Identification of Human Remains Using Mitochondrial DNA Sequencing: Potential Mother-Child Mutational Events

*M.M. Holland, PhD, R.K. Roby, MPH, D.L. Fisher, MPH, J. Ruderman, BS, D.A. Lee, MSFS, C.K. Bryson, MFS, T. Kupferschmid, MFS, R.S. Lofts, Maj, USAF, PhD, and A.J. Eisenberg, PhD**

The Armed Forces DNA Identification Laboratory, The Armed Forces Institute of Pathology, Washington, DC 20306-6000, USA
*Department of Pathology, University of North Texas Health Science Center, Ft. Worth, TX 76107, USA

INTRODUCTION

Determining the primary sequence of DNA is the ultimate form of DNA testing for discriminating between individuals. Various DNA sequencing methodologies are currently available, although the dideoxynucleotide "terminator" chemistry of Sanger is the most widely used in laboratories today (1,2). The development of fluorescence and infrared-based detection methods, coupled with automated DNA sequencing instruments, has greatly improved throughput and efficiency (3-7). Therefore, sequencing technology is readily available to most laboratories, and automation has made the process more cost effective.

At the present time, length variations (RFLP and AmpFLP loci) and sequence site-specific variations (HLA DQα and Polymarker from Roche Molecular Systems) are more than adequate to reliably distinguish between unrelated individuals. Nevertheless, the question often arises in the courtroom whether two alleles with the same length, or the same site-specific sequence, actually have the same primary sequence? There is always a chance that matching alleles will differ at the sequence level. Regardless, an example of such a case where many alleles match at different loci, yet the matching alleles have different sequences, would be an extremely rare event. The more loci tested which result in a match, the more rare the event would become. Therefore, although sequence analysis is available for forensic identification, it is not required to discriminate between individuals. Sequence analysis would be most useful in situations where currently available methods fail. One such area which is receiving attention, and has been generally accepted as a method of human remains identification, is sequence analysis of the mitochondrial DNA (mtDNA) genome (8-12).

Three characteristics of mtDNA are important for human remains identification. First, mtDNA is present in high copy number in each human cell (13). On average, each cell contains hundreds of copies of mtDNA for every copy of nuclear DNA. Second, contrary to classical inheritance patterns, mtDNA is maternally inherited (14). To date, no examples of paternal inheritance of mtDNA in humans has been observed. Finally, mtDNA is polymorphic (15-17). Within the 1200 base pair control region there are two hypervariable segments designated HV1 and HV2. Using the numbering system determined by Anderson (18), most of the polymorphic information is between positions 16024 and 16365 of HV1 and between positions 73 and 340 of HV2. Although sequence analysis of HV1 and HV2 is preferred, techniques for rapid comparison of mtDNA sequence types are being developed for screening prior to sequencing (19,20). Until these are further refined, sequence analysis will be performed without prior screening tests. This manuscript summaries the process of identifying human remains using mtDNA sequence analysis, and describes some of the results from recent research efforts completed in our laboratory. In addition, an example of two polymorphisms between a mother and child will be discussed. The methods of DNA

Advances in Forensic Haemogenetics 5 (c) Springer-Verlag Berlin Heidelberg

extraction, amplification and sequence analysis can be found below in the Materials and Methods section.

MATERIALS AND METHODS

Extraction of DNA From Tissue and Whole Blood

DNA was extracted from tissue and whole blood using the Chelex 100 method (23). Approximately one cubic millimeter of fresh tissue or 3 μl of whole blood were analyzed. The initial wash step was omitted when extracting DNA from tissue.

Extraction of DNA From Skeletal Remains

DNA extraction was performed using a modification of the method described by Holland et al. (10). In brief, approximately 2 grams of compact bone was extensively sanded and cleaned, crushed into small pieces, and the small pieces ground into a powder using either a Micro-Mill grinder (Fisher Scientific, Pittsburgh, PA), or a Mixer-Mill 8000 mixer (Spex Industries Inc., Edison, NJ). The powder was suspended in extraction buffer (10 mM Tris-HCl, pH 8.0, 100 mM NaCl, 2% sodium dodecyl sulfate, and 10 mM Na_2EDTA) in the presence of 0.67 mg/ml Proteinase K and incubated at 56°C for 12-18 hours. An additional 0.67 mg/ml Proteinase K was added and the incubation continued for 3 hours. Insoluble material was pelleted by centrifugation and discarded. The supernatant was purified in three steps: 1) organic extraction of protein using chloroform/phenol/isoamyl alcohol, 2) two n-butanol washes (to remove excess phenol and to reduce the extract volume), and 3) two successive washes of the DNA extract with sterile TE (buffer; 10 mM Tris-HCl, pH 7.6, 1 mM Na_2EDTA) in an Amicon Centricon 100 micro-dialysis column, followed by concentration of the sample to a final working volume of 200-500 μl.

Amplification of Mitochondrial DNA

We have found that for highly degraded DNA extracts, amplification of four overlapping regions of the control region, each approximately 250 bp in length, is necessary in order to obtain substantial sequence information (10,11). For convenience, the two hypervariable regions are amplified in their entirety for good quality DNA extracts.

The PCR master mix recipe for a 50 μl reaction is (all concentrations are final based on a 50 μl reaction) 10 mM Tris-HCl, pH 8.3, 50 mM KCl, 1.5 mM $MgCl_2$, 200 μM each dNTP, 0.4 μM forward amplification primer, 0.4 μM reverse amplification primer, 2.5 units of AmpliTaq® DNA polymerase (Perkin Elmer), QS to 40 μl with sterile deionized water. The PCR Master Mix can be supplemented with 8 μg of DNA grade bovine serum albumin or elevated AmpliTaq® (up to 12.5 units total AmpliTaq®) if inhibition of the amplification occurs. When amplifying good quality DNA, the PCR can be reduced to 25 μl and the master mix can be reduced to 15 μl. The PCR primer sets for mtDNA are listed in Table 1.

The amplification of good quality DNA was performed using the forward primers of Primer Sets 1 and 3, and the reverse primers from Primer Sets 2 and 4. For enhanced success in the DNA sequence analysis of PCR product of good quality DNA, primer R484 (5'-TGA GAT TAG TAG TAT GGG AG-3') can be substituted for the reverse primer of Primer Set 4. Add 10 μl of DNA template (approximately 1-5 ng of total genomic DNA, QS the volume to 10 μl with sterile deionized water) to each tube. Add 40 μl of PCR Master Mix to each tube (15 μl for good quality extracts) and place the tubes in a Perkin Elmer 9600 Thermal Cycler preheated to 96°C. The PCR was performed according to the following parameters; For Primer Set 1 - a) 94°C X 30 seconds; b) 94°C X 30 seconds, 62°C X 30

Advances in Forensic Haemogenetics 5 (c) Springer-Verlag Berlin Heidelberg

seconds, and 72°C X 30 seconds for 38 cycles; and c) 4°C soak. For Primer Sets 2, 3, and 4 - a) 94°C X 30 seconds; b) 94°C X 30 seconds, 56°C X 30 seconds, and 72°C X 30 seconds for 38 cycles; and c) 4°C soak. For good quality templates, the amplification parameters are a) 94°C X 30 seconds; b) 94°C X 20 seconds, 56°C X 20 seconds, and 72°C X 20 seconds for 30 cycles; c) 4°C soak. The PCR product was evaluated on a 1-2% agarose gel. If PCR inhibition occurred and BSA or elevated AmpliTaq® could not overcome the inhibition, no further amplification was attempted and DNA was re-extracted from the specimen source.

Table 1: PCR primers for the amplification of mitochondrial DNA from highly degraded DNA extracts.

Primer Set[a]	Forward Primer Light Strand	Reverse Primer Heavy Strand	Region Amplified
1	5'-TTA ACT CCA CCA TTA GCA CC-3'	5'-TGG CTT TGG AGT TGC AGT TG-3'	15971-16258 HV1[b]
2	5'-TAC TTG ACC ACC TGT AGT AC-3'	5'-CAC GGA GGA TGG TGG TCA AG-3'	16140-16414 HV1[b]
3	5'-CAC CCT ATT AAC CAC TCA CG-3'	5'-TGT GTG GAA AGT GGC TGT GC-3'	15-274 HV2[b]
4	5'-CTC ATC CTA TTA TTT ATC GC-3'	5'-CTG GTT AGG CTG GTG TTA GG-3'	145-389[c] HV2[b]

[a]Primer sequences are based on the numbering system of Anderson et al. (18).
[b]HV1 = Hypervariable Region 1, HV2 = Hypervariable Region 2.
[c]When amplifying good quality DNA extracts, the reverse primer R484 may be used (5'-TGA GAT TAG TAG TAT GGG AG-3').

Automated DNA Sequencing

Purification of the PCR product in a Centricon 100 (Amicon Corporation) spin-dialysis column is necessary before sequence analysis to remove unincorporated PCR primers. Following purification, each sequencing reaction contained 3.2-10 pmol sequencing primer, approximately 1 μg of DNA template, 9.5 μl of Applied Biosystems DyeDeoxy™ Terminator Cycle Sequencing reaction mix (or Prism™ Ready Reaction mix), and QS to 20 μl with sterile deionized water. The sequencing primers for forward and reverse sequencing of mitochondrial DNA are located in Table 2.

The thermal cycling conditions were as follows; a) Place tubes in a Perkin Elmer 9600 Thermal Cycler preheated to 96°C; b) 94°C X 30 seconds; c) 96°C for 15 seconds, 50°C for 5 seconds, 60°C for 2 minutes for 25 cycles; d) 4°C soak (modified with respect to the AB protocol).

The sequencing reaction products were purified using Quick Spin™ Columns from Boehringer Mannheim according to the Applied Biosystems protocol. A volume of 4 μl of loading buffer was added to each purified sample (4 μl of a mixture of 1 μl of 50 mM Na_2EDTA and 5 μl of deionized formamide), the samples were then mixed, allowed to re-suspend for approximately 10 minutes, and loaded within two hours. Samples can be stored for one month at -20°C before adding loading buffer.

Advances in Forensic Haemogenetics 5 (c) Springer-Verlag Berlin Heidelberg

Prior to loading, the reactions were heated at 90°C for 2 minutes, placed on ice for 5 minutes, and then loaded onto an AB 373A DNA Gene Sequencer. Gel electrophoresis was performed according to the AB 373A Protocol. A summary of the conditions is as follows: a) A pre-run electrophoresis step in 1X TBE (89 mM Tris-HCl, pH 8.3, 89 mM Boric Acid, and 2 mM Na_2EDTA) was executed for approximately 15 minutes; b) The samples were loaded, even numbers first. Electrophoresis was performed for approximately 10 minutes; c) The remaining samples were loaded; d) The electrophoresis was performed for 9-14 hours depending upon the amount of sequence desired (400-800 base pairs, respectively). The run parameters were at a constant power of 30 watts, variable readings of 18-21 milliamps and 980-1600 volts, and a temperature of 40°C. The data was collected and analyzed using AB software. Sequence comparisons were made using AB SeqEd (sequence editing) software.

Table 2: DNA sequencing primers for mitochondrial DNA sequence analysis of highly degraded DNA.

Primer Set[a]	Forward Sequencing Primer	Reverse Sequencing Primer	Sequence Obtained (Approximate)
1	F15971 5'-TTA ACT CCA CCA TTA GCA CC-3'	R16255 5'-CTT TGG AGT TGC AGT TGA TG-3'	16024-16225 HV1[b]
2	F16144 5'-TGA CCA CCT GTA GTA CAT AA-3'	R16414 5'-CAC GGA GGA TGG TGG TCA AG-3'	16180-16365 HV1[b]
3	F29 5'-CTC ACG GGA GCT CTC CAT GC-3'	R270 TGG AAA GTG GCT GTG CAG AC-3'	73-230 HV2[b]
4	F155 5'-TAT TTA TCG CAC CTA CGT TC-3'	R381[c] 5'-GCT GGT GTT AGG GTT CTT TG-3'	200-340 HV2[b]

[a]Primer sequences are based on the numbering system of Anderson et al. (18).
[b]HV1=Hypervariable Region 1, HV2=Hypervariable Region 2.
[c]When sequencing PCR product generated from good quality DNA extracts, the reverse primer R484 may be used in both the amplification and sequencing reactions (5'-TGA GAT TAG TAG TAT GGG AG-3').

SUMMARY

Primers reported in the literature (21,22), as well as novel primers (11), were used to amplify the two hypervariable regions (HV1 and HV2), and to generate sequence information from the products (Tables 1 and 2). The sequence is then compared to maternal reference sources to assist in the identification of human remains. Identifications have been made on a broad range of specimens; from aged skeletal fragments of the Vietnam War era to fresh tissue recovered from recent aircraft accidents (10,11). The high copy number of mtDNA allows for the analysis of only a few picograms of DNA. In addition, maternal inheritance allows for the identification of remains when immediate relatives are unavailable for comparison purposes. Distant maternal relatives can provide reference material for direct comparisons. Therefore, when the quality of the evidentiary DNA is not an issue, but

Advances in Forensic Haemogenetics 5 (c) Springer-Verlag Berlin Heidelberg

immediate relatives are not available, an identification can still be made using mtDNA sequencing.

When performing casework, sequence information is confirmed by sequence analysis of both DNA strands, or by multiple sequencing of the same strand. After generating confirmed sequence for both the evidentiary material and maternal references, the sequences are aligned and compared. Polymorphisms are determined with respect to the Anderson sequence (18). Transitions from one purine nucleotide to another purine nucleotide or a pyrimidine nucleotide to another pyrimidine nucleotide (i.e. adenine to guanine or cytosine to thymidine, respectively) are common polymorphisms. Insertions and deletions are also commonly observed. Transversions from a purine nucleotide to a pyrimidine nucleotide, or visa versa, are uncommon.

Once a mtDNA sequence match has been made, determination of the significance is necessary. In order to apply a statistical value to the match, databases have been generated by a number of laboratories (24,25). Our laboratory has a core database of 100 unrelated United States Caucasians. Adding staff, research, and casework, the total database is greater than 140. Within the core 100 Caucasians, we have found 92 unique sequence types, two individuals with one type, two other individuals with a second type, three individuals with a third type, and one individual with the Anderson sequence (18). Polymorphisms were detected at a total of 112 sites within the control region, and 52 of these were unique (only seen once). These data suggest that a significant number of polymorphisms have yet to be revealed. In total, 656 polymorphisms were detected in the 609 base pairs analyzed. Therefore, the average nucleotide diversity for each individual in the database is 1.07%. Overall, the polymorphic character of the control region appears to be excellent.

Combining our database with databases from other laboratories, there are more than 1000 individuals representing a mixture of Caucasians, African Americans, and Afro-Carribians. Most of the databases, however, have not been published. Because the databases have not been analyzed as a whole, there is no accepted statistical model available to calculate the significance of a mtDNA sequence match. Through personal communications, however, a new sequence can be compared to all databases. The frequency of a new sequence may then be reported as one over the size of the combined database. Although this does not take into account confidence limits or the relatively small size of the combined database, a match involving a mtDNA sequence that is unique to the database would strongly support an identification.

The mutation rate of the mtDNA control region is thought to be one polymorphism every 200-400 generations (26). This range was calculated by extrapolating back from sequence information of ancient populations to sequence information from modern populations. It is possible, however, that this approximation is too large. Unpublished data and results from our laboratory suggest that the mutation rate may be higher (25,27). A polymorphism has been observed between individuals separated by 18 generations (25). In addition, a polymorphism giving a heteroplasmic sequence profile was observed in the analysis of the Russian Tsar's remains (12,27). In our laboratory we have observed an example of two polymorphisms in the same individual which may have occurred in a single generation (28).

While completing a case involving skeletal remains, sequence was obtained for three family references (mother, brother, and sister of the deceased). The mtDNA sequences of the mother and brother (son) were identical. However, there were two polymorphisms between the mother and sister (daughter)(Table 3). The sequence analysis was repeated from the original blood sources with the same results. Consequently, the relationship between the

Advances in Forensic Haemogenetics 5 (c) Springer-Verlag Berlin Heidelberg

daughter and mother was questioned. The AmpFLP loci D1S80 (29), HUMFABP (30,31), HUMARA (30,31), and HUMACTBP2 (SE33)(32,33) were used to screen for maternity, and the daughter shared an allele with the mother at all loci. RFLP testing was subsequently performed by Art Eisenberg at the University of North Texas Health Science Center in Ft. Worth, Texas. A total of eight probes were used (D2S44, D14S13, D10S28, D4S139, D17S79, D1S7, D5S110, and D17S26). Again, the daughter shared at least one allele at each locus tested. A maternity index (using seven of the eight loci) of 197,695 was calculated and a 99.999% probability of maternity was determined. Therefore, the relationship of the daughter was no longer an issue. We have not been able to obtain the paternal reference in this case to determine whether this is an example of paternal inheritance. Assuming maternal inheritance, this would be the first example of two polymorphisms between generations found in humans. Nevertheless, given the two additional examples described above, the observed mutation rate of mtDNA may be higher than expected.

Table 3: Polymorphisms Observed for Three Case References

Reference	MtDNA Sequence Polymorphisms[a]
Mother and Son	16221T, 207A[b], 263G, 309.1C, 315.1C
Daughter	16093C[b], 16221T, 263G, 309.1C, 315.1C

[a]The numbering system represents the position where a polymorphisms has occurred and the nucleotide at that position. Numbers followed by a ".1" indicate an insertion.
[b]The positions where the two sequences differ.

In order to establish a mutation rate for the control region of mtDNA, a preliminary study was performed in our laboratory. Seven CEPH (Center for the Study of Human Polymorphisms, #13291-13294, #1331-1333) families were analyzed. On average, sequence information was generated for six children and two parents. The seven families exhibited a total of 58 polymorphisms at 28 different sites, and an average of 9 polymorphisms between each family. In a total of 46 mother-child comparisons no polymorphisms were observed. In addition, there were no examples of paternal inheritance. Additional studies are underway in our laboratory in collaboration with the Dr. Kevin Sullivan at the Home Office in England and Mark Wilson at the FBI Academy in Quantico, Virginia, to further investigate the rate of single-generation and multiple-generation mtDNA mutations.

In conclusion, comparative mtDNA sequence analysis is a powerful method of human remains identification. With the development of more rapid forms of automated sequence analysis, the rate of generating sequence information could be greater than 3000 nucleotide every hour in the near future (34). This would potentially provide an opportunity to sequence both nuclear and mtDNA on a routine basis in forensic laboratories.

REFERENCES

1. Maxam, A. M. and Gilbert, W. (1977) *Proc. Nat. Acad. Sci. USA*, **74**, 560-564.

2. Sanger, F., Nicklen, S., and Coulson, A.R. (1977) *Proc. Nat. Acad. Sci. USA* **74**, 5463-5467.

3. Smith, L.M., Sanders, J.Z., Kaiser R.J., Hughes, P., Dodd, C., Connell, C.R., Heiner, C., Kent, S.B.H., and Hood, L.E. (1986) *Nature* **321**, 674-679.

Advances in Forensic Haemogenetics 5 (c) Springer-Verlag Berlin Heidelberg

4. Ansorge, W., Sproat, B., Stegemann, J., Schwager, C., and Zenke, M. (1987) *Nuc. Acids Res.* **15**, 4593-4602.

5. Connell, C., Fung, S., Heiner, C., Bridgeham, J., Chakerian, V., Heron, E., Jones, B., Menchen, S., Mordan, W., Raff, M., Recknor, M., Smith, L., Springer, J., Woo, S., and Hunkapiller, M. (1987) *BioTechniques* **5**, 342-348.

6. Prober, J.M., Trainor, G.L., Dam, R.J., Hobbs, F.W., Robertson, C.W., Zagursky, R.S., Cocuzza, A.J., Jensen, M.A., and Baumeister, K. (1987) *Science* **238**, 336-341.

7. Brumbaugh, J.A., Middendorf, L.R., Grone, D.L., and Ruth J.L. (1988) *Proc. Nat. Acad. Sci USA* **85**, 5610-5614.

8. Sullivan, K.H., Hopgood, R., and Gill, P. (1992) *Int. J. Legal Med.* **105**, 83-86.

9. Ginther, C, Issel-Tarver, L., and King, M.-C. (1992) *Nature Genomics* **2**, 135-138.

10. Holland, M.M., Fisher, D.L., Mitchell, L.G., Rodriguez, W.C., Canik, J.J., Merril, C.R., and Weedn, V.W. (1993) *J. For. Sci.* **38**, 542-553.

11. Holland, M.M., Fisher, D.L., Roby, R.K., Ruderman, J., Bryson, C., and Weedn, V.W. (1993) Mitochondrial DNA sequence analysis of human remains, *Crime Laboratory Digest*, submitted for publication May 1993.

12. Ivanov, P. (1993) DNA-based identification of the last Russia's royal family remains, The Fourth International Symposium on Human Identification, Scottsdale, AZ, September 29, 1993.

13. Robin, E.D. and Wong, R. (1988) *J. Cell. Phys.*, **136**, 507-513.

14. Giles, R.E., Blanc, H., Cann, H.M., and Wallace, D.C. (1980) *Proc. Nat. Acad. Sci. USA*, **77**, 6715-6719.

15. Aquadro, C.F. and Greenberg, B.D. (1983) *Genetics*, **103**, 287-312.

16. Cann, R.L., Brown, W.M., and Wilson, A.C. (1984) *Genetics*, **106**, 479-499.

17. Cann, R.L., Stoneking, M., and Wilson, A.C. (1987) *Nature*, **325**, 31-36.

18. Anderson, S., Bankier, A.T., Barrell, G.B., et al. (1981) *Nature*, **290**, 457-465.

19. Stoneking, M., Hedgecock, D., Higuchi, R.G., Vigilant, L., and Erlich, H.A. (1991) *Am. J. Hum. Genet.* **48**, 370-382.

20. Pena, S. (1993) Mitochondrial D-loop "signatures" with low-stringency specific primer PCR (LSSP-PCR): A novel approach to identity testing, The Fourth International Symposium on Human Identification, Scottsdale, AZ, September 29, 1993.

Advances in Forensic Haemogenetics 5 (c) Springer-Verlag Berlin Heidelberg

21. Orrego, C. and King, M.-C. (1990) PCR Protocols: A Guide to Methods and Applications, M.A. Innis, D.H. Gelfand, J.J. Sninsky, T.J. White (eds.), Academic Press, London, San Diego, pp. 416-426.

22. Sullivan, K. M., Hopgood, R., Lang, B., and Gill, P. (1991) *Electrophoresis*, **12**, 17-21.

23. Walsh, P. S., Metzger, D. A., Higuchi, R. (1991) *BioTechniques*, **10**, 506-513.

24. Piercy, R., Sullivan, K. M., Benson, N., and Gill, P. (1993) The application of mitochondrial DNA typing to the study of white Caucasian genetic identification, *Int. J. Leg. Med.*, in press.

25. Dr. Charles L. Ginther, The University of California, Berkeley, Department of Biomedical and Environmental Health Sciences, School of Public Health, Berkeley, CA, personal communication.

26. Dr. Mark Stoneking, Department of Anthropology, Pennsylvania State University, University Park, PA, personal communication.

27. Dr. Kevin Sullivan, The Central Research and Support Establishment, The Home Office Forensic Science Service, England.

28. Holland, M. M., Roby, R. K., Ruderman, J., Fisher, D. L., Lee, D. A., Kupferschmid, T., Lofts, R. S., Eisenberg, A. V., A new perspective on the mutation rate of the human mitochondrial DNA genome, manuscript in preparation.

29. Budowle B., Chakraborty R., Giusti A. M., Eisenberg A. J., Allen R. C. (1991) *Am. J. Hum. Genet.*, **48**, 137-144.

30. Edwards A., Civitello A., Hammond H. A., Caskey C. T. (1991) *Am. J. Hum. Genet.* **49**, 746-756.

31. Edwards A., Hammond H. A., Jin L., Caskey C. T., Chakraborty R. (1992) *Genomics*, **12**, 241-253.

32. Polymeropoulos M. H., Rath D. S., Xiao H., Merril C. R. (1992) *Nuc. Acids. Res.*, **20**, 1432.

33. Holland, M. M., Fisher, D. L., Lee, D. A., Bryson, C. K., and Weedn, V. W. (1993) DNA Fingerprinting: State of the Science, S. D. Pena, R. Chakraborty, J. T. Epplen, & A. J. Jeffreys (eds.), Birkhauser Verlag Basel, Switzerland, 267-274.

34. Swerdlow, H., Zhang, J. Z., Chen, D. Y., Harke, H. R., Grey, R., Wu, S., and Dovichi, N. J. (1991) *Anal. Chem.*, **63**, 2835-2841.

Advances in Forensic Haemogenetics 5 (c) Springer-Verlag Berlin Heidelberg