

Michael D. Coble · Peter M. Vallone ·
Rebecca S. Just · Toni M. Diegoli ·
Brion C. Smith · Thomas J. Parsons

Effective strategies for forensic analysis in the mitochondrial DNA coding region

Received: 13 June 2005 / Accepted: 30 August 2005 / Published online: 28 October 2005
© Springer-Verlag 2005

Abstract Recently, it has been recognized that accessing information in the mtDNA coding region can provide additional forensic discrimination with respect to the standard typing of the D-loop region, augmenting the sometimes rather limited forensic power of mtDNA testing. Here, we discuss considerations relating to maximally effective approaches for recovering additional discrimination in the coding region, bearing in mind that (1) DNA quality and quantity in typical mtDNA casework usually restrict the amount of additional sequence that can be obtained, and (2) the need for additional discrimination primarily arises when common HV1/HV2 types are encountered. Most investigators have sought additional discrimination by sequencing short segments of coding region that are thought to be particularly variable. Unfortunately, efforts in this regard have generally failed to appreciate that most variation in the coding region is redundant with information already present in HV1/HV2 and have therefore overvalued the potential of this approach for providing additional discrimination. An alternative single nucleotide polymorphism-based approach (Int J Legal Med 118:137–146, 2004) has been to identify specific bases that provide resolution in specific common HV1/HV2 types (and related sequences). We investigate several highly relevant data sets wherein the latter approach performs appreciably better than sequencing selected short portions of the coding region. This is true even when only synony-

mous variation is targeted to minimize the potential for problems arising from discovery of mutations that have reportedly been related to disease.

Keywords Mitochondrial DNA · Single nucleotide polymorphism · mtDNA coding region · Increased forensic discrimination

Introduction

Recently, a number of investigators have sought to increase the forensic discrimination of mtDNA testing by accessing variation that occurs outside of the HV1/HV2 portions of the control region [2, 4, 9, 16, 18, 20, 25, 26]. Given the range of possible approaches, it is therefore worthwhile to seek maximally effective approaches for targeting additional information in the mtDNA-coding region, keeping clearly in mind the circumstances under which the desire for additional discrimination will typically arise. Other considerations should be addressed as we turn to a highly functional coding region, where mutations are known to directly give rise to a wide range of diseases and variation is regularly proposed to be associated in a more complex manner with other diseases or medically significant phenotypes. Coble et al. [9] and Vallone et al. [26] presented discriminatory single nucleotide polymorphism (SNP) panels and assays over the mtDNA genome (mtGenome) that were selected to avoid polymorphic variants with potential for direct phenotypic effect, i.e., those that cause amino acid replacement or modifications in ribosomal or tRNA structure. Budowle et al. [6] have provided a useful discussion regarding disease association and forensic mtDNA testing (to which we will return later) and sharply criticized the synonymous site selection criterion of Coble et al. [9] and Vallone et al. [26] as a “severe” limitation regarding additional information that can be obtained from the coding region. Superficially, this might seem true, as approximately two thirds of the sites in protein-coding genes would vary nonsynonymously. However, we present here an evaluation that produces a different conclusion

M. D. Coble (✉) · P. M. Vallone
Biotechnology Division,
National Institute of Standards and Technology,
100Bureau Drive Mail Stop 8311,
Gaithersburg, MD 20899, USA
e-mail: Michael.coble@nist.gov
Tel.: +1-301-9754330
Fax: +1-301-9758505

R. S. Just · T. M. Diegoli ·
B. C. Smith · T. J. Parsons
Armed Forces DNA Identification Laboratory,
Rockville, MD, USA

regarding the information cost of omitting nonsynonymous variation: The cost is minor, as long as a carefully thought out strategy for synonymous SNP panel selection is used. The considerations that lead to this conclusion are also highly instructive for directing effective approaches for obtaining additional discrimination in the coding region, completely apart from the question of including or excluding nonsynonymous variation. It is our aim to explore these considerations in some detail as part of a dialog on how best to proceed as the field moves forward in these areas.

Materials and methods

Sequence analysis

The regions of HV1 and HV2 used for this study spans from 16,024 to 16,365 for HV1 and from 73 to 340 for HV2. Seventeen individuals identified as the most common HV1/HV2 group (i.e., H:1; see Coble et al. [9]) were identified by control region sequencing (unpublished) from a set of Caucasian population samples [7]. Seven recently published haplogroup H mtGenomes [1] were also identified from their HV1/HV2 sequences as H:1 individuals (AY738960, AY738966, AY738976, AY738979, AY738987, AY738996, and AY739000) and were combined with the 17 unpublished H:1 sequences (above) for a final set of 24 H:1 individuals. Coding region SNPs using multiplex A (MPA) [9, 26] were determined for the 17 H:1 individuals [14].

Six “high-frequency” coding region fragments selected by Allen and Andreasson [2] as particularly informative for increasing the forensic discrimination of individuals that match in HV1/HV2, with particular relevance to the revised Cambridge Reference Sequence (rCRS) [3, 5], were analyzed for the 24 H:1 individuals. The six regions include the nucleotide positions 2,695–2,794; 4,303–4,343; 8,689–8,779; 10,385–10,484; 12,694–12,783; and 15,777–15,872. The number of nucleotides within each region was extrapolated from the dispensation order (5′–3′) for each fragment in Table 3 of Allen and Andreasson [2]. We targeted equivalent information from the H:1 genomes by polymerase chain reaction (PCR) amplification and cycle sequencing of these mtGenome regions as previously described [refer to the Armed Forces DNA Identification Laboratory (AFDIL) protocol presented in Levin et al. [17]]. Automated fluorescent sequencing was performed either on the Applied Biosystems 3100 (Foster City, CA, USA) instrument and used POP-6 polymer on a 36-cm capillary or on the Applied Biosystems 3730 instrument and used POP-7 polymer on a 50-cm capillary. Sequence determination was confirmed from both strands of the mtDNA molecule. Sequences were aligned with the rCRS and edited using Sequencher Plus 4.1.4Fb19 (GeneCodes, Ann Arbor, MI, USA).

Sequence information from 54 complete haplogroup H mtGenomes [1] was also analyzed for each of the six informative fragments of Allen and Andreasson [2] and the 30 haplogroup H SNPs identified by Coble et al. [9]. The sequences are available at the Web site [http://ipvgen.unipv.](http://ipvgen.unipv.it/docs/projects/torroni_data/torroni_sequences.html)

[it/docs/projects/torroni_data/torroni_sequences.html](http://ipvgen.unipv.it/docs/projects/torroni_data/torroni_sequences.html) or from GenBank accession numbers AY738940–AY739001.

Polymorphism frequencies were calculated by using either mtDB (<http://www.genpat.uu.se/mtDB/>) or the Human Mitochondrial Genome Polymorphism database [22] (http://www.giib.or.jp/mtsnp/index_e.shtml).

Results and discussion

First, we consider that the use of mtDNA testing is typically dictated by the advanced state of degradation and low DNA quantity present in a sample (sometimes, of course, also by the availability of only maternally related reference samples). Therefore, in actual practice, extract quantities are typically limiting and small amplicons are needed. It remains the case that the best first course for the recovery of forensically discriminating information resides in the hypervariable regions of the D-loop, and by the time this information has been pieced together according to forensic standards [10], the scientist is lucky to have small amounts of extract left. Having obtained HV1/HV2 information, one likely outcome is that the sequence is either unique or very rare in a given population database, in which case additional sequence resolution would have very little effect, if any, on the strength of the evidence. Where the desire for additional discrimination most often arises is when one of the few more common HV1/HV2 types is encountered—in US “Caucasians,” for example, this is around 20% of the time [9]. Forensic scientists then find themselves wishing for a strategy for going beyond HV1 and HV2 that will permit the use of the small remaining extract to maximal advantage.

One strategy for discovering additional discriminatory variation in the coding region is to fish for it by sequencing any number of additional fragments; however, as noted above, the size and number of fragments that can be recovered will generally be limited by the sample. With this in mind, a number of reports survey variation in selected coding fragments in random population samples [2, 4, 16, 18, 25] to use the level of variation in the fragment as a guide to its potential for increasing discrimination. However, these studies overlook the vitally important assessment of how this coding region variation maps onto variation already present in HV1/HV2. Unfortunately, the complete linkage of mutations over the mtGenome results in the fact that most randomly encountered coding region mutations are redundant in terms of forensic discrimination. Because of the lower evolutionary rate in the coding region, most mutational variants in the coding region are older than those in the D-loop. Fast-evolving D-loop mutations tend to discriminate within coding region lineage markers, rather than vice versa. It is important to understand that “variable” is not a reliable indicator of “informative” when it comes to additional discrimination in the coding region. This is particularly true of nonsynonymous mutations in protein coding genes.

Presently, 2,064 diverse mtGenomes are listed on the mtDB Web site (<http://www.genpat.uu.se/mtDB/>), allow-

ing examination of these issues. The 13 protein coding genes are composed of approximately 11,391 nucleotides of the mtGenome. In the mtDB data set, 2,358 positions varied among the protein coding genes, only 721 of which (31%) are at nonsynonymous positions. Mutations at either of the first two codon positions are almost always nonsynonymous, and there are 7,594 such positions. Hence, only 10% of these varied. Mutations at the third position are almost always synonymous, and there are 3,797 of them; approximately 40% of the third codon positions varied. In searching for additional discrimination among protein coding genes, we have 1,637 variable synonymous sites and 721 variable nonsynonymous sites. These general considerations indicate that the variation is concentrated at synonymous sites (no surprise to the evolutionary biologist), but should not be interpreted to mean that in avoiding nonsynonymous variation, we are discarding 31% of the additional discrimination that resides in the coding region. When combined with variation in the D-loop, and with synonymous variation, a great majority of nonsynonymous variation is redundant.

In the ND1 gene, as a typical example, 61 nonsynonymous mutations were observed in the mtDB database. However, 30 of these occurred only once in the 2,064 sequences: If you sequence the entire 957-bp ND1 gene, you would encounter one of these discriminatory nonsynonymous private mutations about 1% of the time. What about higher frequency nonsynonymous mutations? Seven nonsynonymous ND1 mutations were found in 1% or more of the mtDB database. Comparing these to published phylogenetic analyses ([11, 13, 15, 19, 23, 24] and references within), all of them could be identified as haplogroup-associated polymorphisms. For example, the T4216C variant (occurring in >10% of the mtDB database) is a diagnostic mutation for individuals belonging to the sister haplogroups J and T (about 20% of the Caucasian population) and is therefore completely redundant with J/T diagnostic mutations in the control region (and elsewhere). Remarkably, of the 114 nonsynonymous variants occurring in the entire mtDB database at a frequency of >1%, for only *three* were we unable to readily ascribe an association with a named haplogroup. There can be homoplasious mutations involving even haplogroup-associated polymorphisms, so not all of these will not be wholly uninformative; nevertheless, the clear take home message is that nonsynonymous variation is a poor repository of nonredundant discriminatory information.

Our considerations so far advance us toward the conclusion that avoiding nonsynonymous variation is not very restrictive to tapping the majority of additional discriminatory information in the mtGenome. However, this leaves aside the question of the actual desirability/necessity for doing so, as well as the discussion of constraints that this requirement may place on how the data are acquired. We will return to these significant issues, but now will consider an alternative approach to sequencing selected “highly variable” portions of the coding region. Coble et al. [9] presented a strategy that sought to discover precisely those sites in the mtGenome that are capable of resolving com-

mon HV1/HV2 types (in US Caucasians). The hypothesis underlying this approach was that mutations in the coding region capable of providing additional discrimination would generally be few and scattered widely; moreover, different sites would resolve different common HV1/HV2 lineages. These initial suppositions proved to be true. Coble et al. [9] sequenced 241 complete mtGenomes from individuals matching 18 common US Caucasian HV1/HV2 types and found variation at approximately 500 sites. Eliminating sites with redundant information and avoiding sites that vary only in a single individual and those that represented nonsynonymous or structural RNA variation, we identified 59 sites that provide greatly increased forensic resolution. Few of the sites, however, were “globally” useful. Common types from more distantly related haplogroups (e.g., H and K) were resolved by entirely different sets of sites, and phylogenetic analysis of relative mutation rates over the coding region showed that the useful sites did not correlate well to the faster evolving sites in the coding region [8]. The result of the approach of Coble et al. [9] and Vallone et al. [26] is a series of SNP sites arranged in such a way that for a given common type—or related type—one can turn to one or two multiplex amplifications involving the small number of SNP sites over the entire mtGenome that give the best chance for resolution for the case at hand.

Two potential limitations associated with the SNP approach described above, compared to sequencing selected coding region fragments, are that it fails to access sporadic private mutations (as it makes no sense to target private polymorphisms as discriminatory SNP markers) and that it purposely avoids nonsynonymous variation (a severe limitation, according to Budowle et al. [6]). We have described advantages of the approach as well. How do these factors balance out in actual practice? Data suitable for comparison are not abundant, but we can investigate the relative performance in two highly relevant situations: (1) with the most common European HV1/HV2 type and (2) within a range of HV1/HV2 sequences from haplogroup H (accounting for approximately 40% of European individuals [24]). For these comparisons, we do not consider the mtGenome data of Coble et al. [9], as these were used to discover the SNPs, and ascertainment bias could cloud conclusions. We turn to an independent set of 24 sequences matching the most common European HV1/HV2 type (termed H:1). For these H:1 sequences, we also sequenced six coding region fragments reported in Allen and Andreasson [2] as being highly informative, covering approximately 520 additional bases. We applied multiplex panel A (MPA) SNP [26] to these samples as well. The nine coding region SNPs of MPA provided, in a single amplification assay, much better resolution than sequencing the six fragments, which did not vary at all in 21 of the 24 sequences (Fig. 1a).

To compare performance within a more general sampling of haplogroup H individuals, we examined the recently published data set of 54 haplogroup H mtGenomes [1]. Sequencing the six mtDNA coding region segments of Allen and Andreasson [2] produced no additional variation

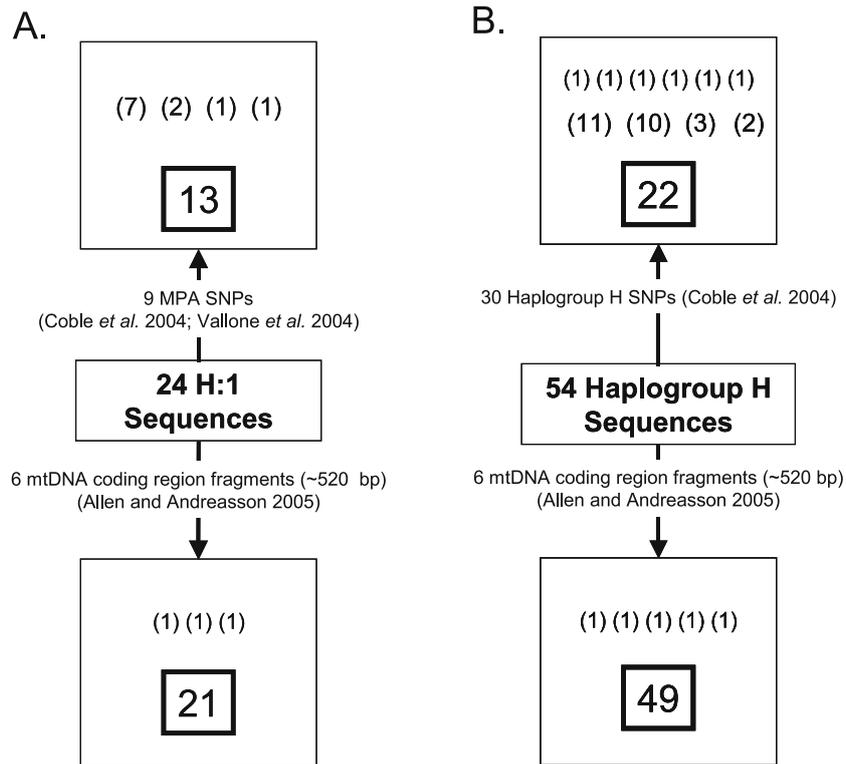


Fig. 1 Forensic discrimination of haplogroup H individuals. Resolution of each group of haplogroup H individuals utilized either multiplex SNPs [9, 26] or a sequencing strategy for highly variable regions in the mtDNA coding region [2]. **a** The application of only coding region information for both strategies. The application of nine multiplex SNPs (*top*) resolved the 24 sequences into five types "two unique types and one type containing 13 individuals. Using six mtDNA coding region fragments (*bottom*) resolved the 24 sequences into four types "three unique and one type containing 21

individuals. **b** Discrimination of 54 haplogroup H mtGenomes [1] utilizing either haplogroup H SNPs identified by Coble *et al.* [9] or a sequencing strategy for highly variable regions in the mtDNA coding region [2]. The application of all 30 haplogroup H SNPs identified by Coble *et al.* [9] resolved the 54 sequences into 11 types "six unique types and one type containing 22 individuals (*top*). Using six coding region mtDNA fragments from Allen and Andreasson [2] resolved the 54 sequences into six types "five unique types and one type containing 49 individuals (*bottom*)

in 49 of these 54 sequences, with four individuals resolved as unique types (Fig. 1b). Using the 30 coding region SNPs designated by Coble *et al.* [9] as useful for discriminating within haplogroup H, on the other hand, separated the 54 haplogroup H mtGenomes into 11 types: 6 occurring only once, with the most frequent type consisting of 22 individuals (Fig. 1b). Evolutionarily, it makes sense that polymorphic sites identified as particularly discriminatory markers for common HV1/HV2 types will also tend to perform well in closely related sequences (which themselves tend to be common, as offshoots of predominant founder lineages). The comparatively good performance of select SNPs in this general population sample indicates that this is, in fact, the case. However, because these 54 H sequences of Achilli *et al.* [1] included individuals with very rare HV1/HV2 types, this comparison significantly undervalues the performance that the SNPs would exhibit under conditions when additional discrimination would actually be needed: when common types are encountered.

We have seen that carefully selected synonymous SNP panels actually provide a significantly better chance for increased forensic resolution than the published approach of Allen and Andreasson [2], where multiple short fragments are sequenced via PyroSequencing. Nonetheless,

some discriminatory information is sacrificed using the conservative criterion of avoiding nonsynonymous variation. Recognizing this, we reviewed the data of Coble *et al.* [9] for nonsynonymous variation that provided further discrimination. Only 32 nonsynonymous SNPs were discovered that were not redundant with synonymous SNP variation. Twenty-three of these varied in only a single individual and were widely scattered, making them amenable to neither SNP assays nor sequencing of short fragments. The remaining nine nonsynonymous SNPs (G17 19A, A1811G, C2772T, C4025T, T4639C, T8433C, G15 323A, G15773A, and A15924G) do contribute a modest increase in the number of haplotypes compared to the eight multiplex panels of Coble *et al.* [9], increasing the total number of types from 112 to 127 (from the 18 HV1/HV2 types among 241 common HV1/HV2 type mtGenomes). As a stand-alone nonsynonymous multiplex panel, these SNPs could in some cases usefully augment the multiplex panels of Coble *et al.* [9].

We now turn to an overdue discussion of the actual desirability of avoiding nonsynonymous protein-coding or structural RNA variation. Budowle *et al.* [6] present a cogent discussion of many of the significant issues. Our perspective is not far from theirs, particularly regarding the

conclusion that the choice to restrict information assayed or reported should be left to individual practicing laboratories, in regard to the application at hand. We have also long been aware that, as a theoretical issue, it is impossible to apply an absolute restriction against the potential for detecting variation that might be associated with medically relevant information and still be left with any DNA testing approach that passes muster. As Budowle et al. [6] point out, mtDNA mutations, whether synonymous or not or whether in the control region or the coding region, are linked to every other mutation in the same lineage (although homoplasious mutations tend to erase the traces of this linkage rather quickly in many instances). That said, we are well served to step away from theory and be guided by the potential for problems to arise in actual practice.

At the AFDIL, mtDNA testing is applied on a large scale in the context of missing persons identification. This work tends to be publicly visible and involves large segments of the general population. Moreover, the US military now has a policy of compulsory submission of a blood sample retained solely for the purposes of DNA identification, which is necessary in the face of military casualty. In this context, issues on genetic privacy are weighed with gravity, and there was concern relating to fullscale sequencing in coding genes where mutations have been directly associated with numerous diseases and other life-history traits (<http://www.mitomap.org/>).

This concern went from theory to reality in one of AFDIL's very first instances of accessing coding region sequence variation for increased discrimination. Sequence data were reported to the reference subject, who out of curiosity surfed the Web and quickly determined from the Mitomap Web site that he or she shared several mutations listed under the heading "Reported Mitochondrial DNA Base Substitution Diseases." This included one mutation associated with an increased risk for an adverse condition that arises later in life. At that point, the real issue facing AFDIL was not whether this genetic risk was in reality high, low, or even nonexistent [12], but whether it was proper and/or desirable for AFDIL as an institution to become involved in these issues. In response and in coordination with careful deliberation by the Armed Forces Institute of Pathology Institutional Review Board, AFDIL adopted the conservative approach of restricting attention (almost exclusively) to variation at synonymous sites in the coding region. This retains essentially an equal footing with accessing variation in the D-loop, which has yet not presented any problems. (The Mitomap site currently lists 240 mutations putatively associated with a disease at non-synonymous and structural RNA sites; it lists none at synonymous sites, and it lists the *one* mutation in the control region that was cited by Budowle et al. [6].)

Conclusions

It may be that the restriction adopted by AFDIL is more conservative than necessary for some laboratories, in the context of their own work. However, in reaching that

determination, it should be useful for laboratories to know that carefully devised strategies based on synonymous SNPs can provide a highly effective means for accessing increased discrimination. In fact, the synonymous SNP strategy described in Coble et al. [9] is quite a bit more effective than the sequence-based strategies presented to date. An obvious initial disadvantage of the SNP approach is the large amount of work required to obtain the necessary knowledge of mtGenome variation to accurately target maximally discriminatory sites. However, once the information is obtained, it remains available for all applications, and work is well advanced to provide this information for common HV1/HV2 types beyond those of European ancestry described in Coble et al. [9].

Although one cannot absolutely avoid medically significant information in any genetic typing system, and that some established typing systems have been shown to have significant associations [6], it seems there remains a general sense that it is best to choose and design systems to avoid these associations when reasonably possible. Some countries, such as Germany, have strict regulations in this regard [18], resulting in the call for disqualification of certain markers [21]. As we work through these issues for mtDNA, it is important to accurately weigh the costs and benefits of various approaches and to seek the most effective ways for accessing additional variation under the most applicable circumstances. The calculus may also change if, for example, new technical capabilities are developed for recovering much larger portions of the entire mtGenome from small amounts of degraded extracts in a cost- and time-effective manner. Supplementary information for this paper can be found at: <http://www.cstl.nist.gov/biotech/strbase/mtDNA.htm>.

Acknowledgements We thank John M. Butler of NIST for discussion and comments on the manuscript. The contribution of the U.S. National Institute of Standards and Technology is not subject to copyright. Certain commercial equipment, instruments, and materials are identified to specify experimental procedures as completely as possible. In no case does such identification imply a recommendation or endorsement by the NIST, nor does it imply that any of the materials, instruments, or equipment identified are necessarily the best available for the purpose. This work was supported by the National Institutes of Justice agreement 2003-IJ-R-029 with the NIST Office of Law Enforcement Standards and by a National Institutes of Justice grant 2000-IJ-CX-K010 to the American Registry of Pathology (T.J.P.). Points of view in this document are those of the authors and do not necessarily represent the official position or policies of the U.S. Department of Justice, the U.S. Department of Defense, or the U.S. Department of the Army.

References

1. Achilli A, Rengo C, Magri C, Battaglia V, Olivieri A, Scozzari R, Cruciani F, Zeviani M, Briem E, Carelli V, Moral P, Dugoujon JM, Roostalu U, Loogvali EL, Kivisild T, Bandelt HJ, Richards M, Villems R, Santachiara-Benerecetti AS, Semino O, Torroni A (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

2. Allen M, Andreasson H (2005) Mitochondrial D-loop and coding sequence analysis using pyrosequencing. *Methods Mol Biol* 297:179–196
3. Anderson S, Bankier AT, Barrell BG, deBruijn MHL, Coulson AR, Drouin J, Eperon IC, Nierlich DP, Roe BA, Sanger F, Schreier PH, Smith AJH, Staden R, Young IG (1981) Sequence and organization of the human mitochondrial genome. *Nature* 290:457–465
4. Andreasson H, Asp A, Alderborn A, Gyllensten U, Allen M (2002) Mitochondrial sequence analysis for forensic identification using pyrosequencing technology. *Biotechniques* 32: 124–133
5. 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
6. Budowle B, Gyllensten U, Chakraborty R, Allen M (2005) Forensic analysis of the mitochondrial coding region and association to disease. *Int J Legal Med* 119:314–315. Epub Apr 21, 2005
7. Butler JM, Schoske R, Vallone PM, Redman JW, Kline MC (2003) Allele frequencies for 15 autosomal STR loci on U. S. Caucasian, African American, and Hispanic populations. *J Forensic Sci* 48:908–911
8. Coble MD (2004) 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. Ph.D. dissertation, The George Washington University. Available as a pdf file at: http://www.cstl.nist.gov/biotech/strbase/pub_pres/Coble2004dis.pdf
9. Coble MD, Just RS, O'Callaghan JE, Letmanyi IH, Peterson CT, Parsons TJ (2004) Single nucleotide polymorphisms over the entire mtDNA genome that increase the forensic power of mtDNA testing in Caucasians. *Int J Legal Med* 118:137–146
10. Edson SM, Ross JR, Coble MD, Parsons TJ, Barritt SM (2004) Naming the dead—confronting the realities of rapid identification of degraded skeletal remains. *Forensic Sci Rev* 16(1):64–89
11. Finnila S, Lehtonen MS, Majamaa K (2001) Phylogenetic network for European mtDNA. *Am J Hum Genet* 68:1475–1484
12. Herrnstadt C, Howell N (2004) An evolutionary perspective on pathogenic mtDNA mutations: haplogroup associations of clinical disorders. *Mitochondrion* 4:791–798
13. Herrnstadt C, Elson JL, Fahy E, Preston G, Turnbull DM, Anderson C, Ghosh SS, Olefsky JM, Beal MF, Davis RE, Howell N (2002) Reduced-median-network analysis of complete mitochondrial DNA coding-region sequences for the major African, Asian, and European haplogroups. *Am J Hum Genet* 70:1152–1171
14. Kline MC, Vallone PM, Redman JW, Duewer DL, Calloway CD, Butler JM (2005) Mitochondrial DNA typing screens with control region and coding region SNPs. *J Forensic Sci* 50:377–385
15. Kong QP, Yao YG, Sun C, Bandelt HJ, Zhu CL, Zhang YP (2003) Phylogeny of East Asian mitochondrial DNA lineages inferred from complete sequences. *Am J Hum Genet* 73:671–666
16. Lee SD, Lee YS, Lee JB (2002) Polymorphism in the mitochondrial cytochrome B gene in Koreans. An additional marker for individual identification. *Int J Legal Med* 116:74–78
17. Levin BC, Holland KA, Hancock DK, Coble M, Parsons TJ, Kienker LJ, Williams DW, Jones M, Richie KL (2003) Comparison of the complete mtDNA genome sequences of human cell lines—HL-60 and GM10742A—from individuals with pro-myelocytic leukemia and Leber hereditary optic neuropathy, respectively, and the inclusion of HL-60 in the NIST human mitochondrial DNA standard reference material—SRM 2392-I. *Mitochondrion* 2:387–400
18. Lutz-Bonengel S, Schmidt U, Schmitt T, Pollak S (2003) Sequence polymorphisms within the human mitochondrial genes MTATP6, MTATP8, and MTND4. *Int J Legal Med* 117:133–142
19. 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
20. Parsons TJ, Coble MD (2001) Increasing the forensic discrimination of mitochondrial DNA testing through analysis of the entire mitochondrial DNA genome. *Croat Med J* 42:304–309
21. Szibor R, Hering S, Edelmann J (2005) The HumARA genotype is linked to spinal and bulbar muscular dystrophy and some further disease risks and should no longer be used as a DNA marker for forensic purposes. *Int J Legal Med* 119 (3):179–180
22. Tanaka M, Takeyasu T, Fuku N, Li-Jun G, Kurata M (2004) Mitochondrial genome single nucleotide polymorphisms and their phenotypes in the Japanese. *Ann N Y Acad Sci* 1011:7–20
23. Tanaka M, Cabrera VM, Gonzalez AM, Larruga JM, Takeyasu T, Fuku N, Guo LJ, Hirose R, Fujita Y, Kurata M, Shinoda K, Umetsu K, Yamada Y, Oshida Y, Sato Y, Hattori N, Mizuno Y, Arai Y, Hirose N, Ohta S, Ogawa O, Tanaka Y, Kawamori R, Shamoto-Nagai M, Maruyama W, Shimokata H, Suzuki R, Shimodaira H (2004b) Mitochondrial genome variation in eastern Asia and the peopling of Japan. *Genome Res* 14:1832–1850
24. 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
25. Tzen CY, Wu TY, Liu HF (2001) Sequence polymorphism in the coding region of mitochondrial genome encompassing position 8389–8865. *Forensic Sci Int* 120:204–209
26. Vallone PM, Just RS, Coble MD, Butler JM, Parsons TJ (2004) A multiplex allele-specific primer extension assay for forensically informative SNPs distributed throughout the mitochondrial genome. *Int J Legal Med* 118:147–157