The Probative Power of DNA Mixtures

The Balancing Act of Justice
Louisiana District Attorneys Association
LDAA 39th Annual Conference
June, 2016
Destin, FL

Mark W Perlin, PhD, MD, PhD
Cybergenetics, Pittsburgh, PA

Cybergenetics

Cybergenetics © 2003-2016

Booker Diggins stays in prison

1987 – aggravated rape, armed robbery
    Booker Diggins sentenced to life

2010 – "bulletproof scientific evidence that he is not the guy"

2012 – rape kit found, tested by crime lab (insufficient DNA)
    Cybergenetics finds match statistic of 38 thousand

New DNA test points to New Orleans rape
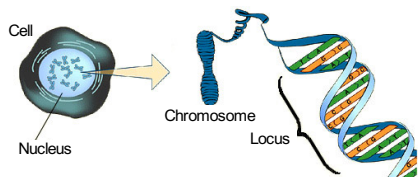convict who was nearly set free
The Times-Picayune

DNA biology

Cell

Nucleus

Chromosome
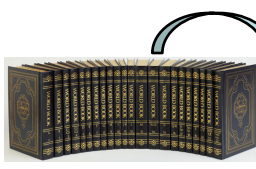
Locus

Short tandem repeat

DNA locus paragraph

Take me out to the ball game
take me out with the crowd
buy me some peanuts and Cracker Jack
I don't care if I never get back
let me
root root root root root root root root root root
for the home team,
if they don't win, it's a shame for it's one, two,
three strikes, you're out
at the old ball game

23 volumes in
cell's
DNA encyclopedia

"root" repeated 10 times, so
allele length is 10 repeats
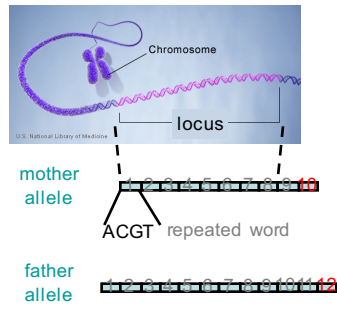

DNA genotype

Chromosome

locus

mother allele

ACGT   repeated word

father allele

A genetic locus has
two DNA sentences,
one from each parent.

An allele is the number
of repeated words.

A genotype at a locus
is a pair of alleles.

10, 12

Many alleles allow for
many many allele pairs.
A person's genotype
is relatively unique.

U.S. National Library of Medicine
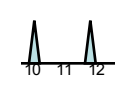

DNA laboratory

Evidence item   Lab →   Evidence data

DNA from one person

10 11 12

Separations
• Extract
• Amplify
• Detect

DNA interpretation

Evidence item →(Lab)→ Evidence data →(Infer)→ Evidence genotype

10, 12

DNA from one person

10 11 12

DNA match statistic

Evidence item →(Lab)→ Evidence data →(Infer)→ Evidence genotype

10, 12

DNA from one person

10 11 12

Compare

Known genotype

10, 12

$$\frac{1}{\text{Prob(coincidental match)}}$$
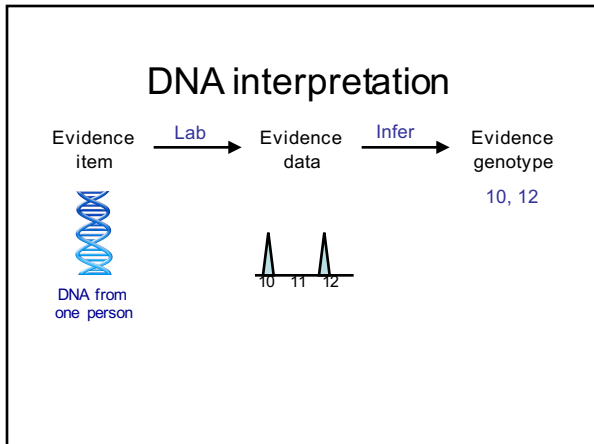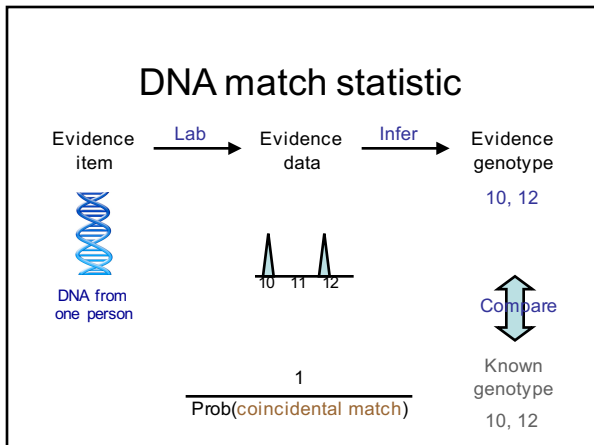
National Academy of Sciences

STRENGTHENING FORENSIC SCIENCE IN THE UNITED STATES — A PATH FORWARD

NATIONAL RESEARCH COUNCIL

"Strengthening Forensic Science: A Path Forward" (2009)

• Human examination bias
• Statistics & reporting
• Underlying scientific basis

DNA Gold Standard

However, ... there may be problems ... with how the DNA was ... interpreted, such as when there are mixed samples
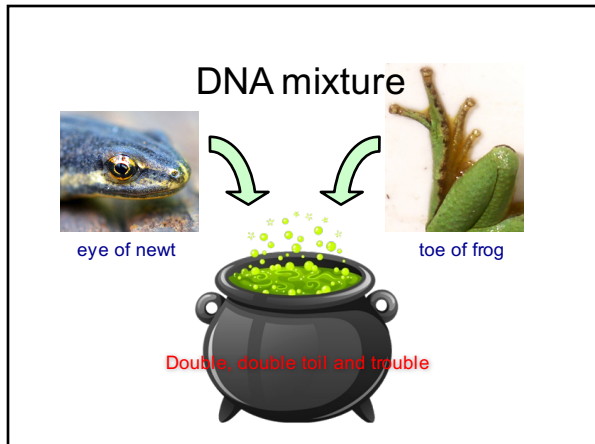
DNA mixture

eye of newt   toe of frog

Double, double toil and trouble

DNA statistic shuts down labs

"National accreditation board suspends all
DNA testing at D.C. crime lab"
The Washington Post *April 27, 2015*
Did not comply with FBI standards

"New protocol leads to reviews of
'mixed DNA' evidence"
The Texas Tribune *September 12, 2015*
24,468 lab tests affected

MIX05: NIST mixture interpretation interlaboratory study.
Butler JM, Kline MC, National Institute of Standards and Technology
Promega's Sixteenth International Symposium on Human Identification, 2005

MIX05: Statistics not reproducible

National Institute of Standards and Technology
Two Contributor Mixture Data, Known Victim

Some Differences in Reporting Statistics

LabID	Kits Used	Caucasians	African Americans	Hispanics
90	ProPlus/Cofiler	1.18E+15	2.13E+14	3.09E+15
34	ProPlus/Cofiler	2.40E+11	7.86E+09	9.80E+10
33	ProPlus/Cofiler	2.94E+08	1.12E+08	1.74E+09
6	ProPlus/Cofiler	40,000,000	3,500,000	280,000,000
9	ProPlus/Cofiler	1.14E+07	1.97E+07	1.54E+06
79	ProPlus/Cofiler	930,000	17,690	1,350,000
16	ProPlus/Cofiler	434,600	31,710	399,100

**Remember that these labs are interpreting
the same MIX05 electropherograms**

213 trillion (14)

31 thousand (4)

MIX13: An interlaboratory study on the present state of DNA mixture interpretation in the U.S.
Coble M, National Institute of Standards and Technology
5th Annual Prescription for Criminal Justice Forensics, Fordham University School of Law, 2014.

MIX13: Statistics falsely include

MIX13 Case 5 Outcomes with Suspect C
(whose genotypes were _not present in the mixture_)

# Labs	Report Conclusions	Reasons given
6	Exclude Suspect C	detailed genotype checks (ID+); TrueAllele negative LR (ID+); assumed major/minor and suspects did not fit (ID+); 3 labs noted Penta E missing allele 15 (PP16HS)
3	Inconclusive with C only (A & B included)	All these labs used PP16HS
21	Inconclusive for A, B, and C	
70	<u>**Include**</u> **& provide CPI statistics**	_All over the road…_

Range of CPI stats for Caucasian population:
FBI allele frequencies: **1 in 9** to **1 in 344,000**

CPI lacks probative value

J Pathol Inform

Research Article

Inclusion probability for DNA mixtures is a subjective one-sided match statistic unrelated to identification information

Mark William Perlin[1]

Forensic crime laboratories have generated CPI statistics on hundreds of thousands of DNA mixture evidence items. However, this commonly used match statistic behaves like a random generator of inclusionary values, following the LLN rather than measuring identification information. A quantitative CPI number adds little meaningful information beyond the analyst's initial qualitative assessment that a person's DNA is included in a mixture. Statistical methods for reporting on DNA mixture evidence should be scientifically validated before they are relied upon by criminal justice.

Subjective interpretation

(1)
Choose, alter, discard, edit, and manipulate the DNA **data signals**

(2)
Compare defendant's **genotype** to edited **data** & decide if he is in the DNA evidence

(3)
If he is "included", then **calculate** a DNA mixture statistic

• Human examination bias
• Statistics & reporting
• Underlying scientific basis

TrueAllele® Casework
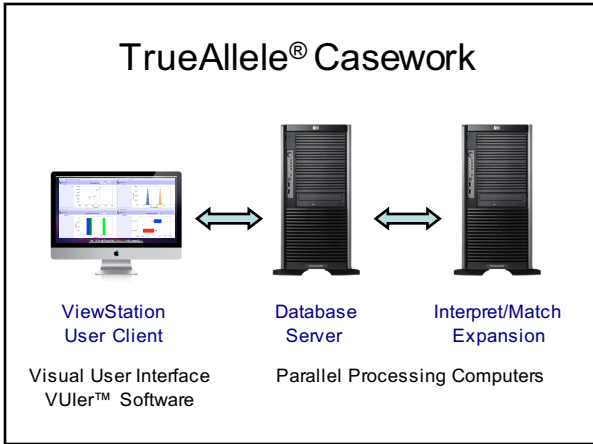
ViewStation
User Client

Database
Server

Interpret/Match
Expansion

Visual User Interface
VUIer™ Software

Parallel Processing Computers

DNA mixture data

Evidence
item

Lab →

Evidence
data

DNA from
two people

10 11 12

DNA mixture interpretation

Evidence
item

Lab →

Evidence
data

Separate →

Evidence
genotypes

10, 12

+

10, 11 @ 20%
11, 11 @ 30%
11, 12 @ 50%

DNA from
two people

10 11 12

DNA match statistic

Evidence item → **Lab** → Evidence data → **Separate** → Contributor genotype

10, 11 @ 20%
11, 11 @ 30%
11, 12 @ 50%

DNA from two people

10 11 12

Compare

Known genotype

11, 12

$$\frac{\text{Prob(evidence match)}}{\text{Prob(coincidental match)}}$$

Mixtures corroborate witnesses

Julia "Zoe" Anderson

Christopher Hutsell

Water bottle

Computers can use all the data

Quantitative peak heights at locus TPOX

peak size

peak height

7

8

11

10

RFU

100
80
60
40
20
0

223   228   233   238   243
Size (bp)

How the computer thinks

Consider every possible genotype solution

Explain the peak pattern

7

8

11

10

A third person's allele pair

One person's allele pair

A second person's allele pair

Better explanation has a higher likelihood

RFU

100 — 80 — 60 — 40 — 20 — 0

223 228 233 238 243
Size (bp)
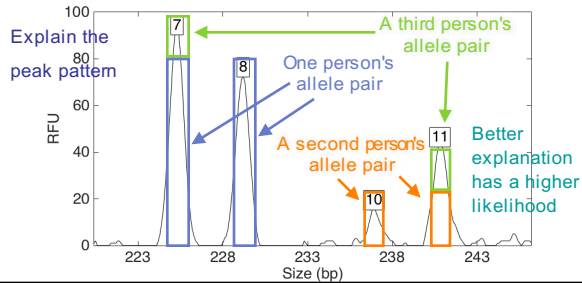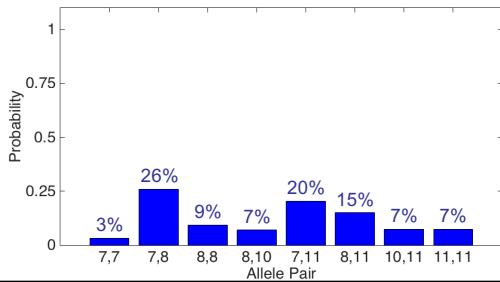
Evidence genotype

Objective genotype determined solely from the DNA data. Never sees a comparison reference.

Probability

1 — 0.75 — 0.5 — 0.25 — 0
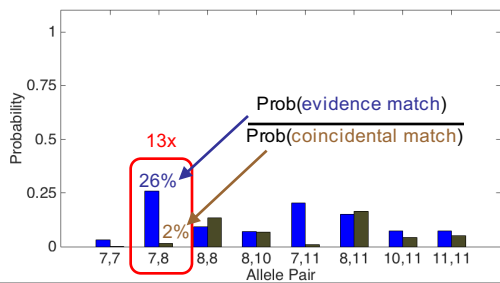
3% | 26% | 9% | 7% | 20% | 15% | 7% | 7%

7,7 7,8 8,8 8,10 7,11 8,11 10,11 11,11
Allele Pair

DNA match information

How much more does the suspect match the evidence than a random person?

Probability

1 — 0.75 — 0.5 — 0.25 — 0

$\dfrac{\text{Prob(evidence match)}}{\text{Prob(coincidental match)}}$

13x

26%

2%

7,7 7,8 8,8 8,10 7,11 8,11 10,11 11,11
Allele Pair

Match information at 15 loci

DNA Locus (top to bottom): CSF1PO, D13S317, D16S539, D18S51, D19S433, D21S11, D2S1338, D3S1358, D5S818, D7S820, D8S1179, FGA, TH01, TPOX, vWA

X-axis: Likelihood Ratio — 0, 10, 20, 30, 40, 50, 60, 70
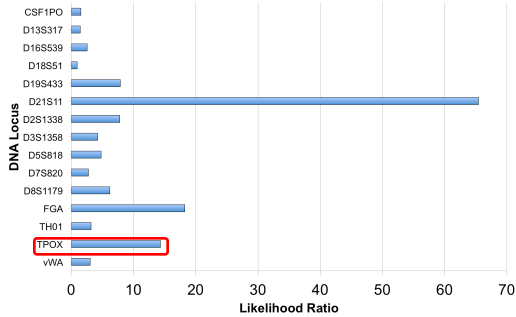
Is the suspect in the evidence?

A match between the water bottle
and Christopher Hutsell is:

19 billion times more probable than
a coincidental match to an unrelated African-American person

74.9 billion times more probable than
a coincidental match to an unrelated Caucasian person

127 billion times more probable than
a coincidental match to an unrelated Hispanic person

Is the victim in the evidence?

A match between the water bottle
and Julia Zoe Anderson is:

889 thousand times more probable than
a coincidental match to an unrelated African-American person

1.11 million times more probable than
a coincidental match to an unrelated Caucasian person

1.75 million times more probable than
a coincidental match to an unrelated Hispanic person

Life sentence

Jury convicts Christopher Hutsell in stabbing death of Indiana transient
The Times-Picayune

A New Orleans jury on Tuesday night (November 17, 2015) convicted Christopher Hutsell in the March 19 stabbing death of Julia "Zoe" Anderson. Hutsell, 38 years old and homeless, faces a mandatory sentence of life in prison on a second-degree murder charge after refusing to plead guilty to manslaughter in exchange for a 40-year sentence.

Mixtures reduce coincidence

Fence jumped as shooters fled the scene.
Handprints swabbed by Ville Platte police.
Crime lab unable to determine DNA match.
District Attorney's Office sent DNA data to
Cybergenetics for statistical analysis.
Ville Platte Gazette

DNA mixture data
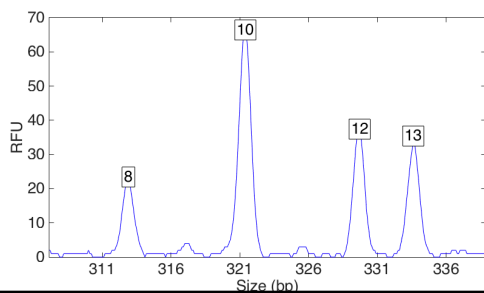
Quantitative peak heights at locus CSF1PO

Match statistics

Exhibit	Description	19 Hilton Wilson	20 Reokenski Thomas	21 Raven Gallow	27 Arlandus Richard
13	post - A	70.9 thousand	335 thousand		
14	post - B	23.1 million	6.28 million		
13 & 14	posts - A & B	157 million	340 million		

Match statistics

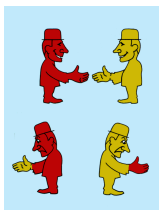| Exhibit	Description	19 Hilton Wilson	20 Reokenski Thomas	21 Raven Gallow	27 Arlandus Richard
13	post - A	4.85	5.53		
14	post - B	7.36	6.80		
13 & 14	posts - A & B	8.20	8.53		

DNA transfer

- How did DNA get there?
- When was DNA deposited?

Multiple mixtures reduce coincidence
Left & right hand mixtures
Both defendants present

Defense: Isn't DNA transfer a possibility?

Expert: Not very likely to happen twice.

Convictions obtained

Guilty verdicts handed down in Thomas-Wilson murder trial
Ville Platte Gazette

TrueAllele statistically matched the defendants' DNA to
the DNA collected from the fence at the crime scene.

After seven days of listening to eye witness and expert testimony,
an Evangeline Parish Jury of 12 found Reokinski Thomas and
Hilton Wilson guilty of first degree murder and attempted first
degree murder (March 16, 2016).

The prosecutors polled the jury after trial –
the DNA was crucial in obtaining the convictions

TrueAllele is reliable

On June 30, 2012, Darius Vicks was waiting for a traffic light to change.
Jordan Key was a passenger in the Baton Rouge car.
Chattley Chesterfield got out of a vehicle behind them, ran up to the car,
and fired three shots into the car, killing Key and injuring Darius Vicks
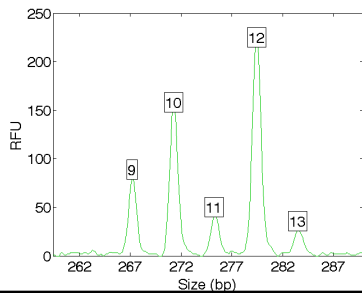
Crime lab: pistol DNA mixture of 3 people

DNA mixture data pattern

Quantitative peak heights at locus D16S539

Older methods use less data

Over threshold, peaks are labeled as allele events

Threshold

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

Under threshold, alleles vanish
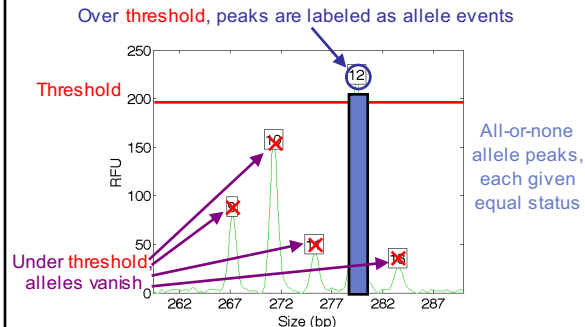
TrueAllele explains all the data

Consider many possible genotype answers

One person's allele pair

Second person's allele pair

Third person's allele pair
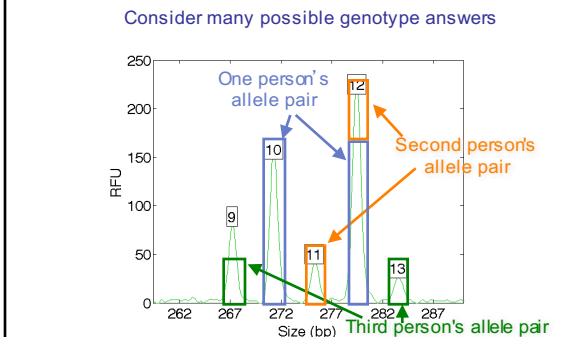
Are the suspects in the evidence?

A match between the **pistol**
and **Chattley Chesterfield** is:
108 thousand times more probable than
a coincidental match to an unrelated Black person

A match between the **pistol**
and **Samuel Nicholas** is:
3,210 times more probable than
a coincidental match to an unrelated Black person

Reliability (La. Rule 702)

Testimony by Experts

A witness who is qualified as an expert by knowledge, skill, experience, training, or education may testify in the form of an opinion or otherwise if:
(1) The expert's scientific, technical, or other specialized knowledge will help the trier of fact to understand the evidence or to determine a fact in issue;
(2) The testimony is based on sufficient facts or data;
(3) The testimony is the product of reliable principles and methods; and
(4) The expert has reliably applied the principles and methods to the facts of the case.

Daubert v. Merrell Dow Pharmaceuticals (1993)

Plaintiff: Bendectin caused birth defects
Defendant: no reliable scientific evidence

Judge as gatekeeper

(1) testable and tested
(2) peer review and publication
(3) known error rate
(4) standards and controls
(5) generally accepted in the relevant community

Peer-review process

Do research

Write paper

Journal editor

Independently & anonymously review paper

Publish paper

Peer-reviewed validation study

Mark W. Perlin,[1] Ph.D., M.D.; Jennifer M. Hornyak,[1] M.S.; Garett Sugimoto,[2] M.S.; and Kevin W.P. Miller,[2] Ph.D.

TrueAllele® Genotype Identification on DNA Mixtures Containing up to Five Unknown Contributors*

ABSTRACT: Computer methods have been developed for mathematically interpreting mixed and low-template DNA. The genotype modeling approach computationally separates out the contributors to a mixture, with uncertainty represented through probability. Comparison of inferred genotypes calculates a likelihood ratio (LR), which measures identification information. This study statistically examined the genotype modeling performance of Cybergenetics TrueAllele® computer system. High- and low-template DNA mixtures of known randomized composition containing 2, 3, 4, and 5 contributors were tested. Sensitivity, specificity, and reproducibility were established through LR quantification in each of these eight groups. Covariance analysis found LR behavior to be relatively invariant to DNA amount or contributor number. Analysis of variance found that consistent solutions were produced, once a sufficient number of contributors were considered. This study demonstrates the reliability of TrueAllele interpretation on complex DNA mixtures of representative casework composition. The results can help predict an information outcome for a DNA mixture analysis.

KEYWORDS: forensic science, DNA mixture, genotype modeling, validation study, likelihood ratio, probabilistic genotyping

---

TrueAllele validation papers

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):103-14.

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):1430-47.

Perlin MW, Belrose JL, Duceman BW. New York State TrueAllele® Casework validation study. *Journal of Forensic Sciences*. 2013;58(6):1458-66.

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.

Validation axes

Sensitivity. The extent to which interpretation identifies the correct person.
Truly include, don't falsely exclude.

Specificity. The extent to which interpretation does not misidentify the wrong person.
Truly exclude, don't falsely include.

Reproducibility. The extent to which interpretation gives the same answer to the same question.
Concordant independent computer runs.

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

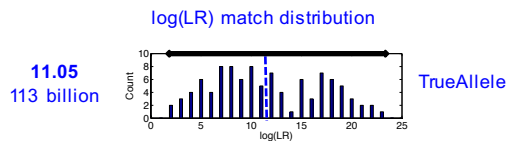## TrueAllele sensitivity

log(LR) match distribution

11.05
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
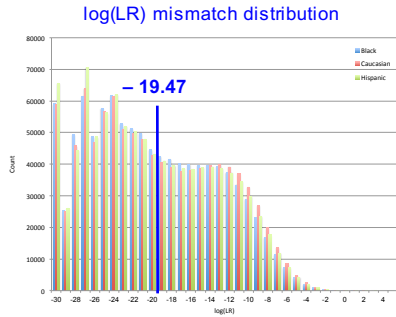
log(LR) mismatch distribution

Reproducibility

The extent to which interpretation
gives
the same answer to the same question

Statistical computing has sampling variation

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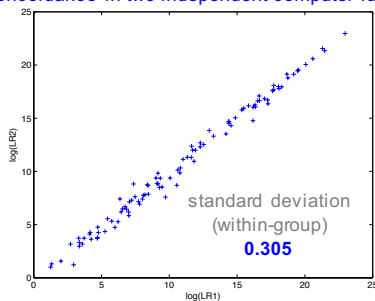on 101 matching genotypes
measure log(LR) variation

TrueAllele reproducibility

Concordance in two independent computer runs

standard deviation
(within-group)
0.305

Comparison of methods

6.83
6.68 million — CPI

2.15
140 — mCPI

11.05
113 billion — TrueAllele
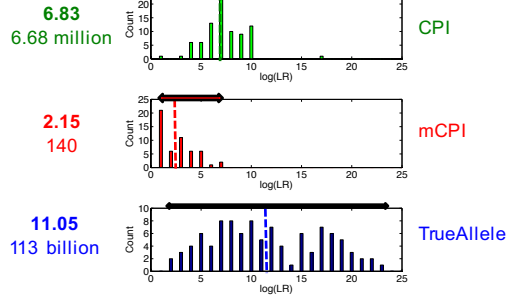
Information varies with amount

Match statistic

DNA amount
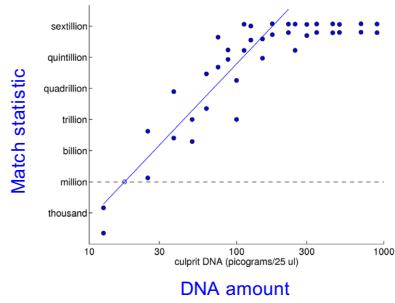
Accuracy of methods

CPI

mCPI

TrueAllele
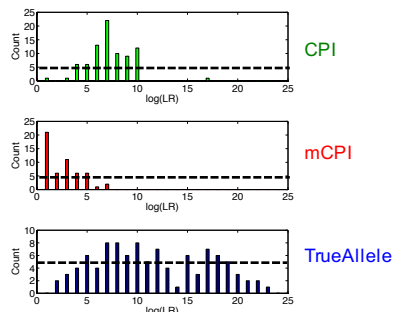
Chesterfield and Dyson convicted

Chesterfield   Nicholas   Dyson

Baton Rouge rapper Samuel "Mista Cain" Nicholas was found not guilty in the 2012 shooting death of an 18-year-old man and wounding another.

The East Baton Rouge Parish jury, however, convicted the accused triggerman, Chattley Chesterfield, 21, and an accomplice, Essence Dyson, 24, each of second-degree murder and aggravated battery.

Resolving limited DNA

- DNA from a .45 cartridge case
- Insufficient data for manual interpretation
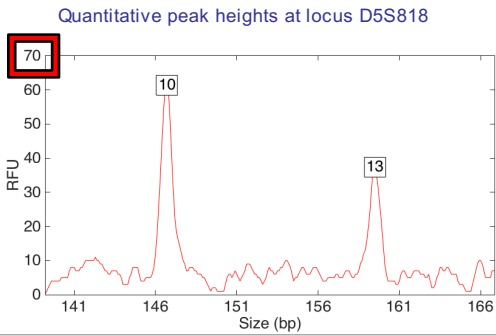- Sent for TrueAllele computer analysis

Small DNA amount

Quantitative peak heights at locus D5S818

Match statistic error rate
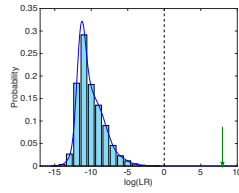
Statistic	Value
mean	-10.1049
std dev	1.7708

LR	log(LR)	Tail probability	One in ...	Ethnicity
93.3 million	7.9697	3.6816×10^{-10}	2.72 billion	African-American

---

Life sentence

Man sentenced to life without parole for 2011 murder
KPLC 7 News, Lake Charles, Louisiana

Corderrius Mitchell, 24, of Baton Rouge, was found guilty of principal to first-degree murder, aggravated flight from an officer and reckless operation charges in the shooting death of 26-year-old Fausto Ortiz-Herrera, Baton Rouge. Witnesses said three men may have tried to rob Herrera and shot him when he tried to run away.

Louisiana outcomes

Crime	Evidence	Defendant	Outcome	Sentence
murder	pants pockets	Brian Williams	guilty plea	
murder	fingernail	Emilio Calderon	guilty	life
murder	.45 shell	Corderrius Mitchell	guilty	life
rape	tampon string	Booker Diggins	guilty	life
murder	pistol	Chattley Chesterfield	guilty	life
murder	fence post	Hilton Wilson	guilty	life
murder	fence post	Reokenski Thomas	guilty	life
molestation	bra cup	Henry Lewis, Jr	guilty	50 years
murder	handcuffs	Robert Barthelemy	guilty	life
murder	water bottle	Christopher Hutsell	guilty	life

TrueAllele in Louisiana

30 cases, 5 trials, 1 Daubert

Parish:
Acadia
Ascension
Baton Rouge
Evangeline
Jefferson
Iberia
Orleans
Sabine
St. James

Crime laboratory:
Acadiana
Baton Rouge

Hardware
Software
Training
Validation
Protocols

Widespread acceptance

Admitted after Frye or Daubert challenge in: California, Indiana, Louisiana, Massachusetts, New York, Ohio, Pennsylvania, South Carolina, Virginia, Australia & United Kingdom

Crime labs use TrueAllele® system in California, Maryland, South Carolina & Virginia

Used in five hundred criminal cases in most of the United States, for both prosecution and defense

WTC DNA data reanalysis

18,000
victim remains

2,700
missing people

match

TrueAllele database

Suspects: S1, S2, S3, S4, S5

Food mart
• gun
• hat

Jewelry
• counter
• safe

Hardware
• safe
• phone

Market
• hat 1
• hat 2
• overalls
• shirt

Convenience
• keys
• tape
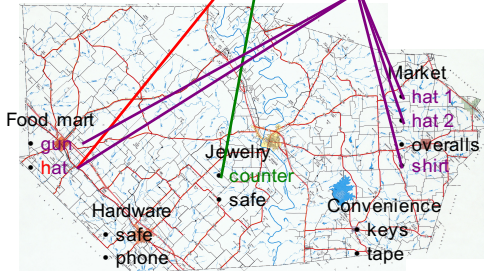
Inconclusive mixture

Crime laboratory DNA report
Crime lab user fee: $5,000

Conclusions:

Item 1 – Swab of textured areas from a handgun

The data indicates that DNA from four (4) or more contributors was obtained from the swab of the handgun. Due to the complexity of the data, no conclusions can be made regarding persons A and B as possible contributors to this mixture.
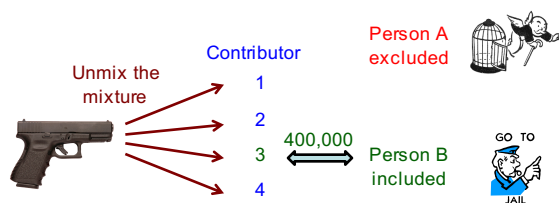

Computer reanalysis

Cybergenetics TrueAllele® report
Match statistics provide information

Unmix the mixture

Contributor
1
2
3
4

400,000

Person A excluded

Person B included

Leading the way

The probative power of DNA evidence
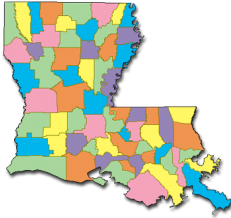
- Accurate
- Objective
- Understandable

More DNA mixture information

http://www.cybgen.com/information

- Courses
- Newsletters
- Newsroom
- Presentations
- Publications
- Webinars

http://www.youtube.com/user/TrueAllele
TrueAllele YouTube channel

Cybergenetics

You Tube

perlin@cybgen.com