Highly Heterozygous STR Markers for Enhanced DNA Mixture Deconvolution

Nicole Novroski[1,2], August E. Woerner[2], Frank R. Wendt[2,3], Magdalena M. Bus[2], Michael D. Coble[2], and Bruce Budowle[2]
[1]Forensic Science Program, Department of Anthropology, University of Toronto Mississauga, 3359 Mississauga Road, Mississauga, Ontario, L5L 1C6, CANADA
[2]Center for Human Identification, University of North Texas Health Science Center, 3500 Camp Bowie Blvd., Fort Worth, Texas 76107, USA
[3]Department of Psychiatry, Yale School of Medicine, 333 Cedar Street, New Haven, Connecticut 06511, USA

De-convolution of complex mixtures can be challenging. Various improvements in polymerase chain reaction coupled with capillary electrophoresis (PCR-CE) and massively parallel sequencing (MPS) chemistries coupled with downstream statistical analyses have been developed and implemented to better resolve two or more person DNA mixtures. However, current genotyping outputs describe STR variation solely based on allele size and do not exploit the full genetic information contained within target markers to distinguish between or among component contributors.

MPS for typing forensically-relevant STR loci has dramatically impacted our abilities to identify allele diversity due to sequence variation within STR repeat and flanking regions. Studies have described STR sequence variation in large population groups and demonstrated that there are enormous amounts of diversity and complexity within the currently utilized STR markers for forensic genetic analysis. However, the literature continues to demonstrate that some of the current core CODIS loci lack repeat or flanking region sequence diversity, minimizing the relative information gain via MPS for these STRs. Thus, novel STRs with increased sequence variation should be sought to facilitate mixture deconvolution.

This presentation will highlight an exploratory study of a multiplex comprised of 73 highly polymorphic STRs from individuals comprising three US populations. Mixture deconvolution capabilities for two-person mixtures were assessed based on complete allele resolution per locus (i.e., four alleles observed) of pairwise mixtures using *in silico* methods. A subset of 20 loci (referred as the 20Plex) were compared to the 20 CODIS core loci on all population samples with full DNA profiles for both multiplexes (n= 443). Based on proportion of loci displaying four alleles, the 20Plex outperformed the CODIS core loci with increases of 82.6% and 89.3% using length-based and sequence-based alleles, respectively. The results of this study indicate that the 20Plex improves upon current mixture deconvolution efforts by employing markers that allow for better allele resolution of component contributors in a mixed DNA sample due to a substantial degree of diversity compared with the current core CODIS STR loci used for forensic identity typing. In turn, incorporation of these STRs into current multiplexes could facilitate probabilistic genotyping, which will increase the number of resolved genotypes in mixed samples being compared to reference and suspect profiles.