A Match Likelihood Ratio for DNA Comparison

American Academy of Forensic Sciences
February, 2009

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

Uncertain Genotype

\[ \text{probability distribution} \]

\[ q(x) \]

\[ r(x) \]

\[ s(x) \]

Match Likelihood Ratio

\[ LR = \frac{Pr(Q=S)}{Pr(Q=R)} \]
Interpreting DNA Evidence

A. Obtain DNA data
B. Infer genotype
   1. Data
   2. Model
   3. Compare
   4. Probability
C. Likelihood ratio

Genotype Inference
1. Data
   • evidence
   • victim
2. Model
   • genotype candidate
   • generate pattern
3. Compare
   • likelihood function (bell curve)
   • product rule for data
4. Probability
   • genotype probability distribution
   • genotype probability = \frac{\text{genotype likelihood}}{\text{sum of all genotype likelihoods}}

Different Methods: Data Used

<table>
<thead>
<tr>
<th>Data Used</th>
<th>inclusion</th>
<th>subtraction</th>
<th>addition</th>
</tr>
</thead>
<tbody>
<tr>
<td>victim profile</td>
<td>NO</td>
<td>YES</td>
<td>YES</td>
</tr>
<tr>
<td>original data</td>
<td>NO</td>
<td>NO</td>
<td>YES</td>
</tr>
</tbody>
</table>
Cybergenetics © 2003-2009

Statistical Inference View
inclusion method vs. likelihood ratio approach

*often robs the items of any probative value* - B. Weir

"usually discards a lot of information compared to the correct likelihood ratio approach" - C. Brenner

"does not use as much of the information included in the data as the LR approach but, conceptually, they are equivalent" - M. Krawczak

"Recommendation 1: The likelihood ratio is the preferred approach to mixture interpretation.* - DNA commission of the International Society of Forensic Genetics

Mixture Case

- DNA from under victim's fingernails
- two contributors to DNA mixture
- 93.3% victim & 6.7% unknown
- 2 ng DNA in 50 ul
- ProfilerPlus + Coffiler STR analysis
- three different mixture interpretations
  1. inclusion
  2. subtraction
  3. addition

D16S539
### Inclusion Method

- discard peak heights
- create uniform peaks
- included allele pairs
- uniform probability

**'Inclusion' LR at D16S539**

<table>
<thead>
<tr>
<th>Genotype Probability Distributions</th>
<th>Match Probabilities</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Q</strong></td>
<td><strong>R</strong></td>
</tr>
<tr>
<td>allele pair</td>
<td>x</td>
</tr>
<tr>
<td>11 11</td>
<td>0.16667</td>
</tr>
<tr>
<td>11 12</td>
<td>0.16667</td>
</tr>
<tr>
<td>11 13</td>
<td>0.16667</td>
</tr>
<tr>
<td>12 11</td>
<td>0.16667</td>
</tr>
<tr>
<td>12 12</td>
<td>0.16667</td>
</tr>
<tr>
<td>12 13</td>
<td>0.16667</td>
</tr>
<tr>
<td>13 11</td>
<td>0.16667</td>
</tr>
<tr>
<td>13 12</td>
<td>0.16667</td>
</tr>
<tr>
<td>13 13</td>
<td>0.16667</td>
</tr>
</tbody>
</table>

Pr(Q=S): 0.16667
Pr(Q=R): 0.10004
Likelihood Ratio: 1.666

### Addition Method

- similar pattern, high likelihood
- dissimilar pattern, low likelihood
'Addition' LR at D16S539

<table>
<thead>
<tr>
<th>allele pair</th>
<th>q(x)</th>
<th>r(x)</th>
<th>s(x)</th>
<th>Pr(Q=S)</th>
<th>Pr(Q=R)</th>
<th>Pr(Q=Q)</th>
<th>Likelihood Ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td>11 11</td>
<td>0.07413</td>
<td>0.72200</td>
<td>0.72200</td>
<td>0.72200</td>
<td>0.06423</td>
<td>0.07400</td>
<td>9.756</td>
</tr>
<tr>
<td>11 12</td>
<td>0.18466</td>
<td>0.11080</td>
<td>0.08896</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>11 13</td>
<td>0.25229</td>
<td>0.02949</td>
<td>0.02669</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>12 12</td>
<td>0.11499</td>
<td>0.00310</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>12 13</td>
<td>0.02800</td>
<td>0.00667</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>13 13</td>
<td>0.25000</td>
<td>0.02669</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Information

Validation
Calibration

Bibliography

- Quantitative STR Peak Information
- Genotype Probability Distributions
- Computer Interpretation of STR Data
- Statistical Modeling and Computation
- Likelihood Ratio Literature
- Mixture Interpretation Admissibility
- Computer Systems for Quantitative DNA Mixture Deconvolution
- TrueAllele Casework Publications

Conclusions: MLR

- A useful tool for determining identification information with uncertain genotypes
- Works well on forensic mixture cases
- Enables quantitative validation, calibration, and comparison of genotype inference methods
- Based on generally accepted scientific principles