THE FLANK, THE WHOLE FLANK, AND NOTHING BUT THE FLANK

L.A. Devesse[1], D.J. Ballard[1], L.B. Davenport[1], and D. Syndercombe Court[1]
[1] King's Forensics, King's College London

The application of massively parallel sequencing (MPS) to forensic genetics has led to improvements in multiple aspects of DNA analysis, however additional complexities are concurrently associated with these advances. In relation to STR analysis, the move to assign alleles using sequence rather than length-based methodologies has highlighted the extent to which previous allelic variation was masked. In this work, a series of samples (n=1000) from five different population groups were genotyped using the MiSeq FGx™ Forensic Genomics System. Sequence variation has been characterised both within and outside STR repeat regions.

At certain loci, the increase in allelic diversity when considering flanking region variation is significant – notably, D16S539 which shows very little repeat region variation but has a 100% increase in the number of distinct alleles observed when accounting for flanking region variants. However, 16 out the 27 loci studied show little to no variation outside of the repeat regions. Of the remaining STRs, D12S391 is the most polymorphic, where the number of alleles increases from 25 length-based to 98 sequence-based alleles, yet only 12% of this increase is provided by flanking region variants. The challenges associated with sequence-based allele nomenclature expand markedly when considering flanking regions, bringing into question the value of these regions for forensic casework – are they worth it? This presentation will illustrate how flanking region variation differs across markers and populations, as well as provide examples of complex relationship cases and how useful (or not) both repeat region and flanking region sequence variants can be.

Key words: MPS, STRs, Variation, Sequencing