Available online at www.sciencedirect.com
ScienceDirect
Forensic Science International: Genetics 1 (2007) 13–19
FSI GENETICS
www.elsevier.com/locate/fsig

Extended guidelines for mtDNA typing of population data in forensic science

Walther Parson [a,*], Hans-Jürgen Bandelt [b]

[a] *Institute of Legal Medicine, Innsbruck Medical University, Muellerstrasse 44, 6020 Innsbruck, Austria*
[b] *Department of Mathematics, University of Hamburg, Bundesstr. 55, 20146 Hamburg, Germany*

Received 16 May 2006; received in revised form 16 November 2006; accepted 19 November 2006

Abstract

Mitochondrial DNA analysis has become a vital niche in forensic science as it constitutes a powerful technique for low quality and low quantity DNA samples. For the forensic field it is important to employ standardized procedures based on scientific grounds, in order to have mtDNA evidence be accepted in court. Here, we modify and extend recommendations that were spelled out previously in the absence of solid knowledge about the worldwide phylogeny. Refinement of those earlier guidelines became necessary in regard to sample selection, amplification and sequencing strategies, as well as a posteriori quality control of mtDNA profiles. The notation of sequence data should thus reflect this growing knowledge.
© 2006 Elsevier Ireland Ltd. All rights reserved.

Keywords: Mitochondrial DNA; Haplogroup; Sequencing; Statistical analysis; Quality control; Network; EMPOP

1. Introduction

MtDNA analysis has found widespread application in forensic medicine in the past 15 years, as it often constitutes the last chance for successful DNA typing when only very limited or severely degraded DNA is present in a sample. In other cases, mtDNA screening is applied as its mode of inheritance allows testing for a putative exclusion scenario in human identification. While recently developed mtDNA screening methods provide a reliable and cost-effective supplementary technique for routine applications, the golden standard still constitutes direct sequencing of PCR products or cloned amplicons. Laboratory guidelines and general considerations were established in order to standardize the mtDNA typing process [1–3]. The increasing body of population data that became available in the past few years improved our understanding of the worldwide mitochondrial phylogeny—but the wealth of data also revealed that some laboratory practice still met difficulties in dealing with mtDNA sequences, both at the sequencing and documentation stages [4–10]. The earlier forensic recommendations did not detail amplification and sequencing schemes, which are, however, crucial for the generation of high-quality sequencing data and a meaningful interpretation of the results. Moreover, a reappraisal of the interpretation rules for length variants in polycytosine tracts, which were previously defined by Refs. [11,12] in a first attempt, became necessary as well.

We briefly discuss the essential steps to take from sampling to sequencing and finally documenting the results. We specifically address laboratory procedures that are crucial for the generation of high quality data. A number of errors can be avoided by applying appropriate quality control procedures a posteriori to the sequencing process. Good laboratory practice demonstrates that the additional analysis of coding-region sites helps to confirm control-region haplotypes and increases our understanding of haplogroup-specific variation [13]. Finally, we draw attention to a modified interpretation of length variants that will be featured in a separate paper [14].

2. Sampling scheme and sample information

Sampling for population databases in the forensic field does not seem to follow any standardized scheme. Most samples would constitute convenience samples of “laboratory staff members, or from blood donor samples with the cooperation of a local blood bank, or from samples from victims and suspects

* Corresponding author. Tel.: +43 512 9003 70640; fax: +43 512 9003 73640.
E-mail address: walther.parson@i-med.ac.at (W. Parson).

1872-4973/$ – see front matter © 2006 Elsevier Ireland Ltd. All rights reserved.
doi:10.1016/j.fsigen.2006.11.003

examined in the course of casework'' [15, p. 44]. Convenience samples are not representative for the targeted national populations or ethnic groups and cannot be regarded as ''random'', inasmuch as they are typically pre-structured in various ways that may or may not influence the mtDNA profiles. Instead, one should strive for avoiding the adverse influence of hidden stratification (social stratum, ethnic origin, etc.), although this seems hard to achieve in practice for mixed (e.g. urban) populations.

The sample information should be as detailed as possible concerning matrilineal ancestry in order to achieve maximum flexibility for future choices of reference databases tailored to a specific forensic case. The broad regional geographic information about the birth place of the maternal grandmother should be provided whenever retrievable. Regional databases should then not include mtDNAs of recent immigrants without identifying the matrilineal origin (within the last two generations), since immigrant groups will be insufficiently represented even in large samples. Reference to specific ethnic or social groupings may be relevant insofar as such strata may partially reflect immigration events of the distant past. Reference to hair or skin colour (or whatever concept of so-called ''race'') of the mtDNA donor, however, is not really the information one should aim at: ''Geographical origin (ancestry) appears to be more relevant than a person's self-identified race'' [16]. In particular, the vague classifier ''Caucasian'' or ''Caucasoid'' should no longer be used to label samples, say, from specific European or recently admixed populations, since such outdated concepts are derived from ill-based preconceptions about human genetic variation; see also Section 6 (Names and Labels) in the American Heritage® Book of English Usage (www.bartleby.com/64/6.html).

In EMPOP, the EDNAP MtDNA Population Database (www.empop.org), metapopulations are distinguished at the most basal level of classification, viz. the sub-Saharan African, West Eurasian (including North African), South Asian, East Asian, Southeast Asian, Oceanian, Native American metapopulations, and the Central Asian population (which originated from mainly West and East Asian groups) as well as other populations (e.g. of the Americas) that were the result of colonization, slave trade, and more recent large-scale immigration. The next level is the national/ethnic group. Future analyses of population stratification will then permit fine-grained sampling schemes that can specifically focus on geographic and social strata within a national state or ethnic group.

3. Sequencing of separate segments in the control region

The mtDNA control region spans over 1122 base pairs from nucleotide position 16,024 (the origin of heavy-strand replication) to position 576. Traditionally, only the first two hypervariable segments (HVS-I and HVS-II, alias HV1 and HV2) of the control region have been targeted for forensic purposes, because a good portion of the differences between two mtDNAs can be found in these two short segments, with maximum ranges of roughly 16024–16,400 and 40–400. In some cases, only HVS-I typing (which is still popular in population genetics) has been exercised, which, however, can no longer be considered to be state of the art. The segment ranges that are sequenced may vary from study to study (depending on the employed primers), but too narrow ranges may miss a number of frequent nucleotide variants that are characteristic for several haplogroups, such as transitional changes at positions 16,051 (e.g. in haplogroup U2), 16,390 (in haplogroups L2, E, N1b1, U3a, etc.), 16,391 (in haplogroup I), 16,399 (in haplogroups L2d, F4a, and U5a1), 64 (in haplogroups A2, L0a1, etc.), 72 (in haplogroup HV0 = pre-V), and 357 (in haplogroup L1b).

In the late 1990s it became evident that there is yet another short stretch in the control region that also harbours notable variation, which has then been called the third hypervariable segment (HVS-III), with approximate range from about 430 until 576, the 3′-end of the control region [17–19]. There were only very few studies that targeted HVS-III besides the other two hypervariable segments [20–22]. HVS-III includes a few sites that make part of the motifs of some Eurasian haplogroups [23–26] and thus increase the discriminative power of forensic analyses, such as the 447G transversion in haplogroup M2, and the transitions at positions 431 (D4b1b), 456 (H5), 461 (M6), 462 (J1), 482 (M3, J1c1), 489 (M, J), 497 (K1a), 499 (B4b, U4′9), and 508 (U2e).

The separate analysis of the two or three hypervariable segments of an mtDNA sample inherently bears the risk of mix-up of these regions between different mtDNAs, leading to artificial recombinants [4,8,27,28]. Typically, the laboratory process of mtDNA analysis involves DNA extraction, amplification, purification of the amplification products, double-strand sequencing, purification of the sequencing products, and preparation of the samples for electrophoresis (e.g. [6,29]). Mix-up can be generated at the bench when handling the tubes or in the course of generating sample sheets for the individual laboratory steps. Each of these errors would result in recombinant mitochondrial haplotypes, consisting of segments that may belong to different haplogroups. Such errors cannot be unmasked by examining the raw data but only through a posteriori investigation using phylogenetic analysis [4,27–29]. A repeated full analysis of each sample (such as recommended for casework analysis [1,30]) would, of course, help minimizing this kind of error, but for the high through-put analysis of population samples this strategy might seem too laborious and too expensive.

4. Sequencing of the entire control region

Instead of performing multiple separate amplifications, it has proven most efficient to amplify the entire control region in a single PCR assay and to sequence with internal primers, because this scheme minimizes the risk of introducing sample mix-up at the stage of the PCR products. Due to the high nucleotide variability within the control region, internal primer annealing sites are prone to mutations which would lower or even inhibit PCR yield. Further, the polycytosine tracts incurred

by mutations at 16,189, 309, 310, and 573 as well as length variants in the dinuclear repeat region between 515 and 524 hinder readability. Therefore, a parsimonious design with only few internal primers (e.g. four, as applied by [31]) does not permit the reading of every single nucleotide from at least two amplicons derived from different primers. This double-reading, however, is necessary to shield against phantom mutations [5,32].

In order to meet the minimal requirement that two independent – preferably forward and reverse – sequence strands cover each control-region position, at least 10 sequencing primers are necessary. Actually, as many as 13 were employed by [33], whereas [34] applied a total of 16 primers for full double-strand sequence coverage of degraded DNA. It has further proven useful to perform the sequencing reactions on two independent PCR products amplified with different primer pairs in order to have potential PCR artefacts signposted and to avoid mis-amplification in the case of mutations at primer binding sites. The sequencing primers need to be chosen in a way that complementary sequence information is provided by each of the two amplicons (Fig. 1). Although such an amplification/ sequencing strategy is effectively performed using liquid automation handling in the laboratory, it can easily be carried out manually on 96-well plates with multi-dispensers and eight-channel pipettes. This primer selection then yields full double-stranded control-region sequences for the vast majority of mtDNA samples, including those which display length heteroplasmy around the notorious positions 16,189, 309, 315, 524, and 573. In rare cases, length heteroplasmy can also be observed in other polynucleotide tracts, such as those in the vicinity of positions 16,258, 16,293, 73, and 455 or in some non-repetitive sequence regions due to insertion and deletion events; such instances may then require additional sequencing primers specifically designed for the positions involved.

5. Allocation to haplogroups: multiplex analysis of coding region sites

Full control-region sequences do not always suffice to determine the location of a sequence in the mtDNA phylogeny. Sites from the coding region can then help to eliminate ambiguities [35]. Herrnstadt et al. [36] have pioneered large-scale sequencing of the coding region, however, without providing any control region information. This additional information, however, has been released for a few samples in a piecemeal fashion [37–41]. In view of the enormous amount of HVS-I sequence data published since the early nineties, there is a particular need for thoroughly linking coding-region motifs to control-region motifs. In this respect, sequencing of some mitochondrial genes without linkage to at least HVS-I [42–44] is of little interest. From available portraits of complete mtDNA phylogenies [24–26,39,41,45–49] one can readily infer which mitochondrial genes or specific coding-region variants are particularly useful to discriminate basal parts of a (sub)-continental mtDNA phylogeny.

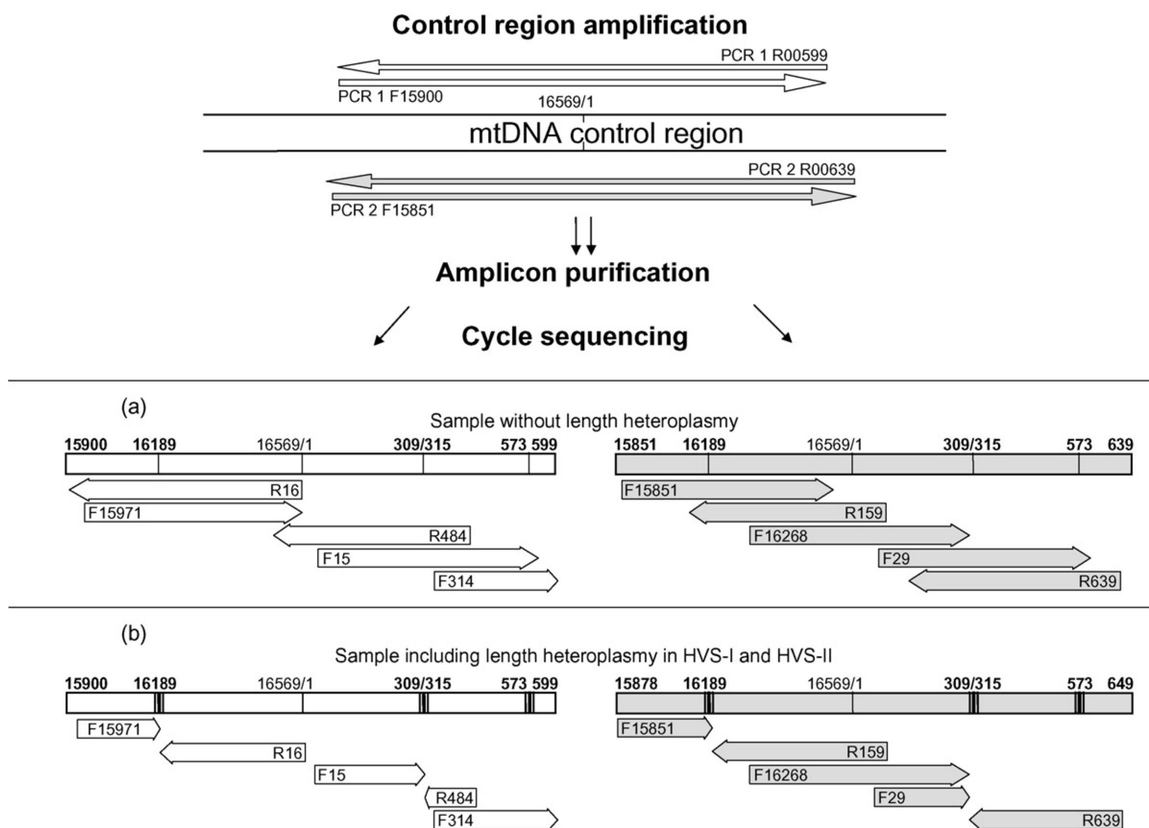

Fig. 1. Amplification and sequencing scheme for full double-strand sequence coverage of the mtDNA control region.

Haplotypes should be allocated to the up-to-date hierarchy of nested haplogroups. To this end, it may be necessary to screen additional positions in the coding region. SNP-based approaches in multiplex format are now becoming increasingly popular in the forensic field [50–56] as they speed up the laboratory process by rapidly eliminating uninteresting samples from the laborious sequence analysis. For defining the haplogroup status of an unknown sample, a too limited set of SNPs may suffer from incorrect default categories and from the circumstance that some positions screened are too variable. The interpretation of unexpected SNP profiles has therefore to be done with caution.

To give an example, assume haplogroup W would be recognized solely by the transitional variant at 8251. This would constitute a poor choice because we now know that this position appears to be among the top-ten highest variable positions of the coding region [13,57]. A better alternative would then be provided by position 1243 or 3505. In addition, it would be preferable to employ also one of the four coding-region variants (say, 11674T) that are characteristic of the super-haplogroup N2 in which W is nested. Similarly, variant 1719A, which has been used for identifying haplogroup X ranks among those top-ten fastest mutations. More problematically, this marker actually seems to define subhaplogroup X2 (or only the major part of it [41]) rather than haplogroup X as a whole.

While the knowledge on (sub)haplogroup-specific coding-region markers is constantly evolving, the forensic scientist faces the challenge that maximum discriminative information is often to be drawn from a sample of limited DNA content and low quality. Such a task requires reliable laboratory technology, which needs to undergo internal validation prior to its use with real-world samples. It is therefore obvious that such technology cannot quite keep up with the pace that new information on loci becomes available. Therefore, SNP-multiplex platforms need to be updated and optimised on a regular basis.

6. Quality control

"A forensic DNA testing laboratory should follow tough quality assurance and quality control guidelines, which are defined by the DNA Advisory Board and the scientific working group on DNA analysis methods, SWGDAM" [58]. We now know, however, that those guidelines are not sufficient. There are multiple sources for errors that can slip into a forensic database. While some of these problems could be eliminated by optimizing the whole sequencing process, the bulk of errors are normally introduced when the data get manually transferred from the sequencer output to the database and printed page. Therefore, any manual transcription should be avoided or necessary scrutiny applied when results are cross-read.

Poor quantitation of primers and/or template DNA leads to top-heavy data in the beginning of a sequence read. This usually goes together with a distinct loss of signal later in the sequence strand and eventually results in a failure to distinguish between signal and noise, which therefore does not permit unequivocal base assignment. Inappropriate instrumental settings can affect electrophoretic DNA migration and may result in ambiguous sequence data. This is particularly true for short and very long DNA fragments. The latter are sometimes found to be poorly resolved and/or of low signal height, which favours oversight of differences relative to the reference sequence. In short fragments, the attached dye has a relatively high impact on the migration rate of the fragments which causes non-rhythmic base-spacing (e.g. A-peaks migrate slower than T-peaks). In addition, homodimer peaks (e.g. AA) tend to be detected as a single broad peak when instrumental settings are far from optimal. Homopolymers also tend to be mis-interpreted for longer DNA fragments where single base-pair resolution is reduced.

We recommended applying alternative primers for re-sequencing of a problematic region in order to avoid repetition of primer-specific artefacts. This is especially true for polycytosine tracts that restrict the interpretable electropherogram to the sequence 5′ of the C-stretch (e.g. 15,971 and 15,851, see Fig. 1). In the presence of long C-tracts an extra internal primer annealing to the polycytosine region can help establishing the electropherogram [59,60], albeit the first nucleotides immediately adjacent to the C-tract may be lost with the sort of sequencing chemistry currently in use.

Any mtDNA data set should undergo routine a posteriori data analysis before submission to print in order to spot unforeseen sequencing and documentation problems. Unusual sequences with suspicious features need another independent amplification for confirmation, and it would be helpful to state this in the publication explicitly. To facilitate a posteriori data evaluation we have implemented a software package in EMPOP (NETWORK; available online), which uses quasi-median network analysis for visualization of mtDNA data sets. An important feature of this method is the filtering option, which highlights parts of the data that should definitely be re-checked by controlling the raw sequence data. This software has proven extremely useful, as it reduces the complexity of a given data set and facilitates spotting unexpected mutations.

7. Statistical analysis

In connection with mtDNA population data, it has been popular to perform some straightforward analyses such as matching analysis, sequence statistics, or calculation of population genetics parameters. To be meaningful, such computations should only be executed on very large databases and not be restricted to single medium-sized data sets. In particular, it does not make much sense to produce these statistics for small samples. A sample of size 100 or even 200 is certainly "small" in this respect [61]. Statistics gleaned from classical population genetics that treat every mtDNA position as if it was an independent allelic marker are not useful as they convey only little information in the case of mtDNA. In contrast to most classical nuclear data, the positions in mtDNA do not constitute independent markers, but instead are linked via the underlying phylogeny, albeit somewhat blurred by recurrent mutations. Therefore, tables of sequence polymorphisms that count the amount of polymorphism at a particular position are at best meaningless for the comparison of different mtDNA

population data sets. For instance, the real transition/transversion rate ratio has to be evaluated along an estimated phylogeny [38], but in forensic genetics this has often been confused with a trivial polymorphism count, which has little to do with the mutational process.

To give another example, any count of how often C is observed at the hypervariable position 16,311 would lump together 'private' mutations (i.e. mutations of a relatively recent origin) together with mutations that occurred deep in the phylogeny and thus belong to the signature of entire haplogroups. For instance, in [21] the fact is highlighted that "nucleotide substitutions at position 16,362 (T to C) in HVR-I have been found with high frequency in Taiwanese". This is not a very meaningful statement because the 16,362 transition ranks among the top five hotspot mutations in HVS-I and belongs to the motif of major haplogroups in East Asia (e.g. D, G, and E) that have different geographic distributions. We conclude that any reasonable statistical analysis of mtDNA data cannot bypass the (estimated) phylogeny.

8. Documentation and notation

It goes without saying that sequencing results for forensic databases have to be made publicly available in electronic form as full sequences aligned to the rCRS and deposited in either GenBank or a public forensic databank, such as EMPOP (www.empop.org) or in the form of mutation-motif profiles as supplementary data on the website of a scientific journal. Home pages of laboratories do not constitute a feasible location for such sequence information, as web addresses and contents may change without warning. The matrix format, i.e. the representation of mtDNA variability in dot tables, has proven to be quite inconvenient for both electronic copying and visual reading and bears a high risk of transcriptional error. Instead, motif lists, which present the haplotypes as profiles (e.g. 16519C, 263G, 315.1C, 523d, 524d) are easier to read and less prone to error. Raw data must be retained and directly linked to the database entry for scrutiny of the data. It is important to maintain mtDNA samples in a deep-freezer for later rechecking and confirmation. For previously published data, errata lists should be made available on a website.

The compact mutation motif profiles should adhere to some standardized notation for enhancing comparison with other profiles. Although a control region profile, such as 73G 263G 315.1C 523d 524d 16093C 16519C, implies an alignment to the rCRS, this does not mean that it was exclusively obtained through binary alignment to the rCRS by employing some formal rules. Rather, the knowledge of the whole mtDNA phylogeny and the inference about the placement of this profile in the mtDNA phylogeny, suggest proper candidate alignments. Even then unique notation cannot always be achieved, so that alternative notations will stay in use. Therefore, (near-) matching analysis needs to explore alternative alignments as well, which could mutually differ up to, say, three seeming mismatches. A focused phylogenetic analysis in the vicinity of a targeted mtDNA profile may therefore be necessary in a number of instances.

As an example consider the two closely related sequences Nai023 and Nai068 from [33], which belong to haplogroup L5a1 [62]: sequence Nai023 was scored as 16183C 16186T 16189C and Nai068 as 16183d 16187T 16189C within the stretch 16,180–16,193. This looks like a difference of three mismatches, but the latter sequence could alternatively be scored as 16183C 16186T 16189C 16193d, yielding only one mismatch. Since the ancestral haplogroup motif in L5a1, however, is 16187T 16189C (within 16,180–16,193), this would rather suggest a scoring of Nai023 and Nai068 as 16183d 16187T 16189C 16193+C and 16183d 16187T 16189C, respectively. In order to decide this ambiguity of alignment, one would need a number of complete mtDNA sequences from haplogroup L5a1.

9. Conclusion

The amount of published sequence data that are flawed is still unacceptably high, also in forensic mtDNA sequencing. While it will be impossible to reduce the error rate to zero, a call for more quality in mtDNA sequencing is mandatory. We have observed that in mtDNA typing the wheel has been reinvented again and again, so to say, including that the same errors are being made over and over again. Some of the above-mentioned suggestions may sound trivial, but in practical work they have proven to increase the quality of a mtDNA population data set.

Acknowledgements

We thank the mtDNA staff at the Institute of Legal Medicine, Innsbruck Medical University Anita Brandstätter, Nina Duftner (currently University of Texas at Austin), Cordula Eichmann, Anna König, Roswitha Mühlmann (now Department for Internal Medicine), Daniela Niederwieser and Bettina Zimmermann for excellent technical work in generating and analyzing thousands of mtDNA sequences.

References

[1] W. Bär, B. Brinkmann, B. Budowle, Á. Carracedo, P. Gill, M. Holland, P.J. Lincoln, W. Mayr, N. Morling, B. Olaisen, P.M. Schneider, G. Tully, M. Wilson, DNA Commission of the International Society for Forensic Genetics: guidelines for mitochondrial DNA typing, Int. J. Legal Med. 113 (2000) 193–196.

[2] Á. Carracedo, W. Bär, P. Lincoln, W. Mayr, N. Morling, B. Olaisen, P. Schneider, B. Budowle, B. Brinkmann, P. Gill, M. Holland, G. Tully, M. Wilson, DNA Commission of the International Society for Forensic Genetics: guidelines for mitochondrial DNA typing, Forensic Sci. Int. 110 (2000) 79–85.

[3] G. Tully, W. Bär, B. Brinkmann, Á. Carracedo, P. Gill, N. Morling, W. Parson, P. Schneider, Considerations by the European DNA profiling (EDNAP) group on the working practices, nomenclature and interpretation of mitochondrial DNA profiles, Forensic Sci. Int. 124 (2001) 83–91.

[4] H.-J. Bandelt, P. Lahermo, M. Richards, V. Macaulay, Detecting errors in mtDNA data by phylogenetic analysis, Int. J. Legal Med. 115 (2001) 64–69.

[5] H.-J. Bandelt, L. Quintana-Murci, A. Salas, V. Macaulay, The fingerprint of phantom mutations in mitochondrial DNA data, Am. J. Hum. Genet. 71 (2002) 1150–1160.

[6] H.-J. Bandelt, W. Parson, Fehlerquellen mitochondrialer DNA-Datensätze und Evaluation der mtDNA-Datenbank ‘‘D-Loop-BASE’’ [Sources of errors in mitochondrial DNA datasets and evaluation of the mtDNA database ‘‘D-Loop-BASE’’], Rechtsmedizin 14 (2004) 251–255.
[7] Á. Carracedo, E. D’Aloja, B. Dupuy, A. Jangblad, M. Karjalainen, C. Lambert, W. Parson, H. Pfeiffer, H. Pfitzinger, M. Sabatier, D. Syndercombe-Court, C. Vide, Reproducibility of mtDNA analysis between laboratories: a report of the European DNA Profiling Group (EDNAP), Forensic Sci. Int. 97 (1998) 165–170.
[8] W. Parson, A. Brandstätter, A. Alonso, N. Brandt, B. Brinkmann, Á. Carracedo, D. Corach, O. Froment, I. Furac, T. Grzybowski, K. Hedberg, C. Keyser-Tracqui, T. Kupiec, S. Lutz-Bonengel, B. Mevag, P. Rafal, H. Schmitter, P. Schneider, D. Syndercombe-Court, E. Sørensen, H. Thew, G. Tully, R. Scheithauer, The EDNAP mitochondrial DNA population database (EMPOP) collaborative exercises: organisation, results and perspectives, Forensic Sci. Int. 139 (2004) 215–226.
[9] A. Salas, L. Prieto, M. Montesino, C. Albarrán, E. Arroyo, M.R. Paredes-Herrera, A.M. Di Lonardo, C. Doutremepuich, I. Fernández-Fernández, A.G. de la Vega, C. Alves, C.M. López, M. López-Soto, J.A. Lorente, A. Picornell, R.M. Espinheira, A. Hernández, A.M. Palacio, M. Espinoza, J.J. Yunis, A. Pérez-Lezaun, J.J. Pestano, J.C. Carril, D. Corach, M.C. Vide, V. Álvarez-Iglesias, M.F. Pinheiro, M.R. Whittle, A. Brehm, J. Gómez, Mitochondrial DNA error prophylaxis: assessing the causes of errors in the GEP’02–03 proficiency testing trial, Forensic Sci. Int. 148 (2005) 191–198.
[10] G. Tully, S.M. Barritt, K. Bender, E. Brignon, C. Capelli, N. Dimo-Simonin, C. Eichmann, C.M. Ernst, C. Lambert, M.-V. Lareu, B. Ludes, B. Mevag, W. Parson, H. Pfeiffer, A. Salas, P.M. Schneider, E. Staalstrom, Results of a collaborative study of the EDNAP group regarding mitochondrial DNA heteroplasmy and segregation in hair shafts, Forensic Sci. Int. 140 (2004) 1–11.
[11] M.R. Wilson, M.W. Allard, K.L. Monson, K.W.P. Miller, B. Budowle, Further discussion of the consistent treatment of length variants in the human mitochondrial DNA control region, Forensic Sci. Commun. 4 (2002) 4.
[12] M.R. Wilson, M.W. Allard, K.L. Monson, K.W. Miller, B. Budowle, Recommendations for consistent treatment of length variants in the human mitochondrial DNA control region, Forensic Sci. Int. 129 (2002) 35–42.
[13] M.D. Coble, The identification of single nucleotide polymorphisms in the entire mitochondrial genome to increase the forensic discrimination of common HV1/HV2 types in the Caucasian population, PhD Thesis, 2004.
[14] H.-J. Bandelt, W. Parson, Consistent treatment of length variants in the human mtDNA control region: a reappraisal, Int. J. Legal Med., 2007, in press.
[15] I.W. Evett, B.S. Weir, Interpreting DNA Evidence: Statistical Genetics for Forensic Scientists, Sinauer Associates, Sunderland, MA, 1998.
[16] S.B. Haga, J.C. Venter, Genetics. FDA races in wrong direction, Science 301 (2003) 466.
[17] S. Lutz, H.-J. Weisser, J. Heizmann, S. Pollak, A third hypervariable region in the human mitochondrial D-loop, Hum. Genet. 101 (1997) 384.
[18] S. Lutz, H.-J. Weisser, J. Heizmann, S. Pollak, Location and frequency of polymorphic positions in the mtDNA control region of individuals from Germany, Int. J. Legal Med. 111 (1998) 67–77; 111 (1998) 286 (erratum); 112 (1999) 145–150 (erratum).
[19] S. Lutz, H. Wittig, H.-J. Weisser, J. Heizmann, A. Junge, N. Dimo-Simonin, W. Parson, J. Edelmann, K. Anslinger, S. Jung, C. Augustin, Is it possible to differentiate mtDNA by means of HVIII in samples that cannot be distinguished by sequencing the HVI and HVII regions? Forensic Sci. Int. 113 (2000) 97–101.
[20] C. Bini, S. Ceccardi, D. Luiselli, G. Ferri, S. Pelotti, C. Colalongo, M. Falconi, G. Pappalardo, Different informativeness of the three hypervariable mitochondrial DNA regions in the population of Bologna (Italy), Forensic Sci. Int. 135 (2003) 48–52.
[21] M.H. Chen, H.M. Lee, C.Y. Tzen, Polymorphism and heteroplasmy of mitochondrial DNA in the D-loop region in Taiwanese, J. Formos. Med. Assoc. 101 (2002) 268–276.
[22] Y.J. Zhang, Q.S. Xu, Z.J. Zheng, H.Y. Lin, J.B. Lee, Haplotype diversity in mitochondrial DNA hypervariable region I, II and III in northeast China Han, Forensic Sci. Int. 149 (2005) 267–269.
[23] A. Achilli, C. Rengo, V. Battaglia, M. Pala, A. Olivieri, S. Fornarino, C. Magri, R. Scozzari, N. Babudri, A.S. Santachiara-Benerecetti, H.-J. Bandelt, O. Semino, A. Torroni, Saami and Berbers—an unexpected mitochondrial DNA link, Am. J. Hum. Genet. 76 (2005) 883–886.
[24] Q.-P. Kong, Y.-G. Yao, C. Sun, H.-J. Bandelt, C.-L. Zhu, Y.-P. Zhang, Phylogeny of East Asian mitochondrial DNA lineages inferred from complete sequences, Am. J. Hum. Genet. 73 (2003) 671–676;
Q.-P. Kong, Y.-G. Yao, C. Sun, H.-J. Bandelt, C.-L. Zhu, Y.-P. Zhang, Phylogeny of East Asian mitochondrial DNA lineages inferred from complete sequences, Am. J. Hum. Genet. 75 (2004) 157, erratum.
[25] M.g. Palanichamy, C. Sun, S. Agrawal, H.-J. Bandelt, Q.-P. Kong, F. Khan, C.-Y. Wang, T.K. Chaudhuri, V. Palla, Y.-P. Zhang, Phylogeny of mitochondrial DNA macrohaplogroup N in India, based on complete sequencing: implications for the peopling of South Asia, Am. J. Hum. Genet. 75 (2004) 966–978.
[26] C. Sun, Q.-P. Kong, M.g. Palanichamy, S. Agrawal, H.-J. Bandelt, Y.-G. Yao, F. Khan, C.-L. Zhu, T.K. Chaudhuri, Y.P. Zhang, The dazzling array of basal branches in the mtDNA macrohaplogroup M from India as inferred from complete genomes, Mol. Biol. Evol. 23 (2006) 683–690.
[27] H.-J. Bandelt, V. Macaulay, M. Richards, Median networks: speedy construction and greedy reduction, one simulation, and two case studies from human mtDNA, Mol. Phylogenet. Evol. 16 (2000) 8–28.
[28] H.-J. Bandelt, A. Salas, S. Lutz-Bonengel, Artificial recombination in forensic mtDNA population databases, Int. J. Legal Med. 118 (2004) 267–273.
[29] H.-J. Bandelt, T. Kivisild, J. Parik, R. Villems, C. Bravi, Y.-G. Yao, A. Brandstätter, W. Parson, Lab-specific mutation processes, in: H.-J. Bandelt, M. Richards, V. Macaulay (Eds.), Human Mitochondrial DNA and the Evolution of *Homo sapiens*, Springer-Verlag, Berlin, Heidelberg, 2006 (Chapter 6).
[30] P. Gill, C. Brenner, B. Brinkmann, B. Budowle, Á. Carracedo, M.A. Jobling, P. de Knijff, M. Kayser, M. Krawczak, W.R. Mayr, N. Morling, B. Olaisen, V. Pascali, M. Prinz, L. Roewer, P.M. Schneider, A. Sajantila, C. Tyler-Smith, DNA commission of the international society of forensic genetics: recommendations on forensic analysis using Y-chromosome STRs, Forensic Sci. Int. 124 (2001) 5–10.
[31] T. Vanecek, F. Vorel, M. Sip, Mitochondrial DNA D-loop hypervariable regions: Czech population data, Int. J. Legal Med. 118 (2004) 14–18.
[32] A. Brandstätter, T. Sänger, S. Lutz-Bonengel, W. Parson, E. Béraud-Colomb, B. Wen, Q.-P. Kong, C.M. Bravi, H.-J. Bandelt, Phantom mutation hotspots in human mitochondrial DNA, Electrophoresis 26 (2005) 3414–3429.
[33] A. Brandstätter, C.T. Peterson, J.A. Irwin, S. Mpoke, D.K. Koech, W. Parson, T.J. Parsons, Mitochondrial DNA control region sequences from Nairobi (Kenya): inferring phylogenetic parameters for the establishment of a forensic database, Int. J. Legal Med. 118 (2004) 294–306.
[34] S.M. Edson, J.P. Ross, M.D. Coble, T.J. Parsons, S.M. Barritt, Naming the dead—confronting the realities of rapid identification of degraded skeletal remains, Forensic Sci. Rev. 16 (2004) 63–90.
[35] M.D. Coble, R.S. Just, J.E. O’Callaghan, I.H. Letmanyi, C.T. Peterson, J.A. Irwin, T.J. Parsons, Single nucleotide polymorphisms over the entire mtDNA genome that increase the power of forensic testing in Caucasians, Int. J. Legal Med. 118 (2004) 137–146.
[36] C. Herrnstadt, J.L. Elson, E. Fahy, G. Preston, D.M. Turnbull, C. Anderson, S.S. Ghosh, J.M. Olefsky, M.F. Beal, R.E. Davis, N. Howell, Reduced-median network analysis of complete mitochondrial DNA coding-region sequences for the major African, Asian, and European haplogroups, Am. J. Hum. Genet. 70 (2002) 1152–1171.
[37] H.-J. Bandelt, C. Herrnstadt, Y.-G. Yao, Q.-P. Kong, T. Kivisild, C. Rengo, R. Scozzari, M. Richards, R. Villems, V. Macaulay, N. Howell, A. Torroni, Y.-P. Zhang, Identification of Native American founder mtDNAs through the analysis of complete mtDNA sequences: some caveats, Ann. Hum. Genet. 67 (2003) 512–524.
[38] H.-J. Bandelt, Q.-P. Kong, M. Richards, V. Macaulay, Estimation of mutation rates and coalescence times: some caveats, in: H.-J. Bandelt, M. Richards, V. Macaulay (Eds.), Human Mitochondrial DNA and the Evolution of *Homo Sapiens*, Springer-Verlag, Berlin, Heidelberg, 2006 (Chapter 4).

[39] D.M. Behar, E. Metspalu, T. Kivisild, A. Achilli, Y. Hadid, S. Tzur, L. Pereira, A. Amorim, L. Quintana-Murci, K. Majamaa, C. Herrnstadt, N. Howell, O. Balanovsky, I. Kutuev, A. Pshenichnov, D. Gurwitz, B. Bonne-Tamir, A. Torroni, R. Villems, K. Skorecki, The matrilineal ancestry of Ashkenazi Jewry: portrait of a recent founder event, Am. J. Hum. Genet. 78 (2006) 487–497.

[40] N. Howell, J.L. Elson, D.M. Turnbull, C. Herrnstadt, African Haplogroup L mtDNA sequences show violations of clock-like evolution, Mol. Biol. Evol. 21 (2004) 1843–1854.

[41] M. Reidla, T. Kivisild, E. Metspalu, K. Kaldma, K. Tambets, H.V. Tolk, J. Parik, E.-L. Loogväli, M.V. Derenko, B. Malyarchuk, M. Bermisheva, S. Zhadanov, E. Pennarun, M. Gubina, M. Golubenko, L. Damba, S. Fedorova, V. Gusar, E. Grechanina, I. Mikerezi, J.P. Moisan, A. Chaventre, E. Khusnutdinova, L. Osipova, V. Stepanov, M. Voevoda, A. Achilli, C. Rengo, O. Rickards, G.F. De Stefano, S. Papiha, L. Beckman, B. Janicijevic, P. Rudan, N. Anagnou, E. Michalodimitrakis, S. Koziel, E. Usanga, T. Geberhiwot, C. Herrnstadt, N. Howell, A. Torroni, R. Villems, Origin and diffusion of mtDNA haplogroup X, Am. J. Hum. Genet. 73 (2003) 1178–1190.

[42] S. Lutz-Bonengel, U. Schmidt, T. Schmitt, S. Pollak, Sequence polymorphisms within the human mitochondrial genes MTATP6, MTATP8 and MTND4, Int. J. Legal Med. 117 (2003) 133–142.

[43] C.Y. Tzen, T.Y. Wu, H.F. Liu, Sequence polymorphism in the coding region of mitochondrial genome encompassing position 8389–8865, Forensic Sci. Int. 120 (2001) 204–209.

[44] U. Szibor, R. Schneider-Stock, C. Augustin, M. Benecke, S. Elias, S. Harada, O. Landt, U. Lass, P.H. Van, K. Schmerbach, H. Wittig, R. Szibor, Variability of the mitochondrial loci nt00073 and nt16519 in populations of Germany, Syria, Cameroon, Japan, Vietnam and Peru—a study using RFLP and Light Cycler technique, Prog. Forensic Genet. 9 (2003) 547–551.

[45] S. Finnilä, M.S. Lehtonen, K. Majamaa, Phylogenetic network for European mtDNA, Am. J. Hum. Genet. 68 (2001) 1475–1484.

[46] J. Friedlaender, T. Schurr, F. Gentz, G. Koki, F. Friedlaender, G. Horvat, P. Babb, S. Cerchio, F. Kaestle, M. Schanfield, R. Deka, R. Yanagihara, D.A. Merriwether, Expanding Southwest Pacific mitochondrial haplogroups P and Q, Mol. Biol. Evol. 22 (2005) 1506–1517.

[47] M. Tanaka, V.M. Cabrera, A.M. González, J.M. Larruga, T. Takeyasu, N. Fuku, L.-J. Guo, R. Hirose, Y. Fujita, M. Kurata, K.-I. Shinoda, K. Umetsu, Y. Yamada, Y. Oshida, Y. Sato, N. Hattori, Y. Mizuno, Y. Arai, N. Hirose, S. Ohta, O. Ogawa, Y. Tanaka, R. Kawamori, M. Shamoto-Nagai, W. Maruyama, H. Shimokata, R. Suzuki, H. Shimodaira, Mitochondrial genome variation in eastern Asia and the peopling of Japan, Genome Res. 14 (2004) 1832–1850.

[48] K. Thangaraj, G. Chaubey, T. Kivisild, A.G. Reddy, V.K. Singh, A.A. Rasalkar, L. Singh, Reconstructing the origin of Andaman Islanders, Science 308 (2005) 996.

[49] J.A. Trejaut, T. Kivisild, J.H. Loo, C.L. Lee, C.L. He, C.J. Hsu, Z.Y. Li, M. Lin, Traces of archaic mitochondrial lineages persist in Austronesian-speaking Formosan populations, PLoS Biol. 3 (2005) e247.

[50] A. Brandstätter, T.J. Parsons, W. Parson, Rapid screening of mtDNA coding region SNPs for the identification of west European Caucasian haplogroups, Int. J. Legal Med. 117 (2003) 291–298.

[51] L. Pereira, M. Richards, A. Goios, A. Alonso, C. Albarrán, O. Garcia, D.M. Behar, M. Gölge, J. Hatina, L. Al-Gazali, D.G. Bradley, V. Macaulay, A. Amorim, Evaluating the forensic informativeness of mtDNA haplogroup H sub-typing on a Eurasian scale, Forensic Sci. Int. 159 (2005) 43–50.

[52] B. Quintáns, V. Álvarez-Iglesias, A. Salas, C. Phillips, M.-V. Lareu, Á. Carracedo, Typing of mitochondrial DNA coding region SNPs of forensic and anthropological interest using SNaPshot minisequencing, Forensic Sci. Int. 140 (2004) 251–257.

[53] K. Umetsu, M. Tanaka, I. Yuasa, N. Saitou, I. Takeyasu, N. Fuku, E. Naito, K. Ago, N. Nakayashiki, A. Miyoshi, S. Kashimura, G. Watanabe, M. Osawa, Multiplex amplified product-length polymorphism analysis for rapid detection of human mitochondrial DNA variations, Electrophoresis 22 (2001) 3533–3538.

[54] K. Umetsu, I. Yuasa, Recent progress in mitochondrial DNA analysis, Leg. Med. (Tokyo) 7 (2005) 259–262.

[55] K. Umetsu, M. Tanaka, I. Yuasa, N. Adachi, A. Miyoshi, S. Kashimura, K.S. Park, Y.H. Wei, G. Watanabe, M. Osawa, Multiplex amplified product-length polymorphism analysis of 36 mitochondrial single-nucleotide polymorphisms for haplogrouping of East Asian populations, Electrophoresis 26 (2005) 91–98.

[56] P.M. Vallone, R.S. Just, M.D. Coble, J.M. Butler, T.J. Parsons, A multiplex allele-specific primer extension assay for forensically informative SNPs distributed throughout the mitochondrial genome, Int. J. Legal. Med. 118 (2004) 147–157.

[57] T. Kivisild, P. Shen, D.P. Wall, B. Do, R. Sung, K. Davis, G. Passarino, P.A. Underhill, C. Scharfe, A. Torroni, R. Scozzari, D. Modiano, A. Coppa, P. de Knijff, M. Feldman, L.L. Cavalli-Sforza, P.J. Oefner, The role of selection in the evolution of human mitochondrial genomes, Genetics 172 (2006) 373–387.

[58] L. Carey, L. Mitnik, Trends in DNA forensic analysis, Electrophoresis 23 (2002) 1386–1397.

[59] W. Parson, T.J. Parsons, R. Scheithauer, M.M. Holland, 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 (1998) 124–132.

[60] R. Szibor, M. Michael, Correct mitochondrial L-strand sequencing after C-stretches, Int. J. Legal Med. 112 (1999) 348–349.

[61] L. Pereira, C. Cunha, A. Amorim, Predicting sampling saturation of mtDNA haplotypes: an application to an enlarged Portuguese database, Int. J. Legal Med. 118 (2004) 132–136.

[62] T. Kivisild, M. Reidla, E. Metspalu, A. Rosa, A. Brehm, E. Pennarun, J. Parik, T. Geberhiwot, E. Usanga, R. Villems, Ethiopian mitochondrial DNA heritage: tracking gene flow across and around the gate of tears, Am. J. Hum. Genet. 75 (2004) 752–770.