

Genotype Information Criteria for Forensic DNA Databases

2019 AAFS Annual Scientific Meeting
February 23, 2018

CODIS Searches

- Allele to Allele Comparison
- High certainty genotypes - good
- Low certainty genotypes?

NDIS Upload Criteria

Specimen Category	Max Alleles per Locus	Moderate Match Estimate
Forensic Mixture	4	1 in 10 million

Falling Through the Cracks

- What about mixtures that have high information but do not meet upload criteria?
- How do you identify mixtures with high match information?

Probabilistic Genotyping

Kullback-Leibler Divergence (KL)

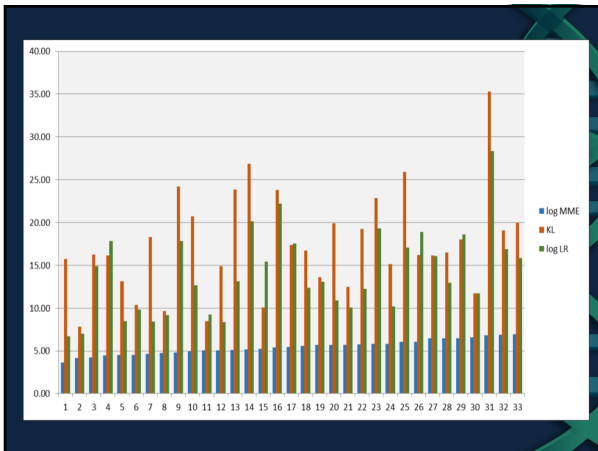
- Measure of information gain in Bayesian inference
- Posterior (after computer inference) vs. prior (before computer inference)

Advantages of KL for Databases

- Estimate of match information (LR) in absence of reference sample
- Calculated over entire profile, not a subset of loci
- Can be used to identify high information mixtures

Methods

1. Process all samples in TrueAllele®
2. Assess KL of inferred genotypes
3. Determine viability for CODIS
 - KL <10 = typically insufficient for CODIS
 - KL ≥10 = sufficient but specimen category varies



Case Example

- Assault case – female victim, male suspect
- Three questioned swabs
- No victim sample submitted

Case Example

Item	Quantity	Amplified?	Result	Weight
#1	~300 pg	Yes	Mixture (2)	57 / 43
#2	~170 pg	Yes	Minimal	N/A
#3	~10 pg	No	N/A	N/A

Case Example

- Examiner focused on Item 1 for CODIS entry
- Manual data review – 14 loci (8 core) entered
- Insufficient MME for NDIS; uploaded to SDIS

SDIS Search #1

- June 2017
 - 28 candidate matches
 - Human review (>1 hour)
 - 25 eliminated
 - 3 candidates requested for additional testing (expanded loci)

SDIS Search #1

- July 2017
 - Expanded loci testing complete
 - Human review
 - All three excluded
- Summary – 28 candidates, 0 hits, ~1.5 hours human time (entry & review)

Additional Testing

- October 2017
 - Victim standard submitted
 - TrueAllele® victim match, log LR 14.8 (630 trillion)
 - Examiner manually refines CODIS profile
 - MME still insufficient for NDIS
 - Uploads to SDIS

SDIS Search #2

- October
 - 2 new candidate matches
 - One eliminated
 - One identified and confirmed
- Summary – 2 candidates, 1 hits, ~15 minutes human time

SDIS Searches Summary

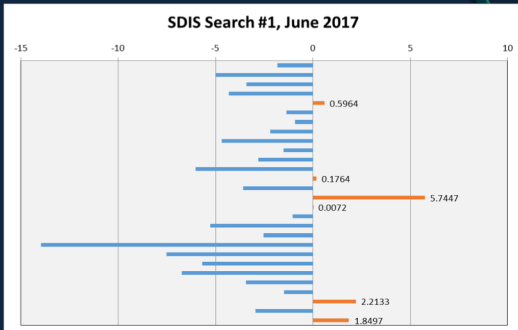
- 30 candidates
- 1 hit
- ~2 hours human analysis/review
- Four months searching

- Offender was not identified in initial search!

Could We Improve?

1. Use TrueAllele® at candidate match review?

TrueAllele® Match Review



Could We Improve?

2. Use TrueAllele® to select alleles for CODIS entry?

TrueAllele® for CODIS Entry

Method	Candidate Matches	Searches	Victim Sample Needed for Hit?	Review Time
Human	30	2	Yes	~2 hrs
TrueAllele®	3	1	No	~15 minutes

TrueAllele® for CODIS Entry

- SC has created Probabilistic Genotyping index
- No MME, max of 8 alleles per locus
- Minimum 7 original core loci

TrueAllele® for CODIS Entry

- Search against single source profiles only
- Mixtures that were previously ineligible for searching
- Several hits to date

Summary

- MME is good predictor for number of matches but not quality of data
- KL is good predictor of match strength
- High information mixture data is not being searched
- Probabilistic mechanism is needed

Thank you!
