DNA biology

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

“ROOT” repeated 10 times, so allele length is 10 repeats

23 volumes in cell’s DNA encyclopedia
DNA genotype

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 allele pairs.

A person's genotype is relatively unique.

DNA evidence interpretation

Evidence item → Lab → Evidence data → Infer → Evidence genotype

DNA from one person → 10,11,12 → Compare → Known genotype 10,12

National Academy of Sciences

“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

<table>
<thead>
<tr>
<th>Lab</th>
<th>2005 (14)</th>
<th>2006 (14)</th>
<th>2007 (14)</th>
<th>2008 (14)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Prob (xx)</td>
<td>2005 (14)</td>
<td>2006 (14)</td>
<td>2007 (14)</td>
<td>2008 (14)</td>
</tr>
<tr>
<td>Prob (yy)</td>
<td>2005 (14)</td>
<td>2006 (14)</td>
<td>2007 (14)</td>
<td>2008 (14)</td>
</tr>
</tbody>
</table>

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)

<table>
<thead>
<tr>
<th>Case</th>
<th>Conclusion</th>
<th>Test Results</th>
<th>Expected Genotypes ( shades of gray)</th>
</tr>
</thead>
<tbody>
<tr>
<td>6</td>
<td>Exclude</td>
<td></td>
<td>Subject C</td>
</tr>
<tr>
<td>3</td>
<td>Inconclusive with C only (A &amp; B included)</td>
<td></td>
<td>All three kits had PPW+IS</td>
</tr>
<tr>
<td>21</td>
<td>Inconclusive</td>
<td></td>
<td></td>
</tr>
<tr>
<td>70</td>
<td>Include &amp; provide CPI statistics</td>
<td></td>
<td>All over the road...</td>
</tr>
</tbody>
</table>

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

Biased DNA workflow
(1) Choose data
(2) Calculate statistic
(3) Person decides
• 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
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


