A MATCH LIKELIHOOD RATIO FOR DNA COMPARISON

Mark W. Perlin, Ph.D., MD, Ph.D.[1], Joseph B. Kadane, Ph.D.[2], Robin W. Cotton, Ph.D.[3]
[1]*Cybergenetics, Pittsburgh, PA*
[2]*Statistics Department, Carnegie Mellon University, Pittsburgh, PA*
[3]*Biomedical Forensic Sciences, Boston University, Boston, MA*

Forensic scientists analyze DNA evidence to determine a strength of association between two biological specimens. When feasible, they use a likelihood ratio (LR) to describe this association, since the LR removes all considerations except of the weight of DNA evidence. LRs help the police in DNA investigation, and the prosecutor with presenting DNA evidence. The classical likelihood ratio is used routinely for identifying individuals (random match probability), kinship (paternity index) and mixtures (combined likelihood ratio).

Recent years have seen the advent of other, equally important, forensic DNA applications. These applications include missing persons, familial search, low copy number, mass disasters, touch DNA and complex mixture interpretation. Each of these approaches entails the analysis of uncertain DNA data, which may include more than one genotype possibility.

We have developed a general approach to constructing a match likelihood ratio (MLR) for all such DNA comparison problems [1]. The forensic scientist first infers a genetic profile, representing genotype uncertainty using probability. Then, genetic profiles are compared using a straightforward sum of products formula that computes the match likelihood ratio.

For example, the standard Common Probability of Inclusion (CPI) mixture statistic can be shown to be a match likelihood ratio. Every allele pair that is included in the allele data becomes a genotype possibility; each such possibility is then assigned an equal genotype probability. The CPI statistic is obtained by substituting these inclusion genotype probabilities into the MLR formula. This approach establishes that CPI is indeed a likelihood ratio, and also illustrates how genotype probabilities are regularly used in current forensic DNA practice.

Inferring genotypes, and then comparing them in a general match likelihood ratio, can be helpful in: (i) performing objective, unbiased comparisons, (ii) using highly informative genotypes inferred from statistical or "expert system" computing, (iii) separating crime scene and suspect DNA workflow processes, (iv) enabling sophisticated computer matching on genotype databases, (v) providing a general framework for current forensic match approaches (as with CPI above), (vi) quantifying the strength of case-to-case matches to link crime scenes, and (vii) validating and comparing different DNA interpretation methods for efficacy and reproducibility.

In this presentation, we introduce and motivate MLR. We give intuitive examples that show how match LR already appears in current DNA analysis, and discuss the inherent limitations of some earlier classical LR methods. We show how MLR transcends these limitations, and handles modern DNA problems, such as missing persons, familial search, low copy number, mass disasters, touch DNA and complex mixture interpretation. The use of MLR may help forensic scientists extract more match information from their current DNA data.

[1] Perlin MW, Kadane JB, Cotton RW. Forensic DNA Inference. In: Seventh International Conference on Forensic Inference and Statistics; 2008 August; Lausanne, Switzerland.