

The Triumph and Tragedy of DNA Evidence

Friendship Village
May, 2017
Pittsburgh, PA

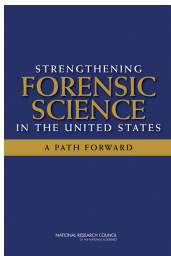
Mark W Perlin, PhD, MD, PhD
Cybergenetics, Pittsburgh, PA



Cybergenetics

Cybergenetics © 2003-2017

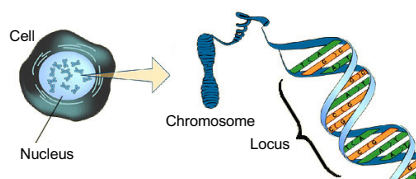
National Academy of Sciences



Trouble in River
City
2009

Where's the
"science" in
Forensic Science?

DNA biology



Short tandem repeat

DNA locus paragraph

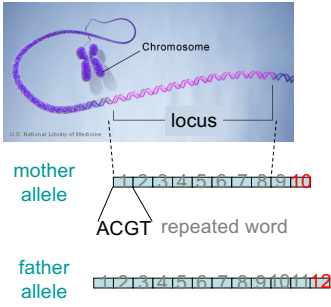


23 volumes in cell's DNA encyclopedia

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
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

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.

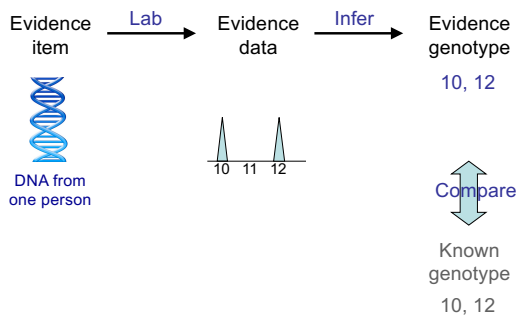
An allele is the number of repeated words.

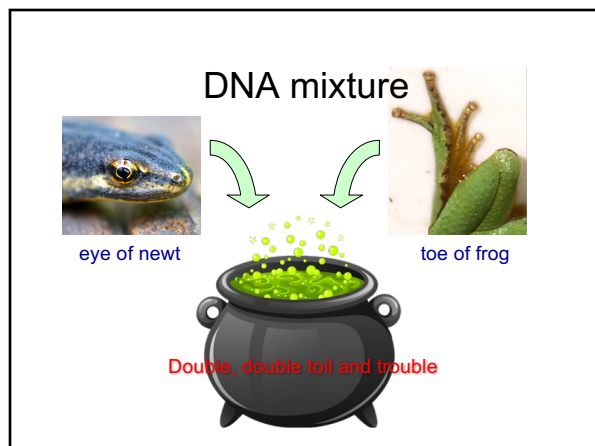
A genotype at a locus is a pair of alleles.

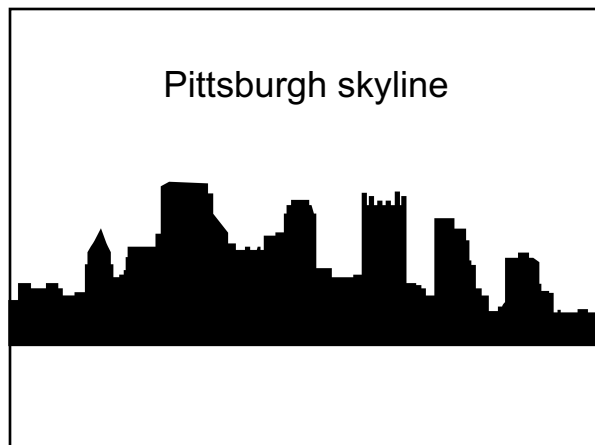
10, 12

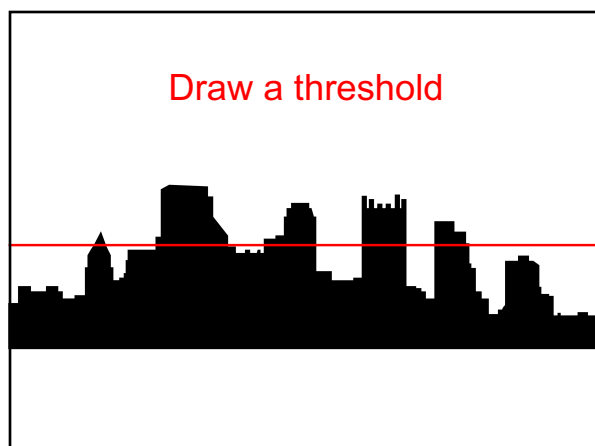
Many alleles allow for many many allele pairs. A person's genotype is relatively unique.

Simple DNA interpretation

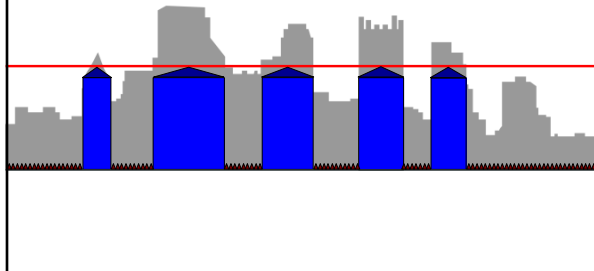




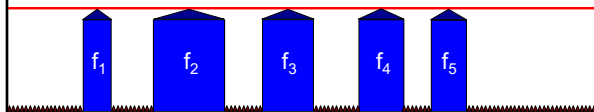




Same heights & vacant lots



Simplify data to interpret mixture



1. Apply threshold to find "alleles"
2. Add allele frequencies ($f_1 + f_2 + \dots$)
3. Square sum, take reciprocal
4. Locus "probability of inclusion" (PI)
5. Multiply locus PI values
6. Combined (CPI) **match statistic**

Unreliable DNA mixture statistics

NIST (Commerce Department) study in 2005
Two contributor mixture data, known victim

Some Differences in Reporting Statistics

LabID	Kits Used	Case1		
		Caucasians	African-Americans	Hispanics
90	ProPlus/Ceiler	1.18E+15	2.13E+14	3.09E+15
34	ProPlus/Ceiler	2.40E+11	7.90E+09	9.80E+10
33	ProPlus/Ceiler	2.94E+08	1.12E+08	1.74E+09
6	ProPlus/Ceiler	40,000,000	3,600,000	280,000,000
9	ProPlus/Ceiler	1.14E+07	1.97E+07	1.54E+08
29	ProPlus/Ceiler	930,000	41,000	1,350,000
16	ProPlus/Ceiler	434,600	31,710	399,100

Remember that these labs are interpreting the same MIX05 electropherograms

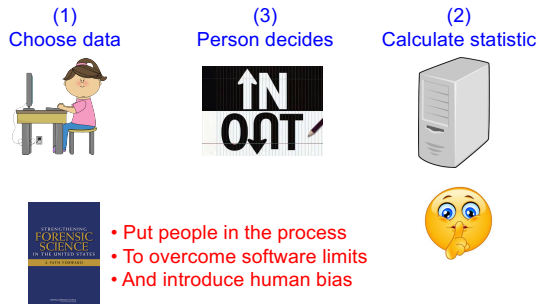
When not
"inconclusive":

213 trillion (14)

31 thousand (4)

Forensic DNA labs put on notice ten years ago

Biased DNA workflow



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

Statistics lack scientific basis

Misled courts for 15 years on countless DNA mixtures

J Pathol Inform

Editor-in-Chief
André F. Pfenninger
Pittsburgh, PA, USA
OPEN ACCESS
HTML format
For article Editorial Board visit: www.jpathinformatics.org/editorialboard.asp

Research Article

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

Mark William Perlin¹

¹Cybergenetics, Pittsburgh, USA

E-mail: ^{*}Dr. Mark William Perlin - perlin@cybgen.com

^{*}Corresponding author

Received 18 July 2015

Accepted 21 September 2015

Published 28 October 2015