B107 Multi-Marker Match Statistics: Combining Results Across Sequence-Based Short Tandem Repeat (STR) and Identity Single Nucleotide Polymorphism (SNP) Markers

Katherine B. Gettings, PhD, National Institute of Standards and Technology, Gaithersburg, MD 20899; Andreas Tillmar, PhD, National Board of Forensic Medicine, Linkoping SE-58758, SWEDEN; Peter M. Vallone, PhD, National Institute of Standards and Technology, Gaithersburg, MD 20899-8314*

Learning Overview: After attending this presentation, attendees will understand the considerations for calculating multi-marker statistics in kinship and routine forensic cases.

Impact on the Forensic Science Community: This presentation will impact the forensic science community by ensuring attendees have an improved understanding of the statistical considerations for combining results across loci/markers, and ForenSeq™ users in attendance will receive actionable guidance.

Analyzing signs of linkage disequilibrium among population samples and father-son pairs will inform decisions regarding statistically combining markers.

Significant progress has been made in the past five years in developing the assays, bioinformatic methods, and population frequency data needed for the implementation of routinely sequencing STR markers. In addition, the forensic community is determining the utility of sequencing identity SNP markers, the smaller targets of which can outperform STRs in degraded samples. When both autosomal STR (auSTR) and autosomal identity SNP genotype data are present for a casework sample and these genotypes are consistent with a person of interest, laboratories are faced with the new challenge of reporting match statistics for a combination of marker types.

At this time, the most likely combination of markers to be genotyped concurrently, due to commercial availability, are the 27 auSTR and 94 Identity Informative SNP loci included in the Verogen ForenSeq™ DNA Signature Prep Kit. In the associated software user guide, the manufacturer indicates that the 121 autosomal genetic identity markers generally meet expectations of independence at the population level but may not be independent for the purposes of kinship analysis.[1] In such cases, they recommend either the use of haplotype frequencies for the pair of loci or performing the calculation using the more informative of the two loci within a particular case.

Tillmar and Phillips discussed the issue of linkage and provided a method for evaluating the effect of linkage in identity and relationship testing.[2] Regarding Linkage Disequilibrium (LD) assessment, this study provides tables of locus pairs within 0.5cM and evaluated these pairs for signs of LD with HapMap data and proxy SNPs as needed (e.g., for all STR loci). For the 121 ForenSeq™ STR and Identity Informative SNP loci, this study reported five pairs located within 0.5cM of one another and none showed signs of LD.

In this study, LD evaluation will be performed between these 121 SNP and STR loci using National Institute of Standards and Technology (NIST) population sequence data for approximately 1,400 samples, including 400 father-son pairs (auSTR data reported).[3] This analysis is representative of the use-case, as it will not require proxy SNPs, and it will be based upon the STR sequences, which are expected to be used for forensic casework calculations. In addition to outlining a method that can be reproduced for other current or future assay/marker combinations, this work will provide refined guidance to ForenSeq™ users on statistically combining results in identity and kinship cases.

Reference(s):

1. *ForenSeq Universal Analysis Software Guide*, Document # VD2018007 Rev. A (June 2018).
2. Tillmar, Andreas O. and Christopher Phillips. Evaluation of the impact of genetic linkage in forensic identity and relationship testing for expanded DNA marker sets. *Forensic Science International: Genetics*. (January 2017) 26:58-65. doi: 10.1016/j.fsigen.2016.10.007.
3. Gettings, Katherine B., Borsuk, Lisa A., Steffen, Carolyn R., Kiesler, Kevin M., and Peter M. Vallone. Sequence-based U.S. population data for 27 autosomal STR loci. *Forensic Science International: Genetics*. (November 2018) 37:106-115. doi: 10.1016/j.fsigen.2018.07.013.

STR, SNP, Linkage Disequilibrium

*Presenting Author