B26 A Custom Relational Database Application to Assist in the Interpretation of Novel mtDNA Sequence Variation

Kimberly A. Sturk, MFS, Jodi A. Irwin, MS, Jennifer E. O'Callaghan, MFS, Jessica L. Saunier, BS, and Michael D. Coble, PhD, Armed Forces DNA Identification Laboratory, 1413 Research Boulevard, Building 101, Rockville, MD 20850; and Thomas J. Parsons, PhD, International Commission on Missing Persons, Alipasina 45 A, Sarajevo, 71000, Bosnia and Herzegovina*

After attending this presentation, attendees will be familiar with a bioinformatics tool that assists in the interpretation of complex mtDNA sequence variation and have a greater appreciation for the extreme variability of the mtDNA control region.

This presentation will impact the forensic community and/or humanity by highlighting the importance of consistent nomenclature for mtDNA forensic evidence and presenting a tool in which to maintain this necessary consistency.

The use of mitochondrial DNA typing in forensic investigations is dependent on the comparison of a questioned haplotype to specific reference databases. Differences from the rCRS[1], as determined by alignment of the unknown sequence to the reference sequence, are queried against population databases in order to determine the relative rarity of the questioned haplotype. It is not uncommon, however, to encounter unusual variants that can be aligned to the rCRS in multiple ways. This can result in identical sequences being interpreted differently, and can ultimately lead to skewed database comparisons, if the alignment and/or nomenclature of specific motifs differs between the unknown sample and the reference database.

Given the frequent occurrence of unusual variants, the Armed Forces DNA Identification Laboratory (AFDIL) has developed an Access-based sequence calling guide to establish consistent nomenclature for internal population database and casework samples. Most of the nomenclature for the unusual variants included in the guide follows published recommendations for the placement of insertions and deletions (indels).[2,3] However, new sequence variations that cannot easily be interpreted with published guidelines are regularly encountered. Thus, additional specific guidelines for motifs such as these have also been included.

The application-based calling guide catalogs previously observed sequence variants and organizes them by region (HVI, HVII, HVIII, etc.). The information is accessed through a graphical user interface (GUI) that presents various options to the users. These menu options provide the user with multiple ways to search the database and provide links to various help documents that describe nomenclature standards and database usage. In general, a user would reference the calling guide with a questioned motif in hand and then search specifically for information that would provide nomenclature guidance for that sequence. The user can decide to search the entire database, or can select a specific region (HVI, HVII, etc.) in which to focus the query. If the user elects to search a region, the application will return all unusual motifs previously observed in that region. Motifs or sequences that best represent the mtDNA variation in the questioned sample can then be identified. In addition, the guide offers a unique search feature that accepts simple text strings and returns the most appropriate nomenclature for the submitted motif. This tool filters through hundreds of potential calls and expedites the search for similar sequences. In all cases, electropherograms are linked to the records/motifs so that users can directly compare the raw sequence data from their questioned sample to the examples in the database. The database also tracks alternative calls that are in use by other laboratories so that potential nomenclature discrepancies can be easily identified.

The AFDIL Sequence Calling Guide is used regularly in practice. It is an additional tool that can be referenced to eliminate the subjective alignment of sequence data and establish greater consistency in the interpretation of unusual sequence variation.

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 United States Department of Defense or the United States Department of the Army.

References:

1 Anderson S, Bankier AT, Barrell BG, de Bruijn MH, Coulson AR, Drouin J, Eperon IC, Nierlich DP, Roe BA, Sanger F, Schreier PH, Smith AJ, Staden R, Young IG (1981) Sequence and organization of the human mitochondrial genome. Nature 290(5806):457-65.

2 Wilson MR, Allard MW, Monson K, Miller KW, Budowle B (2002) Recommendations for consistent treatment of length variants in the human mitochondrial DNA control region. Forensic Sci Int 129: 35-42.

3 Wilson MR, Allard MW, Monson K, Miller KW, Budowle B (2002) Further discussion of the consistent treatment of length variants in the human mitochondrial DNA control region. For Sci Comm 4(4).

Mitochondrial DNA, Nomenclature, Sequence Alignment

 ** Presenting Author*