Forensic Science Communications April 2003 — Volume 5 — Number 2

Standards and Guidelines

Guidelines for Mitochondrial DNA (mtDNA) Nucleotide Sequence Interpretation

Scientific Working Group on DNA Analysis Methods (SWGDAM)

Introduction.......Quality Assurance/Quality Control.......Data Interpretation.......Reporting Results.......Weight of Evidence.......References

Introduction

The interpretation of mitochondrial DNA (mtDNA) nucleotide sequencing results in casework is a matter of professional judgement and expertise. Although there are general guidelines for mtDNA analysis (Budowle et al. 1999; Carracedo et al. 2000; Holland and Parsons 1999; Tully et al. 2001), not every situation can or should be covered by a preset rule. It is important that each laboratory develop and implement written guidelines for the interpretation of analytical results. This document provides a framework for the laboratory to develop mitochondrial DNA (mtDNA) interpretation guidelines. The laboratory's interpretation guidelines must be based on validation studies conducted according to the *Quality Assurance Standards for Forensic DNA Testing and Convicted Offender Databasing Laboratories*. The SWGDAM Validation Guidelines and data from the scientific literature (such as Wilson et al. 1995a) should also be consulted.

Quality Assurance / Quality Control

An essential practice for mtDNA typing is to minimize and monitor contamination within the laboratory. Because of the sensitivity of detection with mtDNA analysis, low levels of exogenous DNA contamination and/or background is sometimes observed. However, reliable results can be obtained using appropriate quality control measures.

Contamination must be monitored, and each laboratory must have a method to define and quantify contamination. Laboratories must determine the maximum allowable threshold for contamination through internal validation studies. The laboratory must have standard operating procedures in place to address contamination (Wilson et al. 1995b). Each laboratory must establish evaluation criteria for controls, including but not limited to a positive control, a negative control, and a reagent blank control. Each of these controls

must be processed through sequencing along with the sample. In addition, typing of laboratory personnel is highly recommended for elimination purposes in order to trace potential sources of contamination.

A positive control is a sample of known mtDNA sequence used to monitor the success of the analysis. The positive control must be processed starting at amplification (e.g., DNA purified from the HL60 cell line is required as a positive control for inclusion of mtDNA forensic data into the Combined DNA Index System [CODIS]).

Reagent blanks and negative controls are used to monitor levels of contamination. Reagent blanks monitor contamination from extraction to final sequence analysis. Negative controls monitor contamination from amplification to final sequence analysis. If the reagent blank and/or the negative control of a particular amplification results in a sequence that is the same as that of the sample, all data for the sample must be rejected. The analysis must be repeated, starting with the extraction of the sample. If contamination in the reagent blank and/or negative control is present above the threshold set by the laboratory, then the data cannot be used for interpretative purposes. The samples must be re-amplified or re-extracted.

Nucleotide sequence data obtained from population database samples must include a minimum of HV1 (positions 16024-16365) and HV2 (positions 73-340). Nuceleotide sequence from known (K) samples should include HV1 (positions 16024-16365) and HV2 (positions 73-340). There are no minimum length requirements for nucleotide sequence data obtained from questioned (Q) samples. Both strands of the amplified product must be sequenced to reduce ambiguities in sequence determination.

Data Interpretation

The laboratory must establish criteria to assign nucleotide base calls to appropriate peaks or bands and to determine whether the results are of sufficient quality for interpretation purposes. The overall quality of the electropherogram data must be assessed. The results must be examined to determine if they meet the laboratory's analytical and interpretation threshold(s) established through internal validation studies. If the overall quality of the electropherogram is not suitable for analysis, the data should be rejected and the sample should be re-extracted, re-amplified, and/or re-sequenced.

A consensus sequence obtained from the sample will be compared to the Revised Cambridge Reference Sequence (rCRS) described by Andrews and co-workers ([Andrews et al. 1999; rCRS replaced the Cambridge Reference Sequence (CRS) described by Anderson and co-workers (Anderson et al. 1981]). Differences between the reference sequence and the sample sequence will be noted as polymorphisms. The nucleotide position and the DNA base difference from the reference will be noted (e.g., 16089 C). (N. B. listing of polymorphisms relative to rCRS is required for inclusion of mtDNA data into CODIS).

Insertions are described by noting the site immediately 3' to the insertion with respect to the light strand of the rCRS followed by a point and a '1' for the first inserted base, with sequential numbering for each inserted base thereafter. With homopolymeric regions, the insertion is placed at the highest-numbered end of the homopolymeric region with respect to rCRS. Insertions should not alter subsequent numbering of the sequence. Variants from rCRS should be coded in accordance with the guidelines proposed by Wilson et al. 2002a and Wilson et al. 2002b.

DNA base call designation should be based on the nomenclature system set forth by the International Union of Pure and Applied Chemistry (IUPAC). At confirmed positions of ambiguity, the following IUPAC codes should be used:

G/T = K	A/C = M
A/G = R	A/G/T = D
G/C = S	A/C/T = H
A/T = W	A/C/G = V
C/T = Y	C/T/G/ = B
A/C/G/T = N	

The conservative approach to listing ambiguities is to call them as 'N'. Deletions should be marked as ' - '.

All relevant sequence traces must be imported into a software program for analysis and alignment. The heavy strand sequences should be reverse-complemented so that the bases are aligned in the light strand orientation. Strands are compared and bases designated. Heteroplasmy is defined as more than one mtDNA type present in an individual that can be detected at an operational level. Heteroplasmy can be observed as point heteroplasmy where two DNA bases are observed at the same nucleotide position. Heteroplasmy can also be seen as length heteroplasmy, which typically is observed as a variation in the number of bases in a homopolymeric stretch of bases (i.e., C-stretch). Each laboratory should define heteroplasmy within the operational limits of the system used for sequencing. When the specimens under consideration differ by a single nucleotide, additional samples should be run in order to attempt to resolve the interpretation.

Long stretches of the same nucleotide are referred to as homopolymeric tracts. In HV1, the homopolymeric C-stretch region typically starts at nucleotide position 16184, in HV2 the homopolymeric tract is found between nucleotide positions 303 to 315. Homopolymeric tracts can differ in length within the same individual. In most cases, no attempt will be made to determine the exact number of bases in an HV1 C-stretch; however, laboratories must develop their own interpretation guidelines for HV2 length variants. A common length variant usually can be determined in the HV2 homopolymeric tract. A length variant alone cannot be used to support an interpretation of exclusion (Stewart et al. 2001).

Reporting Results

The laboratory must define conditions under which the data would lead to the conclusion that an individual can or cannot be eliminated as a possible source of the mtDNA. In addition, laboratories should develop guidelines for evaluation of cases where heteroplasmy may have occurred. This may be accomplished by an examination of the number, position, and nucleotide composition of polymorphic sites.

The following guidelines may be used in most cases:

- **Exclusion**—If there are two or more nucleotide differences between the questioned and known samples, the samples can be excluded as originating from the same person or maternal lineage.

- **Inconclusive**—If there is one nucleotide difference between the questioned and known samples, the result will be inconclusive.

- **Cannot Exclude**—If the sequences from questioned and known samples under comparison have a common base at each position or a common length variant in the HV2 C-stretch, the samples cannot be excluded as originating from the same person or maternal lineage.

Weight of Evidence

The mtDNA profile of a reference sample and an evidence sample that cannot be excluded as potentially originating from the same source can be searched in a population database. The population database(s) used to assess the weight of forensic evidence such as the mtDNA population database or CODIS (Miller and Budowle 2001; Monson et al. 2002) must be documented.

References

Anderson, S., Bankier, A. T., Barrell, B. G., deBruijin, 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., and Young, I. G. Sequence and organization of the human mitochondrial genome, *Nature* (1981) 290:457-465.

Andrews, R. M., Kubacka, I., Chinnery, P. F., Lightowlers, R. N., Turnbull, D. M., and Howell, N. Reanalysis and revision of the Cambridge Reference Sequence for human mitochondrial DNA, *Nature Genetics* (1999) 23:147.

Budowle, B., Wilson, M. R., DiZinno, J. A., Stauffer, C., Fasano, M. A., Holland, M. M., and Monson, K. L. Mitochondrial DNA regions HV1 and HV2 population data, *Forensic Science International* (1999) 103:23-35.

Carracedo, A., Bar, W., Lincoln, P., Mayr, W., Morling, N., Olaisen, B., Schneider, P., Budowle, B., Brinkman, B., Gill, P., Holland, M., Tully, G., and Wilson, M. DNA Commission of the International Society for Forensic Genetics: Guidelines for mitochondrial DNA typing, *Forensic Science International* (2000) 110:79-85.

Holland, M. M. and Parsons, T. J. Mitochondrial DNA sequence analysis: Validation and use for forensic casework, *Forensic Science Review* (1999) 11:22-50.

Miller, K. W. P. and Budowle, B. A compendium of human mitochondrial DNA control region sequences: Development of an international standard forensic database, *Croatian Medical Journal* (2001) 42:315-327.

Monson, K. L., Miller, K. W. P., Wilson, M. R., DiZinno, J. A., and Budowle, B. The mtDNA population database: An integrated software and database resource, *Forensic Science Communications* [Online]. Available: www.fbi.gov/hq/lab/fsc/backissu/april2002/miller1.htm

Stewart, J. E. B., Fisher, C. L., Aagaard, P. J., Wilson, M. R., Isenberg, A. R., Polanskey, D., Pokorak, E., DiZinno, J. A., and Budowle, B. Length variation in HV2 of the human mitochondrial DNA control region, *Journal of Forensic Sciences* (2001) 46:862-870.

Tully, G., Bar, W., Brinkmann, B., Carracedo, A., Gill, P., Morling, N., Parson, W., and Schneider, P. Considerations by the European DNA profiling (EDNAP) group on the working practices, nomenclature, and interpretation of mitochondrial DNA profiles, *Forensic Science International* (2001) 124: 83-91.

Wilson, M. R., DiZinno, J. A., Polanskey, D., Replogle, J., and Budowle, B. Validation of mitochondrial DNA sequencing for forensic casework analysis, *International Journal of Legal Medicine* (1995a) 108:68-74.

Wilson, M. R., Polanskey, D., Butler, J., DiZinno, J. A., Replogle, J., and Budowle, B. Extraction, PCR amplification, and sequencing of mitochondrial DNA from human hair shafts, *Biotechniques* (1995b) 18:662-669.

Wilson, M. R., Allard, M. W., Monson, K., Miller, K. W. P., and Budowle, B. Recommendations for consistent treatment of length variants in the human mitochondrial DNA control region, *Forensic Science International* (2002a) 129:35-42.

Wilson, M. R., Allard, M. W., Monson, K., Miller, K. W. P., and Budowle, B. Further discussion of the consistent treatment of length variants in the human mitochondrial DNA control region, *Forensic Science Communications* [Online]. Available: www.fbi.gov/hq/lab/fsc/backissu/oct2002/wilson.htm

Top of the page

Forensic Science Communications April 2003 — Volume 5 — Number 2