GETTING PAST FIRST BAYES WITH DNA MIXTURES

Dr. Mark W. Perlin, Cybergenetics

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 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 Rev. Thomas Bayes showed us how to update our belief in hypotheses (probability) by examining how well those hypotheses explain observed data (likelihood). Bayes has us 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 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. They may fail to use all peak data or not consider all genotype hypotheses. They can confuse likelihood (chance of data given hypothesis) with probability (chance of hypothesis given data). A beginner will apply complex formulas when a simple ratio would suffice. They may change their 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. Since their expertise lies elsewhere, they made many beginner errors and never got past first Bayes. Errors that appear harmless in an academic setting can prove fatal in a court of law, where accuracy is paramount and cross-examination unforgiving.

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