two randomly selected individuals from a population having identical mtDNA types is 0.6%.

Key words DNA typing · Mitochondrial DNA · Non-coding region · D-loop region · Hypervariable regions · PCR · Sequencing · Population study

Introduction

Human mitochondrial DNA (mtDNA) is a double-stranded closed circular molecule present in 1,000–10,000 copies per cell [1]. The complete nucleotide sequence of the 16,569 bp human mitochondrial DNA was determined in 1981 [2]. Analysis of the genome structure revealed that the human mitochondrial DNA is very compactly organized. It carries a single non-coding region of significant size (1,122 bp) located between the tRNA genes for proline ($tRNA^{Pro}$) and phenylalanine ($tRNA^{Phe}$). This so-called "control region" or "D-loop containing region" contains the origin of replication for the synthesis of the H-strand (O_H) as well as the promoters for both H- (HSP) and L- (LSP) strand transcription [3, 4]. Despite its functional importance, this region is believed to be the most rapidly evolving part of the molecule [5].

Nucleotide substitutions accumulate in the mitochondrial genome with a rate considerably higher than for single-copy nuclear DNA [6]. This is most probably due to the lower efficiency of DNA repair as well as to a higher frequency of DNA replication errors in mitochondrial DNA [7]. Consequently, mtDNA and in particular the non-coding region, is highly polymorphic.

Owing to the high copy number of mtDNA molecules per mitochondrion and to the high number of mitochondria per cell, there is a 10^3 to 10^4-fold molar excess of mitochondrial DNA as compared to genomic DNA. This quantitative consideration as well as the strictly maternal inheritance, the lack of recombination and the high mutation rate of mitochondrial DNA are the main reasons for the use of mitochondrial sequences in forensic medicine [8], population studies [9], molecular evolution [10], anthropology [11] and archaeology [12]. The usage of polymorphic nuclear DNA loci for forensic analyses has been frequently hampered by highly degraded template DNA or very small DNA samples. Several recent studies demonstrated that in these cases the analysis of mitochondrial sequences can yield a higher rate of success [12–19]. Knowledge on the frequencies with which certain mito-

S. Lutz (✉) · H.-J. Weisser · J. Heizmann · S. Pollak
Institut für Rechtsmedizin, Klinikum der Universität Freiburg,
Albertstrasse 9, D-79104 Freiburg, Germany
Fax: +49 (761) 203 6858
e-mail: lutz@sun11.ukl.uni-freiburg.de

Table 1 Sequences of primers used for amplification and sequencing of mitochondrial DNAs in this study. 'L' and 'H' stand for the light and heavy strand of the mtDNA molecule, respectively. The primers are numbered according to the location of their 3′-ends in the reference sequence [2]

Amplification primer	Nucleotide sequence
L15995	5′CGTAAAACGACGGCCAGTGAACTCCACCATTAGCACCCAAAG 3′
H16488	5′GGAAACAGCTATGACCATGAGGAACCAGATGTCGGATACAG 3′
L16221	5′GTAAAACGACGGCCAGTGAACAAGCAAGTACAGCAATCAAC 3′
H259	5′GGAAACAGCTATGACCATGGATGTCTGTGTGGAAAGTGGCT 3′
L182	5′GTAAAACGACGGCCAGTGACGCACCTACGTTCAATATTAC 3′
H619	5′CGGAAACAGCTATGACCATGGGTGATGTGAGCCCCGTCTAA 3′
Sequencing primer	**Nucleotide sequence**
Universal (M13*)	5′GTAAAACGACGGCCAGTGA 3′
Reverse (Rev*)	5′GGAAACAGCTATGACCATG 3′

chondrial DNA sequences occur in a given population is of crucial importance for the application of mitochondrial markers in forensic studies. To further our understanding of this matter we have carried out a population study including 200 unrelated individuals from Germany.

Materials and methods

Samples were taken from unrelated German Caucasians from southern Germany partly as muscle tissue from corpses (n = 149) and partly blood from volunteers (n = 51).

DNA extraction from muscle tissue

DNA was extracted from 0.5 g muscle tissue from corpses by freezing and thawing. The tissue sample was ground and incubated overnight in extraction buffer (10 mM Tris-HCl, pH 8.0, 10 mM Na_2EDTA, pH 8.0, 100 mM NaCl, 20% SDS, 40 mM DTT) with proteinase K (800 µg/ml) at 56°C. After two extractions with phenol/chloroform DNA was precipitated with ethanol and the dried pellet was resuspended in 80 µl bidistilled water. DNA concentrations were determined by gel electrophoresis (0.8% agarose gel) and compared with a defined DNA reference standard using the Bioprofil Bio 1D computer program from LTF-Labortechnik.

DNA extraction from blood samples

EDTA blood samples of 1 ml were taken from volunteers. After freezing, thawing and mixing with an equal volume of 1 × SSC buffer (15 mM sodium citrate, 150 mM NaCl), the nuclei were pelleted by centrifugation (5 min at 8,000 g). After discarding the supernatant, the pellet was washed twice with 1 × SSC buffer and centrifuged again (5 min at 8,000 g). The nuclei were resuspended in 700 µl bidistilled water, 50 µl 3 M NaAc, pH 7.0, 50 µl 10% SDS and 20 µl proteinase K (20 mg/ml) and incubated overnight at 56°C. Purification, extraction and quantitation of DNA were carried out as described for muscle tissue.

Amplification and sequencing of mtDNA

PCR amplification using the primer pairs L15995/H16488, L16221/H259 and L182/H619 yielded fragments of 575 bp, 688 bp and 516 bp, respectively. The primers are numbered according to the location of their 3′-ends in the reference sequence [2] and "L" and "H" designate the light and the heavy strands of the mtDNA molecule, respectively. The L-strand primers carry at the 5′-end a sequence derived from the universal primer (M13*), whereas the H-strand primers carry a 5′-terminal sequence derived from the reverse primer (Rev*) (see Table 1). The resulting three PCR products cover the entire mitochondrial control region (Fig. 1). The PCR products were separated from residual primers and directly sequenced by cycle sequencing. Both strands were sequenced using fluorescence-labeled primers, the sequences of which correspond to the 5′ ends of the PCR primers (Table 1). The 373A DNA Sequencer (Applied Biosystems) was used for separation and detection of the fluorescence-labeled chain termination products.

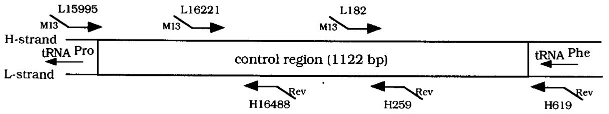

Fig. 1 Amplification and sequencing strategies for the analysis of control regions of human mitochondrial DNAs. The location of all primers used in this study is indicated by the position in the reference sequence [2] that corresponds to the 3′ end of the primer. The orientation of the primers is denoted by arrows. 'L' and 'H' stand for the light and heavy strand of the mtDNA molecule, respectively. 'M13*' and 'Rev*' designate the sequences derived from the universal or reverse primer which were added to the 5′ terminus of the PCR primers to allow for subsequent sequencing with fluorescence-labeled primers

Data analysis

Sequences were aligned and compared with the reference sequence using the Sequence Navigator computer program (Applied Biosystems, Sequence Navigator, version 1.0.1). The results were converted into a Microsoft Excel table (Microsoft Corporation 1995, Microsoft Excel, version 5.0). Each nucleotide substitution, deletion or insertion was registered and summarized in Tables 2, 3, and Figs. 2–4.

Results and discussion

In this study we have determined the nucleotide sequence of the mitochondrial control region from 200 persons who were unrelated in their maternal lineage. Sequence comparison led to the identification of 190 mitochondrial lineages as defined by 202 variable positions. The most frequently occurring lineage was represented by 5 individuals whereas 186 lineages were observed in only one person. Statistical analysis of the number of polymorphic sites per sequenced control region revealed that 61% of the sequences exhibit between four and eight nucleotide substitutions, insertions and/or deletions as compared to the Anderson sequence [2]. Among the 200 samples investigated, we found a single control region carrying the

Table 2 A Nucleotide sequence difference in the major non-coding region of 200 German Caucasian mtDNAs. Sequences were aligned with a reference sequence [2] using the Sequence Navigator computer program (Applied Biosystems, Sequence Navigator, version 1.0.1). The L-strand reading sense is shown. The base number of each marked site is written vertically above the base and corresponds with the base in the reference sequences. Shown are nucleotides at 202 positions on the basis of a comparison of 1192 bp (15996–16569, 1–618) where nucleotide substitutions, deletions, or insertions were observed in at least one individual. The nucleotide in the reference sequence [2] is shown below the order for polymorphic sites. For other sequences, only the differences from the reference sequence are indicated, by letters for base substitutions and by dashes for deletions/insertions. Letters following a number (a–d) indicate additions of nucleotides not found in the reference sequence

	16051	16063	16069	16080	16081	16085	16086	16092	16093	16104	16109	16111	16112	16124	16126	16128	16129	16134	16140	16145	16150	16162	16163	16168	16169	16172	16173	16174	16176	16179	16183	16186	16188	16189	16192	16192a	16193	16201	16205	16209	16213	16214	16215	16216	16221	16222	16223	16224	16230	16231	16234	16239	16245	16249	16255	16256	16259	16261	16263	16265	16269	16270	16278	16286	16288	16290	16291	16292	16293
And.	A	T	C	A	A	C	T	T	T	C	A	C	C	T	T	C	G	C	T	G	C	A	A	C	C	T	C	C	C	C	A	C	C	T	C	–	C	C	C	T	G	C	A	A	C	C	C	T	A	T	C	C	C	T	G	C	C	C	T	A	A	C	C	C	T	C	C	C	A
1			T												C																																																						
2			T												C					A																														C																			
3			T												C					A						C																																T										T	
4	G								C								C																	C																																			
5																																																		C																			
6																	A																														T																						
7																	A																														T																						
8			T												C																																																						
9																																T	C																																T			
10															C				C															C																																			
11																	A									C																					T																						
12																					T																											C																					
13															C								G			C						T		C																																			
14									C						C																																																						
15																	A																														T																						
16			T												C											C																																											
17																											T																				T																				T	
18																		T																																											G								
19															C																																																						
20																																			T																					T						T					T		
21			T												C																																																						
22							C																																									C																					
23							C																																																	T						T				T			
24			T												C					A																														C								T											
25																																																							A								T						
26															C								G									T		C																																			
27																																																C																					
28			T												C																																																T						
29															C																			C																																			
30		C							C						C																																																						
31									C																									C																												T							
32															C								G									T		C																																			
33																																																C																					
34																																															T																				T	
35			T												C																																																						
36													T																					C																																			
37																				A									G																		T																						
38															C																																																						
39			T												C																			C																																			
40																												T						C																												T							
41																																																C					T																
42															C																																																						
43			T												C																																																			T			
44			T												C																																																						
45																	A				A								G																		T																						
46									C																									C	T																											T							
47															C																																																						
48												T																														A												C															
49																	A																	C																																			
50																						G																										C																	C				
51			T												C																																																						
52																G										C																																				T							
53									C						C																																																						
54																																																C																					
55																																		C																													T						
56																																		C																	T																		
57									C																																																		C										
58																																											G																A										
59																																																								T						T	T						
60																																					T																			T						T					T		
61																																																C																					
62															C																																																						
63											C																																																										
64									C																																															T						T							
65																																			T																					T						T							
66																			C																																									G									G
67																	A																																																				
68																																																C																					
69															C																																																						
70	G																					G																																															
71																																																																					
72																						G																										C																					
73																																																C																					
74																																		C		T																												T							
75																																																																					
76								C																																											T																T			
77																																														T																							
78									C																																				T				G																				
79																																			T																			C															
80																																																														T							
81																						G																		C																															
82																																																											C										
83																																																C											C										
84															C																																																						
85																																		C																												T							
86																	A																														T																						
87															C																																																						
88															C																																																						
89																																		C																																			
90																														T																																									
91					G	G			C																																																													
92																	A																																																				
93																																																																					
94																						G																		C																															
95																																												G																									
96																																			T																					T						T							
97																																		C																																			
98																																			T																											T							
99																																																																					

Table 2 A (continued)

	16051	16063	16069	16080	16081	16085	16086	16092	16093	16104	16109	16111	16112	16124	16126	16128	16129	16134	16140	16145	16150	16162	16163	16168	16169	16172	16173	16174	16176	16179	16183	16186	16188	16189	16192	16192a	16193	16201	16205	16209	16213	16214	16215	16216	16221	16222	16223	16224	16230	16231	16234	16239	16245	16249	16255	16256	16259	16261	16263	16265	16269	16270	16278	16286	16288	16290	16291	16292	16293
And.	A	T	C	A	A	C	T	T	T	C	A	C	C	T	T	C	G	C	T	G	C	A	A	C	C	T	C	C	C	C	A	C	C	T	C	–	C	C	C	T	G	C	A	A	C	C	C	T	A	T	C	C	C	T	G	C	C	C	T	A	A	C	C	C	T	C	C	C	
100																																																								T													
101															C																																																							
102								C																															G																														
103																				A																										T																							
104																																																C																					
105																																																								T						T							
106																								T																																													
107																																		C																						T						T							
108																																																																					
109			T												C																																											T											
110																																																																					
111									C																																							C																					
112																																		C																																							
113																										C																																		T													
114																																																																					
115																																																																T						G
116																																						T																																			
117																																																																		C				
118																																																																					
119															C																																											T											
120																																			T							T																															
121																																																																					
122																															C			C																																					G
123																															T		T	C																																					
124										A																																																													
125																									T																																															
126																																																																					
127																																		C																																T			
128																																																																					
129																									T																																															
130																																																				T												T					
131																						G																																																	
132																																																																					
133																																																				T																	
134																						G																																																	
135														C																																																							
136																																																									T												
137																																																																					
138																																									A																																
139																																																																					
140																																																																					
141															C																																											T											
142																																																																					
143												A																																																									
144									C																																																												
145																																																																					G
146																																																																					
147																																																																					
148																																																																					
149																																							G																																		
150																										C																																														
151																																																																					
152																																																																					
153																																																																					
154																																																											C										
155																																																																			T		
156																																																																					
157																						G																																																	
158																																																																					
159																																	A																																								
160																																																																		T			
161																																															T																						
162																																																																					
163																																		C																																							
164																																																																					
165																																																																					
166																																																																					
167																																																														T							
168																																																																					
169																																																C																					
170																																																																					
171																																																																					
172																																																				T																	
173																																																																					
174																																																																					
175																																																																					
176																																																																					
177																																														T																							
178																						G																		C																															
179																																																																					
180																																																																					
181																																													T																								
182																																																																					
183																																																																					
184																																			A																																						
185																																																																					
186																																														T																							
187				G																																																																	
188																																																																					
189																																																																					
190																																																											C										
191																																																																					
192																																																																					
193																																																																					
194																																																																					
195																																																																					
196																																																																					
197																																																																					
198																																																																					
199																																																																					
200																																																																					

Table 2 B

	16294	16296	16298	16299	16301	16303	16304	16311	16316	16319	16320	16324	16325	16327	16343	16354	16355	16356	16357	16362	16366	16368	16390	16391	16398	16399	16401	16426	16480	16482	16519	16527	16536a	00020	00022	00041	00043a	00060	00072	00073	00093	00094	00097	00106	00107	00108	00109	00110	00111	00114	00146	00150	00151	00152	00153	00159	00164	00175	00182	00185	00188	00189	00194	00195	00199	00200	00203	00204	00207
And.	C	C	T	A	C	G	T	T	A	G	C	T	T	C	A	C	C	T	T	T	C	T	G	G	G	A	C	C	A	A	T	C	C	T	T	C	–	T	T	A	A	G	G	G	G	A	G	C	A	C	T	C	C	T	A	T	C	A	C	G	A	A	C	T	T	A	G	T	G
1												C									T		A								C									G																				A	G								
2																																								G												T		C											C				
3																															C									G																		C											
4																				C											C									G														C															
5								C		A																					C									G														C								A/G			C				
6								C											C					A							C									G																						G			C			C	
7																								A							C									G														C											C			C	A
8																					T																			G																				A	G								
9										A								C		C											C									G														C															
10	T	T						C																							C							C		G																								C					
11								C																A							C									G													T												C		A	C	
12																																								G											C																		
13	T		C																							G					C									G																													
14	T	T																													C									G																										G			
15																											A				C									G																									C			C	A
16																																								G																				A	G								
17																											A				C									G																						G	T	C				C	A
18								C										C													C									G														C										C					
19	T	T					C																								C									G																												C	A
20																										G														G																													
21																																								G																				A									
22					T			C																							C									G											C			C										C					
23								C			T															G														G																													
24																																								G												T		C										C					
25	T																														C									G											C				G									C					
26	T																														C									G																								C					
27								C					C																		C									G		A																				G			C				
28																																								G																													
29	T	T	C																												C									G																								C		G			
30																																								G																													
31																						G																		G												T												C				C	
32	T																														C									G														C										C					
33								C																							C									G										T				C															
34																															C									G																						G		C				C	A
35																							A																	G																				A									
36																		C		C											C										G																														
37																															C									G														C															
38	T	T					C																								C									G											C																		
39																																								G																													
40								C																																G											C	T																	
41								C																							C									G																													
42	T	T					C																								C									G																													
43																																								G																				A									
44																																								G																				A									
45																							A																	G													T	C															
46																																								G												T																	
47	T	T					C																								C									G														C															
48														T																										G											C																		
49																	T	C			T										C																																						
50								C																							C									G												T																	
51																																								G																				A									
52										A																					C									G												T																	
53	T						C																								C									G														C															
54								C												C											C									G																								C					
55																															C									G															C									C					
56												C																			C									G																								C					
57								C																							C																																						
58																															C																																						
59																										G														G														C															
60																										G						T																																					
61		T						C																							C									G																								C					
62	T	T					C																								C									G																													
63																															C										G										C		T											C					
64																										G					C									G																													
65																							A			G														G																													
66								C																																																								C					
67																	T	C		C											C																																						
68								C																							C									G											C			C															
69	T	T					C																								C									G																													
70						A																									C									G																													
71			C																												C								C	G																								C					
72								C																							C									G												T																	
73								C																							C									G														C															
74																									A															G												T																	
75																															C												A	–	–	–	–	–	–	T																			
76												C																			C									G																								C					
77																															C																																						
78																															C																																						
79															G																									G												T																	
80																															C									G														C										C					
81																															C									G																	T												
82									G																						C																																						
83																															C									G														C															
84							C																								C									G													T																
85																															C									G												T																	
86																								A							C			G						G																									C			C	
87	T						C																								C									G																		C											
88	T	T					C																								C									G																													
89	T						C																									T																																		G			
90																		C													C									G																								C					
91							C																																																														
92																	T	C		C											C																																						
93															G								A								C									G												T																	
94																															C									G														C															
95			C																												C								C																														
96																										G														G																													
97																															C																				C																		
98																																								G												T								A									
99	T						C				T																																											C															

Table 2B (continued)

	16294	16296	16298	16299	16301	16303	16304	16311	16316	16319	16320	16324	16325	16327	16343	16354	16355	16356	16357	16362	16366	16368	16390	16391	16398	16399	16401	16426	16440	16482	16519	16527	16536a	00020	00022	00041	00043a	00060	00072	00073	00093	00094	00097	00106	00107	00108	00109	00110	00111	00114	00146	00150	00151	00152	00153	00159	00164	00175	00182	00185	00188	00189	00194	00195	00199	00200	00203	00204	00207
And.	C	C	T	A	C	G	T	T	A	G	C	T	T	C	A	C	C	T	T	T	C	T	G	G	G	A	C	C	A	A	T	C	C	T	T	C	–	T	T	A	A	G	G	G	G	A	G	C	A	C	T	C	C	T	A	T	C	A	C	G	A	A	C	T	T	A	G	T	G
100													C																										G																				A									
101	T	T																												C									G																													
102																															C		C																																				
103							C																																																													
104								C																																G											C			C															
105																										G														G																													
106																															C						C																	C															
107																										G														G																													
108																															C									G																								C					
109																																								G																				A		G							
110															G																								G												T																	
111																										G														G																													
112								C																											C																																		
113																															C																															G							
114																															C																																	C					
115																															C																																						
116			C																																																													C					
117																T															C																																						
118																				C																T																																	
119																																																																					
120							C																																																													
121																													G		C																																						
122																															C																							C															
123																															C																																						
124			C																																																	T																
125				G																											C																																						
126								C																								C																																						
127																															C																							C															
128	T						C																																																													
129																														G																																							
130																															C																																						
131																															C									G																													
132								C																								C																																						
133																															C									G																													
134			C																								A																																									
135							C	C																																																													
136																																																						C					T										
137							C																																																C													
138			C																																				C																														
139																															C																																						
140																															C																																						
141																															C																																						
142																															C																							C															
143																															C																																						
144								C																								C																																						
145								C								T																																														C					
146																															C																																						
147							C																																																													
148																															C																																						
149																															C																													A	G								
150								C																																																												
151			C																																				C																														
152								C																								C																				C																		
153																															C																				C																		
154																															C																																						
155																															C																																						
156			C																																				C																														
157																															C									G																													
158															G																																																						
159																															C																																						
160																															C																							C															
161																															C																							C															
162																				C										G																																							
163																		C													C																																						
164							C																																																													
165							C																																																													
166																															C																							C															
167																																																																					
168																		C													C																																						
169								C			T																																																										
170																																																																					
171																															C																																						
172																															C																																						
173			C																												C								C																														
174																																																						C															
175	T						C																																																													
176								C									T																																																			
177																															C																																						
178																															C									G																													
179																															C																																						
180																															C																							C															
181																																																																					
182																															C																																						
183																															C																																						
184								C																																																												
185																															C																																						
186																															C																																						
187																		C																																																		
188																																																																					
189	T																														C																																						
190																																							C																														
191																																																						C															
192																															C																																						
193																															C																																						
194																															C																																						
195																															C																																						
196																															C																																						
197																																																																					
198																																																																					
199																																																																					
200																																																																					

Table 2C

	0208	0215	0217	0225	0226	0227	0228	0235	0239	0242	0250	0256	0257	0258	0260	0263	0269	0271	0272	0285	0291	0295	0299	0303	0309a	0309b	0310a	0315a	0315b	0317	0319	0321	0324	0327	0330	0334	0340	0365	0408	0418	0438	0439a	0456	0462	0466	0477	0482	0489	0494	0497	0498	0499	0508	0513	0514	0516	0522	0523a	0523b	0523c	0523d	0527	0537	0547	0562	0573a	0573b	0573c	0574	0599
And.	T	A	T	G	T	A	G	A	T	C	T	C	A	C	G	A	C	C	A	C	A	C	C	C	–	–	–	–	–	C	T	T	C	C	C	T	C	A	T	C	C	–	C	C	T	T	T	T	C	C	C	G	A	G	C	C	C	–	–	–	–	C	C	A	A	–	–	–	A	A
1							A									G						T			C			C							G									T				C									–													
2		G														G						T					T	C			C																	C						A																
3										T						G						T			C			C																T				C																						
4			C													G					T				C			C									T																G					C	A											
5																G												C		G			G		G																							C	A	C	A									
6											C					G									C	C		C														T																												
7									C		C					G									C			C																														C	A	C	A									
8							A									G					T				C			C															T				C									–													
9								G								G									C			C																													–													
10																G					T				C			C																																										
11											C					G												C																																										
12																G												C												T										T					G			C	A	C	A	G			G					G
13																G									C			C																																										
14																G									C			C																					T														T	T						
15									C		C					G						T						C															T																											
16							A									G									C			C																T				C																						
17																G												C																																										
18																G									C			C																								A																	G	
19																G									C			C										G																																
20																G						T		A	C			C							G						T																													
21							A									G									C			C							G									T				C																						
22																G												C																							–																			
23																G	A											C																																										
24																															C																	C						A																
25				A		G										G												C																																										
26																G									C			C																																										
27																G						T			C			C																																										
28							A									G												C																T				C																		C				
29																G						T						C																																										
30							A									G									C			C																T				C																						
31												A				G												C																																										
32																									C			C																																										
33																G			G									C																						T								C	A	C	A									
34																G						T			C			C																																										
35							A									G												C																T				C																						
36																G									C			C																													–													
37																G									C			C																																										
38																G									C			C																																										
39							A									G									C			C																T				C																						
40																G												C																												T														
41																G									C	C		C		T																				T																				
42																G									C			C				C																																						
43																G						T						C																T			C	C																						
44							A																		C			C																T				C																						
45														A														C																																										
46																G									C			C																														C	A											
47																G									C			C																																										
48																G				T								C					G																																					
49																G									C	C		C																													–													
50																G						T						C																						T																				
51																G									C			C																T				C																						
52																G												C																	C																									
53																G									C			C																																										
54																G									C			C																														C	A											
55				A	C											G												C																																										
56																G									C			C																																										
57																G							A		C			C																		C																								
58																G									C			C							G																							C	A			G								
59																G									C			C																																										
60						G										G												C																																										
61																G												C																						T								C	A											
62																G												C																																						C				
63																G									C			C																																										
64																G									C			C																																										
65																G												C																																						C	C	C		
66																G									C			C																														C	A											
67																G									C			C																													–													
68																G												C																													–													
69																G												C																																										
70																G									C			C																																										
71																G									C			C																																										
72																												C																						T																				
73																G									C			C																																										
74																G												C																																										
75																G									C			C											A																															
76																												C																																										
77																G									C			C		G			G		G																																			
78																G									C			C																																										
79																G									C	C		C																																										
80																G												C																								A																		
81																G									C			C																																										
82							A									G									C			C																		C																								
83																G												C																						T																				
84																G												C																																										
85																G									C			C																																										
86																																																																						
87																G												C																																										
88																	A											C																																										
89																G												C																T																										
90																G												C																								A																		
91																G												C												T																	–													
92																G												C																													–													
93																G												C																																										
94																G												C																																										
95																G									C			C																																										
96																G												C																																										
97																G		G							C			C																																										
98																G												C																																										
99																G												C																T																										

Table 2 C (continued)

	0208	0215	0217	0225	0226	0227	0228	0235	0239	0242	0250	0256	0257	0258	0260	0263	0269	0271	0272	0285	0291	0295	0299	0303	0309a	0309b	0310a	0315a	0315b	0317	0319	0321	0324	0327	0330	0334	0340	0365	0408	0418	0438	0439a	0456	0462	0466	0477	0482	0489	0494	0497	0498	0499	0508	0513	0514	0516	0522	0523a	0523b	0523c	0523d	0527	0537	0547	0562	0573a	0573b	0573c	0574	0599
And.	T	A	T	G	T	A	G	A	T	C	T	C	A	C	G	A	C	C	A	C	A	C	C	C	–	–	–	–	–	C	T	T	C	C	C	T	C	A	T	C	C	–	C	C	T	T	T	T	C	C	C	G	A	G	C	C	C	–	–	–	–	C	C	A	A	–	–	–	A	A
100																G									C			C																																										
101																G												C																																										
102																G												C	C																C																									
103																G									C	C		C															T																											
104																G												C																																										
105																G									C			C																																										
106																G									C			C																																										
107																G												C																																										
108																G									C			C																								A																		
109							A																																																															
110																G												C															T															C	A												
111																G												C																						T																				
112																G									C	C		C																													–													
113																G									C			C																																										
114																G									C			C																		C											–													
115																G									C			C																																										
116																G									C			C																																										
117																G									C			C																																										
118									C							G									C	C		C																																										
119																G									C			C							G																																			
120																G												C												T																														
121																G									C			C											A																															
122																G																																																							
123																G												C																																										
124																G									C			C																																										
125																G									C			C																																										
126																G									C			C						T																																				
127																G												C																																										
128																G									C	C		C															T																											
129									C							G									C			C																																										
130																G									C			C																																										
131																G									C			C																																										
132													G			G												C																		C																								
133																G									C			C																																										
134																G									C	C		C																																										
135																G												C															T																											
136																G												C																													–													
137																G												C															T														–													
138																G									C			C																																										
139																G									C			C																		C																								
140																G									C			C																																						C	C			
141																									C			C																																										
142																G												C																														C	A												
143																G									C			C																																										
144																G												C																																										
145																									C																																															
146																G									C			C																														C	A												
147																G									C			C															T																											
148																G									C			C																																						C				
149																G																																																							
150																G									C			C																																										
151																									C			C								C																																				
152																G												C																																										
153																G									C			C																																										
154																G												C																		C																								
155																G									C			C																																										
156																G									C			C																																										
157																G												C																																										
158																G									C			C																														C	A												
159																G									C			C																																										
160																G												C																																										
161																G												C																																										
162									C							G												C																																										
163																G												C																																										
164																G												C												T																	–													
165																G												C															T														–													
166																G												C																													–													
167															A	G												C																													–													
168																G												C																													–													
169																G												C																																										
170																G									C	C		C		T																																								
171																G									C			C																																										
172																G												C																																										
173																G																																																							
174																G									C	C		C																																										
175																G																																										C	A	C	A									
176																G												C																																										
177																G												C																																										
178																																																																							
179																G									C	C		C																																										
180																G												C																																										
181																G									C			C																																										
182																G									C			C																																										
183																G									C			C																																										
184																G												C																																										
185																G									C			C																																										
186																G												C																																										
187																									C			C																																										
188	C															G									C			C																																										
189																G												C																																										
190																G																														C																								
191																G												C																																										
192																G												C																																										
193																G												C																																										
194																G												C																																										
195																G												C																																										
196																G												C																																										
197																G												C																																										
198																																																																							
199																																																																							
200																																																																							

Table 3 Occurrence and distribution of nucleotide substitutions, deletions and insertions in the hypervariable regions I to III (HV I–III) on the L-strand of the control region of the mitochondrial DNA from 200 maternally unrelated persons from Germany. (Pu: purine-base, Py: pyrimidine-base, A: adenine, C: cytidine, G: guanine, T: thymine)

Mutation type	HVI		HVII		HVIII	
	Number of positions	Total number of mutations	Number of positions	Total number of mutations	Number of positions	Total number of mutations
Substitutions						
Transitions						
Py-Py						
C–T	38	159	12	40	7	30
T–C	23	204	14	104	4	20
Pu-Pu						
A–G	14	31	12	299	3	2
G–A	6	21	8	34	2	3
Total	81	415	46	477	16	55
Transversions						
C–A	5	5	5	6	0	0
A–C	2	2	1	3	0	0
C–G	4	6	4	13	2	3
G–C	1	1	0	0	0	0
T–G	0	0	0	0	0	0
G–T	0	0	0	0	0	0
A–T	1	1	1	3	1	1
T–A	1	1	0	0	0	0
Total	14	16	11	25	3	4
Insertions						
+C	0	0	2	276	1	0
+2C	0	0	2	13	1	0
+3C	0	0	0	1	1	1
+T	1	1	1	0	1	0
+CA	0	0	0	0	1	11
+2CA	0	0	0	0	1	2
Total	1	1	5	290	6	14
Deletions						
–G	0	0	3	3	0	0
–A	0	0	2	2	0	0
–C	0	0	1	1	1	1
–CA	0	0	0	0	1	18
Total	0	0	6	6	2	19

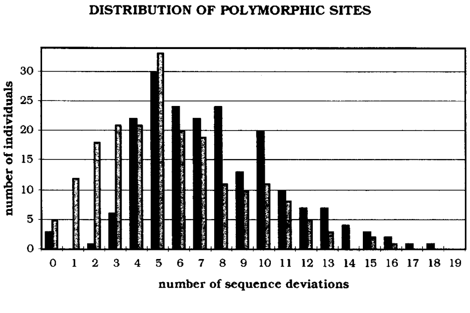

Fig. 2 Number of sequence polymorphisms in the human mitochondrial control region as compared to the Anderson reference sequence [2] (black bars) or a corrected reference sequence (gray bars). X-axis: Number of polymorphisms per control region; Y-axis: Number of individuals carrying a given number of polymorphisms; sample size: n = 200

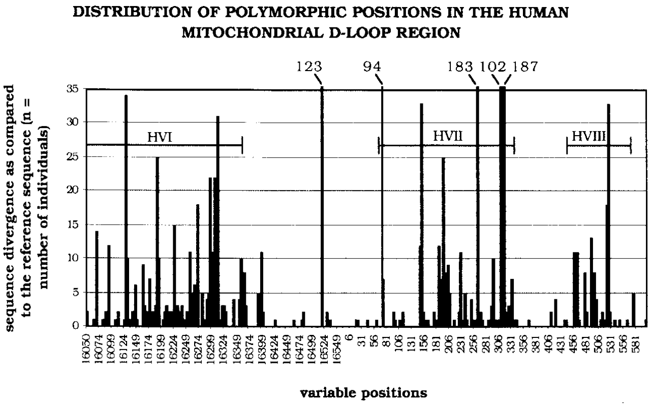

Fig. 3 Number and distribution of polymorphic positions in the human mitochondrial control region among 200 maternally unrelated persons. X-axis: Nucleotide position in the control region. Nomenclature is in accordance with Anderson et al. [2]. Y-axis: Number of individuals who show certain deviations from the Anderson sequence. The scale was limited to n = 35 in order to allow for a better comparison of low values. Values above n = 35 are given above the respective bars

maximum number of 18 deviations from the standard sequence. On the other hand, three sequences were found which showed no change at all and were thus (at least in this region) identical with the Anderson sequence [2]. A

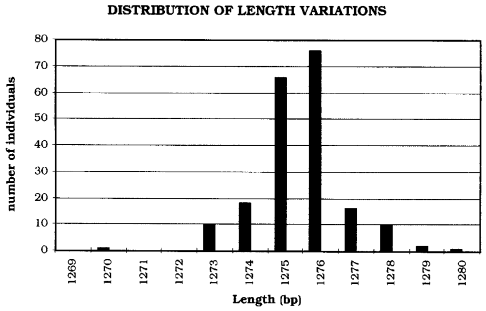

Fig. 4 mtDNA length polymorphisms in the human mitochondrial control region. X-axis: length (in bp); Y-axis: number of individuals; sample size: 200

distribution of the number of deviations per control region is shown in Fig. 2 (black bars). Each sequence was aligned with the original Anderson sequence [2] and all deviations were recorded (Table 2). As compared with our data set, the Anderson sequence is atypical for three nucleotide positions: (I) at position 16,519, cytosine was normally present instead of thymine, (II) at position 263, guanine was normally present instead of adenine, and (III) at positions 311–315 the majority of sequences carried 6 cytosine nucleotides instead of the 5 reported by Anderson. When a consensus sequence corrected at these three positions was used as a reference instead of the Anderson sequence, a slightly different distribution was obtained as indicated in Fig. 2 by gray bars.

The distribution of polymorphic sites across the mitochondrial control region (Fig. 3) clearly shows that the non-coding control region of the human mitochondrial DNA is a highly polymorphic region. Two hypervariable regions have been described at HVI (positions 16024–16365) and HVII (positions 73–340) [7, 20]. Our data provide evidence for the existence of an additional hypervariable region located around positions 438–574. The highest density of polymorphic sites was obtained for hypervariable region I which contains 88 variable positions in total length of 342 bp (26%) and HVII displays 65 mutable sites in 268 bp (24%). HVIII exhibits a slightly lower variability with 25 polymorphic sites in 137 bp (18%). In contrast, the segments located between the hypervariable regions show nucleotide substitution rates of only 7% (positions 16366–16569, 1–72) or 3% (positions 341–437).

A comparison of our sequence data with the Anderson sequence [2] in the region studied here revealed 202 sites where nucleotide substitutions, insertions or deletions had taken place. At 92 positions, the changes are unique i.e. only one of the mtDNAs investigated differed from the reference sequence. At the remaining 110 sites the changes are not unique i.e. there is a base difference in at least two lineages. The majority of polymorphic sites are affected by only one mutation type. At only 14 out of 202 positions (7%) two or more different mutation events were observed.

Table 3 shows a distribution of the sequence polymorphisms among the mutation types for each hypervariable region. Sequence deviations caused by nucleotide substitutions predominate over insertions and deletions (90%). Transitions make up the majority of the nucleotide substitutions (75%). Transversions (15%), insertions (6%) and deletions (4%) were observed with significantly lower frequency. This excessive amount of transitions may indicate that mispairing during replication is the major source of spontaneous mutations in mitochondrial DNA [21]. Among the transitions, pyrimidine substitutions are predominantly found and C–T transitions occur with particular high frequency. Among the transversions there is no obvious preference for a particular type. However, it is striking that on the L-strand not a single G–T change has yet been found.

Owing to the low number of insertion and deletion events in the control region, size variations of the human mitochondrial genome should be limited to very few nucleotides. Indeed length variations usually occur only to a very limited extent in this region of the human mitochondrial genome. Figure 4 shows their distribution in the D-loop containing region of the control regions investigated as compared to the reference sequence [2]. Using the PCR primer pair L15995/H619, 9 different fragment sizes with lengths between 1270 bp and 1280 bp were obtained among the 200 mtDNA control regions. The Anderson sequence yields a PCR product of 1274 bp with these primers.

In the following some of the insertions and deletions that cause length polymorphisms as compared to the Anderson sequence will be described.

Insertions of C residues were described by Torroni et al. [22]. These insertions are located within a stretch of six C's between nt 568 and nt 573 (+C, +2C, +3C), between nt 303 and nt 309 (+C, +2C) or within five C's between nt 311 and 315 (+C, +2C) in the control region. In our investigations, we found in the sequence motif at nt 568–573 three sequences with 7, one sequence with 8, and one sequence with 9 cytidine residues. At positions 303–309, 90 sequences with an insertion of a single C residue and 12 sequences with an insertion of two C residues were identified. Finally, at nt 311–315, 186 control regions with 6 and 1 with 7 cytidine residues were found. When a consensus sequence corrected at positions 311–315 (6 cytidine residues) was used as reference only 13 control regions deviated in this sequence from the standard.

The CA-repeat at position 514–523 (–CA, +CA, +2CA) was described by Bodenteich et al. [23]. Screening the 200 unrelated individuals we found 18 sequences with $(CA)_{5-1}$, 167 with $(CA)_5$, 10 with $(CA)_{5+1}$ and 5 with $(CA)_{5+2}$ repeats.

A deletion at positions 106–111 (–GGAGCA) was described [24, 25] as an mtDNA marker for a Chibcha-speaking Amerindian tribe, the Huetar. This marker is widely used in Amerindian taxonomic research. Surprisingly, in our study we identified a sequence which showed exactly this deletion. A relationship between this family and an Amerindian tribe could be ruled out for at least the last 8 generations (back to the seventeenth century) by pedigree analysis. It therefore appears likely that this mutation arose independently. The 6-bp deletion shares sequence homology with the preceding five base pairs. Thus it seems possible that this length polymorphism was generated through slipped misreplication [26].

The mtDNAs of different individuals do not mix upon fertilization and mtDNAs usually do not recombine. In contrast to genomic DNA, which is partitioned in separate molecules the chromosomes, the mtDNA represents a single locus. It could therefore be envisaged that different mutation events are dependent on each other. For example, the nucleotide substitutions at positions 16294 and 16296 are frequently coupled (in 11 of 22 cases). For this reason, it is not possible to calculate the frequency of the occurrence of a given mitochondrial genotype by simply multiplying the probabilities of the mutations. The mitochondrial control region has to be viewed as a single locus. At best, the hypervariable segments could be treated independently of each other [27]. However, as long as it has not been proved that the individual regions are independent of each other and the mtDNA has to be regarded as a single locus it seems resonable to consider the mutation events of the entire control region when preparing a mtDNA profile.

For the reasons outlined the availability of a large database is central to the calculation of probabilities for the occurrence of identical mitochondrial control regions among unrelated individuals. It thus seems useful to collect sequence data for mitochondrial control regions as already proposed by Miller et al. [27].

From out data set obtained in this study we can estimate the genetic diversity ($h = n(1 - \Sigma\chi^2)/(n - 1)$, when n is the sample size and χ is the frequency of each mtDNA type [28]) of 0.99. The probability of two randomly selected individuals from a population having identical mtDNA types ($P = \Sigma\chi^2$) is 0.6%.

Acknowledgements This work was supported by the Deutsche Forschungsgemeinschaft (Po 579/2–1).

References

1. Bogenhagen D, Clayton DA (1974) The number of mitochondrial deoxyribonucleic acid genomes in mouse L and human HeLa cells. J Biol Chem 249 : 7991–7995
2. 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 organization of the human mitochondrial genome. Nature 290 : 457–464
3. Clayton DA (1982) Replication of animal mitochondrial DNA. Cell 28 : 693–705
4. Clayton DA (1984) Transcription of the mammalian mitochondrial genome. Annu Rev Biochem 53 : 573–594
5. Upholt WB, Dawid IB (1977) Mapping of mtDNA of individual sheep and goats: rapid evolution in the D-loop region. Cell 11 : 571–583
6. Brown WM, George MJ, Wilson AC (1979) Rapid evolution of animal mitochondrial DNA. Proc Natl Acad Sci USA 76 : 1967–1971
7. Wilson MR, Stoneking M, Holland MM, DiZinno JA, Budowle B (1993) Guidelines for the use of mitochondrial DNA sequencing in forensic science. Crime Lab Digest 20 : 68–77
8. Wilson MR, DiZinno J, Polanskey D, Replogle J, Budowle B (1995) Validation of mitochondrial DNA sequencing for forensic casework analysis. Int J Legal Med 108 : 68–74
9. 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
10. Stoneking M (1994) Mitochondrial DNA and human evolution. J Bionerg Biomembr 26 : 251–259
11. Melton T, Stoneking M (1996) Extent of heterogeneity in mitochondrial DNA of ethnic Asian populations. J Forensic Sci 41 : 591–602
12. 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 analyses of the Tyrolean ice man. Science 264 : 1775–1778
13. Ginther C, Issel-Tarver L, King M-C (1992) Identifying individuals by sequencing mitochondrial DNA from teeth. Nat Genet 2 : 135–138
14. Sullivan KM, Hopgood R, Gill P (1992) Identification of human remains by amplification and automated sequencing of mitochondrial DNA. Int J Legal Med 105 : 83–86
15. Holland MM, Fisher DL, Mitchell LG, William MD, Rodriquez WC, Canik JJ, Merril CR, Weedn VW (1993) Mitochondrial DNA sequence analysis of human skeletal remains: identification of remains from the Vietnam war. J Forensic Sci 38 : 542–553
16. Boles TC, Snow CC, Stover ES (1994) Forensic DNA testing on skeletal remains from mass graves: a pilot project in Guatemala. J Forensic Sci 40 : 349–355
17. Gill P, Ivanov PL, Kimpton C, Piercy R, Benson N, Tully G, Evett I, Hagelberg E, Sullivan K (1994) Identification of the remains of the Romanov family by DNA analysis. Nat Genet 6 : 130–135
18. Ivanov PL, Wadhams MJ, Roby RK, Holland MH, Weedn VW, Parsons TJ (1996) Mitochondrial DNA sequence heteroplasmy in the Grand Duke of Russia Georgij Romanov establishes the authenticity of the remains of Tsar Nicholas II. Nat Genet 12 : 417–420
19. Lutz S, Weisser H-J, Heizmann J, Pollak S (1996) mtDNA as a tool for identification of human remains. Int J Legal Med 109 : 205–209
20. Greenberg BD, Newbold JE, Sugino A (1983) Intraspecific nucleotide sequence variability surrounding the origin of replication in human mitochondrial DNA. Gene 21 : 33–49
21. Thomas WK, Beckenbach AT (1989) Variation on Salmonid mitochondrial DNA: evolutionary constraints and mechanisms of substitution. J Mol Evol 29 : 233–245
22. Torroni A, Lott MT, Cabell MF, Chen YS, Lavergne L, Wallace DC (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 5 : 760–776
23. Bodenteich A, Mitchell LG, Polymeropoulos MH, Merril CR (1992) Dinucleotide repeat in the human mitochondrial D-loop. Hum Mol Genet 1 : 140
24. Santos M, Barrantes R (1994) D-loop mt DNA deletion as a unique marker of Chibchan Amerindians. Am J Genet 55 : 413–414
25. Santos M, Ward RH, Barrantes R (1994) mtDNA variations in the Chibcha Amerindian Huetar from Costa Rica. Hum Biol 66 : 963–977
26. Krawczak M, Cooper DN (1991) Gene deletions causing human genetic disease: mechanisms of mutagenesis and the role of the local DNA sequence environment. Hum Genet 86 : 425–441
27. Miller KWP, Dawson JL, Hagelberg E (1996) A concordance of nucleotide substitutions in the first and second hypervariable segments of the human mtDNA control region. Int J Legal Med 109 : 107–113
28. Tajima F (1989) Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. Genetics 123 : 585–595