ANALYSIS OF mtDNA MIXTURES USING NEXT-GENERATION SEQUENCING TECHNOLOGIES AND CUSTOM SOFTWARE

Cassandra D. Calloway, PhD[1,2], Hanna Kim[1], Daniela Cuenca, MS[1,2], Valarie McClain[1,2], Richard E. Green, PhD[3], George Sensabaugh, D. Crim [2,4], and Henry Erlich, PhD[1]

[1]Children's Hospital Oakland & Research Center
[2]University of California, Davis
[3]University of California, Santa Cruz
[4]University of California, Berkeley

Next-generation sequencing technologies have revolutionized the field of genetics and have the potential to make a significant impact to the DNA forensics field. Specifically, the clonal sequencing aspect of these technologies allows for sequencing individual DNA fragments which can be applied to analysis of DNA mixtures often encountered in forensic casework. Current methods which use Sanger sequencing for analysis of mtDNA do not allow for interpretation of mixtures since the peak heights do not accurately represent the individual sequence components of a mixture. Additionally, Sanger sequencing is also limited in detection of minor components of a mixture (10-15%). We present here results from two NGS mtDNA enrichment assays that we developed which demonstrates analyzed using custom software demonstrating the application of two NGS mtDNA enrichment assays for analysis of mixtures.

A duplex PCR assay targeting the mtDNA hypervariable regions I and II (HVI/HVII) was developed using eight sets of 454 MID tagged fusion primers in a combinatorial approach for deep sequencing 64 samples in parallel on a 454 GS Jr. This assay was shown to be highly sensitive for sequencing limited DNA amounts (~100 mtDNA copies) and detecting mixtures with low level variants (~1%) as well as heteroplasmy. In addition, "complex" mixtures (≥3 contributors) were successfully sequenced and analyzed using custom software. Currently, commercial software is limited for the analysis of mixtures as these softwares only report the frequency of variants detected and not the frequency of the distinct sequence haplotypes. We have developed and tested an algorithm, '*hap-summary.pl',* which reports the frequency of the detected sequence haplotypes allowing for analysis of complex mixtures. We have also developed a solution phase sequence capture and NGS assay for targeted enrichment and deep sequencing of the entire mitochondrial genome for increased discrimination power. Using this Sequence Capture NGS assay, 100% sequence coverage of the mitochondrial genome with an ~80% on target rate was achieved as well as detection of a minor component in mixture. Custom software was also developed and used for the alignment and mixture analysis of the entire mitochondrial genome and results are presented here.