

Tomasz Grzybowski

The Ludwik Rydygier
University School of Medical
Sciences, Forensic Medicine
Institute, Bydgoszcz, Poland

Extremely high levels of human mitochondrial DNA heteroplasmy in single hair roots

For many years it has been assumed that the vast majority of mitochondrial genomes of a single individual are identical, both in the same tissue and within different tissues. Incidences of heteroplasmy (*i.e.*, the occurrence of two or more codominating types of molecules within the mitochondrial DNA population of the same individual) were thought to be extremely rare. This study strongly supports the thesis that heteroplasmy is a principle, rather than an exception, in mitochondrial DNA genetics. During direct sequencing of the first hypervariable segment of the human mitochondrial control region (HV1) in 100 single hair roots obtained from 35 individuals, 24 different heteroplasmic positions were identified. Unusually high levels of heteroplasmy (up to six positions in the HV1 region) were encountered in two individuals. Two individuals related in maternal lineage shared the same heteroplasmic positions. Moreover, highly variable levels of heteroplasmy were observed even among roots from the same individual. The most probable mechanisms involved in generating so many mismatches are mutations occurring presumably in the female germline, followed by differential segregation of mitotypes during the development of individual hairs. Generally, heteroplasmy complicates sequence comparisons in mitochondrial DNA testing performed for forensic purposes, but in some cases it can substantially increase the discriminating power of the analysis.*

Keywords: Mitochondrial DNA / Heteroplasmy / Human identification

EL 3771

The human mitochondrial DNA (mtDNA) is a 16 569 bp closed, circular molecule that has been sequenced completely, with all genes mapped [1]. Variation in mtDNA has been used extensively, not only to draw inferences in phylogenetics, population biology, and molecular archaeology [2–5], but also to individualize evidentiary material for forensic purposes [6–8]. The properties of mtDNA that make it valuable for both evolutionary and human identification studies include the high copy number (more than 5000 copies per cell), maternal mode of inheritance, and rapid rate of evolution. For many years it has been generally assumed that the vast majority of the mitochondrial genomes of a single individual are identical [9], both in the same tissue and within different tissues [10]. Events of heteroplasmy (*i.e.*, the occurrence of two or more types of molecules within the mtDNA population of the same individual) were thought to be associated with extremely rare diseases only [11–13]. However, recent reports of intra-individual sequence variability of the human mtDNA control region have called this belief into question and shed light on the fundamental principles of

mtDNA genetics [14–16]. Irrespective of its biological mechanisms, heteroplasmy, with its practical consequences, has been one of the most frequently discussed topics in the forensic community in the last few years. Instances where heteroplasmy is encountered in various types of tissues have been recorded extensively [17, 18]. Since one of the most frequent applications of mtDNA sequencing in forensics is the analysis of hair [19–23], this study aims to estimate the level of mtDNA heteroplasmy within single hair roots.

The first hypervariable segment (HV1) of the human mitochondrial control region was sequenced in 100 single hair roots obtained from 33 unrelated individuals and 2 individuals related in maternal lineage (brother and sister). Between two and six roots were analyzed from each individual (each root in a separate sequencing reaction). The study also involved an actual forensic case coming from the routine practice of the Forensic Medicine Institute in Bydgoszcz. Both HV1 and the second hypervariable segment (HV2) were sequenced in the casework hair sample and respective reference samples (blood and four hair roots obtained from the suspect).

DNA was isolated by digestion with proteinase K at 56°C overnight (extraction buffer contained 10 mM Tris, 100 mM NaCl, 39 mM DTT, 10 mM EDTA, and 2% SDS) followed

Correspondence: Dr. Tomasz Grzybowski, Forensic Medicine Institute, The Ludwik Rydygier University School of Medical Sciences, ul. M. Curie-Skłodowskiej 9, 85-094 Bydgoszcz, Poland
E-mail: tgrzyb@box43.gnet.pl
Fax: +48-52-341-12-49

Abbreviations: **HV1**, the first hypervariable segment; **HV2**, the second hypervariable segment

* Presented at the 5th International DNA Fingerprinting Conference, Grahamstown, South Africa, January 17–23, 1999

by a three-step organic extraction procedure involving phenol: chloroform: isoamyl alcohol and chloroform, with an additional chloroform extraction. DNA samples were then subjected to ultrafiltration with Microcon-100 micro-concentrators (Amicon, Beverly, MA, USA).

PCR amplification of the entire noncoding region (30 cycles) was performed using 20–80 ng of total genomic DNA in a 25 μ L reaction volume comprising 1 \times Promega buffer, 1.5 mM MgCl₂, 200 μ M of each of the dNTPs (Promega, Madison, WI, USA), 1.5 U of *Taq* polymerase (Promega), and 1 μ M of primers L15926 and H00580. Each cycle comprised 20 s at 94°C, 30 s at 50°C, and 2.5 min at 72°C (Thermal Cycler 9600; Perkin Elmer, Norwalk, CT, USA). The resultant D-loop amplification product was diluted 1000-fold and 4 μ L aliquots were added to an array of second-round, nested PCR reactions (32 cycles) to generate sufficient DNA templates for sequencing. This was carried out under the conditions given above, except when the reaction volume was 50 μ L, in which case extension time was 1.5 min, enzyme concentration was 1 U, and the primer concentrations were 0.1 μ M. To generate templates for sequencing both strands of the HV1 segment, the primer sets L15997/M13(-21)H16401 and M13(-21)L15997/H16401 were used in separate amplification reactions. Similarly, the primer sets L00029/M13(-21)H00408 and H00408/M13(-21)L00029 were used to generate templates for sequencing both strands of the HV2 segment. Both primer sequences and nomenclature were used according to Sullivan *et al.* [6]. Negative controls were prepared for both the DNA extraction and the amplification process. PCR products were purified by ultrafiltration (Microcon 100; Amicon) and sequenced directly from both strands with (-21)M13 primer using Dye Primer Cycle Sequencing Kit (Perkin Elmer) according to the manufacturer's protocol. Sequencing reactions were performed twice for the samples where heteroplasmy was encountered. Sequencing products were separated in a 4% PAG gel on the ABI Prism™ 377 DNA Sequencer. Data were analyzed automatically using DNA Sequencing Analysis and Sequence Navigator programs (Perkin Elmer).

During direct sequencing of the HV1 in 35 individuals, 13 persons with heteroplasmic point mutations were encountered. Twenty-four different heteroplasmic positions were identified. These were predominantly T/C transitions (19 positions), although A/G variations were also found (five positions). No transversions were observed. Out of 13 individuals in whom heteroplasmic mutations were identified, as many as seven displayed multiple heteroplasmy, with the number of heteroplasmic positions ranging from 2 to 6 in a single hair root (Fig. 1). Moreover, highly variable levels of heteroplasmy were observed, even among

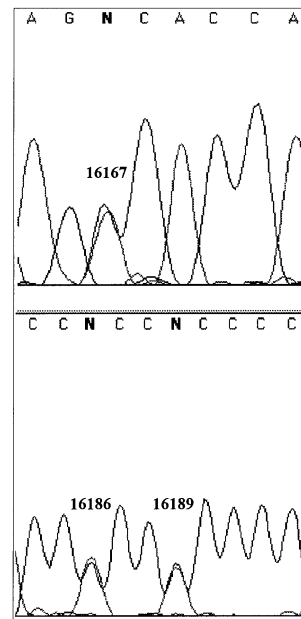


Figure 1. Electropherogram illustrating examples of mtDNA heteroplasmy in single hair roots. Two panels display the light strand sequence from two separate hair samples obtained from unrelated individuals. The upper panel shows T/C heteroplasmy at position 16167 in the HV1 region, whereas at the lower multiple T/C heteroplasmy occurs at positions 16186 and 16189.

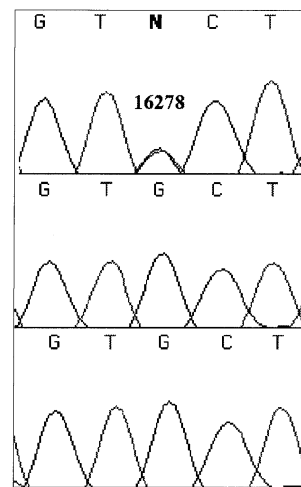


Figure 2. The sequencing results from three separate hair samples from the same individual. Position 16278 shows A/G heteroplasmy in one hair (the top panel), whereas the two others are homoplasmic (the bottom panels). The heavy strand sequence is shown.

roots from the same individual. For instance, sequence data obtained from three separate hair roots from the same individual showed heteroplasmy at position 16278 in only one of them, whereas two others were homoplasmic (Fig. 2). Interestingly, some positions seemed to be subject to a higher rate of heteroplasmy than other sites; positions 16126, 16294, 16296, and 16311 were identified as “hot spots” where heteroplasmy had been encountered more than once in unrelated individuals. Two individuals related in maternal lineage shared the same heteroplasmic positions. These results are summarized in Table 1. An example of heteroplasmy from actual casework (Table 2) shows an exact match of heteroplasmic position between evidentiary hair and one of the reference hair samples.

Table 3. Heteroplasmy levels observed in this study relative to the results of other authors for various types of tissue

Reference	Sample size	Type of tissue	Level of heteroplasmy	Technique of analysis
Ivanov <i>et al.</i> [17]	Tsar Nicolas II identification case – four members of Hessian family lineage	Bone Teeth Hair Soft tissues Blood	Detected in two individuals	Direct sequencing of PCR-amplified DNA
Bendall <i>et al.</i> [15]	180 twin pairs	Blood	Detected in 4 pairs	Sequencing of cloned DNA
Parsons <i>et al.</i> [33]	134 independent mtDNA lineages (357 individuals)	Blood T and B cell lines	Detected in 5 individuals from 3 lineages	Direct sequencing of PCR-amplified DNA
Jazin <i>et al.</i> [16]	3 individuals	Brain Blood	Detected in brain samples from 3 individuals Not detected in blood	DGGE Sequencing of cloned DNA
Hühne <i>et al.</i> [21]	154 individuals (77 mother-child pairs)	Blood	Detected in 7 individuals	Direct sequencing of PCR-amplified DNA
Wilson <i>et al.</i> [18]	3 individuals from a maternal lineage	Blood Buccal cells Hair roots Hair shafts	Detected in all types of tissue analyzed Degree of heteroplasmy differed significantly from hair to hair	Direct sequencing of PCR-amplified DNA
Bendall <i>et al.</i> [24]	9 individuals from a maternal lineage	Blood Buccal cells Several hair roots from one individual	Detected in 3 individuals at similar levels in both blood and buccal cells Highly variable levels detected in hair roots	Sequencing of cloned DNA
Comas <i>et al.</i> [14]	Population survey, sample size unknown	Hair roots	Detected in one individual at two different positions	Direct sequencing of PCR-amplified DNA
Hühne <i>et al.</i> [22]	10 individuals	Blood Saliva 150 hair shafts	Not detected	Direct sequencing of PCR-amplified DNA
Sullivan <i>et al.</i> [23]	One individual	Saliva 12 hair shafts	Detected in three hair shafts	Direct sequencing of PCR-amplified DNA
This study	33 unrelated individuals 2 individuals related in maternal lineage	100 hair roots	Detected in 13 individuals Multiple heteroplasmy (from 2 to 6 positions) observed in 6 individuals Degree of heteroplasmy differed significantly between hairs from the same individual	Direct sequencing of PCR-amplified DNA

MtDNA heteroplasmy appears to be incomparably more frequent in hair than in other tissues. In the course of this study, heteroplasmic mixtures were found which involved more than 1 or 2 base positions already reported by other authors.

Transient mtDNA heteroplasmy has several theoretical explanations, some more probable than others. Since mtDNA sequences are frequently inserted into the nuclear genome [27], where they evolve as nuclear pseudogenes, some authors speculate that heteroplasmy could be the result of coamplification of actual mtDNA and nucleus-embedded mtDNA sequences [28]. To date, a single nuclear pseudogene corresponding to a human control region has been mapped to chromosome 11 [29]. However, because mtDNA sequences greatly predominate in copy number over their nuclear counterparts, the

coamplification scenario is unlikely in the majority of cases. In practice, amplification of the nuclear pseudogene together with the mtDNA control region was observed only in two instances, where specific conditions existed that caused the preferential amplification of the pseudogene [30]. The second mechanism that could be considered as responsible for complex mutational events, as observed in this study, would be incomplete maternal transmission. Theoretically, biparental inheritance of mtDNA may provide a good explanation for the existence of heteroplasmic mixtures at multiple positions, as there

are on average eight differences in control region sequences between two randomly selected Caucasian individuals [31]. However, while low levels of paternally inherited mtDNA molecules were indeed detected in interspecific mouse hybrids by Gyllensten *et al.* [32], to date there is no evidence for biparental mode of inheritance of mtDNA in humans [33]. Moreover, paternal contributions, even if they occurred with minimal frequency, would be undetectable by direct sequencing of PCR-amplified DNA. One may also consider somatic mutations to be involved in generating so many mismatches in a particular type of tissue. These would be in some ways correlated with the stage of development of that tissue (*i.e.*, hair) as it was in the case of high levels of heteroplasmy observed by Jazin *et al.* [16] in human brain. Thus, heteroplasmy would be a natural state for particular tissue types.

Although somatic mutations occurring at particular positions identified as heteroplasmic “hot spots” can not be absolutely excluded as mechanisms for generating heteroplasmy in hair, the fact that the same heteroplasmic positions were encountered in two individuals related in maternal lineage strongly suggests that (at least) these individuals inherited heteroplasmy from their mother. Thus, mutation(s) leading to heteroplasmy would presumably occur in the female germline. Although there is much evidence that heteroplasmy does occur in the germline [15, 17, 34], the molecular basis for its maintenance remains elusive. In mammals, the so-called bottleneck mechanism is widely believed to be responsible for maintaining homoplasmy [15, 35, 36]. The bottleneck is described as an event occurring during oogenesis whereby a small number or even a single mtDNA molecule is selectively sampled from a larger population for transmission and amplification, ultimately populating the organism. According to the bottleneck theory, heteroplasmy is an intermediate state in which new mutations are in the process of rapid shift toward homoplasmy within very few generations. Although the bottleneck is generally proposed as the mechanism involved in segregation of mtDNA types during oogenesis, one cannot exclude the possibility that a similar bottleneck might occur during postzygotic development, leading to differential segregation of mtDNA variants in various tissues. Perhaps the size of such a developmental bottleneck, varying stochastically at subsequent mitotic divisions, is also responsible for variable levels of heteroplasmy segregating differentially among individual hairs.

Irrespective of its biological meaning, extremely high levels of heteroplasmy in single hairs have important implications for the forensic interpretation of mtDNA sequence data. Since there is no sequence homogeneity between hair and other tissues, reference hair will always be

required to identify evidentiary hair samples. Moreover, due to the differential level of heteroplasmy observed from hair to hair, more reference samples should be analyzed before a final conclusion can be reached. There is no doubt that the occurrence of heteroplasmy makes sequence comparisons more complicated. On the other hand, however, there are some cases where heteroplasmy can increase the discriminating power of the analysis. One such instance, when there was an exact match of heteroplasmic position between one of the reference hairs obtained from the suspect and the questioned hair found at the crime scene, has been shown in this study. All unambiguous bases were the same for both evidentiary and reference material. A search for the reference hair profile was conducted in the mtDNA sequence database of 200 Polish Caucasians and no match was found. This suggested that the profile was relatively rare. However, if there was not an exact match of heteroplasmic position between evidentiary hair and one of the reference hairs, one could only state that the suspect could not be excluded as a potential source of the evidentiary sample. Since one of the reference hairs of the suspect and the questioned sample did match at all unambiguous positions and shared heteroplasmy at the same position, this represented simultaneous occurrence of additional unlikely events and, as such, strengthened an association between the crime scene sample and the suspect. In fact, in this case the presence of heteroplasmy strengthened the mtDNA evidence.

In conclusion, extremely high levels of mtDNA heteroplasmy observed in the course of this study strongly support the thesis that heteroplasmy is a principle rather than an exception in mtDNA genetics. Even though relatively little information exists on its biological meaning, it is unquestionably involved in mechanisms of segregation and fixation of mtDNA mutations and can influence the procedure of human identification on the mtDNA level. Family studies are currently underway to determine whether and to what extent heteroplasmic mutational changes are stable within a particular maternal lineage. Perhaps such studies will throw more light on the question of mechanisms of mtDNA segregation in humans.

Received September 29, 1999

References

- [1] Anderson, S., Bankier, A. T., Borel, B. G., de Bruijn, M. H. L., Coulson, A. R., Drouin, J., Eperon, I. C., Nierlich, D. P., Roe, B. A., Sanger, F., Schreier, P. H., Smith, A. J. H., Staden, R., Young, I. G., *Nature* 1981, 290, 457–465.
- [2] Leblanc, C., Richard, O., Kloareg, B., Viehmann, S., Zetsche, K., Boyen, C., *Curr. Genet.* 1997, 31, 193–207.
- [3] Stoneking, M., *J. Bioenerg. Biomembr.* 1994, 26, 251–259.

- [4] van Holst Pellekaan, S. M., Frommer, M., Sved, J. A., Boettcher, B., *Am. J. Hum. Genet.* 1998, **62**, 435–449.
- [5] Hagelberg, E., in: Hermann, B., Hummel, S., (Eds.), *Ancient DNA*, Springer-Verlag, New York 1994, pp. 195–204.
- [6] Sullivan, K. M., Hopgood, R., Gill, P., *Int. J. Legal Med.* 1992, **105**, 83–86.
- [7] Wilson, M. r., DiZinno, J. A., Polanskey, D., Replogle, J., Budowle, B., *Int. J. Legal Med.* 1995, **108**, 68–74.
- [8] Lutz, S., Weisser, H. J., Heizmann, J., Pollak, S., *Int. J. Legal Med.* 1996, **109**, 205–209.
- [9] Monnat, R. J., Loeb, L. A., *Proc. Natl. Acad. Sci. USA* 1985, **82**, 2895–2899.
- [10] Monnat, R. J., Reay, D. T., *Gene* 1986, **43**, 205–211.
- [11] Bidooki, S. K., Johnson, M. A., Chrzanowska-Lightowlers, Z., Bindoff, L. A., Lightowlers, R. N., *Am. J. Hum. Genet.* 1997, **60**, 1430–1438.
- [12] Holme, E., Tulinius, M. H., Larsson, N. g., Oldfors, A., *Biochim. Biophys. Acta* 1995, **1271**, 249–252.
- [13] Lightowlers, R. N., Chinnery, P. F., Turnbull, D. M., Howell, N., *Trends Genet.* 1997, **13**, 450–455.
- [14] Comas, D., Pääbo, S., Bertranpetit, J., *Genome Res.* 1995, **5**, 89–90.
- [15] Bendall, K. E., Macaulay, V. A., Baker, J. R., Sykes, B. C., *Am. J. Hum. Genet.* 1996, **59**, 1276–1287.
- [16] Jazin, E. E., Cavelier, L., Eriksson, I., Orelund, L., Gyllenstein, U., *Proc. Natl. Acad. Sci. USA* 1996, **93**, 12382–12387.
- [17] Ivanov, P. L., Wadhams, M. J., Roby, R. K., Holland, M. M., Weedn, V. W., Parsons, T. J., *Nature Genet.* 1996, **12**, 417–420.
- [18] Wilson, M. r., Polanskey, D., Replogle, J., DiZinno, J. A., Budowle, B., *Hum. Genet.* 1997, **100**, 167–171.
- [19] Wilson, M. R., Polanskey, D., Butler, J., DiZinno, J. A., Replogle, J., Budowle, B., *BioTechniques* 1995, **18**, 662–669.
- [20] Miścicka-Sliwka, D., Grzybowski, T., Woźniak, M., Czarny, J., *Proceedings from the Eighth International Symposium on Human Identification 1997*, Promega Corporation, Madison, WI, USA 1997, pp. 132–142.
- [21] Hühne, J., Pfeiffer, H., Brinkmann, B., *Int. J. Legal Med.* 1998, **112**, 27–30.
- [22] Hühne, J., Pfeiffer, H., Waterkamp, K., Brinkmann, B., *Int. J. Legal Med.* 1999, **112**, 172–175.
- [23] Sullivan, K. M., Alliston-Greiner, F. I. A., Archampong, R., Piercy, G., Tully, P., Gill, P., Lloyd-Davies, C., *Proceedings from the Seventh International Symposium on Human Identification 1996*, Promega Corporation, Madison, WI, USA 1996, pp. 126–130.
- [24] Bendall, K. E., Macaulay, V. A., Sykes, B. C., *Am. J. Hum. Genet.* 1997, **61**, 1303–1308.
- [25] Lutz, S., Weisser, H. J., Heizmann, J., Pollak, S., *Int. J. Legal Med.* 1998, **111**, 67–77.
- [26] Jazin, E., Soodyall, H., Jalonen, P., Lindholm, E., Stoneking, M., Gyllenstein, U., Parsons, T. J., Holland, M. M., *Nature Genet.* 1998, **18**, 109–111.
- [27] Sorenson, M. D., Fleischer, R. C., *Proc. Natl. Acad. Sci. USA* 1996, **93**, 15239–15243.
- [28] Parfait, B., Rustin, P., Munnich, A., Rotig, A., *Biochem. Biophys. Res. Commun.* 1998, **247**, 57–59.
- [29] Zischler, H., Geisert, H., von Haeseler, A., Pääbo, S., *Nature* 1995, **378**, 489–492.
- [30] Morgan, M. A., Parsons, T. J., Holland, M. M., *Proceedings from the Eight International Symposium on Human Identification 1997*, Promega Corporation, Madison, WI, USA 1997, p. 128.
- [31] Holland, M. M., Parsons, T. J., *Forens. Sci. Rev.* 1999, **11**, 21–50.
- [32] Gyllenstein, U., Wharton, D., Josefsson, A., Wilson, A. C., *Nature* 1991, **352**, 255–257.
- [33] Parsons, T. J., Muniec, D. S., Sullivan, K., Woodyatt, N., Alliston-Greiner, R., Wilson, M. R., Berry, D. L., Holland, K. A., Weedn, V. W., Gill, P., Holland, M. M., *Nature Genet.* 1997, **15**, 363–367.
- [34] Howell, N., Kubacka, I., Mackey, D. A., *Am. J. Hum. Genet.* 1996, **59**, 501–509.
- [35] Marchington, D. R., Hartshorne, G. M., Barlow, D., Poulton, J., *Am. J. Hum. Genet.* 1997, **60**, 408–416.
- [36] Marchington, D. R., Macaulay, V., Hartshorne, G. M., Barlow, D., Poulton, J., *Am. J. Hum. Genet.* 1998, **63**, 769–775.