Int J Legal Med (2008) 122:199–204
DOI 10.1007/s00414-007-0207-1

ORIGINAL ARTICLE

Italian mitochondrial DNA database: results of a collaborative exercise and proficiency testing

Chiara Turchi · Loredana Buscemi · Carlo Previderè · Pierangela Grignani · Anita Brandstätter · Alessandro Achilli · Walther Parson · Adriano Tagliabracci · Ge.F.I. Group

Received: 6 October 2006 / Accepted: 6 September 2007 / Published online: 20 October 2007
© Springer-Verlag 2007

Abstract This work is a review of a collaborative exercise on mtDNA analysis undertaken by the Italian working group (Ge.F.I.). A total of 593 samples from 11 forensic genetic laboratories were subjected to hypervariable region (HVS-I/HVS-II) sequence analysis. The raw lane data were sent to MtDNA Population Database (EMPOP) for an independent evaluation. For the inclusion of data for the Italian database, quality assurance procedures were applied to the control region profiles. Only eight laboratories with a final population sample of 395 subjects passed the quality conformance test. Control region haplogroup (hg) assignments were confirmed by restriction fragment length polymorphism (RFLP) typing of the most common European hg-diagnostic sites. A total of 306 unique haplotypes derived from the combined analysis of control and coding region polymorphisms were found; the most common haplotype —CRS, 263, 309.1C, 315.1C/¬7025 AluI– was shared by 20 subjects. The majority of mtDNAs detected in the Italian population fell into the most common west Eurasian hgs: R0a (0.76%), HV (4.81%), H (38.99%), HV0 (3.55%), J (7.85%), T (13.42%), U (11.65%), K (10.13%), I (1.52%), X (2.78%), and W (1.01%).

Keywords Collaborative exercise · MtDNA · Control region · Coding region · Italian database · Forensic

Ge.F.I. Group: Milena Alù, University of Modena and Reggio Emilia, Modena; Giovanni Beduschi, University of Modena and Reggio Emilia, Modena; Carla Bini, University of Bologna, Bologna; Ilaria Boschi, Catholic University Rome, Rome; Eugenia Carnevali, University of Perugia, Terni; Sarah Gino, University of Torino, Torino; Laura Giunti, University of Florence, Florence; Massimo Lancia, University of Perugia, Terni; Vincenzo Pascali, Catholic University Rome, Rome; Susi Pelotti, University of Bologna, Bologna; Silvano Presciuttini, University of Pisa, Pisa; Ugo Ricci, University of Florence, Florence; Carlo Robino, University of Torino, Torino.

Electronic supplementary material The online version of this article (doi:10.1007/s00414-007-0207-1) contains supplementary material, which is available to authorized users.

C. Turchi · L. Buscemi · A. Tagliabracci (✉)
Department of Neuroscience, Section of Legal Medicine,
Università Politecnica delle Marche,
60020 Ancona, Italy
e-mail: a.tagliabracci@univpm.it

C. Previderè · P. Grignani
Department of Legal Medicine and Public Health,
University of Pavia,
Pavia, Italy

A. Brandstätter · W. Parson
Institute of Legal Medicine, Innsbruck Medical University,
Innsbruck, Austria

A. Achilli
Department of Genetic and Microbiology, University of Pavia,
Pavia, Italy

Introduction

Worldwide forensic genetic laboratories currently use the analysis of the two hypervariable regions HVS-I/HVS-II of the mtDNA control region, especially for human identification cases in which low quantity and poor quality of nuclear DNA do not give conclusive results (for a review see, e.g., [1]). A large and high-quality reference database for different populations is essential to estimate the frequency of a given mtDNA haplotype to obtain the probability of a random match.

It has been criticized that published forensic databases are prone to contain errors [2–6], which are more frequently

caused by data transcription and tabulation errors (i.e., clerical errors) than by errors during the laboratory processing of the samples [7, 8].

The aim of MtDNA Population Database (EMPOP) (www.empop.org) is collecting mtDNA sequences from various populations worldwide and assuring a high quality of the data by avoiding the manual transcription of raw sequence data [9, 10].

Our approach to insure the robustness of our mtDNA data was a combined analysis of control region polymorphisms with coding region haplogroup (hg)-specific restriction sites [11–13]; in addition, it was suggested that a posteriori phylogenetic analysis should be a valuable tool in discovering erroneous mtDNA types [2, 7, 14, 15].

Forensic scientists emphasize the need to follow up guidelines on mtDNA typing [16–19], to undergo periodical external quality control by means of collaborative exercises.

The present work is a review of a collaborative exercise on mtDNA analysis undertaken by the Italian Forensic Geneticists (Ge.F.I.) group of the International Society for Forensic Genetics (ISFG). The exercise was developed in three main steps: (1) sequencing of HVS-I and HVS-II regions from three blood stains and from at least 50 subjects for each participating laboratory, (2) transmission of the raw lane data in electronic form to the Institute of Legal Medicine in Innsbruck, Austria, for an independent evaluation and for incorporation into EMPOP and (3) restriction fragment length polymorphism (RFLP) analysis of SNPs diagnostic for major European hgs.

Materials and methods

DNA samples, extraction and quantification

A total of 11 laboratories (Labs 1–11), all belonging to centers for forensic genetics in northern and central Italy, participated in the project. Blood samples and buccal swabs were collected from 593 locally resident and maternally unrelated subjects, Italian ancestry traced for two generations.

Participating laboratories used different methods for DNA extraction and quantification (Supplementary Table S1).

Control region sequencing

The two strands of hypervariable regions HVS-I and HVS-II were amplified and sequenced following Holland and Parson [20, 21]. Sequencing was performed using the BigDye terminator and primer chemistry (Supplementary Table S1). Sequences were aligned and compared to the revised Cambridge reference sequence (rCRS; [22, 23]).

Sequencing analysis and data quality control

Participating laboratories were requested to assure the availability of the DNA template and the sequence raw lane data for a potential double-check.

All laboratories had to fulfill the following requirements: (1) correct typing of three blood stains, one of which was used as a positive control, which were supplied by the pilot laboratory; (2) transmission of the mtDNA profiles to the pilot laboratory for a concerted tabulation; (3) a successive revision of the data in the combined data table to check possible mistakes in the course of transcription; (4) three samples among outlier haplotypes were selected from each lab by a Network analysis and were independently analyzed in another laboratory for a double-check; (5) transmission of the raw sequence data to the Institute of Legal Medicine in Innsbruck for an independent evaluation following EMPOP criteria.

Haplogroup assignment: RFLP analysis

The first hg assignment was done on the basis of HVS-I/HVS-II variations [24] and was subsequently confirmed by restriction endonuclease analysis of eight polymerase chain reaction (PCR)-amplified mtDNA fragments. If the initial hg assignments based on control region polymorphisms were not confirmed by RFLP analysis, the samples were typed on the basis of hierarchical mtDNA RFLP scheme [13, 25].

RFLP analysis was performed in two laboratories participating in the collaborative exercise.

The computer program Primer Express (Version 1.5, AB) was used for primer design (Table 1). PCR conditions were initial denaturation at 94°C for 5 min, followed by 35 cycles of denaturation at 95°C for 1 min, annealing at 55°C for 1 min and extension at 72°C for 1 min. The seven restriction endonucleases, *Acc*I, *Alu*I, *Dde*I, *Hae*III, *Hinf*I, *Mse*I, and *Nla*III were targeted to screen a restricted set of RFLP sites (Table 1), which were diagnostic for the most frequent west Eurasian mtDNA hgs.

Restriction fragments were checked by 2% agarose (Ultra Pure Agarose Electrophoresis Grade, Gibco BRL, Invitrogen, UK) gel electrophoresis, stained with ethidium bromide and visualized by ultraviolet (UV)-induced fluorescence.

HV0 classification into subhaplogroup HV0a was based on the gain of the 15904-*Mse*I site, while the remainder HV0* would group together those HV0 lineages that do not belong to HV0a [25]. Subhaplogroups of the hg U were determined on the basis of control region polymorphisms [31]. The remaining samples assigned to hgs D [27], L3f [32], N2a [33], and N1b [28] on the basis of HVS-I motifs classification were investigated using PCR conditions suggested by Torroni A. (Department of Genetic and

Table 1 Amplification primers and characteristic restriction sites for RFLP analysis in the main western European mtDNA haplogroups

Amplification primers[a] 5′ to 3′		Polymorphic restriction sites[b]	Hg
6924–6947 FW	GCCCTAGGATTCATCTTTCTTTTC	**−7025 *Alu*I**	H [26]
7836–7859 RV	CGTTGACCTCGTCTGTTATGTAAA		
14360–14381 FW	CACAGCACCAATCCTACCTCCA	+11718 *Hae*III; **−14766 *Mse*I**	HV [13, 27]
14830–14849 RV	AGCCGAAGTTTCATCATGCG	**+14465 *Acc*I**	X [13]
9920–9939 FW	CGCCTGATACTGGCATTTTG	**+10032 AluI**, −10397 ***Alu*I**	I [13]
10628–10649 RV	AGGCACAATATTGGCTAAGAGG	+4216 *Nla*III; **+10394 *Dde*I**	J [13]
		+10394*Dde*I; +12308 *Hinf*I	K [13]
		+10397 *Alu*I; −10871 *Mn*II	M [11, 13]
		+4216 *Nla*III; **−10394 *Dde*I**	T [13]
15698–15718 FW	CGCCCACTAAGCCAATCACTT	**+4577 *Nla*III**; **+11718 *Hae*III**;	HV0[c] [25]
16019–16040 RV	GCTTCCCCATGAAAGAACAGAG	−14766 *Mse*I	
11477–11503 FW	GGCTATGGTATAATACGCCTCACACTC	+**11718 *Hae*III**; **+14766 *Mse*I**	R0[c] [28]
11720–11750 RV*	TTGCTAGGCAGAATAGTAATGAGGATGTAGG[d]		
12104–12124 FW	CTCAACCCCGACATCATTACC	**−10394 *Dde*I**; +12308 *Hinf*I	U [11, 13]
12309–12338 RV*	ATTACTTTTATTTGGAGTTGCACCAAGATT[d]		
3935–3954 FW	GCTTCAACATCGAATACGCC	**−4577 *Nla*III**; +11718 *Hae*III; −14766	V [11, 13]
4632–4653 RV	TGAGGAAATACTTGATGGCAGC	*Mse*I; **+15904 *Mse*I**	
8854–8873 FW	GCGGGCACAGTGATTATAGG	**−8994 *Hae*III**	W [11, 13]
9228–9253 RV	CATGGGCTGGGTTTTACTATATGATA		

[a] Primers are numbered according to Anderson et al. [22] and Andrews et al. [23].
[b] Haplogroup diagnostic markers are in boldface; other polymorphic restrictions sites are haplogroup related.
[c] Haplogroups HV0 and R0 were previously know as pre*V and preHV, respectively [30].
[d] The underlined nucleotide indicates the mismatch G instead of A. Mismatched primers were used in order to screen for an A to G transition at: (1) np 11719 in ND4 gene that generates a *Hae*III site when the mutation is present [29]; (2) np 12308 in the tRNA^{Leu} gene that generates a *Hinf*I site when the mutation is present [11].

Microbiology, University of Pavia, personal communication) at the following RFLP sites: ¬5176 *Alu*I for hg D; RFLP sites:–10871 *Mnl*I for hg L3f; RFLP sites:–12696 *Mbo*II for hg N2a; RFLP sites:–11362 *Alu*I, for hg N1b1.

In addition, some samples with HVS-I motifs 16224–16311 and 16126–16294, respectively, proposed for hgs K and T [13], but not confirmed by *Dde*I at np 10394, were investigated for supplementary RFLP sites: *Hha*I at np 9053, the absence of which was typical of hg K and *Ava*II at np 13367, the absence of which characterized hg T [13].

Statistical tests

Reduced-median-network analysis was conducted on each laboratory sample using the program NETWORK 4.1 [34]. Population tests of hg frequency heterogeneity among each laboratory sample were performed with ARLEQUIN 3 [35]. Other statistical analyses were performed with Excel.

Results and discussion

The original database of the Ge.F.I. collaborative exercise included 593 individuals from central and northern Italy typed for HVS-I and HVS-II by 11 different laboratories. Three laboratories were excluded from the exercise (two laboratories failed the proficiency test (Labs 3, 5), and one lab did not archive the sampled DNAs (Lab 7) thus reducing the sample size to 437.

As a further quality control, it was agreed that three samples from each laboratory had to be retyped by a different laboratory; to select these samples, the eight remaining population subsamples were individually submitted to network analysis, and the most outlying haplotypes were determined, i.e., those separated from the closest haplotype by the largest number of base substitutions. Re-sequencing revealed that 6 out of 24 sequences contained differences from the original submission by one or more base substitutions; in one case (Lab 4), all three haplotypes were different.

The results of the first phase of the exercise emphasized the need to perform a more stringent quality control: EMPOP processing and RFLP analysis.

The EMPOP analysis protocol consisted of two independent evaluations of the raw data; special attention was given to point and length heteroplasmy issues, concerning both interpretation and nomenclature topics. For the polycytosine-stretch length heteroplasmies in HVS-I and HVS-II, the “dominant mtDNA type” was recorded in the database, i.e., the haplotype with the number of cytosine residues present at higher proportions as estimated from sequence electropherograms (Fig. 1). All samples were

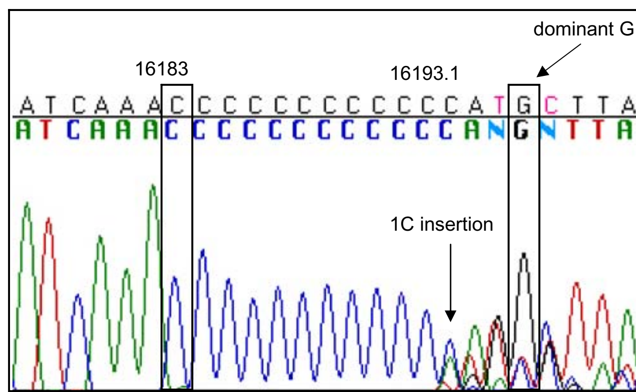

Fig. 1 Example of length heteroplasmy polycytosine stretches in HVS-I interpretation considering the "dominant type". The transversion of A to C at position 16183 together with T to C transition at 16189 produce different C-stretch lengths. The majority molecule was inferred by checking the highest "virtual" G peak of the ATG motif downstream of the C-stretch. The highest virtual G peaks corresponds to the dominant type, and the C-stretch ends two nucleotides before that majority G peak. Using this method, the C-stretch in this sample was annotated as 16193.1 C, as only one insertion was considered.

scrutinized for artificial recombination. Twenty-eight samples were suspicious because of odd combinations of HVS-I and HVS-II motifs. In 22 cases, it was possible to go back to the DNA to re-sequence the sample. In all 22 samples, the initial results were confirmed. In six cases, re-sequencing was not possible due to lack of DNA; therefore, we decided to skip these six samples from the dataset.

After the EMPOP evaluation, 35 among the 437 samples (8.0%) were excluded because of low sequence quality, lack of complementary strands (11 cases) or artificial recombinations (6 cases). Thus, 402 individual haplotypes could be compared with the original data produced by each individual laboratory. A surprisingly high number of discrepancies were found; overall 71 disagreeing haplotypes (17.6%) were found, encompassing 101 individual errors (Supplementary Table S2). In 21 cases, a nonexisting mutation had been reported, whereas the remaining discrepancies were due to overseen polymorphisms. Most discrepancies (66.3%) were due to table transcription errors (clerical errors), erroneous interpretation of C-stretches in cases of length heteroplasmy concerned 15.8 and 10.9% in HVS-I and HVS-II, respectively, and 6.9% were omissions of point heteroplasmic polymorphisms. It was observed that the highest occurrence of phantom mutations concerns one laboratory which used Dye primer chemistry on LICOR-4200 instrument.

Lastly, our stepwise check procedure included hg assignments by coding region RFLP analysis. As seven samples could not be typed by RFLP analysis, probably due to degraded DNA, the final sample incorporated into the EMPOP database included 395 individual haplotypes from eight laboratories.

The resulting haplotypes, their frequency distributions, and their hg classifications are available at www.gefi-forensicDNA.it (Supplementary Table S3).

The 395 sequences consisted of 332 different haplotypes, 306 of which were unique. The most frequent haplotypes were "CRS, 263, 309.1C, 315.1C/–7025 *Alu*I", observed in 20 individuals, and "CRS, 263, 315.1C/–7025 *Alu*I", observed in 14 individuals; in other European populations, these two haplotypes are also the most frequent, although their frequency is inverted [21, 36–38]. Among the other nonunique haplotypes, one was present in four subjects, 5 were shared by three subjects each, and 18 were shared by two subjects each.

We wanted to compare hg assignments based on sequence variation in the two hypervariable regions of the mtDNA control region [24] with those based on RFLP analysis of SNPs in the mtDNA coding region (Supplementary Table S4). In total, for 343 haplotypes (86.8%), the hg assignments matched between the two methods; in addition, RFLP analysis allowed a subclassification of a further 22 haplotypes from 178 of cluster H (19 HV, 2 R0a and 1 HV0*) and a subdivision of hg V into V (8) and HV0* (5). For 12 haplotypes (3.0%), RFLP analysis yielded a different hg affiliation than control region typing: six control region hg H identifications were reclassified as hg U, two hg I and two hg U assignments were all identified as hg H samples, one presumptive W sample was assigned to L3f, and one putative X sample was identified as L1b by RFLP analysis. In conclusion, the RFLP typing changed the hg assignment of 52 subjects (13.1%). Maybe some of these errors in initial and final assignment was due to a poor experience of forensic laboratories in classification by HVS-I/HVS-II status. Moreover, the hg cannot unambiguously be identified based on the two hypervariable segments (HVS-I/HVS-II) of the control region. These sequences are characterized by higher base substitution rates than the coding region and evolve more rapidly [39], causing the appearance of homoplasy due to recurrent and parallel mutations or reversions.

A drawback of the procedure of "cleaning" a database is that it may introduce a bias, for example, if samples that are more difficult to be sequenced and interpreted are preferentially excluded. One such case may be the 16189C transition, which produces length heteroplasmy and results in a blurred sequence after the C-stretch; reading such sequences may be difficult, and one could be inclined to discard them. The occurrence of length heteroplasmy after the T16189C transition is observed in west Eurasian samples in approximately 15% [21, 40]; a significant departure from this value would suggest a bias in reporting population sample composition. As our database collected samples from various Italian regions, it was checked for possible bias by examining the 16189C frequency in each individual

laboratory database without any other substitution in the polycytosine C-stretch (16184–16193), a situation that would lead to a blurred signal. The total frequency of these haplotypes was 41 out of 395 (10.4%) ranging from 0–14.6% among laboratories, and the contingency table χ^2-test was not significant for heterogeneity (χ^2 8.83, 7 *df*, P=0.27). The absence of sampling bias in one subpopulation sample was assured by the involved laboratory, indicating that the observed situation reflected the real structure.

A more general test of population differentiation was conducted on the hg frequencies determined separately in the eight subsamples (subsample size 29–73). The four most frequent hg (H, T, K, and J) included 70.4% of the entire sample and were present in all subsamples; no evidence of heterogeneity was found using ARLEQUIN exact test of population differentiation, both considering the global sample and all 28 pair-wise comparisons (global test: P=0.160±0.069). Thus, the regions of central and northern Italy may be considered homogeneous for the frequency distribution of mtDNA hg.

The majority of mtDNA sequences detected in the Italian population falls into the most common west Eurasian hgs [11–13, 25, 27, 28], listed according to phylogenetic relationships: R0a, HV, H, HV0*, V, J, T, U, K, I, X, W.

The subhaplogroup R0a [30, 41] was found in this sample at a very low frequency of 0.76%.

Hg HV sequences were found at relatively high frequencies of 4.81%, with 0.25% for both subhaplogroups HV1 and HV2 [42].

Hg H encompassed 38.99% of Italians, and thus constituted, like in other European population samples [41], the most frequent hg. For this reason, a definition of hg H subclades could be extremely useful for forensic investigations [41, 43].

The members of HV0 harbor the HVS-I motif 16298 in association with RFLP motifs +4577 *Nla*III and +11718 *Hae*III, while the gain of 15904 *Mse*I defines the clade HV0a [25] that in turn contains V. The members of HV0a* would then group together those HV0a lineages that do not belong to V, which were not found in this Italian population sample, while the paragroup HV0* was found at a frequency of 1.52%, comparable with those observed in Italy by Torroni et al. [25]. Eight haplotypes were identified as belonging to hg V [11, 13, 27].

Hg J [12, 13], encompasses 7.85% of the Italian west Eurasians.

Hg T was the second most frequent hg in the Italian population sample, with a frequency of 13.42% [13].

Of this Italian sample, 11.65% belong to hg U; a deeper analysis of CR polymorphisms allowed a separation into subhaplogroups U1a, U1b, U2e, U3, U3a, U4, U5a, U5a1, U5b, U6a, U7, and U8b [31].

Hg K was found at a frequency of 10.13%.

Hg I, X, and W were found in this study at frequencies of 1.52, 2.78, and 1.01%, respectively. Hg N2a (0.25%) and N1b (1.27%) were observed at low frequencies, according to other west European studies [13, 28]; the occurrence of 152, 16145, 16176G, 16223, 16390 polymorphisms in these samples, in combination with a site loss at 11362 *Alu*I led to a classification into subhaplogroup N1b1 [44].

In addition, a single African-specific sequence type, belonging to hg L3f [31], was found in the Italian samples.

Two Asian hg [45, 46] were also discovered: hg M (0.25%) and D (0.75%).

Two haplotypes did not fit into the described hg and were classified as others (n.d.).

In conclusion, the Ge.F.I collaborative exercise permitted a high-quality mtDNA Italian population database useful for forensic purposes to be set up. Forensic mtDNA databases are usually lacking in coding region polymorphism resolution, and to our knowledge, this is the first forensic mtDNA database that reports the polymorphisms of the two hypervariable regions of the control region together with hg-defining coding region polymorphisms.

The project consumed much attention and time, needed for stringent quality control. Despite all that participating laboratories had to consolidate experience in forensic investigations, a remarkable number of errors were found in several steps of the collaborative exercise; this indicates that mtDNA cannot be introduced as routine casework marker without adequate experience in sequence interpretation (i.e., heteroplasmic conditions) and transcription (i.e., clerical errors) and without exact knowledge of reference nomenclature and mtDNA phylogenetic structure.

Acknowledgments This study was supported by a grant from the Ministero per l'Istruzione, l'Universita` e la Ricerca (MIUR) (grant no. 2005064329_002).

References

1. Budowle B, Allard MW, Mark R, Chakraborty R (2003) Forensics and mitochondrial DNA: applications, debates and foundations. Annu Rev Genomics Hum Genet 4:119–141
2. Bandelt H-J, Lahermo P, Richards M, Macaulay V (2001) Detecting errors in mtDNA data by phylogenetic analysis. Int J Legal Med 115:64–69
3. Röhl A, Brinkmann B, Forster L, Forster P (2001) An annotated mtDNA database. Int J Legal Med 115:29–39
4. Bandelt H-J, Quintana-Murci L, Salas A, Macaulay V (2002) The fingerprint of phantom mutations in mitochondrial DNA data. Am J Hum Genet 71:1150–1160
5. Forster P (2003) To err is human. Ann Hum Genet 67:2–4
6. Dennis C (2003) Error reports threaten to unravel databases of mitochondrial DNA. Nature 421:773–774
7. Budowle B, Polanskey D, Allard MW, Chakraborty R (2004) Addressing the use of phylogenetics for identification of sequences in error in the SWGDAM mitochondrial DNA database. J Forensic Sci 49:1256–1261

8. Parson W, Brandstätter A, Alonso A et al (2004) The EDNAP mitochondrial DNA population database (EMPOP) collaborative exercises: organisation, results and perspectives. Forensic Sci Int 139:215–226
9. Brandstätter A, Niederstätter H, Pavlic M, Grubwieser P, Parson W (2006) Generating population data for the EMPOP database—an overview of the mtDNA sequencing and data evaluation processes considering 273 Austrian control region sequences as example. Forensic Sci Int 166: 164–175
10. Brandstätter A, Niederstätter H, Pavlic M, Grubwieser P, Parson W (2006) Generating population data for the EMPOP database—An overview of the mtDNA sequencing and data evaluation processes considering 273 Austrian control region sequences as example. Int J Legal Med 120:310–314
11. Torroni A, Huoponen K, Francalacci P et al (1996) Classification of European mtDNAs from an analysis of three European populations. Genetics 144:1835–1850
12. Richards MB, Macaulay VA, Bandelt H-J, Sykes B (1998) Phylogeography of mitochondrial DNA in western Europe. Ann Hum Genet 62:241–260
13. Macaulay V, Richards M, Hickey E et al (1999) The emerging tree of west Eurasian mtDNAs: a synthesis of control-region sequences and RFLPs. Am J Hum Genet 64:232–249
14. Bandelt H-J, Salas A, Lutz-Bonengel S (2004) Artificial recombination in forensic mtDNA population databases. Int J Legal Med 118:267–273
15. Yao YG, Bravi CM, Bandelt HJ (2004) Call for mtDNA data quality in forensic science. Forensic Sci Int 141:1–6
16. Bär W, Brinkmann B, Budowle B et al (2000) DNA Commission of the International Society for Forensic Genetics: guidelines for mitochondrial DNA typing. Int J Legal Med 113:193–196
17. Tully G, Bär W, Brinkmann B, Carracedo A, Gill P, Morling N, Parson W, Schneider P (2001) Considerations by the European DNA profiling (EDNAP) group on the working practices, nomenclature and interpretation of mitochondrial DNA profiles. Forensic Sci Int 124:83–91
18. Wilson MR, Allard MW, Monson K, Miller KW, Budowle B (2002) Recommendations for consistent treatment of length variants in the human mitochondrial DNA control region. Forensic Sci Int 141:1–6
19. Scientific Working Group on DNA Analysis (2003) Guidelines for mitochondrial DNA (mtDNA) nucleotide sequence interpretation. Forensic Science Communications [On Line]. (April 2003). http://www.fbi.gov/hq/lab/fsc/backissu/april2003/swgdammitodna.htm
20. Holland MM, Roby RK et al (1994) Identification of human remains using mitochondrial DNA sequencing: potential mother-child mutational events. In: Bar W, Fiori A, Rossi U (eds) Advances in Forensic Haemogenetics 5. Springer, Berlin Heidelberg, pp 399–406
21. 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
22. Anderson S, Bankier AT, Barrell BG et al (1981) Sequence and organization of the human mitochondrial genome. Nature 290:457–465
23. Andrews RM, Kubacka I, Chinnery PF, Lightowlers RN, Turnbull DM, Howell N (1999) Reanalysis and revision of the Cambridge reference sequence for human mitochondrial DNA. Nat Genet 23:147
24. Allard MW, Miller K, Wilson M, Monson K, Budowle B (2002) Characterization of the Caucasian haplogroups present in the SWGDAM forensic mtDNA dataset for 1771 human control region sequences. Scientific Working Group on DNA Analysis Methods. J Forensic Sci 47:1215–1223
25. Torroni A, Bandelt HJ, Macaulay V et al (2001) A signal, from human mtDNA, of postglacial recolonization in Europe. Am J Hum Genet 69:844–852
26. Torroni A, Lott MT, Cabell MF, Chen YS, Lavergne L, Wallace DC (1994) MtDNA and the origin of west-Eurasians. 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 55:760–776
27. Torroni A, Bandelt HJ, D'Urbano L et al (1998) MtDNA analysis reveals a major late paleolithic population expansion from southwestern to northeastern Europe. Am J Hum Genet 62:1137–1152
28. Richards M, Macaulay V, Hickey E et al (2000) Tracing European founder lineages in the near eastern mtDNA pool. Am J Hum Genet 67:1251–1276
29. Saillard J, Magalhaes PJ, Schwartz M, Rosenberg T, Norby S (2000) Mitochondrial DNA variant 11719G is a marker for the mtDNA haplogroup cluster HV. Hum Biol 72:1065–1068
30. Torroni A, Achilli A, Macaulay V, Richards M, Bandelt H-J (2006) Harvesting the fruit of the human mtDNA tree. Trends Genet 22:339–345
31. Achilli A, Rengo C, Battaglia V et al (2005) Saami and Berbers—an unexpected mitochondrial DNA link. Am J Hum Genet 76:883–886
32. Salas A, Richards M, De la Fe T et al (2002) The making of the African mtDNA landscape. Am J Hum Genet 71:1082–1111
33. Metspalu M, Kivisild T, Metspalu E et al (2004) Most of the extant mtDNA boundaries in south and southwest Asia were likely shaped during the initial settlement of Eurasia by anatomically modern humans. BMC Genet 31:5–26
34. Bandelt H-J, Forster P, Sykes BC, Richards MB (1995) Mitochondrial portraits of human populations using median networks. Genetics 141:743–753
35. Excoffier L, Laval G, Schneider S (2005) Arlequin ver. 3.0: An integrated software package for population genetics data analysis. Evolutionary Bioinformatics Online 1:47–50
36. Crespillo M, Luque JA, Paredes M, Fernández R, Ramírez E, Valverde JL (2000) Mitochondrial DNA sequences for 118 individuals from northeastern Spain. Int J Legal Med 114:130–132
37. Baasner A, Schäfer C, Junge A, Madea B (1998) Polymorphic sites in human mitochondrial DNA control region sequences: population data and maternal inheritance. Forensic Sci Int 98:169–178
38. Malyarchuk BA, Grzybowski T, Derenko MV, Czarny J, Wozniak M, Miscicka-Sliwka D (2002) Mitochondrial DNA variability in Poles and Russians. Ann Hum Genet 66:261–283
39. Howell N, Kubacka I, Mackey DA (1996) How rapidly does the human mitochondrial genome evolve? Am J Hum Genet 59:501–509
40. Brandstätter A, Niederstätter H, Parson W (2004) Monitoring the inheritance of heteroplasmy by computer-assisted detection of mixed basecalls in the entire human mitochondrial DNA control region. Int J Legal Med 118:47–54
41. Achilli A, Rengo C, Magri C et al (2004) The molecular dissection of mtDNA haplogroup H confirms that the Franco-Cantabrian glacial refuge was a major source for the European gene pool. Am J Hum Genet 75:910–918
42. Quintana-Murci L, Chaix R, Wells S et al (2004) Where west meets east: the complex mtDNA landscape of the southwest and central Asian corridor. Am J Hum Genet 74:827–845
43. Grignani P, Peloso G, Achilli A et al (2005) Subtyping mtDNA haplogroup H by SnaPshot minisequencing and its application in forensic individual identification. Int J Legal Med 7:1–6
44. Palanichamy MG, Sun C, Agrawal S et al (2004) Phylogeny of mitochondrial DNA macrohaplogroup N in India, based on complete sequencing: implications for the peopling of south Asia. Am J Hum Genet 75:966–978
45. Ballinger SW, Schurr TG, Torroni A et al (1992) Southeast Asian mitochondrial DNA analysis reveals genetic continuity of ancient mongoloid migrations. Genetics 130:139–152
46. Yao YG, Kong QP, Bandelt HJ, Kivisild T, Zhang YP (2002) Phylogeographic differentiation of mitochondrial DNA in Han Chinese. Am J Hum Genet 70:635–651