MPS ANALYSIS PRESENTS A PARADIGM SHIFT IN THE METROLOGICAL TREATMENT OF FORENSIC DNA DATA

Brian Young[1], Daniel Hellwig2, Daniel Hellwig[2], Luigi Armogida[1]
[1]Nichevision Forensics,
[2]Intermountain Forensics

Targeted massively parallel sequencing (PCR-MPS) is an emerging platform for the analysis of forensic DNA, the advantages of which have been well described. Less appreciated are the metrological implications of the switch from length-based to sequence-based alleles. It is natural to view sequence- based STR alleles as variants of length-based alleles where allele number (a numeric property) remains the principal measure and DNA sequence (a nominal property) is treated as an accessory feature. This is a narrow view that constrains the evaluation of PCR-MPS data. The power of sequence-based alleles for the analysis of casework samples can be more fully exploited when sequences are treated as nominal property values.

Treating sequence-based alleles by their nominal property values enables cleaner classifications of alleles, stutter, and non-stutter artifacts. Complex and compound alleles may produce multiple stutter artifacts that numerically are all “N-1”, but nevertheless are distinct by sequence. Non-stutter artifacts generated by the Verogen platform consist mainly of nucleotide substitutions. These artifacts are the same length as the parent alleles from which they derived, as well as from isometric alleles that differ by a SNP polymorphism. We describe effective sequence-centric analysis for reliably discriminating erroneous nucleotide substitution variants from true nucleotide substation polymorphisms. We describe the principal methods of sequence-based analysis where standard numeric methods are replaced with string methods such as hashing, regular expression (REGEX), and comparisons based strictly on string equivalence. We also describe calculation of typing precision and accuracy for sequence-based alleles when validating or comparing PCR-MPS kits.

Addressing STR alleles by their nominal property values is consistent with a paradigm shift occurring throughout science with respect to measurement of nominal property values. Medical science in particular, is notable in that many examinations involve nominal properties (i.e., blood group, tumor type). The draft 4th edition of the International Vocabulary of Metrology (VIM) adds an entirely new section dedicated to nominal properties and examinations. We describe how forensic DNA analysis of sequence-based alleles fits into this framework, and we illustrate all examples with analyses from an operational casework laboratory.