Int J Legal Med (2001) 114 : 229–231

ORIGINAL ARTICLE

F. Cali · M. G. Le Roux · R. D'Anna · A. Flugy
G. De Leo · V. Chiavetta · G. F. Ayala · V. Romano

MtDNA control region and RFLP data for Sicily and France

Received: 10 February 2000 / Accepted: 13 May 2000

Abstract The forensic application of mtDNA typing requires large databases which are regionally well defined. To further this aim, we have typed mtDNA in a sample of 111 French and 106 Sicilians. The French were typed for both hypervariable segments (HVR1 and HVR2) of the mtDNA control region, whereas the Sicilians were only typed for HVR1, but in addition for the coding region RFLP markers for mtDNA groups H, I, J, K, L, M, T, U, V and X. In both samples, the predominant sequence type by far was the Cambridge reference sequence. Comparing HVR1 sequences, we found that the French sample was twice as diverse as the Sicilian sample as measured by sequence matches. A further set of sequence match comparisons including the French, Sicilian, and the published British mtDNA samples, demonstrate that sequence matching probabilities within samples differ by less than a factor of 2 from the matching probabilities between samples.

Keywords MtDNA · RFLP typing · Population study

Supplementary material Data on the mtDNA HVR1 and HVR2 sequences for 111 unrelated French individuals (Table S1) and mtDNA HVR1 sequences and RFLP typing for 106 Sicilians from Castellammare del Golfo (Table S2) are available in electronic form on Springer-Verlag's server at http://link.springer.de/link/service/journals/00414/index.htm.

F. Cali · V. Chiavetta · G. F. Ayala · V. Romano
OASI Istituto per la Ricerca sul Ritardo Mentale e l'Involuzione Cerebrale (IRCCS), Via Conte Ruggero 73,
94018 Troina (Enna), Italy

M. G. Le Roux
Institut de Biologie, Centre Hospitalier Universitaire de Nantes, Quai Moncousu, Hotel Dieu, 44093 Nantes, France

R. D'Anna · A. Flugy · G. De Leo · V. Romano (✉)
Dipartimento di Biopatologia e Metodologie Biomediche, Facoltà di Medicina e Chirurgia,
Università degli Studi di Palermo, Via Divisi 83,
90133 Palermo, Italy

Introduction

For forensic and medico-legal purposes, the typing of mitochondrial DNA (mtDNA) is less useful than autosomal typing because many maternally related individuals will have identical mtDNA types, which limits the discrimination power. However, mtDNA typing is frequently the only option when the DNA sample is degraded, as there are thousands of mtDNA copies for every set of nuclear chromosomes. Having obtained an mtDNA sequence from a sample, the question arises which mtDNA database to choose in order to determine the frequency of a given sequence. In order to assist in this decision, we have typed mtDNA in a sample of 111 French and 106 Sicilians. These important European regions have not previously been adequately covered in the literature. We sequenced the first hypervariable region (HVR1) for both samples. In addition, we determined hypervariable region 2 (HVR2) for the French, and for the Sicilians we performed restriction enzyme digestion of the complete mtDNA molecule for the major mtDNA groups found in Europe, H, I, J, K, L, M, T, U, V and X according to Macaulay et al. (1999). We then compared the utility of the additional typing (HVR2 vs RFLP typing). In another analysis, we compared the Sicilian, French, and the published British (Piercy et al. 1993) mtDNA HVR1 sequences in order to investigate whether matching probabilities are greater when regional databases are used as opposed to mixed databases.

Subjects and methods

We have typed mtDNA in a sample of 111 French (HVR1 and HVR2) from different parts of France, and 106 Sicilians (HVR1 and RFLPs) whose grandparents were born in Castellammare del Golfo. As the population of Castellammare del Golfo numbers 15,000 inhabitants, the samples can safely be assumed to be maternally unrelated in the grandparents generation. For the Sicilians, HVR1 was amplified by PCR using the following primers:

– np15409: 5′-CCCTTACTACACAATCAAAG-3′
– np16543: 5′-CGTGTGGGCTATTTAGGC-3′

PCR was performed in a 50 µl volume containing 50 ng of genomic DNA, 1 U Taq DNA-polymerase (Perkin Elmer, USA), 5 µl reaction buffer 10 × (20 mM Tris-HCl pH 8, 100 mM KCl, 0.1 mM EDTA, 1 mM DTT, 50% glycerol, 0.5% Tween 20, 0.5% Nonidet P40, 1.5 mM $MgCl_2$), 0.2 mM of each dNTP and 0.2 µM of each primer. Cycling conditions used for the PCR were as follows: 96 °C for 1 min, 50° C for 1 min, 72 °C for 1 min 30 s for 30 cycles. The following primers were used to sequence both strands of DNA:

– F15978: 5′-CACCATTAGCACCCAAAGCT-3′
– R16543: 5′-CGTGTGGGCTATTTAGGC-3′

using the ABI PRISM BigDye Terminator Cycle Sequencing Ready Reaction Kit (PE–Applied Biosystems, Milan, Italy). The sequence for HVR1 was determined in the range np 16025 to at least np 16410 for all Sicilian sequences. We furthermore performed restriction enzyme digestion of the complete mtDNA molecule for the major European mtDNA groups H, I, J, K, L, M, T, U, V and X according to Torroni et al. (1996).

For the French samples, HVR1 was amplified by PCR using the following primers (5′ end in parentheses):

– F15971: 5′-TTAACTCCACCATTAGCACC-3′
– R16420: 5′-TGATTTCACGGAGGATGGTG-3′

For HVR2, the primers were

– F00015: 5′-CACCCTATTAACCACTCACG-3′
– R00389: 5′-CTGGTTAGGCTGGTGTTAGG-3′

PCR was performed in a 50 µl volume containing 1 ng of genomic DNA, 1.5 U Ampli Taq Gold DNA-polymerase (Perkin Elmer, USA), 5 µl reaction buffer 10 × (10 mM Tris-HCl pH 8.3, 50 mM KCl, 1.5 mM $MgCl_2$), 177 mg/ml BSA, 0.2 mM of each dNTP and 0.2 µM of each primer. Cycling conditions used for PCR were as follows: 94 °C for 1 min, 53 °C for 1 min, and 72 °C for 1 min for 38 cycles. Primers used for the sequencing of the PCR products were the same as for the amplification. Sequencing was performed on an ABI373 using ABI PRISM BigDye Terminator Cycle Sequencing Ready Reaction Kit (PE–Applied Biosystems). The sequence for HVR1 was determined in the range np 16025 to at least np 16391 for all French sequences, and additionally from np 72 to np 340 for the French. The longer readability of the Sicilian HVR1 sequences was due to the differing sequencing protocol.

The ratios of pair-wise sequence matching likelihood within samples (mw) and between samples (mb) were calculated as in Brinkmann et al. (1999) and Pfeiffer et al. (1999). Minimum and maximum estimates for the mw values were calculated on the basis of assuming population frequencies of 0 and 1/n, respectively, for types only occurring once in the sample (Brinkmann et al. 1999; Pfeiffer et al. 1999).

Results

The French sequencing results are shown in Table S1, and the Sicilian sequencing results, together with the RFLP typing, are shown in Table S2. Considering only HVR1 from np 16025 to np 16391, the most common sequence in both samples was the Cambridge reference sequence (CRS, Anderson et al. 1981), found in 20/111 French and 30/106 Sicilians. No other sequence was found more than 5 times. The within-sample matching probability is therefore dominated by the CRS, and is 0.0436 for the French and 0.0917 for the Sicilians (Table 1). In other words, it is twice as likely to find a matching pair within the Sicilian than within the French sample. In order to determine the most effective way of reducing the relatively high matching probabilities, we calculated the matching probabilities including HVR2 (France) or including RFLP typing (Sicily). By including HVR2, the matching probability for France dropped markedly to 0.0132 and the most common type is reduced to a frequency of 6/111. In contrast the RFLP typing does not differentiate a single Sicilian sequence match, so that the Sicilian matching probability is not reduced at all.

Table 1 MtDNA HVR1 sequence matching probabilities within and between Britain, France, and Sicily (*n* numbers of individuals, mw_{max} maximum matching probability within the population, mw_{min} minimum matching probability within the population, mb_{min} minimum matching probability between two populations)

Matching probability	Britain (B) (*n* = 100)	France (F) (*n* = 111)	Sicily (S) (*n* = 106)
mw_{max}	0.0334	0.0436	0.0917
mw_{min}	117/4950	216/6105	462/5565
mb_{min}	B-F: 0.0297 B-S: 0.0414	F-B: 0.0297 F-S: 0.0528	S-B: 0.0414 S-F: 0.0528
mw_{max}/mb_{min}	B/F: 1.125 B/S: 0.807	F/B: 1.468 F/S: 0.826	S/B: 2.215 S/F: 1.737
mw_{min}/mb_{min}	B/F: 0.795 B/S: 0.570	F/B: 1.192 F/S: 0.670	S/B: 2.005 S/F: 1.572

In a different analysis, we used the approach of Brinkmann et al. (1999) and Pfeiffer et al. (1999) to determine how many more times it is likely to find a sequence match within each sample than between the samples of Sicily, France and Britain. An upper and a lower estimate of each ratio is given in Table 1. The main result is that it is approximately equally likely (mw/mb = 1) or even slightly less likely (mw/mb < 1) to draw a French-French or British-British sequence match for HVR1 (np 16025 to np 16391) than a French-British, British-Sicilian or French-Sicilian sequence match. In contrast, it is more likely to draw a Sicilian-Sicilian match than a Sicilian-Other match (mw/mb = 1.6–2.2). The mw/mb ratios are evidently dominated by the frequency of the Cambridge reference sequence (14/100 British, 20/111 French, and 30/106 Sicilians). The high frequency of the CRS is also a feature in other parts of Sicily (unpublished data).

Discussion

The comparison between HVR2 typing and RFLP typing demonstrates that HVR2 typing is greatly superior when the aim is to improve discrimination power with the minimum wastage of template sequence. RFLP typing is often used for population genetic research (e.g. Torroni et al. 1998), and there may be cases in forensic applications where the geographic origin of maternal descent of a sample (or set of a samples) is required. In that case RFLP typing should only be performed if HVR1 and HVR2 sequencing did not provide a conclusive result.

Concerning the matching probabilities within and between the British, French, and Sicilian samples, it may at first glance appear surprising that the probability of drawing matching pairs is similar (within a factor of 2), whether sequences are drawn from one sample or from different samples. This result is in accordance with Pfeiffer et al.

(1999) who compared their own northwest German data with Austrian data from Parson et al. (1998) and south German data from Lutz et al. (1998, 1999). Although it is clear that most branches in the evolutionary tree of the mtDNA molecule are specific to certain regions of the world, even within Europe (e.g. Richards et al. 1996; Torroni et al. 1998), it does not follow that we can detect these geographical patterns simply by searching for exact matches with a given high-resolution mtDNA sequence (in our case, the mtDNA control region). The rapid mutation rate of the mtDNA control region and the rapid genetic drift of mtDNA (4 times faster than for an autosomal DNA) ensures that precise matches between two samples do not survive for long, even though the same underlying mtDNA evolutionary branch may still be present in the two samples. Our matching statistics do not discriminate a near miss from a complete miss. In order to find out the most relevant mtDNA database to compare with a given forensic sample, we suggest that it is either necessary to continue enlarging the current mtDNA databases for different regions of the world, or to identify the most relevant database from sequence similarities.

Acknowledgements The expert technical assistance of Rosanna Galati and Concetta Santoro is acknowledged. We thank an anonymous reviewer for valuable comments. This work was supported by Progetto di Ricerca Corrente del Ministero della Sanità, Roma.

References

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

Brinkmann C, Forster P, Schürenkamp M, Horst J, Rolf B, Brinkmann B (1999) Human Y-chromosomal STR haplotypes in a Kurdish population sample. Int J Legal Med 112: 181–183

Lutz S, Weisser H-J, Heizmann J, Pollak S (1998) Location and frequency of polymorphic positions in the mtDNA control region of individuals from Germany. Int J Legal Med 111: 67–77

Lutz S, Weisser H-J, Heizmann J, Pollak S (1999) Location and frequency of polymorphic positions in the mtDNA control region of individuals from Germany (Erratum). Int J Legal Med 112: 145–150

Macaulay V, Richards M, Hickey E, Vega E, Cruciani F, Guida V, Scozzari R, Bonne-Tamir B, Sykes B, Torroni A (1999) The emerging tree of West Eurasian mtDNAs: a synthesis of control-region sequences and RFLPs. Am J Hum Genet 64: 232–249

Parson W, Parsons TJ, Scheithauer R, Holland MM (1998) Population data for 101 Austrian Caucasian mitochondrial d-loop sequences: application of mtDNA sequence analysis to a forensic case. Int J Legal Med 111: 124–132

Pfeiffer H, Brinkmann B, Hühne J, Rolf B, Morris AA, Steighner R, Holland MM, Forster P (1999) Expanding the forensic German mitochondrial DNA control region database: genetic diversity as a function of sample size and microgeography. Int J Legal Med 112: 291–298

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

Richards M, Corte-Real H, Forster P, Macaulay V, Wilkinson-Herbots H, Demaine A, Paphia S, et al, (1996) Palaeolithic and neolithic lineages in the mitochondrial gene pool. Am J Hum Genet 59: 185–203

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

Torroni A, Bandelt H-J, D'Urbano L, Lahermo P, Moral P, Sellitto D, Rengo C, Forster P, Savontaus M-L, Bonné-Tamir B, Scozzari R (1998) MtDNA analysis reveals a major late palaeolithic population expansion from southwestern to northeastern Europe. Am J Hum Genet 62: 1137–1152