A77 Casework Validation of Genetic Calculator Mixture Interpretation

Mark W. Perlin, PhD, Cybergenetics, 160 North Craig Street, Suite 210, Pittsburgh, PA 15213; and Barry W. Duceman, PhD, NY State Police Crime Lab, Forensic Identification Center, 1220 Washington Avenue, Building 30, Albany, NY 12226*

After attending this presentation, attendees will better understand how to conduct a DNA mixture validation study, how to measure the efficacy and reproducibility of any DNA interpretation method, and why computer interpretation of DNA evidence can be more informative than manual review.

The presentation will impact the forensic science community by enabling practitioners to conduct DNA mixture validation studies on interpretation methods that they would like to present in court.

Interpreting DNA mixtures can be challenging. With the advent of statistical computing, one can reproducibly infer consistent, highly informative results. Such reliable mixture inference is critical for the admissibility of scientific evidence. This paper establishes the *efficacy* of computer-based genetic calculator mixture interpretation by comparing inferred match information on adjudicated mixture cases relative to currently used manual methods. It also demonstrates the *reproducibility* of the computer's results.

The key mixture interpretation task is inferring a questioned genotype of an unknown contributor. When there is uncertainty in an inferred genotype, allele pairs are assigned a probability distribution that describes this uncertainty. Different mixture interpretation methods may infer different genotype distributions.

A genetic calculator provides a statistical computer approach that infers genotypes by hypothesizing all feasible solutions, comparing these with observed STR peak height data, and assigning higher probabilities to genotype hypotheses that better fit the data. Two quantitative inference methods were examined:

- TA1, which uses a known victim genotype to help infer the other unknown contributor, and
- TA2 that does not use a victim genotype, but instead infers two unknown contributors.

There are also qualitative list-based inclusion methods that apply peak thresholds:

- CLR, which uses a known victim genotype, and
- CPI, a qualitative approach that does not use a victim genotype. The Likelihood Ratio (LR) is the generally accepted forensic science measure of match rarity. The LR gives the probability of a match between the evidence genotype and a suspect, relative to a match with a random person. The data-inferred evidence genotypes above (TA1, TA2, CLR, and CPI) each produce a LR match statistic when their probability distribution is substituted into a generic LR match formula.

The efficacy of the genetic calculator was determined by comparing its LR match information to other methods. In particular, the LR logarithm (i.e., order of magnitude, or powers of ten) was determined on eight adjudicated cases for the two unknown TA2 computer method, and compared with that of the reported CPI value. Whereas the average log (LR) information for CPI was 7 (LR = 10 million to one), the average match information on these same cases with TA2 was 13 (LR = 10 trillion). This shows a six order of magnitude improvement when using genetic calculator method TA2 relative to manual CPI.

Relative efficacy was also assessed when the victim profile was known, and just one unknown contributor was inferred. The average log (LR) match information reported on eight adjudicated CLR cases was 13 (10 trillion). The average genetic calculator TA1 match information on these same cases was 18 (quintillion), a five order of magnitude improvement. Thus, for both one and two unknown contributors, the genetic calculator mixture interpretation method is more informative than the manual CPI and CLR match statistics.

Reproducibility was measured on these sixteen mixture cases by obtaining duplicate computer solutions for each case. The average match information deviation between the two independent solutions was under half a log (LR) unit.

From this study it is concluded that a genetic calculator can provide reliable mixture interpretation. Specifically, when inferring either one or two unknown contributor genotypes, the genetic calculator is effective relative to current manual methods. Moreover, we quantified the genetic calculator's interpretation reproducibility using match information. The genetic calculator has already been admitted into evidence in a *Frye* jurisdiction. This validation study (assessing efficacy and reproducibility) establishes the genetic calculator's reliability for the additional prongs of *Daubert*.

DNA Mixture, Validation Study, Computer Interpretation

 ** Presenting Author*