

Letter to the Editor

## High levels of mitochondrial DNA heteroplasmy in human hairs by Budowle et al.

Peter D'Eustachio\*

*Department of Biochemistry, NYU School of Medicine, 550 First Avenue, New York, NY 10016, USA*

Received 31 May 2002; accepted 26 July 2002

The small size of the mitochondrial (mt) genome and its presence in many copies per cell make it useful for human DNA typing when only small amounts of material are available or when the material is badly degraded [1]. Despite its small size, and the fact that it is normally transmitted only maternally and without recombination, substantial DNA sequence polymorphism has accumulated in the mt genome, especially in two segments of the control region: HVI and HVII. The mtDNA sequence recovered from an item of evidence can thus be used to exclude many individuals as potential sources of that evidence. Conversely, if an evidence sequence matches a known reference sequence, the frequency of unrelated individuals who match the evidence sequence can be estimated by reference to a human mtDNA population database.

The high copy number and high mutation rate that make mtDNA useful for forensic DNA typing also lead to mtDNA sequence heteroplasmy, the presence of more than one mtDNA sequence in an individual. Initially thought not to occur in the mitochondria of normal individuals [2], heteroplasmy is now known to be widespread. Major unresolved issues include the molecular mechanisms responsible for the occurrence of heteroplasmy to different extents in different tissues, and the possibility that heteroplasmy levels in an individual might vary with age [3–6]. Hairs are often strongly affected, so that different individual hairs from one person can differ in mtDNA sequence by one or more bases, or can show mixed sequences [7–10].

The article by Budowle et al. in the 28 March edition of *Forensic Science International* 126 (2002) 30–33 raises two issues in the course of evaluating the recent observation by Grzybowski [11,12] of very high levels of heteroplasmy in human hairs. The first issue is that detection of heteroplasmy is dependent on the exact laboratory procedures used to

analyze mtDNA. The second is that interpretation of observed instances of heteroplasmy is dependent on models of mtDNA mutation and of human population genetics.

Considering the relationship between assay conditions and amounts and kinds of heteroplasmy found, Budowle et al. note that both the amount of template DNA (20–80 ng) and the number of PCR cycles (30 + 32, in a nested PCR strategy) in Grzybowski's studies [11,12] differ from those used by the United States FBI Laboratory (0.1 ng target amount of DNA; 36 PCR cycles [13]). However, in validation studies carried out by Allen et al. [14], accurate mtDNA sequence information was reliably obtained with up to 33 ng of template DNA (the largest amount tested) in a 25 + 32-cycle nested PCR strategy, and in studies carried out to validate the British FSS protocol, accurate mtDNA sequence information was obtained with up to 200 ng of template in a nested PCR strategy with up to 32 cycles in the first step and up to 32 in the second [15,16]. Grzybowski's assay conditions thus fall within the range validated by several laboratories for forensic mtDNA typing.

At the same time, detection of heteroplasmy clearly is assay dependent. Such dependency was described by Sullivan et al. [7]. Calloway et al. [3] observed that their ability to detect heteroplasmy varied both with assay conditions and with the particular mtDNA sequence variants under study. Indeed, a recent inter-laboratory exercise provides a striking demonstration of the range of results possible when current forensic mtDNA typing protocols are applied to genuinely heteroplasmic hairs [10].

Continuing their discussion of experimental conditions that might confound mtDNA typing, Budowle et al. suggest that nuclear pseudogenes might be the source of some of the DNA sequences observed by Grzybowski [11]. Amplification of nuclear DNA corresponding to the HVI sequence has been reported, albeit only in samples depleted of mtDNA or containing mtDNA mutated so as to reduce its affinity for mt HVI PCR primers [17]. In contrast, amplification of nuclear DNA corresponding to mtDNA coding sequences such as

\* Tel.: +1-212-263-5779; fax: +1-212-263-8166.

E-mail address: deustp01@med.nyu.edu (P. D'Eustachio).

Table 1  
Human nuclear DNA sequences related to mt HVI, HVII, CO1, and CO2

Query sequence	Matching sequences					
	Source	Accession number	Chromosome	Expected	Length	Percent identity
HVI (16024–16324)	Contigs	No significant matches				
	HTGS	AC107937.2	11	1e-93	238	93
HVII (31–323)	Contigs	NT_024862.7	17	8e-59	207	89
	HTGS	AC091134.3	17	4e-56	208	88
	HTGS	AC068619.4	17	1e-53	208	88
	HTGS	AC107937.2	11	3e-07	31	31
CO1 (5904–7445)	Contigs	NT_007412.8	6	<e-145	1542	98
	Contigs	NT_023115.8	5	<e-145	1055	89
	Contigs	NT_011613.8	X	<e-145	750	93
	Contigs	NT_010707.8	17	<e-145	628	96
	Contigs	NT_024862.7	17	<e-145	1563	83
	Contigs	NT_010164.8	14	<e-145	705	92
	Contigs	NT_008218.8	14	<e-145	705	92
	Contigs	NT_005354.8	2	e-145	701	84
	Contigs	NT_015120.7	2	e-117	272	94
	Contigs	NT_023290.4	5	e-107	825	81
	Contigs	NT_023943.8	9	e-105	196	100
	Contigs	NT_008413.8	9	2e-97	984	80
	Contigs	NT_030106.3	11	3e-71	163	96
	Contigs	NT_011515.6	21	1e-49	180	89
	Contigs	NT_011878.7	Y	1e-27	180	89
	Contigs	NT_011520.8	22	3e-16	47	100
	Contigs	NT_024871.5	17	3e-10	41	97
	HTGS	AL359496.30	6	<e-145	1542	98
	HTGS	AC021965.3	?	<e-145	1055	89
	HTGS	AL359973.11	X	<e-145	750	93
	HTGS	AC091154.3	17	<e-145	628	96
	HTGS	AC073480.3	17	<e-145	628	96
	HTGS	AC091134.3	17	<e-145	1563	83
	HTGS	AC068619.4	17	<e-145	1563	83
	HTGS	AC090166.3	17	<e-145	628	95
	HTGS	AC053536.2	17	<e-145	628	95
	HTGS	AL132988.4	14	<e-145	705	92
	HTGS	AC067849.6	8	<e-145	362	97
	HTGS	AC009234.3	2	e-145	701	84
	HTGS	AC009975.9	2	e-117	272	94
	HTGS	AC019336.5	5	e-106	825	81
	HTGS	AL591968.4	9	e-105	196	100
	HTGS	AL161450.14	9	<e-145	1514	77
	HTGS	AP003461.2	11	2e-86	1035	79
	HTGS	AC058808.1	11	2e-86	1035	79
	HTGS	AP000763.4	11	7e-71	163	96
	HTGS	AC025238.6	11	7e-71	163	96
	HTGS	AC078949.1	15	7e-71	163	96
	HTGS	AC103748.1	15	2e-49	180	89
	HTGS	AJ239326.3	21	2e-49	180	89
HTGS	AC009952.4	Y	2e-27	66	100	
HTGS	AL079295.1	22	5e-16	47	100	
HTGS	AC087530.5	17	5e-10	41	97	
HTGS	AC023493.6	17	5e-10	41	97	

Table 1 (Continued)

Query sequence	Matching sequences					
	Source	Accession number	Chromosome	Expected	Length	Percent identity
CO2 (7586–8269)	Contigs	NT_007412.8	6	<e-145	684	98
	Contigs	NT_008218.8	8	<e-145	684	96
	Contigs	NT_023115.8	5	<e-145	685	90
	Contigs	NT_024862.7	17	e-126	671	83
	Contigs	NT_024989.4	18	3e-91	192	97
	HTGS	AL359496.30	6	<e-145	684	98
	HTGS	AC067849.6	8	<e-145	684	96
	HTGS	AC021965.3	?	<e-145	685	90
	HTGS	AC091134.3	17	e-131	670	84
	HTGS	AC068619.4	17	e-126	671	83
	HTGS	AC104765.5	8	6e-91	192	97
	HTGS	AP001386.3	18	6e-91	192	97

The NCBI Human Genome BLAST server [33] was used to search the human nuclear genomic DNA sequence, in two parts, the Homo sapiens genomic contigs sequences, posted 8 February 2002, and the Human HTGS database, posted 7 May 2002, for matches to mtDNA sequences HVI, HVII, CO1, and CO2, taken from the Cambridge Mitochondrial DNA Concordance [34] and the MITOMAP list of mtDNA function locations [35]. The inclusive base numbers of each mt sequence are shown in parentheses. Each nuclear DNA match is identified by its GenBank accession number, the identity of the chromosome from which the sequence is derived, the number of times that a match of the observed quality would be expected to occur by chance in the database, the length of DNA sequence involved in the match, and the percent identity of the mt and nuclear DNA sequences over this span.

cytochrome oxidase (CO) occurs under a variety of conditions [18,19]. Approximately 98% of the human nuclear DNA sequence is known [20], so the possible existence of nuclear DNA sequences related to any given mtDNA sequence can be tested directly by computer analysis. BLAST [21] searches of both the part of the human genome assembled into long-range “contigs” and the part that has not yet been assembled (“HTGS”) found one nuclear sequence similar to the mt HVI control region (Table 1). Four nuclear sequences matched HVII, 44 matched CO1, and 12 matched CO2. Nuclear matches to mt CO genes are not only more numerous, but of better quality. The single HVI match spans 238 bases, and the longest stretch of identity between the nuclear and mt sequences is 59 bases. The best CO1 match spans 1542 bases and includes several stretches of over 150 bases that match perfectly. These analyses do not support the notion that nuclear DNA is a good source of template molecules for HVI PCR reactions.

Another concern that Budowle et al. raise about Grzybowski’s assay procedure relates to contamination. The control assays that were reported [12] indeed do not exclude the possibility that “heteroplasmy” was due to exogenous tissue or DNA closely associated with the hairs that were typed. However, Wilson et al. [22] showed that the ultrasonic cleaning methods used both by the FBI [13,23] and by Grzybowski [12], in contrast to the mechanical cleaning method used by the FSS [15,16], consistently remove such contaminants, although FBI personnel have testified that the latter method can produce valid results in casework [24].

Turning to the second general area of concern, that of interpreting potential instances of heteroplasmy in terms of

what is known about human molecular biology and population genetics, Budowle et al. suggest that many of the sites associated with heteroplasmy in Grzybowski’s work [11] are implausible because these sites are not polymorphic in the SWGDAM mtDNA human population database [25]. This suggestion makes the assumption that heteroplasmy in somatic tissues should parallel population-level polymorphism. This assumption has recently been examined by Stoneking [26], who reviewed published data for de novo germline mutations in human pedigrees, somatic mutations in tumors, and site heteroplasmies, and compared the distributions of these variants to the distribution of variable sites in the mtDNA control region determined from human population mtDNA sequencing studies. Stoneking’s analysis suggested that germline, tumor, and heteroplasmic mtDNA mutations all occur preferentially at variable sites in the human mtDNA control region. Comparison of Grzybowski’s “hot spots”—HVI residues 16126, 16294, 16296, and 16311—with catalogs of variable sites based on human population data [27–31] shows that all four sites vary at rates above the average for HVI as a whole. Three of them—16126, 16294, and 16311—are consistently identified as hypervariable. It is noteworthy that the heteroplasmy of Tsar Nicholas II and his brother [32] affected a residue, 16169, that is about as variable in these population surveys as residue 16296. The relative lack of population variability at residue 16169 has not previously been cited as a reason to doubt the validity of the Russian mtDNA studies. While available samples of both mutational events and heteroplasmic mtDNA sequences are small, these comparisons suggest that it

may be the SWGDAM database that does not accurately reflect human mtDNA sequence variability.

Even if some of Grzybowski's specific results cannot be independently verified, several general conclusions about heteroplasmy seem warranted. First, heteroplasmy is common in hairs, can affect multiple bases within the HVI and HVII sequences of an individual, and can assort variably, leading to a mosaic presentation when multiple hairs of one individual are compared. More extensive surveys will be needed to determine how common heteroplasmy is, and the range of forms it can take. Second, heteroplasmy is readily detected by current forensic mtDNA typing procedures, but the various procedures differ significantly in their ability to reveal heteroplasmy in any given case. Further standardization of mtDNA typing procedures may be needed. Third, uncertainties about the relationship of the molecular events that give rise to heteroplasmy observed in somatic tissues and the events that give rise to the polymorphisms in HVI and HVII sequences catalogued in population databases remain an obstacle to the reliable, objective use of these databases to interpret mtDNA typing results.

## References

- [1] A. 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.
- [2] B. Budowle, D.E. Adams, C.C. Comey, C.R. Merrill, Mitochondrial DNA: a possible genetic material suitable for forensic analysis, in: H.C. Lee, R.E. Gaensslen (Eds.), *Advances in Forensic Sciences*, 1990, pp. 76–97.
- [3] C.D. Calloway, R.L. Reynolds, G.L. Herrin, W.W. Anderson, The frequency of heteroplasmy in the HVII region of mtDNA differs across tissue types and increases with age, *Am. J. Hum. Genet.* 66 (2000) 1384–1397.
- [4] Y. Michikawa, F. Mazzucchelli, N. Bresolin, G. Scarlato, G. Attardi, Aging-dependent large accumulation of point mutations in the human mtDNA control region for replication, *Science* 286 (1999) 774–779.
- [5] L.A. Tully, T.J. Parsons, R.J. Steighner, M.H. Holland, M.A. Marino, V.L. Prenger, A sensitive denaturing gradient-gel electrophoresis assay reveals a high frequency of heteroplasmy in hypervariable region 1 of the human mtDNA control region, *Am. J. Hum. Genet.* 67 (2000) 432–443.
- [6] M. Lagerström-Fermér, C. Olsson, L. Forsgren, A.-C. Syvänen, Heteroplasmy of the human mtDNA control region remains constant during life, *Am. J. Hum. Genet.* 68 (2001) 1299–1301.
- [7] K.M. Sullivan, R. Allison-Greiner, F.I.A. Archampong, R. Piercy, G. Tully, P. Gill, C. Lloyd-Davies, A single difference in mtDNA control region sequence observed between hair shaft and reference samples from a single donor, in: *Proceedings of the Seventh International Symposium on Human Identification*, 1996, <http://www.promega.com/geneticidproc/>.
- [8] M.R. Wilson, D. Polansky, J. Replogle, J.A. DiZinno, B. Budowle, A family exhibiting heteroplasmy in the human mitochondrial DNA control region reveals both somatic mosaicism and pronounced segregation of mitotypes, *Hum. Genet.* 100 (1997) 167–171.
- [9] A. Salas, M.V. Lareu, A. Carracedo, Heteroplasmy in mtDNA and the weight of evidence in forensic mtDNA analysis: a case report, *Int. J. Leg. Med.* 114 (2001) 186–190.
- [10] A. Alonso, A. Salas, C. Albarrán, E. Arroyo, A. Castro, M. Crespillo, A.M. di Lonardo, M.V. Lareu, C. López Cubría, M. López Soto, J.A. Lorente, M. Montesino Semper, A. Palacio, M. Paredes, L. Pereira, A. Pérez Lezaun, J. Pestano Brito, A. Sala, M. Conceição Vide, M. Whittler, 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 (2002) 1–7.
- [11] T. Grzybowski, Extremely high levels of human mitochondrial DNA heteroplasmy in single hair roots, *Electrophoresis* 21 (2000) 548–553 (erratum after p. 1034).
- [12] T. Grzybowski, Reply, *Electrophoresis* 22 (2001) 181–182.
- [13] FBI Laboratory, DNA Unit II, *Mitochondrial DNA Sequencing Protocol*, 1998.
- [14] M. Allen, A.-S. Engström, S. Meyers, O. Handt, T. Saldeen, A. von Haeseler, S. Pääbo, Mitochondrial DNA sequencing of shed hairs and saliva on robbery caps: sensitivity and matching probabilities, *J. Forensic Sci.* 43 (1998) 453–464.
- [15] K.M. Sullivan, R. Hopgood, P. Gill, Identification of human remains by amplification and automated sequencing of mitochondrial DNA, *Int. J. Leg. Med.* 105 (1992) 83–86.
- [16] R. Hopgood, K.M. Sullivan, P. Gill, Strategies for automated sequencing of human mitochondrial DNA directly from PCR products, *BioTechniques* 13 (1992) 82–92.
- [17] M.A. Morgan, T.J. Parsons, M.H. Holland, Amplification of human nuclear pseudogenes derived from mitochondrial DNA: a problem for mitochondrial DNA identity testing? in: *Proceedings of the Eighth International Symposium on Human Identification*, 1997, <http://www.promega.com/geneticidproc/>.
- [18] M. Hirano, A. Shtilbans, R. Mayeux, M.N. Davidson, S. DiMauro, J.A. Knowles, E.A. Schon, Apparent mtDNA heteroplasmy in Alzheimer's disease patients and in normals due to PCR amplification of nucleus-embedded mtDNA pseudogenes, *Proc. Natl. Acad. Sci. U.S.A.* 94 (1997) 14894–14899.
- [19] D.C. Wallace, C. Stuard, D. Murdock, T. Schurr, M.D. Brown, Ancient mtDNA sequences in the human nuclear genome: a potential source of errors in identifying pathogenic mutations, *Proc. Natl. Acad. Sci. U.S.A.* 94 (1997) 14900–14905.
- [20] NCBI Human Genome Sequencing Progress, 2002, <http://www.ncbi.nlm.nih.gov/genome/seq/page.cgi?F=HsProgress.shtml&ORG=Hs>.
- [21] S.F. Altschul, W. Gish, W. Miller, E.W. Myers, D.J. Lipman, Basic local alignment search tool, *J. Mol. Biol.* 215 (1990) 403–410.
- [22] M.R. Wilson, J.A. DiZinno, D. Polansky, J. Replogle, B. Budowle, Validation of mitochondrial DNA sequencing for forensic casework analysis, *Int. J. Leg. Med.* 108 (1995) 68–74.

- [23] A.R. Isenberg, J.M. Moore, Mitochondrial DNA analysis at the FBI Laboratory, *Forensic Sci. Commun.* 1 (1999), <http://www.fbi.gov/hq/lab/fsc/backissu/july1999/dnalist.htm>.
- [24] R.V. Murrin, 181 D.L.R., (4th), 320, 1999, <http://www.denverda.org/legalResource/murrin.htm>.
- [25] K.L. Monson, K.W.P. Miller, M.R. Wilson, J.A. DiZinno, B. Budowle, The mtDNA population database: an integrated software and database resource for forensic comparison, *Forensic Sci. Commun.* 4 (2002), <http://www.fbi.gov/hq/lab/fsc/current/backissu.htm>.
- [26] M. Stoneking, Hypervariable sites in the mtDNA control region are mutational hotspots, *Am. J. Hum. Genet.* 67 (2000) 1029–1032.
- [27] L. Excoffier, Z. Yang, Substitution rate variation among sites in mitochondrial hypervariable region I of humans and chimpanzees, *Mol. Biol. Evol.* 16 (1999) 1357–1368.
- [28] M. Hasegawa, A. Di Rienzo, T.D. Kocher, A.C. Wilson, Toward a more accurate time scale for the human mitochondrial DNA tree, *J. Mol. Evol.* 37 (1993) 347–354.
- [29] S. Meyer, G. Weiss, A. von Haeseler, Pattern of nucleotide substitution and rate heterogeneity in the hypervariable regions I and II of human mtDNA, *Genetics* 152 (1999) 1103–1110.
- [30] G. Pesole, C. Saccone, A novel method for estimating substitution rate variation among sites in a large dataset of homologous DNA sequences, *Genetics* 157 (2001) 859–865.
- [31] J. Wakeley, Substitution rate variation among sites in hypervariable region I of human mitochondrial DNA, *J. Mol. Evol.* 37 (1993) 613–623.
- [32] P.L. Ivanov, M.J. Wadhams, R.K. Roby, M.M. Holland, V.W. Weedn, T.J. Parsons, Mitochondrial DNA sequence heteroplasmy in the Grand Duke of Russia Georgij Romanov establishes the authenticity of the remains of Tsar Nicholas II, *Nat. Genet.* 12 (1996) 417–420.
- [33] NCBI Human Genome BLAST page, 2002, <http://www.ncbi.nlm.nih.gov/genome/seq/page.cgi?F=HsBlast.html&&ORG=Hs>.
- [34] Cambridge Mitochondrial DNA Concordance, 2002, <http://shelob.bioanth.cam.ac.uk/mtDNA/anderson.html>.
- [35] MITOMAP, Mitochondrial DNA Function Locations, 2002, <http://www.gen.emory.edu/cgi-bin/MITOMAP/bin/tbl1gen.pl>.