INTERPRETING Y CHROMOSOME STR HAPLOTYPE MIXTURE

Jianye Ge[1,2,3,] Bruce Budowle [2,3], Ranajit Chakraborty[1]

[1] Center for Genome Information, Department of Environmental Health, University of Cincinnati College of Medicine, Cincinnati, OH, 45267, USA

[2] Institute of Forensics and Investigative Genetics, University of North Texas Health Science Center, Fort Worth, Texas 76107

[3] Department of Forensic and Investigative Genetics, University of North Texas Health Science Center, Ft Worth, Texas 76107, USA

Mixture interpretation is a challenging problem in forensic DNA analyses. The interpretation of Y short tandem repeat (STR) haplotype mixtures, due to a lack of recombination), differs somewhat from that of the autosomal DNA markers and is more complex. We describe approaches for calculating the Probability of Exclusion (PE) and Likelihood Ratio (LR) methods to interpret Y STR mixture evidence with population substructure incorporated. For a mixture sample, first, all possible contributor haplotypes in a reference database are listed as a candidate list. The PE is the complement of the summation of the frequencies of haplotypes in the candidate list. The LR method compares the probabilities of the evidence given alternative hypotheses. The hypotheses are possible explanations for the mixture results. The maximum number of contributors is based on the candidate list and the computing complexity is polynomial. Mixtures were simulated by combining two or three 16 Y STR marker haplotypes derived from the US forensic Y STR database. The average PE was related to the size of database. With a database comprised of 500 haplotypes an average PE value of at least 0.995 can be obtained for two-person mixtures. The PE decreases with an increasing number of contributors to the mixture. Using the total sample population, the average number of candidate haplotypes of two-person mixtures is 3.73 and 95% of mixtures have less than or equal to 10 candidate haplotypes. More than 98.7% of two-person mixtures can only be explained by the haplotype combinations that mixtures are composed of two haplotypes. The average number of candidate haplotypes is generally higher for three-person mixtures, but the proportion of self-explained mixture is lower. A small proportion of three-person mixture can appear as only two haplotypes (ignoring quantitation).