Int J Legal Med (2000) 114 : 130–132

SHORT COMMUNICATION

M. Crespillo · J. A. Luque · M. Paredes
R. Fernández · E. Ramírez · J. L. Valverde

Mitochondrial DNA sequences for 118 individuals from northeastern Spain

Received: 12 July 1999 / Accepted: 11 April 2000

Abstract A population database was generated from 118 unrelated Caucasoid individuals living in Spain. Sequence polymorphisms of the mitochondrial DNA (mtDNA) control region, hypervariable regions I and II (HVRI and HVRII) were determined using the polymerase chain reaction (PCR) and direct sequencing. A total of 102 different sequences were found as defined by 105 variable positions. The most common sequence occurred six times, and this sequence is also the most frequent in other European populations such as Austria, Germany and Britain. The mean pair-wise difference for the two HVR regions taken together was 7.74. The study revealed that transitions made up the majority of the variations (88%), whereas we observed a significantly lower frequency of transversions (8%). Also one individual in this study was observed with two positions of heteroplasmy at nucleotides 150 (C/T) and 153 (G/A). A statistical estimate of the results for this population showed a genetic diversity of 0.99. The probability of two random individuals showing identical mtDNA haplotypes is 1.3%. In order to use the mtDNA analysis in forensic casework, we consider that it is of crucial importance to know the frequency of the different sequences of mtDNA, and this data base study could be a useful tool to statistically evaluate the results.

Key words Mitochondrial DNA · Polymerase chain reaction · Forensic DNA typing · Population study · Automatic sequencing

Supplementary material Data on the mtDNA haplotypes found in the population from northeastern Spain (Table S1) are available in electronic form on Springer-Verlag's server at http://link.springer.de/link/service/journals/00414/index.htm

M. Crespillo (✉) · J. A. Luque · M. Paredes · R. Fernández
E. Ramírez · J. L. Valverde
Sección de Biología. Instituto Nacional de Toxicología.
Ministerio de Justicia, Merced 1, 08002 Barcelona, Spain
e-mail: biolog@bcn.inaltox.es; Fax: +34-93-3182530

Introduction

The important development of the polymerase chain reaction (PCR) [1], which enables samples with small amounts or even degraded DNA to be typed, and the high number of mtDNA copies per cell [2] that allows an increase in the probability of obtaining a DNA typing result, make the analysis and sequencing of mtDNA a powerful and validated method and provide forensic science with a method to type samples which contain minimal amounts of genomic DNA, such as bones, hair shafts and faeces [3, 4, 5, 6].

During the last 3 years, we have used mtDNA analysis in our laboratory in routine forensic casework where conventional DNA typing has not worked. These cases mainly involved typing naturally shed hairs or hairs without roots and to confirm the identification of highly decomposed bodies and skeletal remains.

The present study presents a mtDNA database for the regions HVRI and HVRII, from unrelated Caucasoid individuals living in Spain. This database is being used to evaluate the results of mtDNA analysis in our laboratory.

Materials and methods

DNA was obtained from casework reference samples (blood, saliva, hair) of 118 unrelated Caucasoid individuals living in the northeast of Spain, and was extracted by a standard phenol-chloroform method [7]. The mtDNA regions subjected to analysis were 16024–16365 (HVRI) and 73–340 (HVRII). Amplification reactions were performed as described previously [8] employing a GeneAmp System 2400 thermal cycler (Perkin-Elmer, Norwalk, Conn.).

The amplified products were purified by filtration in Centricon-100 tubes (Amicon ,Beverly, Mass.) following the manufacture's recommendations. Sequencing fragments were generated using the Taq DyeDeoxy Terminator Cycle Sequencing Ready Reaction kit (Perkin-Elmer, Foster City, Calif.) and primers identical to those employed for amplification, according to the protocol provided with the kit. DNA products were purified by ethanol precipitation and the fragments were analysed using an automated DNA sequencer ABI PRISM 310 (PE/ABD). Each template was sequenced in both directions and the consensus sequence was aligned and compared with the reference sequence [9] using the software SeqEd V. 1.0.3 (PE/ABD).

The analysis of data was performed using Microsoft access 2.0. The pair-wise comparison was calculated with an access basic routine developed in our laboratory.

We calculated the probability of genetic identity using the formula $p = \sum\chi^2$ where χ is the frequency of each mtDNA haplotype and the genetic diversity employing the algorithm $h = n(1 - \sum\chi^2)/(n - 1)$, where n is the sample size and χ is the frequency of each mtDNA haplotype [10].

Results and discussion

Table S1 shows the distribution of sequence polymorphisms compared to the Anderson sequence.

For the HVRI region, 61 variable sites were detected and 84 different haplotypes were observed in 118 individuals, of which 73 were unique and the rest were found in at least 2 individuals. In this region 17% of the haplotypes presented were identical with the Anderson sequence which is similar to other European Caucasian databases e.g. 20% for Austria [11] and 14% for Britain [12]. For the HVRII region, 44 variable sites were detected and 66 different haplotypes were observed in 118 individuals of which 51 were unique. A total of 102 different haplotypes were determined from both regions I and II as defined by 105 variable positions (Table S1).

The mean pair-wise differences for our data were 4.34 for the region HVRI, 3.39 for the region HVRII, and 7.74 for the two regions taken together (Table 1). This data is similar to other Caucasian populations e.g. 8.45 for British [12] or 8.38 for French [13].

Table 2 shows the distribution of nucleotide substitutions. The majority of mutations in both regions were transitions (88%) which is in agreement with previously published data [14, 15], and a lower number of transversions were obtained (8%), whereas the insertions and deletions were rather rare. We detected three positions presenting length heteroplasmy. All deletions observed (three different positions) were detected in the same individual (individual 43) at positions 249, 290 and 291 in the region HVRII. Our analysis revealed that the majority of observed transitions were pyrimidines, with a pyrimidine to purine ratio of 2.3:1.

The majority of polymorphic sites presented only 1 mutation type, however 2 different mutation events were observed at 3 out of 105 positions (16075, 16126, and 185).

Homopolymeric tracts of poly-C bases occurred in both HVRI and HVRII regions, and a length heteroplasmy was often observed due to the presence of several populations of mtDNA molecules differing in the number of cytosines. When a transition from thymine to cytosine at position 16189 was present, a polycytosine stretch occurred in the region HVRI [16]. However if an additional transition from cytosine to thymine takes place between the positions 16184–16193 no length heteroplasmy sequences were observed, except for one individual (individual 116) who presented a double transition at positions 16189 and 16193, from cytosine to thymine and from thymine to cytosine, respectively. In this particular case we found a length heteroplasmy sequence between positions 16183–16191. Our data showed that approximately 16% of the individuals presented a mixture of different length variants for a heteroplasmic situation in region HVRI. Likewise, in the HVRII region an insertion of at least one cytosine (compared with the reference Anderson sequence) sometimes occurred between positions 303–310. We found a large number of sequences (41%) in which the number of cytosines present was impossible to clearly define, and this sequence alteration resulted in a length heteroplasmy.

Table 1 The mean pair-wise differences, number of different haplotypes, variable sites and sequences that were found once for HVRI, HVRII and both region taken together

	HVRI	HVRII	HVRI + HVRII
Number of different haplotypes	84	66	102
Number of variable sites	61	44	105
Number of unique haplotypes	73	51	94
Mean pair-wise differences	4.34	3.39	7.74

Table 2 Observed nucleotide substitutions, deletions, insertions and heteroplasmic events found in the HVRI and II on the L-strand of the control region of mitochondrial DNA for 118 Spanish Caucasian from northeastern Spain. (*Pu* purine base, *Py* pyrimidine, *A* adenine, *C* cytosine, *G* guanine, *T* thymine)

Mutation type	Number of position	Total number of mutations
Transitions		
Py-Py		
C-T	34	158
T-C	31	189
Pu-Pu		
A-G	17	217
G-A	11	44
Total	93	608
Transversions		
A-C	2	14
C-G	1	1
G-C	1	1
T-G	2	2
G-T	1	3
A-T	1	1
Total	8	22
Insertions		
+C	3	50
+2C	1	12
+3C	1	1
Total	5	63
Deletions		
–A	3	3
Total	3	3
Heteroplasmy		
Position	2	2
Length	3	68

Table 3 Haplotypes observed in the Spanish Caucasian population that were found at least in two individuals

Haplotypes	*n*
263(G), 315.1(C)	6
263(G), 309.1(C), 315.1(C)	5
16189(C), 263(G), 315.1(C)	3
16069(T), 16126(C), 16278(T), 73(G), 185(A), 188(G), 228(A), 263(G), 295(T), 309.1(C), 315.1(C)	2
16145(A), 16222(T), 263(G), 315.1(C)	2
263(G), 309.1(C), 309.2(C), 315.1(C)	2
16069(T), 16126(C), 16193(T), 73(G), 150(T), 152(C), 263(G), 295(T), 315.1(C)	2
16093(C), 16224(C), 16311(C), 73(G), 195(C), 263(G), 315.1(C)	2

In this study we found one individual (individual 73) who presented two heteroplasmic positions in region HVRII. A position was considered as heteroplasmic if we observed a secondary peak of more than about 40% peak height below the primary peak, and could be confirmed in the reverse sequencing reaction [17]. A mixture of cytosine and thymine was detected at position 150 while at position 153 we observed adenine and guanine residues. A second extraction, amplification and cycle sequencing of the sample confirmed this heteroplasmy, although it could not be determined in our study whether this double mutation was inherited from the maternal lineage.

Table 3 shows the haplotypes observed in at least two individuals in our population. The most frequent haplotype, in concordance with other Caucasian populations, was the sequence defined by the transition of an adenine at position 263 to guanine and the insertion of a cytosine between position 311 and 315. In our study this sequence was found in approximately 5% and is also the main haplotype in Austrians (3%) [11] and British Caucasians (4%) [12].

We calculated a probability of genetic identity of 0.013 and a genetic diversity of 0.99 using the formulae previously described. These results suggest that sequence polymorphism of the mtDNA control region would be very useful in forensic practice to confirm individual identification.

References

1. Saiki RK, Scharf S, Faloona F, Mullis KB, Horn GT, Erlich HA, Arnheim N (1985) Enzymatic amplification of β-globin genomic sequences and restriction site analysis for diagnosis of sickle cell anemia. Science 230: 1350–1354
2. Robin DE, Wong R (1988) Mitochondrial DNA molecules and virtual number of mitochondria per cell in mammalian cells. J Cell Physiol 136: 507–513
3. Holland MM, Fisher DL, Mitchell LG, William MD, Rodriguez 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
4. Wilson MR, Polanskey D, Butler J, DiZinno JA, Replogle J, Budowle B (1995) Extraction, PCR amplification and sequencing of mitochondrial DNA from human hair shafts. Biotechniques 18: 662–669
5. Hopwood AJ, Mannucci A, Sullivan KM (1996) DNA typing from human faeces. Int J Legal Med 108: 237–243
6. Holland MM, Parsons TJ (1999) Mitochondrial DNA sequence analysis-validation and use for forensic casework. Forensic Sci Rev 11: 22–50
7. Sambrook J, Fritsch EF, Maniatis T (1989) Molecular cloning: a laboratory manual, 2nd edn. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY
8. Wilson MR, DiZinno JA, Polanskey D, Replogle J, Budowle B (1995) Validation of mitochondrial DNA sequencing for forensic casework analysis. Int J Legal Med 108: 68–74
9. Anderson S, Bankier AT, Barrel BG, Bruijin MHL de, Coulson AR, Drouin IC, Eperon I C, Nierlick DP, Roe BA, Sanger F, Schreider PH, Smith AJH, Staden R, Young IG (1981) Sequence and organization of the mitochondrial genome. Nature 290: 457–465
10. Tajima F (1989) Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. Genetics 123: 585–595
11. Parson W, Parsons TJ, Scheithauer R, Holland MM (1998) Population data for 101 Austrian Caucasian mitochondrial DNA d-loop sequences: application of mtDNA sequence analysis to a forensic case. Int J Legal Med 111: 124–132
12. 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
13. Rousselet F, Mangin P (1998) Mitochondrial DNA polymorphisms: a study of 50 French Caucasian individuals and application to forensic casework. Int J Legal Med 111: 292–298
14. Aquadro CF, Greenberg BD (1983) Human mitochondrial DNA variation and evolution: analysis of nucleotide sequences from seven individuals. Genetics 103: 287–312
15. Vigilant L, Stoneking M, Harpending H, Hawkes K, Wilson AC (1991) African populations and the evolution of human mitochondrial DNA. Science 253: 1503–1507
16. Bendall KE, Sykes BC (1995) Length heteroplasmy in the first hypervariable segment of the human mtDNA control region. Am J Hum Genet 57: 248–256
17. Hühne J, Pfeiffer H, Brinkmann B (1998) Heteroplasmic substitution in the mitochondrial DNA control region in mother and child samples. Int J Legal Med 112: 27–30