COMBINING MTDNA SEQUENCE INFORMATION WITH SINGLE AMINO ACID POLYMORPHISM (SAP) ANALYSIS OFFERS GREATER DISCRIMINATION FOR SHED HAIRS

Emily C. Ashe[1], Glendon J. Parker[2], Mark. F. Kavlick[3]
[1]Visiting Scientist Program, Counterterrorism and Forensic Science Research Unit, Federal Bureau of Investigation Laboratory Division
[2]University of California, Davis
[3]Counterterrorism and Forensic Science Research Unit, Federal Bureau of Investigation Laboratory Division

Shed hairs are a common form of evidence encountered in crime scenes and missing persons cases and are often the only biological evidence available. Unfortunately, hair shafts rarely contain detectable quantities of nuclear DNA (nDNA), requiring analysts to turn to mitochondrial DNA (mtDNA) for examination. Currently, the FBI Laboratory sequences the control region (CR), a highly variable non-coding 1,121 base pair segment of the mitochondrial genome. While the information obtained from mtDNA can be probative, it is not without drawbacks. Due to the maternal inheritance of mtDNA, there is not enough variability among close family members to use mtDNA as a method for identification. Alternatively, recent studies have examined the utility of SAPs, single amino acid polymorphisms, obtained from protein analysis. Since SAPs are produced by non-synonymous SNPs, they indirectly provide information on nDNA. However, similar to mtDNA analysis, SAP profiles do not permit identification due to frequency estimates ranging from 1/1.001 to 1/8,437.

This study explores combining mtDNA and SAP analyses in pursuance of reducing frequency estimates and improving discrimination power. mtDNA from 57 hair shaft samples was extracted, amplified, and then sequenced via Sanger sequencing. Amplification of the whole control region was accomplished for 28% of all samples, while hypervariable regions 1 and 2 were achieved for the other 72%. Sequences were analyzed using Sequencher™ 5.4.5 software and then interpreted through the use of EMPOP (www.empop.org). The most conservative frequency estimate was determined using the Clopper Pearson confidence interval for haplotypes appearing in the database, while EMPOP's CI from Zero Population function was utilized for those that did not. The average frequency estimate when querying the database's European population was 1/214 (range, 1/13 to 1/2,416). The mechanisms of maternal mtDNA and autosomal SNP inheritance are biologically independent. If statistical independence is assumed and the mtDNA frequencies are combined with imputed SNP profile probabilities then frequency estimates averaged 1/859 with ranges of 1/33 to $1/1.9 \times 10^7$, thus improving the power of discrimination for some specimens. While the combination of mtDNA CR and SAP analyses still doesn't yield a power of discrimination high enough for identification purposes, and correlation patterns of mtDNA haplotype and protein-based imputed SNP profiles need to be elucidated, this approach does improve forensic DNA analysis of shed hair.