

# Scientific Validation of Mixture Interpretation Methods

17th International Symposium on Human Identification  
Sponsored by the Promega Corporation  
October, 2006

Mark W Perlin, PhD, MD, PhD  
Pittsburgh, PA USA  
perlin@cybgen.com



Cybergenetics

Cybergenetics © 2003-2006

---

---

---

---

---

---

---

---

## Admissible Evidence

**FRE 702: reliable**

Data  
Method  
Application

**Frye (1923)**  
General acceptance

**Daubert (1993)**  
General acceptance  
Testable  
Error rate  
Peer review  
Standards

---

---

---

---

---

---

---

---

## DNA Mixture Interpretation

different labs, different methods

same data, different profiles

concordance study not possible –  
no direct profile comparison

need general validation approach –  
for Frye & Daubert requirements

---

---

---

---

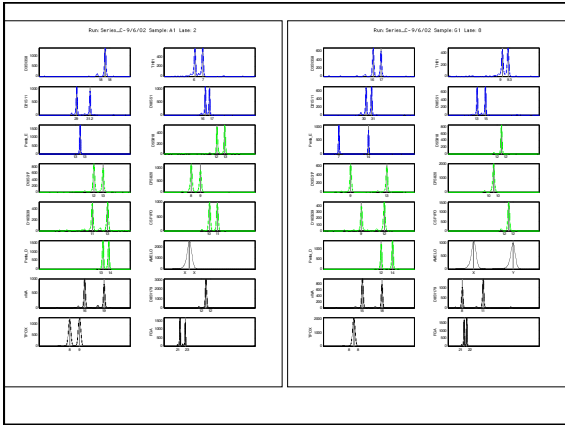
---

---

---

---






---

---

---

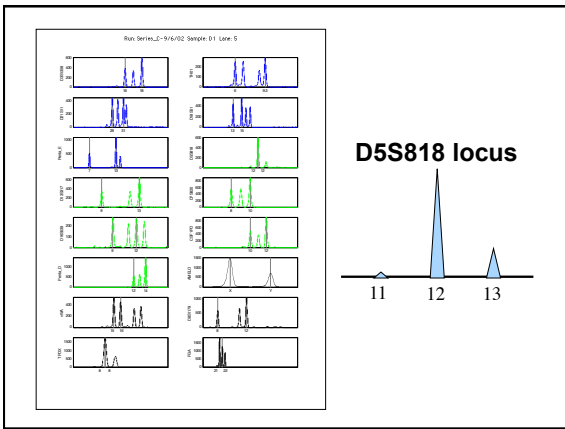
---

---

---

---

---




---

---

---

---

---

---

---

---

**Compare 5 Mixture Reviews**  
 Same data, different interpretations

- **Government**  
 Reviewer A  
 Reviewer B
- **Private Lab**  
 Reviewer A  
 Reviewer B
- **Computer**  
 TrueAllele Review

---

---

---

---

---

---

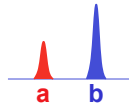
---

---

### Conservative: Government Review

Avoid overcalling the results

uncertain data



Allele 1: b  
Allele 2: anything

Report 0, 1 or 2 alleles

---

---

---

---

---

---

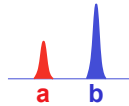
---

---

### Aggressive: Private Lab Review

Try ruling out unlikely combinations

uncertain data



Allele calls:  
1. a b  
2. b b

Report list of allele calls

---

---

---

---

---

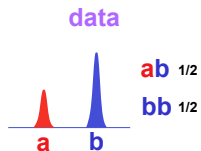
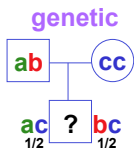
---

---

---

### Objective: TrueAllele Computer Review

Preserve match information



Genotypes with probability

---

---

---

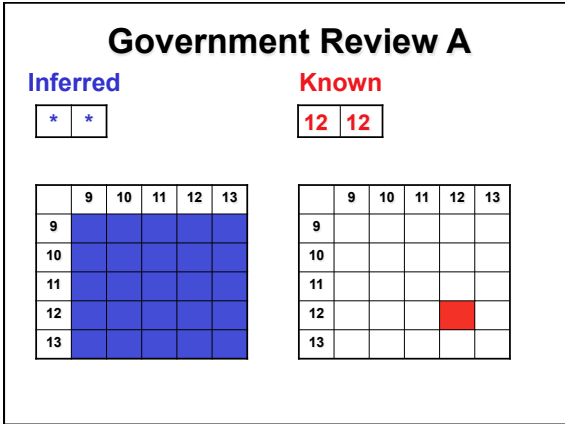
---

---

---

---

---




---

---

---

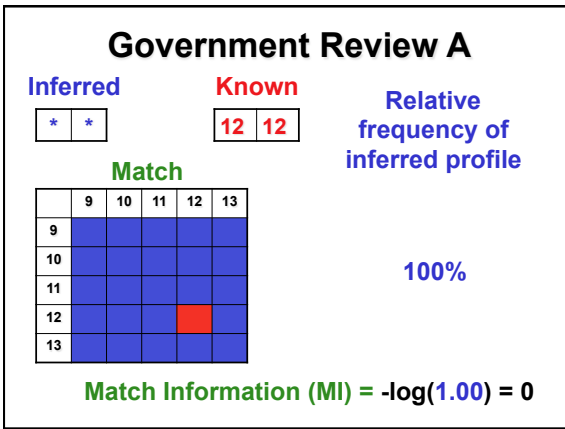
---

---

---

---

---




---

---

---

---

---

---

---

---

### Match Information

$$\text{Match LR} = \frac{\text{Prob}(\text{random match})}{\text{Prob}(\text{specific match})}$$
 ~ random match probability  
 ~ relative frequency

**Compare against the true profile**

**Logarithm measures information**  
**match information (MI) =  $-\log(\text{LR})$**

---

---

---

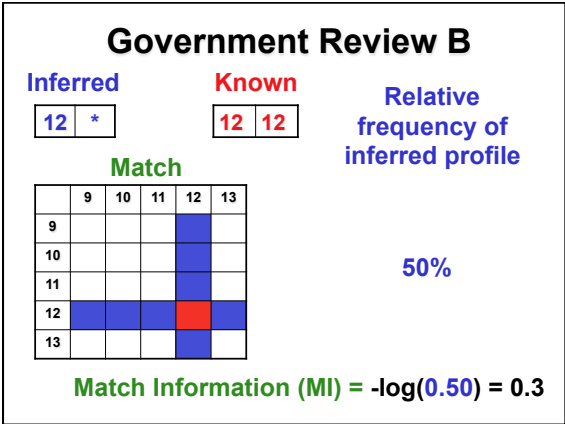
---

---

---

---

---




---

---

---

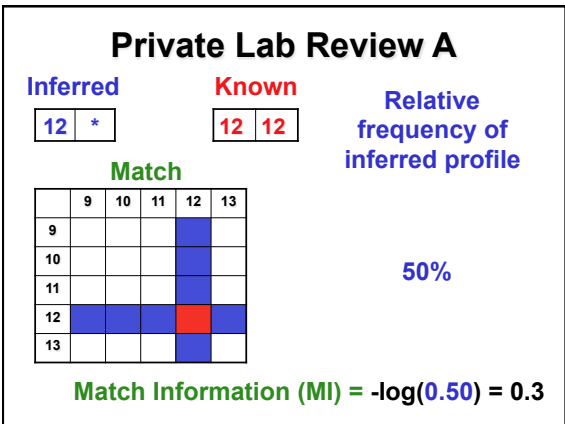
---

---

---

---

---




---

---

---

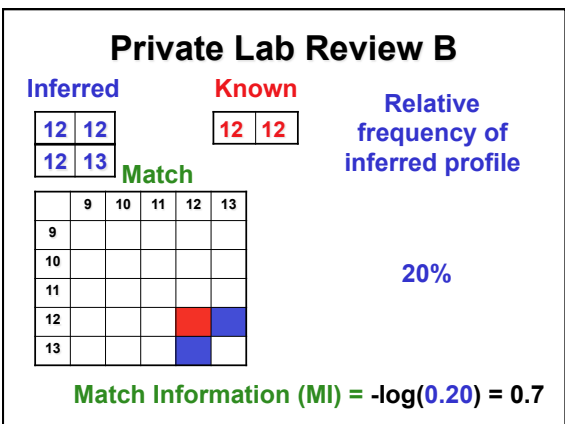
---

---

---

---

---




---

---

---

---

---

---

---

---

## TrueAllele® Computer Review

**Inferred**  
12 12
**Known**  
12 12
**Relative frequency of inferred profile**

**Match**

	9	10	11	12	13
9					
10					
11					
12					
13					

10%

Match Information (MI) =  $-\log(0.10) = 1$

---

---

---

---

---

---

---

---

## Profile Match Information

Add up the match information (MI) from every locus in the profile

$$MI_{profile} = \sum_{locus} MI_{locus}$$


---

---

---

---

---

---

---

---

The screenshot shows the TrueAllele System 3 web interface. The main content area displays a match strength report for a profile with a strength of 17.501. The report includes a table with columns for source, contrib, target, case, label, and strength. The source is 'CM', contrib is '2', target is 'NIST', case is 'G', and label is '17.501'. Below the table, there is a list of loci and their corresponding strengths, such as AMELO (0.305), CSF1PO (1.048), D13S317 (1.799), D16S539 (1.060), D18S51 (1.540), D21S11 (1.462), D3S1358 (0.948), D5S818 (0.928), D7S820 (1.038), D8S1179 (2.918), FGA (1.291), TH01 (1.232), TPOX (0.758), and vWA (1.260). The interface also features a sidebar with navigation options like User, Report, and Technology, and a top navigation bar with tabs for report, search, select, and review.

---

---

---

---

---

---

---

---

## Scientific Validation of Mixture Interpretation

One mixture, four different answers

- Accuracy
- Precision
- Reproducibility

---

---

---

---

---

---

---

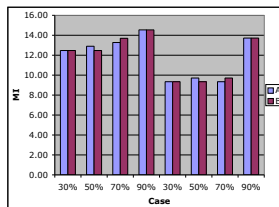
---

### Accuracy



Digging in the right place?

Is it the true profile?




---

---

---

---

---

---

---

---

### Accuracy



Digging in the right place?

Is it the true profile?

average

$$\frac{1}{n} \sum_{i=1}^n MI_i$$

case	A	B
30%	12.47	12.47
50%	12.90	12.47
70%	13.27	13.69
90%	14.54	14.54
30%	9.34	9.34
50%	9.71	9.34
70%	9.34	9.71
90%	13.72	13.72
Average	11.91	

---

---

---

---

---

---

---

---

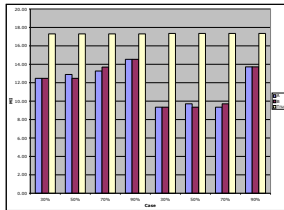


## Precision



How deep can we go?

How close to true profile?




---

---

---

---

---

---

---

---

## Precision



How deep can we go?

How close to true profile?

average difference

$$\frac{1}{n} \sum_{i=1}^n (MI_i - MI_i^{true})$$

case	A	B	True	A-True	B-True
30%	12.47	12.47	17.30	-4.83	-4.83
50%	12.90	12.47	17.30	-4.40	-4.83
70%	13.27	13.69	17.30	-4.03	-3.61
90%	14.54	14.54	17.30	-2.76	-2.76

Average difference -5.41

---

---

---

---

---

---

---

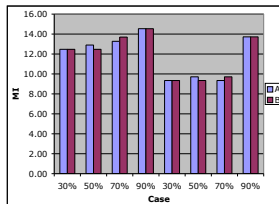
---

## Reproducibility



Do it again the same way?

How repeatable is our answer?




---

---

---

---

---

---

---

---

## Reproducibility



Do it again the same way?

How repeatable is our answer?

standard deviation

$$\frac{1}{N} \sum_{\text{case}=1}^m \sum_{i=1}^{n_{\text{case}}} (MI_{\text{case},i} - \overline{MI}_{\text{case}})^2$$

case	A	B	m	(A-m) <sup>2</sup>	(B-m) <sup>2</sup>
30%	12.47	12.47	12.47	0.000	0.000
50%	12.90	12.47	12.68	0.045	0.045
70%	13.27	13.69	13.48	0.045	0.045
90%	14.54	14.54	14.54	0.000	0.000
30%	9.34	9.34	9.34	0.000	0.000
50%	9.71	9.34	9.53	0.034	0.034
70%	9.34	9.71	9.53	0.034	0.034
90%	13.72	13.72	13.72	0.000	0.000
sum of squares			SS	0.314	
average SS			SS/N	0.0197	
standard deviation			sqrt(SS/N)	0.1402	

---

---

---

---

---

---

---

---

---

---

## Validation Results

	Accuracy	Precision	Reproducibility
Government	11.91	-5.41	0.140

---

---

---

---

---

---

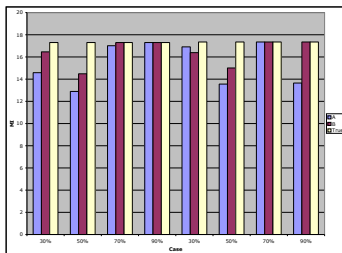
---

---

---

---

## Private Lab Data




---

---

---

---

---

---

---

---

---

---

## Validation Results

	<u>Accuracy</u>	<u>Precision</u>	<u>Reproducibility</u>
<b>Government</b>	11.91	-5.41	0.140
<b>Private Lab</b>	15.93	-1.39	0.833

Tradeoff?

---

---

---

---

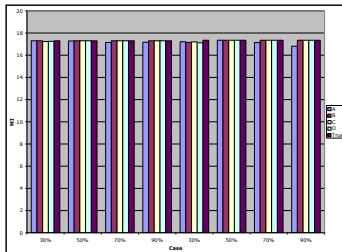
---

---

---

---

## TrueAllele Data




---

---

---

---

---

---

---

---

## Validation Results

	<u>Accuracy</u>	<u>Precision</u>	<u>Reproducibility</u>
<b>Government</b>	11.91	-5.41	0.140
<b>Private Lab</b>	15.93	-1.39	0.833
<b>TrueAllele</b>	17.24	-0.09	0.095

---

---

---

---

---

---

---

---

## Scientific Validation of Mixture Interpretation Methods

- law requires admissible evidence
- mixture interpretation not validated
- *match information* (MI) comparison
- simple spreadsheet calculation – accuracy, precision, reproducibility



Cybergenetics

---

---

---

---

---

---

---

---