Forensic Science International: Genetics 6 (2012) e182–e184

Contents lists available at SciVerse ScienceDirect

Forensic Science International: Genetics

journal homepage: www.elsevier.com/locate/fsig

Letter to the Editor

A cautionary note on switching mitochondrial DNA reference sequences in forensic genetics

Dear Editor,

The first human mitochondrial DNA (mtDNA) sequence was produced in 1981 from an individual of European descent [1]. Since then, this sequence has been known as the Cambridge Reference Sequence (CRS) with a total length of 16,569 base pairs. As is common practice in other fields of genome research, this first mitochondrial genome (mtGenome) served as reference for the scientific community, relative to which other mtDNA haplotypes were reported. Eighteen years later the CRS was re-sequenced and corrected at 10 positions (3423T, 4985A, 9559C, 11335C, 13702C, 14199T, 14272C, 14365C, 14368C, and 14766C) to form the revised Cambridge Reference Sequence (rCRS) [2]. The new analysis also revealed that this mtGenome consists of only 16,568 nucleotides, as a base at position 3107 was mistakenly reported in the CRS. Instead of redefining all nucleotide positions downstream of 3107, this position is indicated in the rCRS as a deletion (unfortunately often indicated as "N", which is reserved for any base in the IUP-code [3]). Thus, the numbering system employed for the CRS and the body of established data can continuously be used with the rCRS. More than 15,500 mtGenomes and well over 150,000 (partial) control region sequences (including databases) have been published to date (http://www.phylotree.org/, Phylotree Build 14 [4]), in which the CRS and the rCRS have been cited 5603 and 968 times, respectively (http://apps.isiknowledge.com/; queried on May 2012).

In a recent study, the switch to a new reference sequence, the so-called Reconstructed Sapiens Reference Sequence (RSRS), has been proposed [5]. This ancestral reconstructed sequence represents the deepest root in the known human mtDNA phylogeny at the base of the split of haplogroups L0 and L1′2′3′4′5′6 after combining sequence information from all available mtGenomes from *Homo neanderthalensis* and novel human mtGenomes. The authors believe that the switch would solve misunderstandings and problems associated with the existing nomenclature relative to the rCRS, which belongs to the recently coalescing European haplogroup H2a2a1. In the following we briefly review developments in forensic mitochondrial genetics and discuss possible implications of the proposed switch.

Mitochondrial DNA is highly abundant in cells compared to nuclear DNA (nDNA) with increased typing success rates for analysis of highly degraded samples and also hair shafts that often do not harbor detectable amounts of nDNA. To assess the significance of a match between two mtDNA haplotypes (e.g. from the crime scene and from a suspect), mtDNA databases have been developed. Earlier, the difference-coded haplotypes (with respect to the rCRS) were directly compared to the haplotypes in the mtDNA databases with the risk that multiple different alignments of the same sequence led to biased results [6]. The field addressed this potential drawback by developing alignment-free query engines [7] to guarantee that all haplotypes can be detected in a search regardless of the alignment employed. This way, the established notation scheme relative to the rCRS could be maintained.

Considerations on the working practices, nomenclature and interpretation of mtDNA profiles were established already more than ten years ago by the European DNA profiling (EDNAP) group [8]. In the past decade the forensic community has moved towards a phylogenetic approach to analyzing mtDNA for several reasons (see [9] for a review). Due to its maternal mode of inheritance, mtDNA passes through the matrilineage as a single haplotypic block, where new germ-line mutations accumulate over existing haplotypes. Reconstructing evolutionary patterns of variation is straightforward in mtDNA sequences compared to autosomal markers where phylogenies are much more complex due to recombination. Thus, the mtDNA phylogeny is very useful for differentiating between patterns of variation that are natural from those that may be artificially generated in the laboratory (due to artificial recombination through unintentional sample mix-up, contamination, phantom mutations, etc. . .). Screening for patterns of variation in mtDNA datasets that are inconsistent or unexpected with respect to the established phylogeny has been very useful for the reduction of sequencing and interpretation errors (e.g. [10–15]). The forensic community is particularly attuned to the detrimental impact of technical errors, given that mtDNA testing has become an essential tool for serving the public in criminal, missing persons and disaster victim identifications. The phylogenetic concept of high-quality mtDNA analysis has been driven by the establishment of EMPOP (www.empop.org; [16]), a freely accessible collection of mtDNA haplotypes for forensic evaluation and an important resource of phylogenetically based tools to aid quality control efforts (e.g. [7,17–19]).

In contrast to other fields of research, the forensic community strives for a high degree of consistency between laboratories and countries when it comes to technical and interpretation methods, as the derived results provide evidence for legal proceedings in criminal and civil cases. Therefore, forensic researchers are active in discussing best practice protocols and guidelines that aid the international harmonization process. This includes recommendation letters and collaborative multi-centric exercises [20–26] in order to provide more powerful solutions to typical problems affecting forensic casework as well as detect sources of errors in forensic laboratories. Also, it is essential for the forensic field that novel scientific achievements can be used by practitioners to conduct their legal services.

The above-mentioned characteristics demonstrate that a sudden switch from an established reference (the rCRS) to a newly proposed standard would require a well-concerted process accompanied by rigorous validation of the new approach. As an immediate response to the new proposal, the following issues would need to be addressed in closer detail.

1872-4973/$ – see front matter
http://dx.doi.org/10.1016/j.fsigen.2012.06.015

The generation, manipulation, and analysis of mtDNA sequence data is inherently complex, particularly with challenging forensic samples that often need to be amplified in separate overlapping amplicons. A consequence of this is the requirement for a high degree of user knowledge of mtDNA variation in order to be aware of patterns of variation and site-specific mutation rates. Simply put, the current nomenclature system relative to the rCRS is the language with which expert users understand and communicate mtDNA sequence variation, and modifications would be extremely disruptive to expert familiarity and discourse.

In addition, reporting haplotypes in an unfamiliar format will most likely lead to an increased risk of error at various stages. It has been repeatedly demonstrated that transcription and clerical errors are the main sources of error in proficiency tests ([21–28], GEDNAP blind trial C. Hohoff, personal communication) and occur frequently in population and clinical studies ([12,13,29–31], and as experience from EMPOP QC tests as outlined in [18,32]). It is important to learn from past experience where the simple transition between the incorrect CRS to the corrected rCRS led to a number of errors in the literature [13,27,33] and the fact that both nomenclatures were living together for a large period of time.

There are vast resources of mtDNA population databases and analytical tools that have been established according to the rCRS. Forensic casework requires comparison of mtDNA haplotypes to population databases, and a switch in reference system would require transformation of all mtDNA databases currently in use. Informatic tools and strategies for haplogroup assessment and quality control, based on sequence motifs, would likewise all need to be revised.

Forensic reports would need to be adapted to the new nomenclature. Previously issued casework reports would require clarification, as different reported results would be produced at different times for the identical haplotype. Forensic geneticists are well aware of the difficulties that such actions have in court.

Forensic geneticists are also active in reporting mtDNA population data in forensic and non-forensic journals. This data is frequently used in forensic casework for database searches. The fact that two different nomenclatures would live together in the literature and that databases would both complicate knowledge exchange, could lead to misunderstandings when (manual) database queries are executed.

Conversion of mtDNA nomenclature based on an inferred ancestral haplotype, although containing some attractive conceptual features [5], could lead to some dramatic practical consequences for forensic applications. It is important for the forensic community to avoid or at least minimize negative consequences that the use of a double nomenclature would have in the field [5]. One of the key benefits of any nomenclature system is to ensure that the scientific community speaks a common language that permits easy comparison of results across laboratories, disciplines, and time periods. At this stage of development the benefit of switching to the RSRS does not seem to justify the cost and effort nor the risk of misunderstanding, confusion and error.

Therefore, while waiting for an international consensus regarding all the spheres of mtDNA research, for the time being we believe that the forensic field should keep the rCRS as the standard for nomenclature in forensic DNA studies and casework.

Acknowledgments

This study received support from the Ministerio de Ciencia e Innovación (SAF2008-02971 and SAF2011-26983) (A.S.) and the Austrian Science Fund (P22880-B12) (WP). LG is supported by a grant from CAPES/Brasil. The opinions and assertions contained herein are solely those of the authors and are not to be construed as official or as views of the US Department of Defense or the US Department of the Army.

References

[1] S. Anderson, A.T. Bankier, B.G. Barrell, M.H.L. de Bruijn, A.R. Coulson, J. Drouin, I.C. Eperon, D.P. Nierlich, B.A. Roe, F. Sanger, P.H. Schreier, A.J.H. Smith, R. Staden, I.G. Young, Sequence and organization of the human mitochondrial genome, Nature 290 (1981) 457–465.
[2] R.M. Andrews, I. Kubacka, P.F. Chinnery, R.N. Lightowlers, D.M. Turnbull, N. Howell, Reanalysis and revision of the Cambridge reference sequence for human mitochondrial DNA, Nat. Genet. 23 (2) (1999) 147.
[3] A. Cornish-Bowden, Nomenclature for incompletely specified bases in nucleic acid sequences: recommendations 1984, Nucl. Acids Res. 13 (9) (1985) 3021–3030.
[4] M. van Oven, Revision of the mtDNA tree and corresponding haplogroup nomenclature, Proc. Natl. Acad. Sci. USA 107 (11) (2010) E38–E39 (author reply e40-31).
[5] D.M. Behar, M. van Oven, S. Rosset, M. Metspalu, E.L. Loogvali, N.M. Silva, T. Kivisild, A. Torroni, R. Villems, A "copernican" reassessment of the human mitochondrial DNA tree from its root, Am. J. Hum. Genet. 90 (4) (2012) 675–684.
[6] H.-J. Bandelt, W. Parson, Consistent treatment of length variants in the human mtDNA control region: a reappraisal, Int. J. Legal Med. 122 (1) (2008) 11–21.
[7] A. Röck, J. Irwin, A. Dür, T. Parsons, W. Parson, SAM: string-based sequence search algorithm for mitochondrial DNA database queries, Forensic Sci. Int. Genet. 5 (2) (2011) 126–132.
[8] 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 (1) (2001) 83–91.
[9] A. Salas, H.-J. Bandelt, V. Macaulay, M.B. Richards, Phylogeographic investigations: the role of trees in forensic genetics, Forensic Sci. Int. 168 (1) (2007) 1–13.
[10] H.-J. Bandelt, A. Salas, C.M. Bravi, Problems in FBI mtDNA database, Science 305 (5689) (2004) 1402–1404.
[11] H.-J. Bandelt, A. Salas, S. Lutz-Bonengel, Artificial recombination in forensic mtDNA population databases, Int. J. Legal Med. 118 (5) (2004) 267–273.
[12] A. Salas, Á. Carracedo, V. Macaulay, M. Richards, H.-J. Bandelt, A practical guide to mitochondrial DNA error prevention in clinical, forensic, and population genetics, Biochem. Biophys. Res. Commun. 335 (3) (2005) 891–899.
[13] Y.-G. Yao, A. Salas, I. Logan, H.-J. Bandelt, mtDNA data mining in GenBank needs surveying, Am. J. Hum. Genet. 85 (6) (2009) 929–933 (author reply 933).
[14] Q.-P. Kong, H.-J. Bandelt, C. Sun, Y.-G. Yao, A. Salas, A. Achilli, C.-Y. Wang, L. Zhong, C.-L. Zhu, S.-F. Wu, A. Torroni, Y.-P. Zhang, Updating the East Asian mtDNA phylogeny: a prerequisite for the identification of pathogenic mutations, Hum. Mol. Genet. 15 (13) (2006) 2076–2086.
[15] Q.-P. Kong, A. Salas, C. Sun, N. Fuku, M. Tanaka, L. Zhong, C.-Y. Wang, Y.-G. Yao, H.-J. Bandelt, Distilling artificial recombinants from large sets of complete mtDNA genomes, PLoS One 3 (8) (2008) e3016.
[16] W. Parson, A. Dür, EMPOP—a forensic mtDNA database, Forensic Sci. Int. Genet. 1 (2) (2007) 88–92.
[17] 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 (2) (2000) 79–85.
[18] W. Parson, L. Roewer, Publication of population data of linearly inherited DNA markers in the International Journal of Legal Medicine, Int. J. Legal Med. 124 (5) (2010) 505–509.
[19] B. Zimmermann, A. Rock, G. Huber, T. Kramer, P.M. Schneider, W. Parson, Application of a west Eurasian-specific filter for quasi-median network analysis: sharpening the blade for mtDNA error detection, Forensic Sci. Int. Genet. 5 (2) (2011) 133–137.
[20] 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 (1) (2004) 1–11.
[21] A. Alonso, A. Salas, C. Albarrán, E. Arroyo, A. Castro, M. Crespillo, A.M. di Lonardo, M.V. Lareu, C.L. Cubria, M.L. Soto, J.A. Lorente, M.M. Semper, A. Palacio, M. Paredes, L. Pereira, A.P. Lezaun, J.P. Brito, A. Sala, M.C. Vide, M. Whittle, J.J. Yunis, J. Gómez, Results of the 1999–2000 collaborative exercise and proficiency testing program on mitochondrial DNA of the GEP-ISFG: an inter-laboratory study of the observed variability in the heteroplasmy level of hair from the same donor, Forensic Sci. Int. 125 (1) (2002) 1–7.
[22] M. Crespillo, M.R. Paredes, L. Prieto, M. Montesino, A. Salas, C. Albarrán, V. Álvarez-Iglesias, A. Amorin, G. Berniell-Lee, A. Brehm, J.C. Carril, D. Corach, N. Cuevas, A.M. Di Lonardo, C. Doutremepuich, R.M. Espinheira, M. Espinoza, F. Gómez, A. González, A. Hernández, M. Hidalgo, M. Jimenez, F.P. Leite, A.M. López, M. López-Soto, J.A. Lorente, S. Pagano, A.M. Palacio, J.J. Pestano, M.F. Pinheiro, E. Raimondi, M.M. Ramon, F. Tovar, L. Vidal-Rioja, M.C. Vide, M.R. Whittle, J.J. Yunis, J. Garcia-Hirschfel, Results of the 2003–2004 GEP-ISFG collaborative study on mitochondrial DNA: focus on the mtDNA profile of a mixed semen-saliva stain, Forensic Sci. Int. 160 (2–3) (2006) 157–167.
[23] M. Montesino, A. Salas, M. Crespillo, C. Albarrán, A. Alonso, V. Álvarez-Iglesias, J.A. Cano, M. Carvalho, D. Corach, C. Cruz, A. Di Lonardo, R. Espinheira, M.J. Farfán, S. Filippini, J. García-Hirschfeld, A. Hernández, G. Lima, C.M. López-Cubría, M. López-Soto, S. Pagano, M. Paredes, M.F. Pinheiro, A.M. Rodríguez-Monge, A. Sala, S.

Sóñora, D.R. Sumita, M.C. Vide, M.R. Whittle, A. Zurita, L. Prieto, Analysis of body fluid mixtures by mtDNA sequencing: an inter-laboratory study of the GEP-ISFG working group, Forensic Sci. Int. 168 (1) (2007) 42–56.
[24] L. Prieto, A. Alonso, C. Alves, M. Crespillo, M. Montesino, A. Picornell, A. Brehm, J.L. Ramírez, M.R. Whittle, M.J. Anjos, I. Boschi, J. Buj, M. Cerezo, S. Cardoso, R. Cicarelli, D. Comas, D. Corach, C. Doutremepuich, R.M. Espinheira, I. Fernández-Fernández, S. Filippini, J. Garcia-Hirschfeld, A. González, B. Heinrichs, A. Hernández, F.P.N. Leite, R.P. Lizarazo, A.M. López-Parra, M. López-Soto, J.A. Lorente, B. Mechoso, I. Navarro, S. Pagano, J.J. Pestano, J. Puente, E. Raimondi, A. Rodríguez-Quesada, M.F. Terra-Pinheiro, L. Vidal-Rioja, C. Vullo, A. Salas, 2006 GEP-ISFG Q1 collaborative exercise on mtDNA. Reflections about interpretation, artefacts, and DNA mixtures, Forensic Sci. Int. Genet. 2 (2) (2008) 126–133.
[25] L. Prieto, M. Montesino, A. Salas, A. Alonso, C. Albarrán, S. Álvarez, M. Crespillo, A.M. Di Lonardo, C. Doutremepuich, I. Fernández-Fernández, A.G. de la Vega, L. Gusmão, C.M. López, M. López-Soto, J.A. Lorente, M. Malaghini, C.A. Martínez, N.M. Modesti, A.M. Palacio, M. Paredes, S.D. Pena, A. Pérez-Lezaun, J.J. Pestano, J. Puente, A. Sala, M. Vide, M.R. Whittle, J.J. Yunis, J. Gómez, The 2000–2001 GEP-ISFG Collaborative Exercise on mtDNA: assessing the cause of unsuccessful mtDNA PCR amplification of hair shaft samples, Forensic Sci. Int. 134 (1) (2003) 46–53.
[26] 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 (2–3) (2005) 191–198.
[27] H.-J. Bandelt, Y.-G. Yao, C.M. Bravi, A. Salas, T. Kivisild, Median network analysis of defectively sequenced entire mitochondrial genomes from early and contemporary disease studies, J. Hum. Genet. 54 (3) (2009) 174–181.
[28] C. Turchi, L. Buscemi, C. Previdere, P. Grignani, A. Brandstatter, A. Achilli, W. Parson, A. Tagliabracci, Italian mitochondrial DNA database: results of a collaborative exercise and proficiency testing, Int. J. Legal Med. 122 (3) (2008) 199–204.
[29] H.-J. Bandelt, A. Salas, Contamination and sample mix-up can best explain some patterns of mtDNA instabilities in buccal cells and oral squamous cell carcinoma, BMC Cancer 9 (1) (2009) 113.
[30] H.-J. Bandelt, Y.-G. Yao, A. Salas, T. Kivisild, C.M. Bravi, High penetrance of sequencing errors and interpretative shortcomings in mtDNA sequence analysis of LHON patients, Biochem. Biophys. Res. Commun. 352 (2) (2007) 283–291.
[31] A. Salas, Y.-G. Yao, V. Macaulay, A. Vega, A. Carracedo, H.-J. Bandelt, A critical reassessment of the role of mitochondria in tumorigenesis, PLoS Med. 2 (11) (2005) e296.
[32] Á. Carracedo, J.M. Butler, L. Gusmão, W. Parson, L. Roewer, P.M. Schneider, Publication of population data for forensic purposes, Forensic Sci. Int. Genet. 4 (3) (2010) 145–147.
[33] H.-J. Bandelt, A. Olivieri, C. Bravi, Y.-G. Yao, A. Torroni, A. Salas, 'Distorted' mitochondrial DNA sequences in schizophrenic patients, Eur. J. Hum. Genet. 15 (4) (2007) 400–404 (author reply 402–404).

Antonio Salas*
Unidade de Xenética, Departamento de Anatomía Patolóxica e Ciencias Forenses, and Instituto de Medicina Legal, Facultade de Medicina, Universidade de Santiago de Compostela, 15876 Galicia, Spain

Michael Coble
National Institute of Standards and Technology, Gaithersburg, MD 20850-8314, USA

Stijn Desmyter
National Institute of Criminalistics and Criminology, Vilvoordsesteenweg 100, 1120 Brussels, Belgium

Tomasz Grzybowski
Ludwik Rydygier Collegium Medicum, Department of Molecular and Forensic Genetics, The Nicolaus Copernicus University, M. Curie-Skłodowskiej Str. 9, 85-094 Bydgoszcz, Poland

Leonor Gusmão[a,b]
[a]*IPATIMUP, Institute of Molecular Pathology and Immunology of University of Porto, Portugal*
[b]*Laboratório de Genética Humana e Médica, Instituto de Ciências Biológicas, Universidade Federeal do Pará, Brazil*

Carsten Hohoff
Institute für Forensische Genetik GmbH, Im Derdel 8, 48161 Münster, Germany

Mitchell M. Holland
Forensic Science Program, The Pennsylvania State University, University Park, PA 16802, USA

Jodi A. Irwin
Armed Forces DNA Identification Laboratory, Armed Forces Medical Examiner System, Dover Air Force Base, DE, USA

Tomasz Kupiec
Section of Forensic Genetics, Institute of Forensic Research, Krakow, Poland

Hwan-Young Lee
Department of Forensic Medicine, Yonsei University College of Medicine, 250 Seongsanno, Seodaemun-Gu, Seoul, South Korea

Bertrand Ludes
Laboratory of Molecular Anthropology, Institute of Legal Medecine, Strasbourg University, 67000 Strasbourg, France

Sabine Lutz-Bonengel
Institute of Legal Medicine, Freiburg University Medical Center, Albertstrasse 9, D-79104 Freiburg, Germany

Terry Melton
Mitotyping Technologies, 2565 Park Center Boulevard, Suite 200, State College, PA 16801, USA

Thomas J. Parsons
International Commission on Missing Persons, Alipasina 45A, 71000 Sarajevo, Bosnia and Herzegovina

Heidi Pfeiffer
Institute of Legal Medicine, University of Münster, Münster, Germany

Lourdes Prieto
Comisaría General de Policía Científica, University Institute of Research in Forensic Sciences (IUICP), Madrid, Spain

Adriano Tagliabracci
Department of Biomedical Sciences and Public Health, Section of Legal Medicine, Università Politecnica delle Marche, 60020 Ancona, Italy

Walther Parson**
Institute of Legal Medicine, Innsbruck Medical University, Innsbruck, Austria

*Corresponding author
E-mail address: antonio.salas@usc.es (A. Salas)

**Corresponding author
E-mail address: walther.parson@i-med.ac.at (W. Parson)

18 May 2012