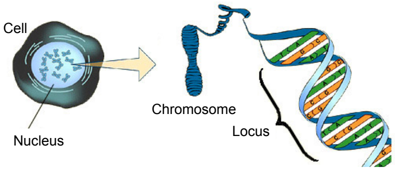
DNA: TrueAllele®
Statistical Analysis, Probabilistic Genotyping

Indiana Prosecuting Attorneys Council
IPAC Winter Conference
December, 2016
Indianapolis, IN

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

Cybergenetics

DNA biology



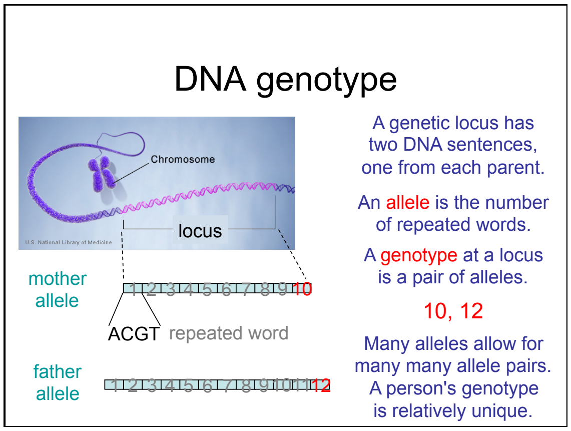
Short tandem repeat

23 volumes in cell's
DNA encyclopedia

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

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



DNA genotype

Chromosome
U.S. National Library of Medicine
locus
mother allele
1 2 3 4 5 6 7 8 9 10
ACGT repeated word
father allele
1 2 3 4 5 6 7 8 9 10 11 12

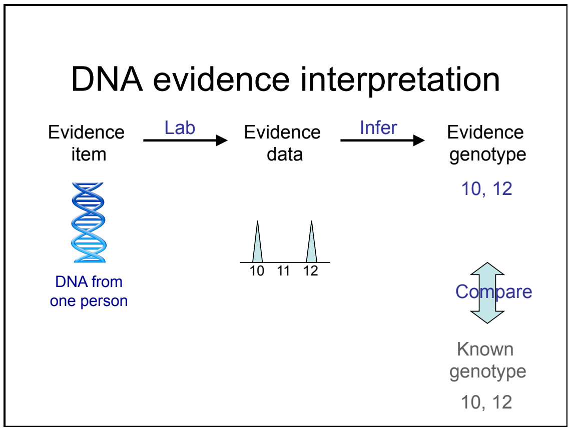
A genetic locus has two DNA sentences, one from each parent.

An allele is the number of repeated words.

A genotype at a locus is a pair of alleles.

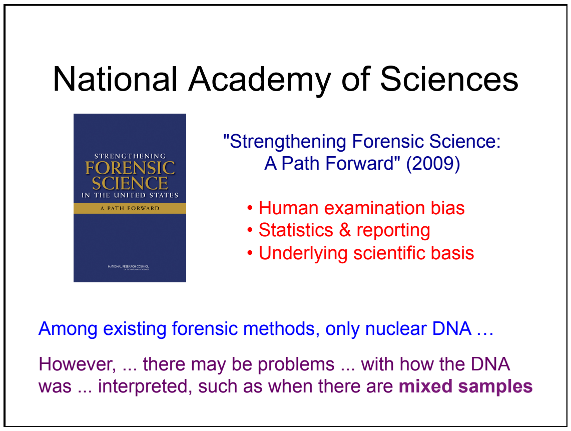
10, 12

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



DNA evidence interpretation

Evidence item
Lab
Evidence data
Infer
Evidence genotype
10, 12
DNA from one person
10 11 12
Compare
Known genotype
10, 12



National Academy of Sciences

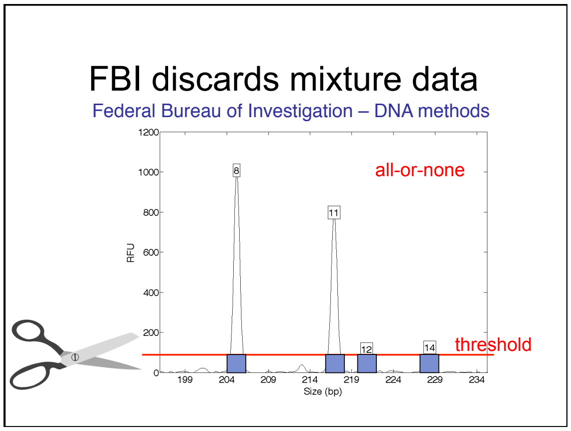
STRENGTHENING FORENSIC SCIENCE IN THE UNITED STATES
A PATH FORWARD

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

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

Among existing forensic methods, only nuclear DNA …

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



FBI discards mixture data

Federal Bureau of Investigation – DNA methods

Unreliable DNA mixture statistics

National Institute of Standards and Technology (NIST)
2005 Commerce Department study
Two contributor mixture data, known victim

Some Differences in Reporting Statistics

		Case1		
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	7.00E+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+08
79	ProPlus/Cofiler	930,000	47,900	1,350,000
16	ProPlus/Cofiler	434,600	31,710	399,100

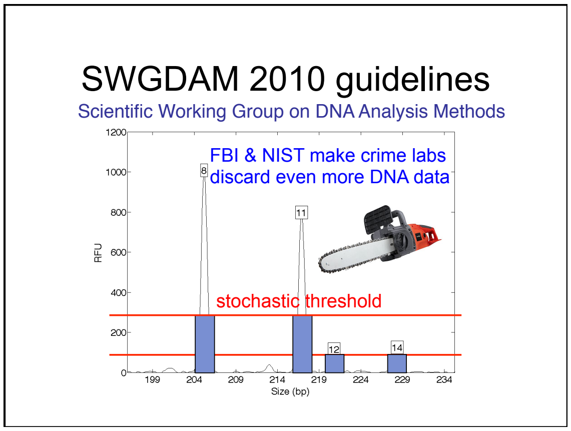
Remember that these labs are interpreting the same MIX05 electropherograms

Most “inconclusive.”
Some gave statistics:

213 trillion (14)

31 thousand (4)

Forensic DNA labs put on notice ten years ago



SWGDAM 2010 guidelines

Scientific Working Group on DNA Analysis Methods

FBI & NIST make crime labs discard even more DNA data

Falsely identify innocent people

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	**Include & 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**

Mixture statistics shut down labs

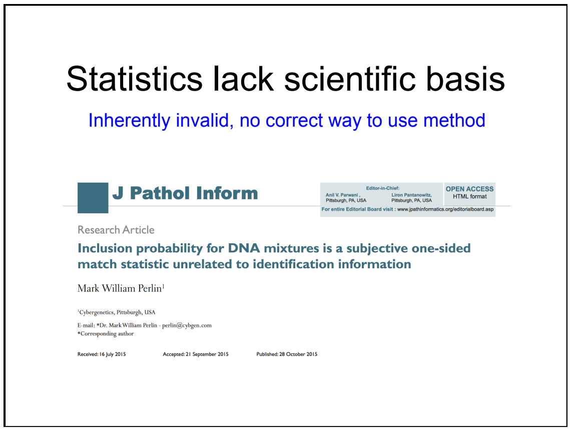
"National accreditation board suspends all DNA testing at D.C. crime lab"
The Washington Post *April 27, 2015*
Did not comply with FBI standards

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

Mixture statistics block CODIS

CODIS handles only simple DNA
Most DNA evidence is mixtures
Less than 10% uploaded to CODIS

Investigative DNA failure
Can't solve crimes
Prosecutor, defense, police, innocence



Statistics lack scientific basis

Inherently invalid, no correct way to use method

J Pathol Inform

Editor-in-Chief:
Anil V. Parwani, Pittsburgh, PA, USA
Liron Pantanowitz, Pittsburgh, PA, USA
For entire Editorial Board visit : www.jpathinformatics.org/editorialboard.asp

OPEN ACCESS
HTML format

Research Article

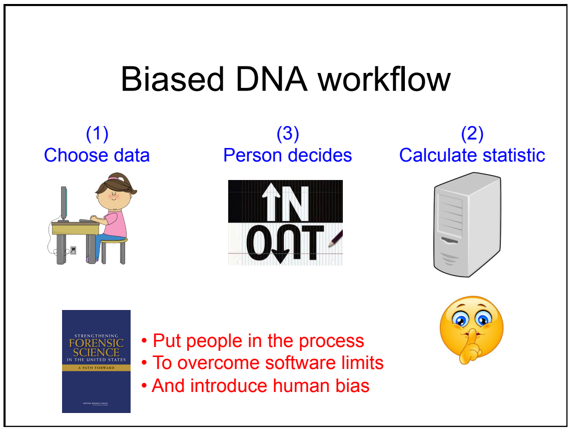
Inclusion probability for DNA mixtures is a subjective one-sided match statistic unrelated to identification information

Mark William Perlin[1]

[1]Cybergenetics, Pittsburgh, USA

E-mail: *Dr. Mark William Perlin - perlin@cybgen.com
*Corresponding author

Received: 16 July 2015   Accepted: 21 September 2015   Published: 28 October 2015



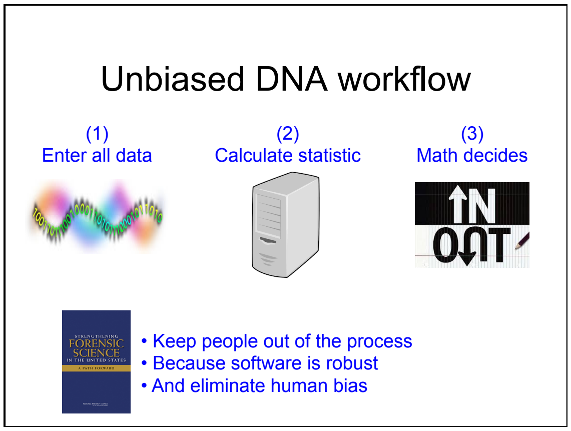
Biased DNA workflow

(1)
Choose data

(3)
Person decides

(2)
Calculate statistic

- Put people in the process
- To overcome software limits
- And introduce human bias



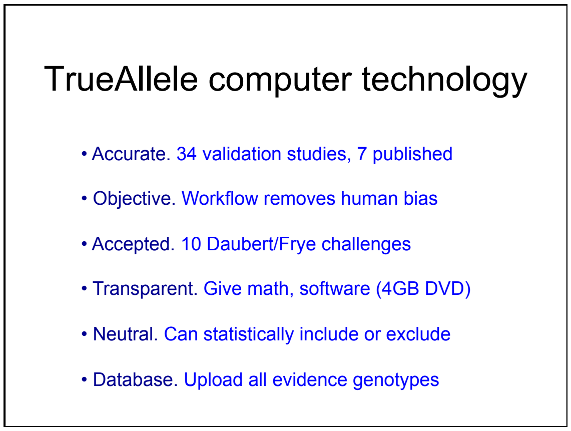
Unbiased DNA workflow

(1)
Enter all data

(2)
Calculate statistic

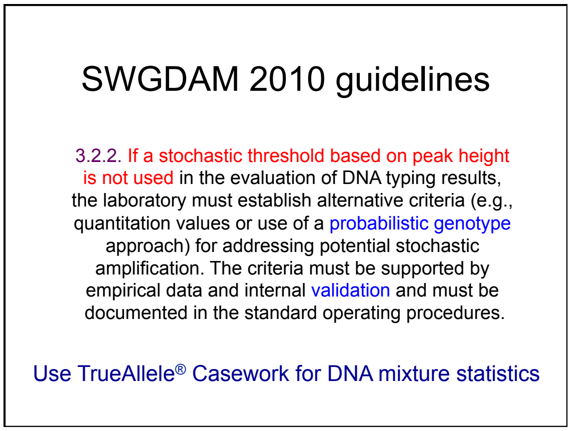
(3)
Math decides

- Keep people out of the process
- Because software is robust
- And eliminate human bias



TrueAllele computer technology

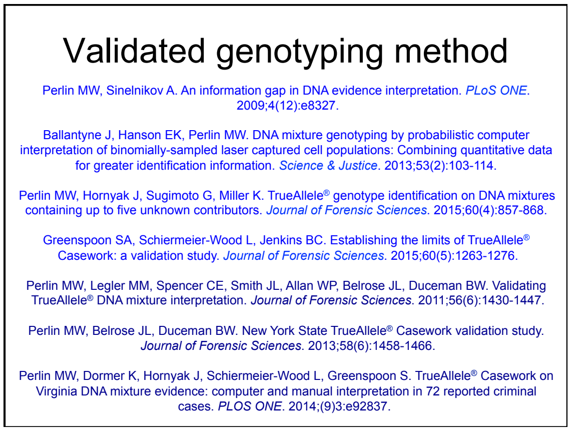
- Accurate. 34 validation studies, 7 published
- Objective. Workflow removes human bias
- Accepted. 10 Daubert/Frye challenges
- Transparent. Give math, software (4GB DVD)
- Neutral. Can statistically include or exclude
- Database. Upload all evidence genotypes



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



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.