Forensic Applications of GeneChip® Mitochondrial Resequencing Arrays

Linda Strausbaughı[1], John Schienman[1], Dione Kampa Bailey[2], Richard Chiles[2], Sean Walsh[2], Giulia Kennedy[2], Josh Suhl[1], Igor Ovchinnikov1, Lu Li[1], Daniel Renstrom[1], Jennifer Tye[2], Anna Berdine[2], Carl Ladd[3] and Henry C. Lee[3] Center for Applied Genetics & Technology, University of Connecticut College of Liberal Arts and Sciences, Storrs, CT [2] Affymetrix, Inc., Santa Clara, CA[3] Connecticut State Forensic Sciences Laboratory, Meriden, CT

Mitochondrial DNA (mtDNA) typing based on the sequences of Hypervariable Regions I and II (HVI, HVII) has proven to be a valuable tool for certain forensic applications, especially when evidentiary DNA is limited in quantity and/or quality. One drawback to typing based on HVI and HVII is that the resolution is relatively low. Several recent studies have demonstrated improvements in resolution by incorporating additional mtDNA sequences outside of the hypervariable region.

We have explored the use of microarray resequencing as an alternative to targeted SNP assays and standard automated sequencing for the acquisition of larger amounts of mtDNA sequence information. The built-in redundancy inherent in the array tiling strategy that provides sequence confirmation, as well as the speed of data acquisition and analysis, make this an attractive method for mtDNA analysis. In this study, we resequenced mtDNA from a set of 14 individuals from our Connecticut Geographic and Genetic History Collection (CT) who represent different degrees of polymorphism in HVI and different ethnicities. Of particular interest: four of the individuals have HVI sequences identical to the RCRS; two differ from each other by one SNP; and one person with highly polymorphic mtDNA has 11 differences in HVI from the RCRS. As little as 5 ηg of genomic DNA (per PCR reaction) extracted from buccal swabs provided high quality results. DNA stored as long as 2 years also produced high quality results and all SNPs identified from the reference sequences were identical in the duplicates. Call rates from the resequencing arrays were very high. >98% of the 16,543 bases in the mt genome were called for all 14 samples. For 11 of the 14 mtDNAs, there was 100% concordance between sequences derived from arrays and those from automated sequencers. For 2 of the 14 samples, the same SNP in HVI was miscalled compared to standard sequencing (16362 C called as T).

The individual who is highly polymorphic relative to RCRS proved problematic for the version of the resequencing chip used in this study: the presence of multiple SNPs in close proximity resulted in "N" calls. The poly-C/G track proved challenging for both microarray and standard methods; in microarray resequencing, most "N" calls are C bases in poly-C stretches. Resequencing to capture the full mtDNA sequence proved considerably more effective than hypervariable region analysis alone in mitotype resolution. The level of variation in the hypervariable region appears to be linearly related to the variation in the rest of the mtDNA. However, even CT samples with little or no variation in the hypervariable region still have significant variation in the coding region relative to the RCRS. Acquisition of the full mtDNA sequence by microarray resequencing resolved all four of the individuals who had HVIs identical to the RCRS; in

all cases there were multiple differences in pairwise comparisons. Resolution of the two individuals who had nearly identical HVIs (but different from the RCRS) was also improved by the greater amount of sequence provided by resequencing. The results of this study suggest a number of strategies for improving microarray-derived mtDNA sequences and the use of this technology in forensic applications.