A65 Maximizing mtDNA Testing Potential With the Generation of High-Quality mtGenome Reference Data

Rebecca S. Just, MFS, 1413 Research Blvd, Bldg 101, 2nd Fl, Rockville, MD 20850; Melissa K. Scheible, MFS, 15245 Shady Grove Rd, Ste 335, Rockville, MD 20850; Spence A. Fast, MS, Kimberly S. Andreaggi, MFS, and Jennifer L. Higginbotham, MFS, 115 Purple Heart Dr, Dover AFB, DE 19902; Elizabeth A. Lyons, MFS, 7320 N Canal Rd, Lansing, MI 48913; Jocelyn M. Bush, MSFS, Michelle A. Peck, MFS, and Joseph D. Ring, MS, 115 Purple Heart Dr, Dover AFB, DE 19902; Toni M. Diegoli, MFS, 115 Purple Heart Dr, Dover AFB, DE 19904; Alexander W. Röck, PhD, and Walther Parson, PhD, Muellerstraße 44, Innsbruck, AUSTRIA; and Jodi A. Irwin, PhD, 1413 Research Blvd, Bldg 101, 2nd Fl, Rockville, MD 20850*

The goals of this presentation are to inform attendees of the need for complete mtGenome reference data for forensic applications, and to report an automated, high-throughput laboratory processing workflow for the generation of high-quality mtGenome haplotypes.

This presentation will impact the forensic science community by developing mtGenome reference data, which meet the highest forensic standards, and making the data publicly available via a web-based search engine to permit their use for forensic casework.

Mitochondrial DNA (mtDNA) testing in the forensic context requires appropriate, high-quality population databases for estimating the rarity of questioned haplotypes and, in turn, the strength of the mtDNA evidence. Since 2003, the Armed Forces DNA Identification Laboratory has been systematically generating mtDNA data to augment available reference population data and, ultimately, to improve the framework upon which forensic mtDNA typing is based. These data; however, and indeed all available forensic mtDNA reference databases (SWGDAM, EMPOP), only include information from the control region (CR). While this information is obviously strengthening the foundation upon which current mtDNA identification efforts are based, these data do not adequately prepare the field for the recent and rapid advancements in mtDNA typing technologies that will soon facilitate the acquisition of entire mitochondrial genome (mtGenome) information from forensic specimens. Novel assays that quickly and easily access mtDNA coding region data for increased discrimination are now available in the form of single nucleotide polymorphism assays, sequence specific oligonucleotide strips, mass spectrometry instrumentation, and next generation sequencing technologies.

Currently; however, there is a lack of appropriate, randomly-sampled and high-quality entire mtGenome reference data suitable for forensic comparisons. Thus, this funded project intends to: (1) increase the large-scale availability of high-quality entire mtDNA genome reference population data; and, (2) improve the information technology infrastructure required to access/search mtGenome data and employ them in forensic casework. The specific goals and objectives of this large-scale databasing effort are the development of 450 complete, high-quality mtGenomes spanning three U.S. population groups, and structure and query modifications to the publicly-available EMPOP database.

To assure the generation of the highest quality mtGenome profiles, a laboratory processing workflow in which nearly all pipetting steps — from initial sample placement through sequence detection — are performed robotically, and employing a rigorous data review process. Amplification of the complete mtGenome is achieved via eight overlapping fragments, with a total of 11 samples (and the appropriate negative controls) amplified per 96-well plate. Each mtGenome is then sequenced in 135 reactions, providing redundant and overlapping forward and reverse sequence coverage across the entire molecule. This optimized, highly automated protocol reduces overall data generation costs, hands-on laboratory time and — most importantly — opportunities for human error by substantially decreasing the number of manual production steps and the extent of sample reprocessing necessary to construct complete mtGenome haplotypes.[1] The data review process follows a strategy previously used for the production of high-quality mtDNA CR sequences, which includes raw data review by no fewer than three scientists, and entirely electronic data transfer with two additional profile reviews.[2] To further assure data reliability, completed mtGenome haplotypes are compared to PhyloTree to confirm phylogenetic consistency across the eight amplicons.[3] In addition, private mutations, heteroplasmies, and transversions are re-reviewed in the raw data.

The presentation will describe the application of this workflow to the development of more than 200 complete mtGenomes from anonymized blood serum samples as part of the NIJ-funded databasing effort. The workflow reliably produced high-quality data from DNA inputs down to at least 10 pg, and the majority of samples did not require any manual reprocessing to generate complete haplotypes. The efficacy of automated processing combined with a rigorous data review strategy in preventing errors with this multi-amplicon protocol was evident from the absence of problems detected at the stage of phylogenetic data evaluation. Ultimately, this project will provide the forensic community with reliable, complete mtGenome reference data and a means to access, search, and use the data in forensic casework.

References:

1. Lyons *et al*., Poster presentation, AAFS 2011, Chicago, IL.
2. Irwin *et al*., *Forensic Sci Intl: Genet* (2007) 1(2):154-7.
3. Van Oven and Kayser, *Hum Mutat* (2009) 30:E386-E394.

 ** Presenting Author*

Mitochondrial DNA, Reference Data, Data Quality

 ** Presenting Author*