UNVEILING THE HIDDEN HAPLOTYPIC VARIATION IN OUR FORENSICALLY RELEVANT GENETIC MARKERS

Jonathan King, UNTHSC

The advent of massively parallel sequencing (MPS) technologies has expanded the capabilities to increase the discrimination power of currently used markers, identify novel markers, and increase marker multiplexing capacity. Within the flanking region of STR repeats and single nucleotide polymorphisms (SNPs) is substantial variation that has been untapped by standard capillary electrophoresis methods. Even more recent MPS-based marker panels contain markers that have yet to be characterized flanking-region variation. The value of these haplotypes is that they may improve match probabilities from low-template and/or degraded DNA samples and may provide value for mixture interpretation (as they do not generate stutter). The ForenSeq Signature Preparation Kit contains 165 SNP amplicons for ancestry- (aiSNPs), identity- (iiSNPs), and phenotype-inference (piSNPs). In this study, 714 individuals from four major populations were assessed using STRait Razor v2s to determine the level of diversity in the flanking regions of these amplicons. The results show that nearly 70% of loci showed some level of flanking region variation with 22 iiSNPs and 8 aiSNPs categorized as microhaplotypes in this study. The heterozygosities of these microhaplotypes approached, and in one instance surpassed, those of some core STR loci. Also, the impact of the flanking region on other forensic parameters (e.g., power of exclusion and power of discrimination) was examined. The probability of detecting a mixture (i.e., observing ≥ 3 haplotypes) using these loci alone was 0.9952. The decrease in random match probabilities for the full regions over the target iiSNPs was found to be significant. When combining the iiSNPs with the autosomal STRs, the combined match probabilities ranged from 6.40×10^{-73} (ASN) to 1.02×10^{-79} (AFA). These data support that when challenged samples yield partial profiles, the random match probability may still be quite low due to the extra variation adjacent residing in the same amplicon as the target loci. The implications of profile interpretation (degraded and mixtures) will be discussed.