A152 Getting Past First Bayes With DNA Mixtures

Mark W. Perlin, PhD, MD, Cybergenetics, 160 N Craig Street, Ste 210, Pittsburgh, PA 15213*

After attending this presentation, attendees will better understand the scientific foundation that underlies the DNA match statistic. Common misconceptions, and ways to avoid them, will be discussed.

This presentation will impact the forensic science community by enabling practitioners to more confidently present accurate DNA mixture results in the classroom and in court.

DNA mixtures are a prevalent form of biological evidence. A mixture contains DNA from two or more contributors. There are usually multiple genotype explanations for the observed Short Tandem Repeat (STR) data. Forensic scientists must understand genotype mixture inference in order to give accurate DNA mixture testimony in court.

Fortunately, Bayes theorem provides a robust framework for genotype inference and match. Over 250 years ago, the Reverend Thomas Bayes showed how to update the belief in hypotheses (probability) by examining how well those hypotheses explain observed data (likelihood). Bayes has individuals use all the data, and consider all hypotheses.

Bayesian genotype inference (for each contributor at every genetic locus) begins with a *prior* belief that the chance of observing an allele pair before seeing data is proportional to its population prevalence. Careful examination of STR data then uses a *likelihood* function to concentrate probability on those genotype values that best explain the laboratory data. This objectively inferred genotype associates a *posterior* probability with every allele pair, multiplying prior probability and likelihood.

A DNA match statistic assesses the strength of match between evidence and reference genotypes, relative to coincidence. This Bayesian Likelihood Ratio (LR) weighs two competing hypotheses — either the reference individual contributed DNA to the evidence, or he did not — based on the observed STR data.

Bayesian beginners often make mistakes. Beginners may fail to use all peak data, or not consider all genotype hypotheses. Bayesian beginners can confuse likelihood (chance of data given hypothesis) with probability (chance of hypothesis given data). A beginner may apply complex formulas when a simple ratio would suffice. Bayesian beginners may change assumptions in mid-step and suggest meaningless comparisons.

On April 12, 2013, the National Institute of Standards and Technology (NIST) Applied Genetics Group gave a full-day webinar on DNA mixture interpretation. The NIST group presented genotype and LR results from Bayesian software. While attendees correctly computed the final LR match numbers, some beginner errors were made. DNA interpretation missteps that appear harmless in a pedagogical setting can prove damaging in a court of law, where accuracy is paramount and cross-examination unforgiving.

This presentation reviews the basic principles of Bayesian DNA mixture interpretation. The NIST webinar experience provides teaching points that can help beginners avoid common mistakes. The corrections provided to NIST after the webinar are used here to highlight key interpretation steps. With some Bayesic training, DNA analysts can accurately testify about mixture results and get past first Bayes.

Bayes Theorem, Genotype Inference, Likelihood Ratio

 ** Presenting Author*