Int J Legal Med (2005) 119: 314–315
DOI 10.1007/s00414-005-0543-y

LETTER TO THE EDITOR

B. Budowle · U. Gyllensten · R. Chakraborty · M. Allen

Forensic analysis of the mitochondrial coding region and association to disease

Received: 1 December 2004 / Accepted: 23 March 2005 / Published online: 21 April 2005

Coble et al. [3] and Vallone et al. [5] have proposed that forensic analyses of the coding region in the mitochondrial DNA (mtDNA) should be restricted to synonymous substitutions only to avoid detection of changes with phenotypic expression. They suggest that sequencing strategies for forensic analyses of the coding region of the mtDNA genome should be avoided due to potential risk of obtaining personal disease state information and that only SNP-based systems should be employed. We disagree with this proposition as applying such a strict criterion is not well thought-out and also would severely hamper the use of mtDNA in forensic testing. There are a number of reasons not to limit the analyses in the way suggested by these authors.

First, most of the disorders associated with mtDNA mutations are uncommon or rare and have variable penetrance. At the mitomap website (http://www.mitomap.org/), over 80 mtDNA substitutions have been reported as associated with disease. Of these, 17 are confirmed by different groups as associated with LHON, NARP, and MELAS, among other diseases. The frequency of these substitutions among a collection of 746 coding region sequences (control samples) from the mtDB website (http://www.genpat.uu.se/mtDB/) is for 3460A (0/746), 4216C (93/746), 4917G (52/746), 11778A (1/746), 14482G (0/746), 14484C (3/746), 8993C/G (0/746), 9176C (0/746), 10158C (0/746), 10191C (0/746), 11777C (0/746), 12706C (0/746), 14459A (1/746), 14487C (0/746), 13513A (0/746), and 13514G (0/746). There are also over 100 rRNA or tRNA mutations associated with MELAS, deafness, and MERRF, among other diseases, of which 21 are confirmed. The frequency of these positions in the mtDB database is for 1555G (2/746), 1606A (0/746), 3243G (0/746), 3256T (0/746), 3260G (0/746), 3271C (0/746), 3291C (0/746), 3302G (0/746), 3303T (0/746), 4298A (0/746), 4332A (0/746), 5537insT (0/746), 5814C (4/746), 7445G (0/746), 7471insC (0/746), 7511C (0/746), 8356C (0/746), 10010C (0/746), 12147A (0/746), 12315A (0/746), and 14709C (1/746). Thus, most of these substitutions are not found in the mtDB database. Although the two substitutions, 4216C (93/746) and 4917G (52/746), are quite frequent in the mtDB website database, these are more frequent among LHON patients than in controls. These variants are at best predisposing factors for development of LHON but not primary LHON mutations. Also, some diseases are likely to require a contribution from nuclear genes or environmental factors to develop. Consequently, it is difficult to limit the number of mtDNA substitutions that are truly associated with disease, of which there are very few.

Second, by limiting the analysis only to synonymous polymorphisms that cannot have any phenotypic effect, a large part of the polymorphic positions (and thus forensically informative) would be excluded. In the mitomap database, 927 mtDNA coding region polymorphisms are reported. Of these, 517 are synonymous changes, resulting in that 44% of the reported variation would be excluded in forensic analysis. This may be an underestimate of the proportion of sites that would be excluded, given that the polymorphisms reported in the mitomap database do not represent information obtained from complete mtDNA genomes, but rather specific regions of interest compiled from different studies. As a further example of information loss, if one were to type only the most frequent SNPs in the non-coding hypervariable regions (such as those described by Gabriel et al. [4]) in lieu of sequencing, only 15% of the variation would be obtained.

Third, even if only synonymous changes are being used, since the mtDNA effectively is one locus, such polymorphisms would still be linked to functional polymorphisms in

B. Budowle
FBI Laboratory,
Quantico, VA, 22135, USA

U. Gyllensten · M. Allen (✉)
Department of Genetics and Pathology,
Uppsala University,
751 85 Uppsala, Sweden
e-mail: marie.allen@genpat.uu.se
Fax: +46-18-4714808

R. Chakraborty
Department of Environmental Health,
University of Cincinnati,
Ohio, 45267, USA

other regions of the molecule and therefore indirectly yield information relevant to disease associations. To use only synonymous variants that are in complete equilibrium with every known or presumed polymorphism associated to disease is not only practically impossible, but would also severely limit the number of sites available for forensic analysis. In addition, certain diseases have been associated to entire mtDNA haplogroups rather than specific sites [1, 6]. The intention of selecting only positions not associated to disease would also need to reject sequencing of the d-loop, as there are some d-loop positions that are known to be associated with disease. For example, the 16189C substitution is associated to type 2 diabetes and cardiomyopathy (see http://www.mitomap.org). The relative risk is low for carriers of 16189C to develop these diseases, but the example points to the fact that there are many sites that are associated with some disease. Indeed, if Coble et al. and Vallone et al. were to follow their proposition, they would have to forego sequencing of the non-coding region of human mtDNA for identity testing. To the best of our knowledge, the Armed Forces DNA Identification Laboratory still employs sequencing of the non-coding region of the mtDNA for identification of human remains.

Fourth, the use of markers associated with disease is not new in the area of forensic genetics. Before the use of DNA technology, hemoglobin variants, such as the protein encoded by the sickle cell allele, were typed for identity purposes. Alpha-1 antitrypsin, used in paternity testing, has variants that are associated with a number of diseases, such as emphysema. Further, one of the first kits developed for forensic DNA analysis used the polymorphic locus HLA DQA1, with alleles that are in strong disequilibrium with other HLA loci and associated to a number of diseases. For example, the DR3/DR4 genotype is strongly associated with type 1 diabetes. This represents one of the first described and strongest associations between a genetic marker and a disease. Nevertheless, DQA1 typing was performed for many years in forensic cases without causing any ethical concerns. However, despite the fact that this association has been known for a number of years, this information has not been disclosed to the individuals in question. The data, to date, clearly demonstrate that any predictive value (and thus the concern for any possible privacy concerns) is small at best. A similar situation occurs in clinical tissue typing for transplantation where family members are tested and non-paternity is revealed. This information is not disclosed to the family as paternity testing is not the aim of the analysis. Even markers used in routine analysis today have been associated with disease.

Finally, associations of mtDNA variants with disease are very heterogeneous and, in many cases, not adequately predictive of disease co-morbidity. In our opinion, the perceived benefit of excluding these positions is questionable. The result would be to severely compromise the information that can be obtained in forensic DNA analysis. However, we all need to be sensitive to potential bias that may arise and consider reasonable approaches. If restrictions are to be made, other strategies can be chosen that will not limit the use of mtDNA analysis in forensics so severely. One approach could be to redact and not disclose the positions that have been confirmed as associated with disease (i.e., the 38 positions mentioned above for the moment—although this may be overly reactive). If the data at a redacted site would exclude a suspect, the exclusion could be stated, but the site(s) would remain redacted. This approach would not impact population databases since sample source is anonymous. An alternative, less restricting strategy would be to exclude and not disclose heteroplasmic positions obtained in a specific analysis. This would prevent revealing a majority of the disease-associated variants, as homoplasmy often is incompatible with life. Among the few diseases that are associated with homoplasmic mutations, environmental risk factors or nuclear gene mutations are likely to be required for disease development as the penetrance is quite variable [2].

Addressing the best approach or deciding which sites should be redacted could be left to individual practicing laboratories. However, to ensure a thorough review of the issue and to have some consistency of practices (taking into account country-specific legal/privacy restrictions), it would be better to develop recommendations under the auspices of a forensic working group, such as the ISFG and/or SWGDAM.

References

1. Brown MD, Starikovskaya E, Derbeneva O et al (2002) The role of mtDNA background in disease expression: a new primary LHON mutation associated with Western Eurasian haplogroup. J Hum Genet 110:130–138
2. Carelli V, Giordano C, d'Amati G (2003) Pathogenic expression of homoplasmic mtDNA mutations needs a complex nuclear–mitochondrial interaction. Trends Genet 19:257–262
3. Coble MD, Just RS, O'Callaghan JE, Letmanyi IH, Peterson CT, Irwin JA, Parsons TJ (2004) Single nucleotide polymorphisms over the entire mtDNA genome that increase the power of forensic testing in Caucasians. Int J Leg Med 118:137–146
4. Gabriel MN, Calloway CD, Reynolds RL, Primorac D (2003) Identification of human remains by immobilized sequence-specific oligonucleotide probe analysis of mtDNA hypervariable regions I and II. Croat Med J 44:293–298
5. 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 Leg Med 118:147–157
6. van der Walt JM, Dementieva YA, Martin ER, Scott WK et al (2004) Analysis of European mitochondrial haplogroups with Alzheimer disease risk. Neurosci Lett 15(365):28–32