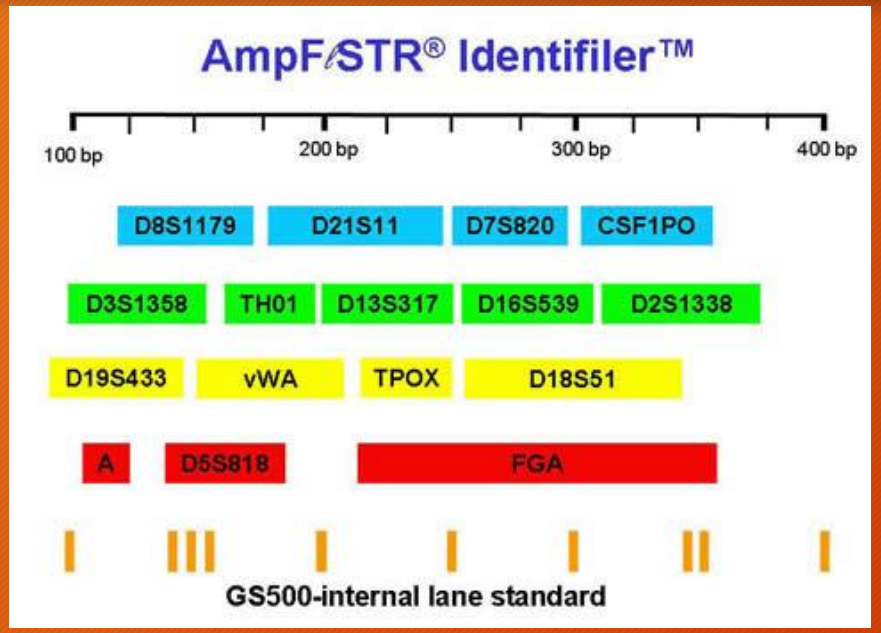
STRmix™ Software

Commercial probabilistic software for establishing the probability that an individual is included or excluded in a DNA mixture

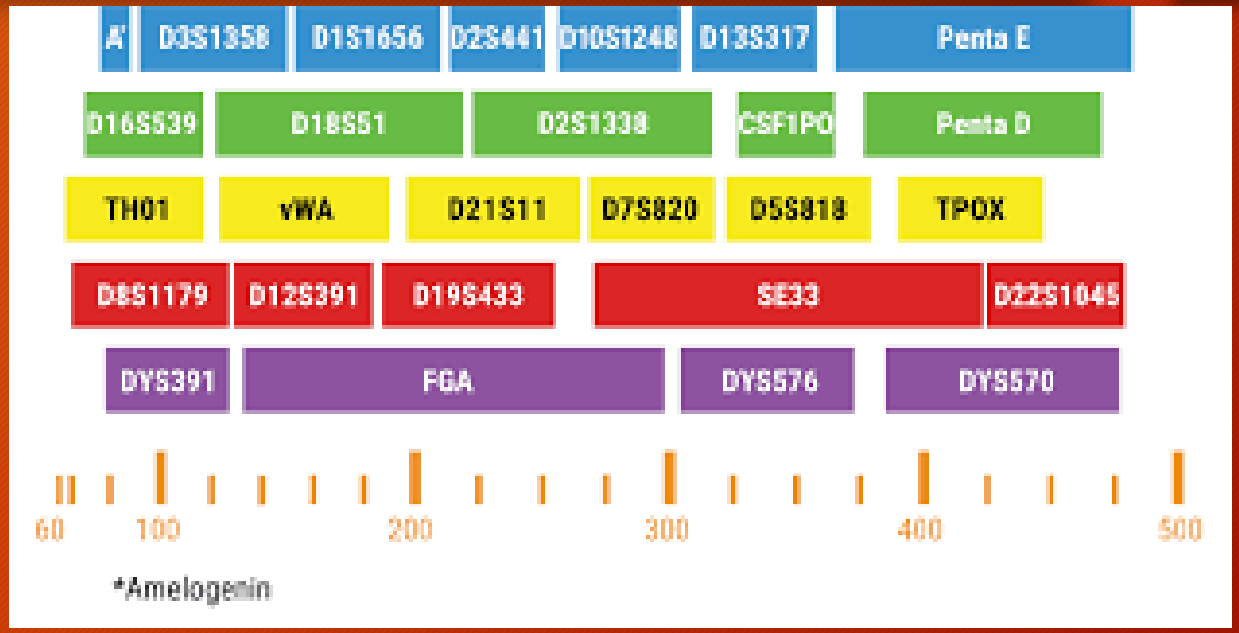
STRmix™ Software

- uses analytical thresholds
- models dropout (missing data)
- adds peak heights and some modeling
- limited likelihood explanation of data
- requires calibration of key variables such as number of contributors
- solves easy mixtures quite well
- good CPI replacement
- used and accepted by many courts in the United States
- commercially available and testable

Identifiler v Powerplex Fusion STR Kits



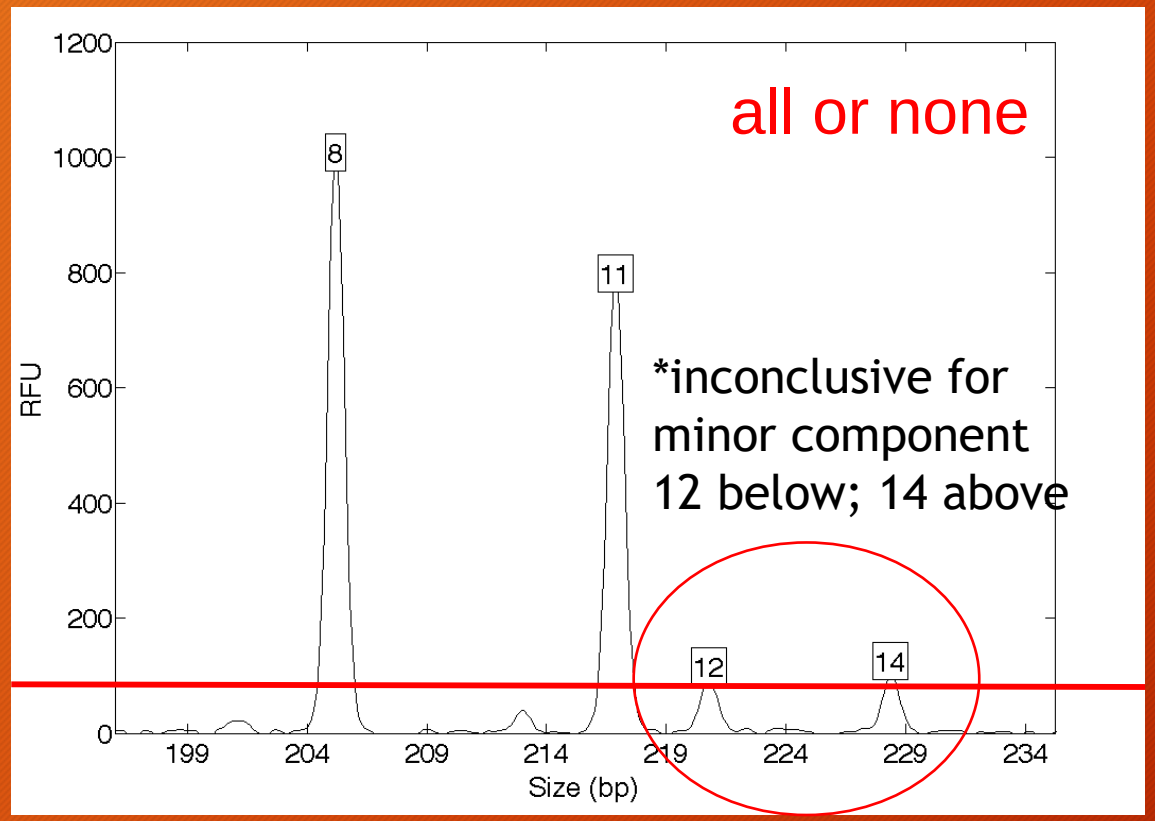
- 15 autosomal loci
- amelogenin sex typing locus
- 5 color multiplex for human identification



- 27 loci total
- 23 autosomal loci
- 3 Y-STRs
- amelogenin sex typing locus
- 6 color multiplex for human identification

Analytical Threshold & CPI

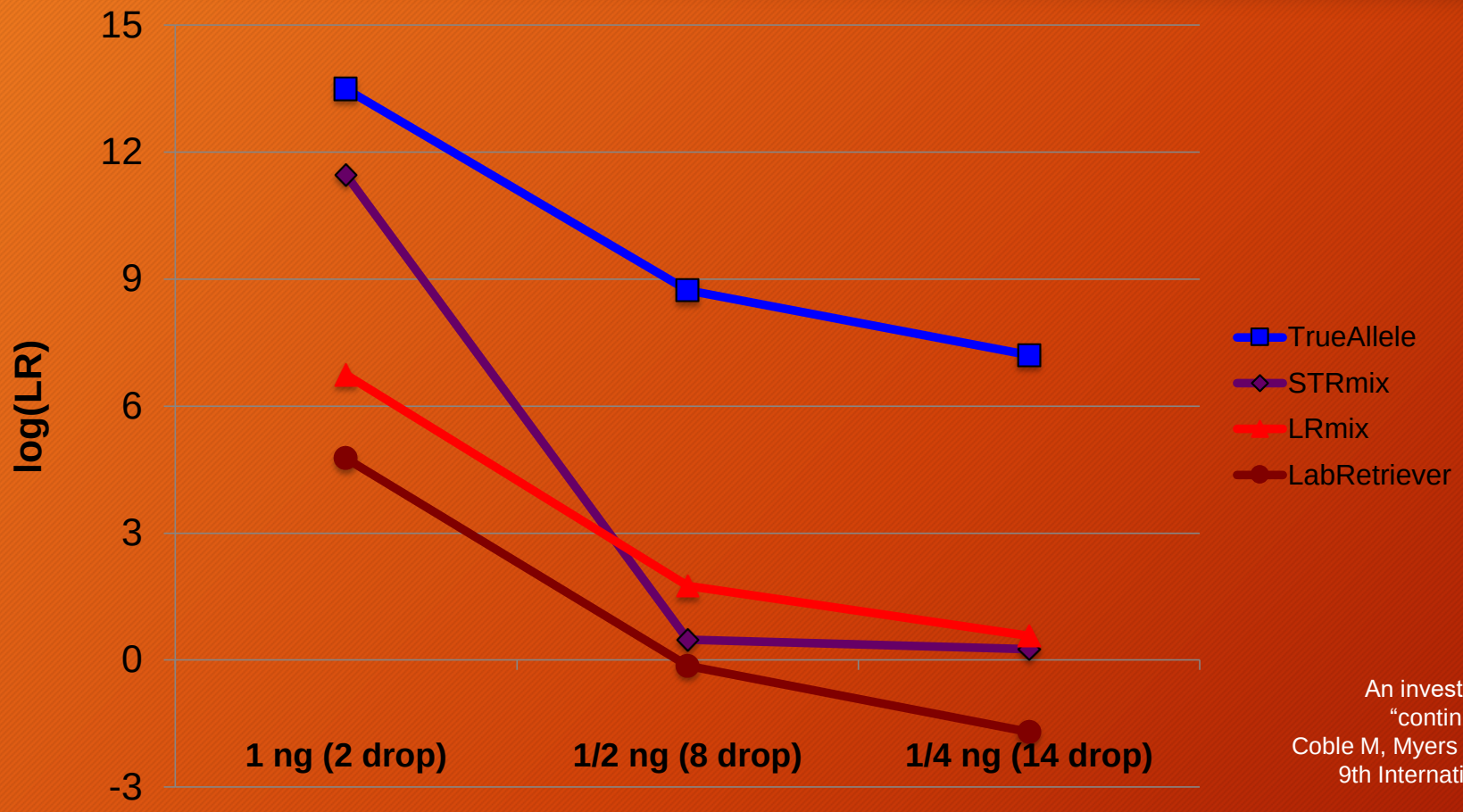
likelihood



1
2
3
4
5
6
7
8
9
10

- Analytical thresholds are laboratory dependent by policy
- Data that is below the analytical threshold is not included in the comparison to the evidence
- Data that is below the analytical threshold is not included in the match statistics
- Minor components can be inconclusive due to data drop-out at various loci (*)
- Laboratories vary as to level of disclosure of potential additional contributors in reports and testimony
- Laboratories vary as to ease of acquiring electronic data to re-analyze for number of contributors

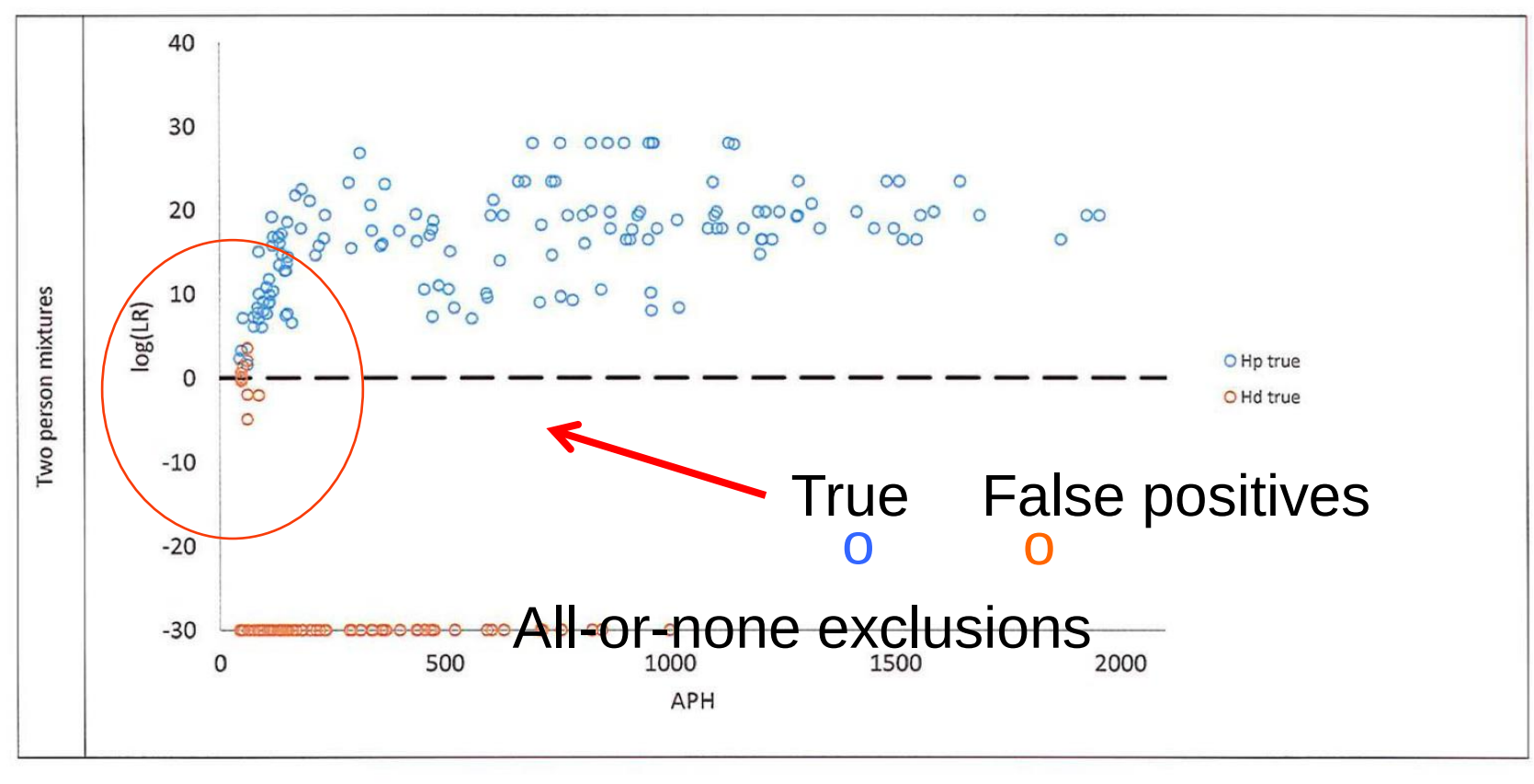
STRmix Threshold Impact – Software Comparison



- Analytical threshold 50 RFU
- Simple 2 person mixture
- With 10% minor contributor
- As DNA concentration decreases, so does likelihood ratio due to increasing loss of data
- TrueAllele uses no analytical threshold so the effect of data loss is less but also results in higher LR or inclusion if the defendant's DNA is present
- Scientific accuracy study

An investigation of software programs using "semi-continuous" and "continuous" methods for complex DNA mixture interpretation.
Coble M, Myers S, Klaver J, Kloosterman A, Leiden University, The Netherlands,
9th International Conference on Forensic Inference and Statistics, 2014.

Low-level Mixture Validation (2015, District of Columbia)



- How well does the software model the DNA mixture data?
- As with many probabilistic genotyping software programs, false positive LR values can occur in the 0-10 value range
- This is due to allele sharing and masking in unrelated individuals in the population (false positive error rate)
- Error also exists based on analytical threshold and drop-out modeling/data loss
- Laboratory policies calculate number of contributors without using data below the analytical threshold

A Comparison of FST and STRmix™ Software

FST

- Not commercially available
- "Black box software"
- Does not model peak heights
- Models "drop-out" incorrectly due to validation with pristine DNA (empirical study)
- Does not account for stutter
- Uses analytical thresholds
- Old technology eliminated 2017
- Denied in Frye admissibility hearing in combination with LCN (low copy number) tests

STRmix™

- Commercially available
- Models peak heights
- Models "drop-out"
- Accounts for stutter
- Uses analytical thresholds
- Accepted technology
- Utilized in many states including ME, NY, TX, MI
- Denied in NY for an individual case; NY application was not validated by the laboratory prior to court (*People of the State of New York v. Oral Hillary*)

STRmix™ Software

- Be aware number of contributors to the DNA mixture should be checked based on data below the analytical threshold for full disclosure
- Be aware that the statistical calculations should be made with the new amended allele frequency databases of the Federal Bureau of Investigation (FBI)
- Be aware that different hypotheses can be modeled by the software, not just the one presented in the case file
- Be aware that the number of contributors in the likelihood ratio do not need to be the same in the numerator (Hp) and the denominator (Hd)
- The simplest manner to look at the variance due to different hypothesis building is to run multiple analyses (for example, analyze with and without an analytical threshold; analyze with multiple contributors if there is data below the analytical threshold)
- Regardless of the likelihood ratio statistics, context of DNA to the case can be defensible
- Regardless of the likelihood ratio statistics, genetic relatives and high degree of allele sharing cannot be modeled effectively
- Regardless of the likelihood ratio statistics, software models are based on input parameters and the assessment for those parameters must be correct for accurate results