Slide 1

Virginia TrueAllele® Validation Study: Casework Comparison

Presented at AAFS, February, 2013
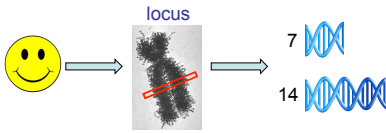**Published in PLOS ONE, March, 2014**

Mark W Perlin, PhD, MD, PhD
Kiersten Dormer, MS and Jennifer Hornyak, MS
Cybergenetics, Pittsburgh, PA
Lisa Schiermeier-Wood, MS and Susan Greenspoon, PhD
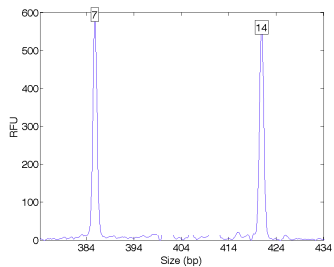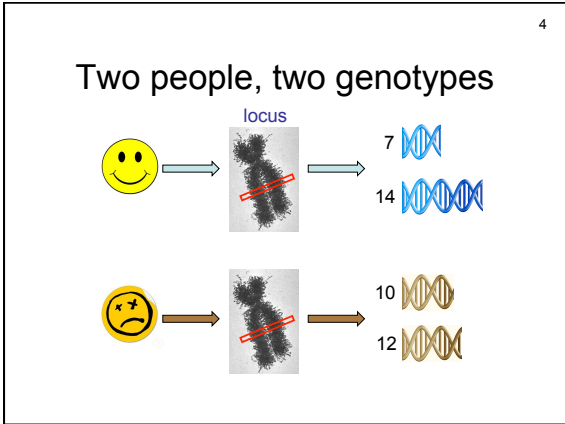Department of Forensic Science, Richmond, VA

Cybergenetics
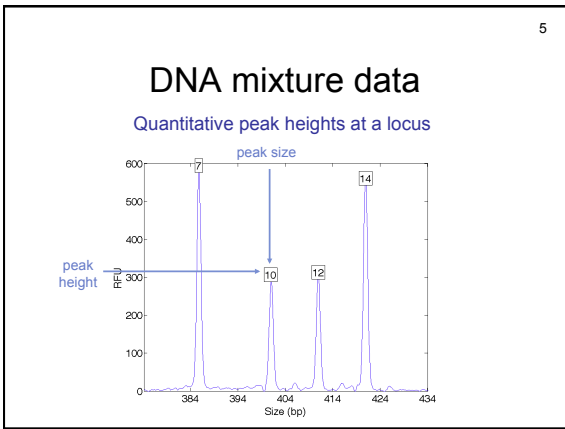

Slide 2

One person, one genotype

locus

7

14

Slide 3

DNA data

One or two allele peaks at a locus

Two people, two genotypes

locus

7

14

10

12

DNA mixture data

Quantitative peak heights at a locus

peak size

peak height

7

14

10

12

RFU

384   394   404   414   424   434

Size (bp)

Data summary – "alleles"

Over threshold, peaks are labeled as allele events

All-or-none allele peaks, each given equal status

Threshold

7

14

10

12

RFU

384   394   404   414   424   434

Size (bp)

Allele Pair
7, 7
7, 10
7, 12
7, 14
10, 10
10%10, 12
10, 14
12, 12
12, 14
14, 14

CPI information

6.83 *(2.22)*
6.68 million

CPI

Combined probability of inclusion

SWGDAM 2010 guidelines

Higher threshold for human review

Under threshold, alleles less used

Threshold

Allele Pair
7, 7
7, 10
7, 12
7, 14
10, 10
0% 10, 12
10, 14
12, 12
12, 14
14, 14

Modified CPI information

6.83 *(2.22)*
6.68 million

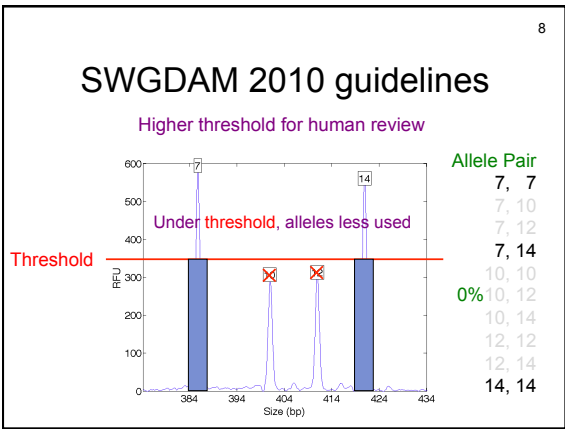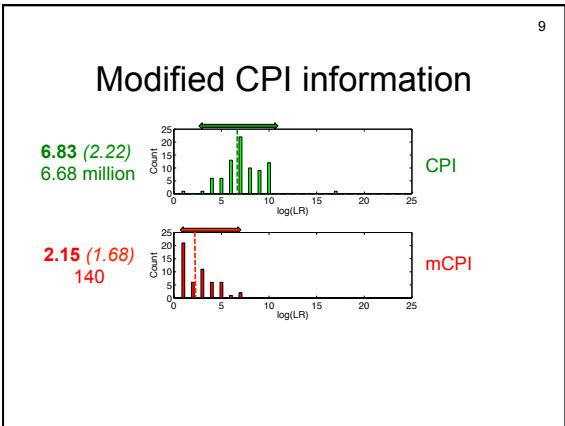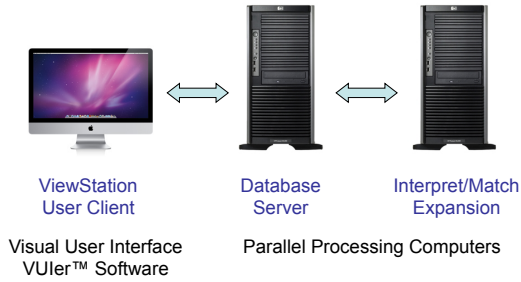CPI

2.15 *(1.68)*
140

mCPI

SWGDAM 2010 guidelines

3.2.2. If a stochastic threshold based on peak height is not used in the evaluation of DNA typing results, the laboratory must establish alternative criteria (e.g., quantitation values or use of a probabilistic genotype approach) for addressing potential stochastic amplification. The criteria must be supported by empirical data and internal validation and must be documented in the standard operating procedures.

Use TrueAllele® Casework for DNA mixture statistics

TrueAllele Casework

ViewStation
User Client

Database
Server

Interpret/Match
Expansion

Visual User Interface
VUIer™ Software

Parallel Processing Computers

Validated genotyping method

Perlin MW, Sinelnikov A. An information gap in DNA evidence interpretation. *PLoS ONE*. 2009;4(12):e8327.

Ballantyne J, Hanson EK, Perlin MW. DNA mixture genotyping by probabilistic computer interpretation of binomially-sampled laser captured cell populations: Combining quantitative data for greater identification information. *Science & Justice*. 2013;53(2):103-114.

Perlin MW, Hornyak J, Sugimoto G, Miller K. TrueAllele® genotype identification on DNA mixtures containing up to five unknown contributors. *Journal of Forensic Sciences*. 2015;60(4):857-868.

Greenspoon SA, Schiermeier-Wood L, Jenkins BC. Establishing the limits of TrueAllele® Casework: a validation study. *Journal of Forensic Sciences*. 2015;60(5):1263-1276.

Perlin MW, Legler MM, Spencer CE, Smith JL, Allan WP, Belrose JL, Duceman BW. Validating TrueAllele® DNA mixture interpretation. *Journal of Forensic Sciences*. 2011;56(6):1430-1447.

Perlin MW, Belrose JL, Duceman BW. New York State TrueAllele® Casework validation study. *Journal of Forensic Sciences*. 2013;58(6):1458-1466.

Perlin MW, Dormer K, Hornyak J, Schiermeier-Wood L, Greenspoon S. TrueAllele® Casework on Virginia DNA mixture evidence: computer and manual interpretation in 72 reported criminal cases. *PLOS ONE*. 2014;(9)3:e92837.

TrueAllele reinterpretation

The Washington Post

Virginia reevaluates DNA evidence in 375 cases
July 16, 2011

"Mixture cases are their own little nightmare," says
William Vosburgh, director of the D.C. police's crime
lab. "It gets really tricky in a hurry."

"If you show 10 colleagues a mixture,
you will probably end up with 10 different answers"
Dr. Peter Gill, Human Identification E-Symposium, 2005
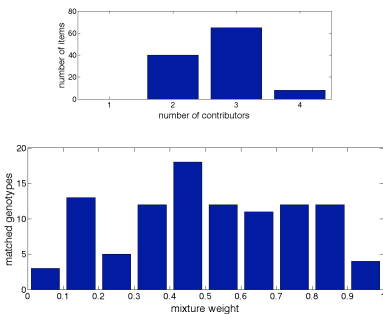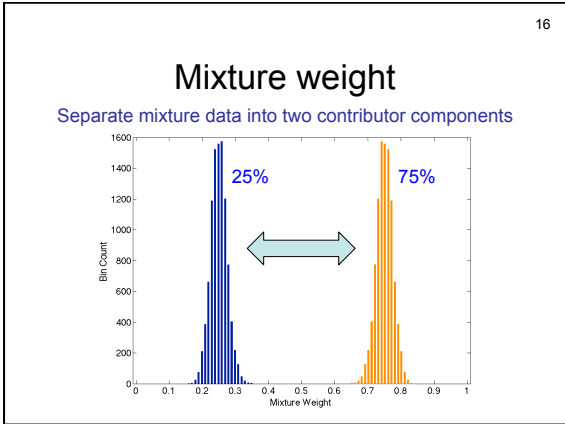

Report mixture statistics

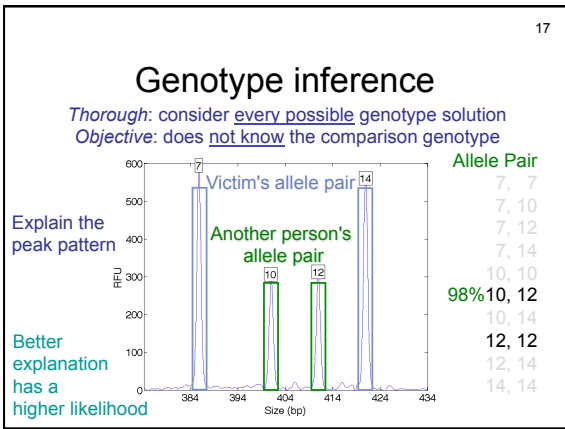- 72 criminal cases
- 92 evidence items
- 111 genotype comparisons

<u>Criminal offense</u>
- 18 homicide
- 12 robbery
- 6 sexual assault
- 20 weapon

DNA mixture distribution

Mixture weight

Separate mixture data into two contributor components

25%   75%

Genotype inference

Thorough: consider <u>every possible</u> genotype solution
Objective: does <u>not know</u> the comparison genotype

Allele Pair

Explain the
peak pattern

Victim's allele pair

Another person's
allele pair

Better
explanation
has a
higher likelihood

7,   7
7, 10
7, 12
7, 14
10, 10
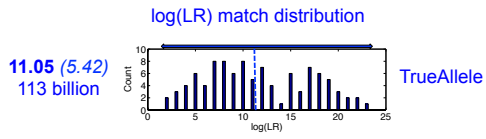98% 10, 12
10, 14
12, 12
12, 14
14, 14

Sensitivity

The extent to which interpretation
identifies the correct person

True DNA mixture inclusions

101 reported genotype matches
82 with DNA statistic over a million

TrueAllele sensitivity

log(LR) match distribution
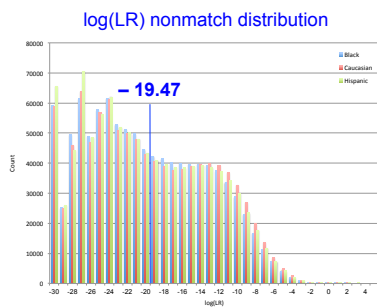
11.05 *(5.42)*
113 billion

TrueAllele

Specificity

The extent to which interpretation does
not misidentify the wrong person

True exclusions, without false inclusions

101 matching genotypes x 10,000 random references
x 3 ethnic populations,
for over 1,000,000 nonmatching comparisons

TrueAllele specificity

log(LR) nonmatch distribution

− 19.47

False positives

in over 1,000,000 comparisons per group

Tail distribution	Black	Caucasian	Hispanic
0	39	32	29
1	8	11	9
2	2	1	1
3	0	0	1
log(LR) > 0	49	44	40

false positive rate is under 1 in 20,000 (0.005%)
for LR > 100, rate is 1 in 1,000,000 (0.0001)%
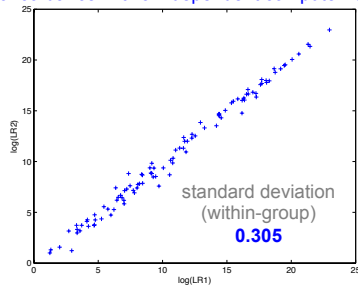

Reproducibility

The extent to which interpretation gives
the same answer to the same question
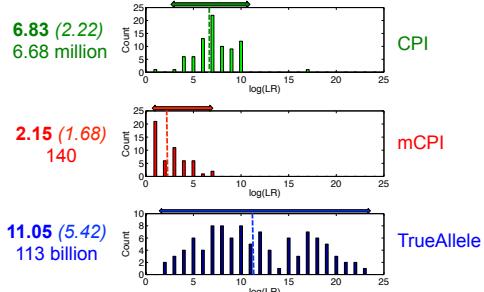
MCMC computing has sampling variation

duplicate computer runs
on 101 matching genotypes
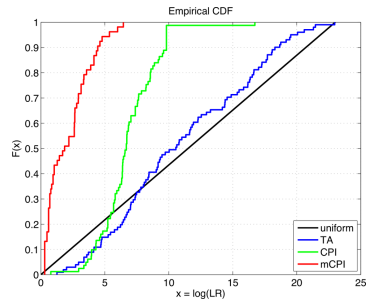
measure log(LR) variation

TrueAllele reproducibility

Concordance in two independent computer runs

standard deviation
(within-group)
0.305

Cybergenetics © 2003-2017

Comparison

6.83 *(2.22)*
6.68 million — CPI

2.15 *(1.68)*
140 — mCPI

11.05 *(5.42)*
113 billion — TrueAllele

(Count vs log(LR) histograms)

Accuracy

Empirical CDF

F(x) vs x = log(LR)

Legend: uniform, TA, CPI, mCPI

Conservative results

Five matches, TrueAllele less than CPI.
Ten comparisons, no statistical support:

Interpretation Method			Data Observations					
TrueAllele	CPI	mCPI	allele dropout	allele overlap	low peaks	peak imbalance	infeasible mixture	infeasible pattern
-10.64			3	4	1			1
-6.52			4	3	1			1
-5.05			4	3	1	1		1
-4.87				3		1	1	1
-4.86	3.48		4		1		1	
-3.22	6.04	6.34		2			1	1
-2.99	4.23		2		1		1	1
-2.18			2		1		1	
-1.41	4.08		1		1		1	
-0.67	2.95	0.60	1	2	1			

TrueAllele Virginia outcomes

144 cases analyzed
72 case reports – 10 trials

City	Court	Charge	Sentence
Richmond	Federal	Weapon	50 years
Alexandria	Federal	Bank robbery	90 years
Quantico	Military	Rape	3 years
Chesapeake	State	Robbery	26 years
Arlington	State	Molestation	22 years
Richmond	State	Homicide	35 years
Fairfax	State	Abduction	33 years
Norfolk	State	Homicide	8 years
Charlottesville	State	Homicide	15 years
Hampton	State	Home invasion	5 years

TrueAllele today

Invented math & algorithms	20 years
Developed computer systems	15 years
Support users and workflow	10 laboratories
Routinely used in casework	7 crime labs
Validate system reliability	34 studies
Educate the community	50 talks
Train & certify analysts	200 students
Go to court for admissibility	10 rulings
Testify about LR results	50 trials
Educate lawyers and laymen	1,000 people
Make the ideas understandable	500 cases, 37 states

Conclusions

TrueAllele Casework DNA mixture interpretation is:

A reliable method
• objective
• sensitive
• specific
• reproducible
• accurate

TrueAllele computer genotyping is
more effective than human review