DISCRIMINATION OF ALLELE AND ARTIFACT BY THE SEQUENCE IDENTIFIER TOOL, MIXTUREACE, AS DEMONSTRATED ON DATA DERIVED FROM THE FORENSEQ™ SIGNATURE PREP KIT

Vishakha Sharma, Amber Khan, Brian Young, Luigi Armogida, Elisa Wurmbach
New York Office of the Medical Examiner

Current identification methods that are based on DNA analysis utilize capillary electrophoresis for size separation of STR loci and disregard sequence specific information, which could improve individual identification and mixture deconvolution significantly.

However, data derived from massively parallel sequencing (MPS) are complex, as artifacts include stutter (sequences that vary in length by one or more repeat units), as well as sequencing errors (sequences that show substitutions, insertions and/or deletions compared to true alleles).

MixtureAce is a sequence identifier tool embedded in ArmedXpert (NicheVision) that can recognize DNA sequences and was used to analyze data generated by the ForenSeq™ Signature Prep kit.

Data analysis included nine experimental runs. The following samples were assessed: (i) standard run following the recommendation of the manufacturer; (ii) sensitivity run; and (iii) testing degraded DNA as well as challenging samples derived from bones and teeth; and (iv) mixture experiments. The mixed DNA samples were assessed by testing over 85 samples of two-person mixtures with ratios ranging from 1:1 to 1:50, as well as 29 samples of three to six-person mixtures with various ratios. ForenSeq™ tests over 150 loci, including Amelogenin, 26 autosomal STRs, 24 Y-STRs, and 7 X-STRs which were informative for the estimation of the number of contributors and their sexes.

MixtureAce was tested with an analytical threshold of a read depth of nine (singletons were disregarded). It is a fast tool performing notably well in distinguishing between signal (allele) and noise (artifacts).

Mixed data showed fewer artifacts compared to single source samples. From two-person mixtures, genotypes of the minor contributors were obtained down to ratio of 1:50.
MixtureAce settings can be further adjusted with regard to the utilized sequencing length (flanking regions) and in recognition of various artifacts (e.g. minus three stutter or sequence errors that contain two or more nucleotides).
Results of greater detail will be presented.