Evolution of DNA Mixture Interpretation
(DNA Advisory Board and pre-2010 SWGDAM Guidelines)
1. Historical preamble showcasing ancient technology
2. Science of mixture analysis
3. Community consensus efforts
“Those who cannot learn from history are doomed to repeat it”

George Santayana
The Life of Reason (1905-1906)
Tasteful Statistics Jokes

• A statistician is a professional who diligently collects facts and data and then carefully draws confusions about them

• You can lie with statistics but even better without

• Statistics means you never have to say you’re certain (wrong)
Seriously True Precepts

• All models are wrong. Some are useful.
  – George Box

• There are no facts, only interpretations.
  – Frederick Nietzsche
“Mixtures for Newbies”

1. Recognize Mixture
2. Infer Genotype(s)
3. Attach Weight with Stats
Pre-DNA Era
Conventional Markers

ABO

- Victim is ‘B secretor’
- Vaginal swab with semen: ‘AB’
- Conclusion: mixture present with the semen donor being an A or AB
- Thus 27% + 6% = 33% of the Hong Kong Chinese male population cannot be excluded as donors of the semen stain
Enzymes: PGM

For a single genetic marker system
To **recognize** a mixture (no assumption of the presence of a particular donor) need genetic marker with $\geq 3$ alleles (e.g., EAP, Gc, PGM)
To **eliminate** a proportion of the population as potential contributors need $\geq 4$ alleles (only 1 system..PGM)

For a multi-locus genetic marker system
To **recognize** a mixture (no assumption of the presence of a particular donor) need at least one genetic marker with $\geq 3$ alleles
To **eliminate** a proportion of the population as potential contributors need loci with at least one common allele missing from the mixture.
Evidence = PGM 2+2-1-

PGM Genotypes INCLUDED:
2+2+, 2-2-, 1-1-, 2+2-, 2+1-, 2-1-

= 3.1 + 0.8 + 1.8 + 3.1 + 4.6 + 2.8

= 16.2

Thus 16.2% of the population cannot be excluded (i.e. included) as potential doors of the stain

<table>
<thead>
<tr>
<th>PGM Genotype</th>
<th>Observed Frequency</th>
<th>Allele Frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td>1+</td>
<td>35.7</td>
<td>0.60</td>
</tr>
<tr>
<td>1-</td>
<td>1.8</td>
<td>0.13</td>
</tr>
<tr>
<td>1+1-</td>
<td>15.9</td>
<td></td>
</tr>
<tr>
<td>2+</td>
<td>3.0</td>
<td>0.17</td>
</tr>
<tr>
<td>2-</td>
<td>0.8</td>
<td>0.10</td>
</tr>
<tr>
<td>2+2-</td>
<td>3.1</td>
<td></td>
</tr>
<tr>
<td>2+1+</td>
<td>19.7</td>
<td></td>
</tr>
<tr>
<td>2+1-</td>
<td>4.6</td>
<td></td>
</tr>
<tr>
<td>2-1+</td>
<td>11.8</td>
<td></td>
</tr>
<tr>
<td>2-1-</td>
<td>2.8</td>
<td></td>
</tr>
</tbody>
</table>
However, 2+2+, 2-2-, 1-1-, 2+2-, 2+1-, 2-1-
= \( p_{2+}^2 + p_{2-}^2 + p_{1-}^2 + 2p_{2+}p_{2-} + 2p_{2+}p_{1-} + 2p_{2-}p_{1-} \)

[Where allele frequencies are \( p_{2+} \), \( p_{2-} \), \( p_{1+} \), \( p_{1-} \) and \( (p_{2+} + p_{2-} + p_{1+} + p_{1-}) = 1 \)]

= \( (p_{2+} + p_{2-} + p_{1-})^2 \) = Probability of Inclusion = PI
= \( \text{RMNE} = 0.16 = 16\% \)

[Probability of Exclusion = 1 - PI = 1 - 0.16 = 0.84 = 84\%]
The DNA Profiling Era

National Library of Medicine
Blake, E., JFS 37 1992, 700-26

Adams, JFS 36 19911284-1298

**FIG. 10**—DNA profile results for the genetic locus DIS7 for stains containing blood from two sources (Lanes 6, 7, and 8). The mixed stains contained 25 μL of blood from the female donor and 50 μL of blood from the male donor. Lanes 1, 5, and 9 contain size markers. Lane 2 is the K562 cell line control. Lane 3 and 4 are the female and male donor controls, respectively.

**FIG. 2**—A 50-ng mixture: DQα 1.1, 4 DNA and DQα 2.3 DNA were mixed in the proportions indicated above. For each sample, a total of 50 ng of this DNA mixture was added to the PCR mix. The samples were amplified for 32 cycles, and DQα typing was performed as described in Materials and Methods. As the quantity of DNA corresponding to the minor component genotype is decreased relative to the major component genotype, the resulting dot intensity for the minor component decreases relative to the major component.
1995

OJ Simpson grimaces as he tries on a glove during his murder trial. Photograph: PA
### Profiles from Bundy

<table>
<thead>
<tr>
<th>Item</th>
<th>Description</th>
<th>Number of loci</th>
<th>RFLP</th>
<th>PCR</th>
<th>Not excluded</th>
</tr>
</thead>
<tbody>
<tr>
<td>42</td>
<td>NB - pool</td>
<td>1</td>
<td>NB</td>
<td></td>
<td></td>
</tr>
<tr>
<td>47</td>
<td>1st drop by victims</td>
<td>7</td>
<td>OS</td>
<td></td>
<td></td>
</tr>
<tr>
<td>48</td>
<td>Bundy walk</td>
<td>6</td>
<td>OS</td>
<td></td>
<td></td>
</tr>
<tr>
<td>49</td>
<td>Bundy walk</td>
<td>7</td>
<td>OS</td>
<td></td>
<td></td>
</tr>
<tr>
<td>50</td>
<td>Bundy walk</td>
<td>7</td>
<td>OS</td>
<td></td>
<td></td>
</tr>
<tr>
<td>52</td>
<td>Bundy walk</td>
<td>5</td>
<td>OS</td>
<td></td>
<td></td>
</tr>
<tr>
<td>56</td>
<td>Shoe print</td>
<td>5</td>
<td>NB</td>
<td></td>
<td></td>
</tr>
<tr>
<td>78</td>
<td>RG boot drop</td>
<td>5</td>
<td>NB</td>
<td></td>
<td></td>
</tr>
<tr>
<td>84</td>
<td>NB nails</td>
<td>7</td>
<td>OS</td>
<td></td>
<td></td>
</tr>
<tr>
<td>115</td>
<td>Rear gate</td>
<td>2</td>
<td>OS</td>
<td></td>
<td></td>
</tr>
<tr>
<td>116</td>
<td>Rear gate</td>
<td>2</td>
<td>OS</td>
<td></td>
<td></td>
</tr>
<tr>
<td>117</td>
<td>Rear gate</td>
<td>9</td>
<td>OS</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

### Profiles from the Rockingham glove

<table>
<thead>
<tr>
<th>Item</th>
<th>Description</th>
<th>Number of loci</th>
<th>RFLP</th>
<th>PCR</th>
<th>Not excluded</th>
</tr>
</thead>
<tbody>
<tr>
<td>9</td>
<td>Inside/back of wrist</td>
<td>5</td>
<td>NB, RG</td>
<td></td>
<td></td>
</tr>
<tr>
<td>9:G1</td>
<td>Inside/back index finger</td>
<td>2</td>
<td>NB, RG</td>
<td></td>
<td></td>
</tr>
<tr>
<td>9:G2</td>
<td>Inside/side middle finger</td>
<td>5</td>
<td>NB, RG</td>
<td></td>
<td></td>
</tr>
<tr>
<td>9:G3</td>
<td>Inside-back ring finger</td>
<td>8</td>
<td>RB</td>
<td></td>
<td></td>
</tr>
<tr>
<td>9:G4</td>
<td>Inside-back of hand</td>
<td>5</td>
<td>NB, RG</td>
<td></td>
<td></td>
</tr>
<tr>
<td>9:G9</td>
<td>Inside-by wrist notch</td>
<td>2</td>
<td>RG</td>
<td></td>
<td></td>
</tr>
<tr>
<td>9:G10</td>
<td>Inside-by wrist notch</td>
<td>2</td>
<td>RG, OS</td>
<td></td>
<td></td>
</tr>
<tr>
<td>9:G11</td>
<td>Outside-near wrist notch</td>
<td>1</td>
<td>NB, RG, OS</td>
<td></td>
<td></td>
</tr>
<tr>
<td>9:G12</td>
<td>Outside-near wrist notch</td>
<td>1</td>
<td>NB, RG</td>
<td></td>
<td></td>
</tr>
<tr>
<td>9:G13</td>
<td>Stitching on wrist notch</td>
<td>1</td>
<td>NB, RG, OS</td>
<td></td>
<td></td>
</tr>
<tr>
<td>9:G14</td>
<td>Inside-back of cuff edge</td>
<td>1</td>
<td>NB, RG</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

### Profiles from the Ford Bronco

<table>
<thead>
<tr>
<th>Item</th>
<th>Description</th>
<th>Number of loci</th>
<th>RFLP</th>
<th>PCR</th>
<th>Not excluded</th>
</tr>
</thead>
<tbody>
<tr>
<td>23</td>
<td>Driver door interior</td>
<td>1</td>
<td>OS</td>
<td></td>
<td></td>
</tr>
<tr>
<td>24</td>
<td>Instrument panel</td>
<td>1</td>
<td>OS</td>
<td></td>
<td></td>
</tr>
<tr>
<td>25</td>
<td>Driver side carpet</td>
<td>2</td>
<td>OS</td>
<td></td>
<td></td>
</tr>
<tr>
<td>29</td>
<td>Steering wheel</td>
<td>6</td>
<td>OS, NB</td>
<td></td>
<td></td>
</tr>
<tr>
<td>30</td>
<td>Centre console</td>
<td>2</td>
<td>OS</td>
<td></td>
<td></td>
</tr>
<tr>
<td>31</td>
<td>Centre console</td>
<td>2</td>
<td>OS, RG</td>
<td></td>
<td></td>
</tr>
<tr>
<td>34</td>
<td>Centre console</td>
<td>1</td>
<td>OS</td>
<td></td>
<td></td>
</tr>
<tr>
<td>293</td>
<td>Driver side wall</td>
<td>2</td>
<td>NB</td>
<td></td>
<td></td>
</tr>
<tr>
<td>303</td>
<td>Centre console 4*</td>
<td>2</td>
<td>OS, NB, RG</td>
<td></td>
<td></td>
</tr>
<tr>
<td>304</td>
<td>Centre console 4*</td>
<td>2</td>
<td>OS, NB, RG</td>
<td></td>
<td></td>
</tr>
<tr>
<td>305</td>
<td>Centre console 4*</td>
<td>2</td>
<td>OS, NB, RG</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Item</th>
<th>Description</th>
<th>Number of loci</th>
<th>RFLP</th>
<th>PCR</th>
<th>Not excluded</th>
</tr>
</thead>
<tbody>
<tr>
<td>13</td>
<td>Ankle area 42A-1</td>
<td>14</td>
<td>7</td>
<td>NB</td>
<td></td>
</tr>
<tr>
<td>14</td>
<td>Leg-opposite 42A-1</td>
<td>2</td>
<td>OS</td>
<td></td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>Leg-same side as 42A-1</td>
<td>9</td>
<td>2</td>
<td>OS</td>
<td></td>
</tr>
<tr>
<td>309</td>
<td>Upper toe region 42A</td>
<td>2</td>
<td>NB</td>
<td></td>
<td></td>
</tr>
<tr>
<td>310</td>
<td>Near ankle 42B</td>
<td>2</td>
<td>RB</td>
<td></td>
<td></td>
</tr>
<tr>
<td>311</td>
<td>Near ankle 42B</td>
<td>2</td>
<td>RB</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Combined 303, 304, 305. OS and RG not excluded.

“it is only the manipulation of uncertainty that interests us. We are not concerned with the matter that is uncertain. Thus we do not study the mechanism of rain; only whether it will rain.”

The Science of Mixtures

Early Evangelism: Ian Evett, Home Office FSS

  - Use of LR instead of coincidence probability as a logical framework for assessment of probity of genetic evidence
  - Single source


“a two-trace transfer problem” (1)

Uses conventional markers and phenotype (not allele) frequencies

Example 1 (single donor)

• A crime has been committed by a man who left a bloodstain at the scene.
• The blood is typed using a polymorphic genetic marker that has a number of distinct phenotypes $\gamma_1$, $\gamma_2$, $\gamma_3$,... that occur in the general population with relative frequencies $q_1$, $q_2$, $q_3$.

- The scientist has evidence $F$
  - the bloodstain is typed as $\gamma_1$
  - the suspect is typed as $\gamma_1$
- $C$: the suspect committed the crime
- $\sim C$: the suspect did not commit the crime
Example 2 (two donors)

- A crime has been committed by two men who both left a bloodstain at the scene.
  - The scientist has evidence F
    - the bloodstain is typed as $\gamma_1$ and $\gamma_2$
    - the suspect is typed as $\gamma_1$
  - $C$: the suspect was one of two men who committed the crime
  - $\sim C$: the suspect was not one of the two men who committed the crime

- Likelihood Ratio (LR) = \( \frac{\text{Probability of } F \text{ given } C \text{ is true}}{\text{Probability of } F \text{ given } \sim C \text{ is true}} \)
“a two-trace transfer problem” (3)

- LR (example 1) = 1/ q₁
- LR (example 2) = (1 x q₂)/2 q₁ q₂ = 1/ 2q₁

- Thus the LR is ½ that of the single stain case (evidence is less probative)
- If n different bloodstains of types γ₁, γ₂, γ₃...γₙ, and a suspect of type γ₁ then the LR = 1/nq₁ (i.e. reduction in LR dependent upon no. of different donors and, with the exception of γ₁, not on their relative phenotype frequencies)
- If q is greater than 0.5 then the LR would be less than 1!
  - Evidence reduces support for C versus ~C
interpreting single locus profiles of DNA mixtures

- Uses DNA markers and band (allele) frequencies
- Example 1 (four allele mixture, abcd with suspect possessing two of them, b and c)
  - \( LR = \frac{2f_a f_b}{24f_a f_b f_c f_d} = \frac{1}{12} f_b f_c \)
    - One sixth of the LR obtained if only one assailant and bands b and c only
  - \( LR = \frac{1}{24} f_a f_b f_c f_d \)
    - If second suspect is arrested and he has a and b alleles
- Example 2 (three allele mixture abc with suspect possessing two of them, b and c)
  - \( LR = \frac{(f_a + 2f_b + 2f_c)}{12f_b f_c (f_a + f_b + f_c)} \)
- As no of alleles increases the LR evaluation becomes “quite complicated”
- Evidential strength falls rapidly with increasing numbers of alleles
“taking account of peak areas when interpreting mixed DNA profiles”

- Conceptual paper that establishes logical framework for taking into account peak areas when interpreting mixed DNA profiles
- Use of peak area data and mixing ratios permits the ranking of different LRs that individually evaluate all possible combinations of genotypes present in the mixture
- Need computer programs
Analysis and interpretation of mixed forensic stains using DNA STR profiling

T.M. Clayton\textsuperscript{a,*}, J.P. Whitaker\textsuperscript{a}, R. Sparkes\textsuperscript{b}, P. Gill\textsuperscript{b}

\textsuperscript{a}Forensic Science Service, Wetherby Laboratory, Sandbeck Way, Audby Lane, Wetherby, West Yorkshire LS22 4DN, UK
\textsuperscript{b}Forensic Science Service, Priory House, Gooch Street North, Birmingham B56QQ, UK

Received 13 May 1997; received in revised form 9 October 1997; accepted 27 October 1997
Analysis and interpretation of mixed forensic stains using DNA STR profile

• Step 1: identify the presence of a mixture
• Step 2: identify the number of contributors
• Step 3: determine the approximate ‘ratio' of the components in the mixture
• Step 4: determine the possible pairwise combinations for the components of the mixture
• Step 5: compare the resultant profiles for the possible components of the mixture with those from the reference samples
• **Step 1: identify the presence of a mixture**
  - Extra bands
    - Also should distinguish true second donor alleles from stutter, chromosomal abnormalities, pullup, n + 1 bands
  - Allele peak asymmetry
    - Also should distinguish true second donor alleles from differential amplification of the alleles (e.g. stochastic effects and primer binding site mutations)

• **Step 2: identify the number of contributors**
  - Maximum alleles at a locus is 4 for two person mixture
  - 5 or 6 alleles indicative of three or more contributors
  - Experience indicates majority of mixtures encountered in casework are two person mixtures
- Step 3: determine the approximate ‘ratio' of the components in the mixture

<table>
<thead>
<tr>
<th>Mixture ratio</th>
<th>Dosage of alleles observed</th>
<th>Ratio of peak areas X:Y</th>
</tr>
</thead>
<tbody>
<tr>
<td>Male (XY)</td>
<td>Female (XX)</td>
<td>X</td>
</tr>
<tr>
<td>10</td>
<td>1</td>
<td>12</td>
</tr>
<tr>
<td>5</td>
<td>1</td>
<td>7</td>
</tr>
<tr>
<td>4</td>
<td>1</td>
<td>6</td>
</tr>
<tr>
<td>3</td>
<td>1</td>
<td>5</td>
</tr>
<tr>
<td>2</td>
<td>1</td>
<td>4</td>
</tr>
<tr>
<td>1</td>
<td>1</td>
<td>3</td>
</tr>
<tr>
<td>1</td>
<td>2</td>
<td>5</td>
</tr>
<tr>
<td>1</td>
<td>3</td>
<td>7</td>
</tr>
<tr>
<td>1</td>
<td>4</td>
<td>9</td>
</tr>
<tr>
<td>1</td>
<td>5</td>
<td>11</td>
</tr>
<tr>
<td>1</td>
<td>10</td>
<td>21</td>
</tr>
</tbody>
</table>
Step 4: determine the possible pairwise combinations for the components of the mixture

<table>
<thead>
<tr>
<th>Four alleles (a,b,c,d)</th>
<th>Three alleles (a,b,c)</th>
<th>Two alleles (a,b)</th>
</tr>
</thead>
<tbody>
<tr>
<td>a,b c,d</td>
<td>a,a b,c</td>
<td>a,a a,b</td>
</tr>
<tr>
<td>a,c b,d</td>
<td>b,b a,c</td>
<td>a,b a,b</td>
</tr>
<tr>
<td>a,d b,c</td>
<td>c,c a,b</td>
<td>a,a b,b</td>
</tr>
<tr>
<td>c,d a,b</td>
<td>a,b a,c</td>
<td>a,b b,b</td>
</tr>
<tr>
<td>b,d a,c</td>
<td>b,c a,c</td>
<td>a,b a,a</td>
</tr>
<tr>
<td>b,c a,d</td>
<td>a,b b,c</td>
<td>b,b a,a</td>
</tr>
<tr>
<td>b,c a,a</td>
<td>b,b a,b</td>
<td>b,b a,b</td>
</tr>
<tr>
<td>a,b c,c</td>
<td>a,b a,b</td>
<td></td>
</tr>
<tr>
<td>a,c b,b</td>
<td>a,b b,b</td>
<td></td>
</tr>
<tr>
<td>a,b c,c</td>
<td>a,b a,b</td>
<td></td>
</tr>
</tbody>
</table>

Key: bold entries represent reciprocal combinations.

“Using the quantitative information drawn from the peak areas in the profile and the approximate ratio of the mixture, some of the pairwise possibilities can then be discounted.”
• Step 5: compare the resultant profiles for the possible components of the mixture with those from the reference samples
  – If the profiles from the suspect's reference sample matches one or other of the alternatives, then that person cannot be eliminated as a possible contributor of one component of the mixed stain.
  – If the factual circumstances of a case are such that the profile from the donor of the sample might also be anticipated, then one might expect this individual's profile to complete the match and account for all of the remaining alleles.
Later Evangelism: Bruce Weir, Dept Statistics, NCSU (now Univ Washington, WA)

• 1997- ”Interpreting DNA mixtures” (Weir, Triggs, Stowell, Walsh, Buckleton, J For Sci 1997 42 p213-222)
  - refines and expands the LR concept and provides how-to formulations

• 1999- ”Interpreting DNA mixtures in structured populations” (Curran, Triggs, Buckleton, Weir J For Sci 1997 44 p987-995)
  - effects of population structure
  - role of evolution in shaping the probabilities of sets of profiles
  - Accounts for the information contained in the profiles of people who are declared not to have contributed to the evidence profile
• Post modern evangelism: taking into account PCR artifacts

• 1998- “Interpretation of simple mixtures of when artefacts such as stutters are present-with special reference to multiplex STRs used by the Forensic Science Service” (Gill, Sparkes, Buckleton, For Sci Int 1998 95 p213-224)

• 2009- “Interpreting low template DNA profiles” (Balding, Buckleton, For Sci Int Genet 2009 4 p1-10)

• 2010-“A universal strategy to interpret DNA profiles that does not require a definition of low-copy-number” (Gill, Buckleton, For Sci Int Genet 2010 4 p221-7)
Meanwhile back at the RMNE Ranch

- 1993 “Forensic inference from genetic markers”
  - how to calculate PE (PI)
2009

Mixture Interpretation: Defining the Relevant Features for Guidelines for the Assessment of Mixed DNA Profiles in Forensic Casework*

ABSTRACT: Currently in the United States there is little direction for what constitutes sufficient guidelines for DNA mixture interpretation. While a standardized approach is not possible or desirable, more definition is necessary to ensure reliable interpretation of results is carried out. In addition, qualified DNA examiners should be able to review reports and understand the assumptions made by the analyst who generated and interpreted the data. A framework for organizing guidelines is proposed, with a focus on the management of DNA mixture interpretation.

*Acknowledgments: The authors would like to acknowledge the contributions of John R. S. Staudenmaier, Ph.D., and Dr. Brian O. G. Scott, Ph.D., for their insightful comments and feedback on early drafts of this manuscript.

Formalized the Use of Stochastic Threshold!
Defining the Relevant Features for Guidelines for the Assessment of Mixed DNA Profiles

• “A standardized mixture interpretation protocol is not recommended or possible”
• Authors clearly prefer the random match probability and Probability of Inclusion (RMNE) (1-PE) approach instead of LR
  – “convey to the trier of fact the probative value of the evidence in a straightforward fashion”
Current evangelism: Quantitative Data Modeling  (Mark Perlin, Cybergenetics)

Hierarchical Bayesian Model with MCMC Solution

- standard approach in modern science
- describes uncertainty using probability
- the "new calculus"
- replaces hard calculus with easy computing
- can solve virtually any problem
- well-suited to interpreting DNA evidence

From Mark Perlin
Generally Accepted Method

Software Solutions for Mixture Deconvolution?

- **Linear Mixture Analysis (LMA)**
  - *Part of TrueAllele system* developed by Mark Perlin and Cybergenetics

- **Least Squares Deconvolution (LSD)**
  - Described by T. Wang (University of Tennessee) at Oct 2002 Promega meeting

- **PENDULUM**
  - *Part of FSS i-3 software suite*

- **NYCOCME**
  - Statistical tool for mixture analysis using LRs and incorporating Pr (drop-in and drop-out) and LTDNA samples
Mixed DNA Profile

**Frequentist approach**
- **Method 1:** Exclusion probability
  - Random Man Not Excluded (RMNE)

**Bayesian approaches**
- Qualitative data
  - **Method 2:** Qualitative approach
  - Likelihood Ratio Approach
- Quantitative and qualitative data
  - **Method 3:** Binary model Approach
  - **Method 4:** Continuous model MCMC

Figure 7.1 from Tim Clayton and John Buckleton, Chapter 7 “Mixtures” in *Forensic DNA Evidence Interpretation* (2005) CRC Press
Community Effort and Diktats!
Technical Working Group on DNA Analysis Methods (TWGDAM)

TWGDAM (1989) – Crime Lab Digest 16(2):40-59
  Kearney et al. “Guidelines for a quality assurance program for DNA restriction fragment length polymorphism analysis” - SILENT on MIXTURES

TWGDAM (1991) – Crime Lab Digest 18(2):44-75
  Kearney et al. “Guidelines for a quality assurance program for DNA analysis”

  Budowle et al. “Guidelines for a quality assurance program for DNA analysis”
TWGDAM (1991) – “Guidelines for a quality assurance program for DNA analysis”

- **4. Validation**
  - 4.1.5.5 Mixed Specimen studies-investigate the ability of the system to detect the components of mixed specimens and define the limitations of the system

- **7. Analytical Procedures**
  - 7.1 Sample Evaluation and Preparation
    - 7.1.2 When semen is identified, a method of differential extraction should be employed and, when appropriate, each of the DNA fractions typed
TWGDAM (1995) “Guidelines for a quality assurance program for DNA analysis”

Identical to 1991 Guidelines

- **4. Validation**
  - **4.1.5.5** Mixed Specimen studies—investigate the ability of the system to detect the components of mixed specimens and define the limitations of the system

- **7. Analytical Procedures**
  - **7.1 Sample Evaluation and Preparation**
    - **7.1.2** When semen is identified, a method of differential extraction should be employed and, when appropriate, each of the DNA fractions typed
National Research Council Reports

- 1992
  “If a suspects pattern is found within the mixed pattern, the appropriate frequency to assign such a ‘match’ is the sum of the frequencies of all the genotypes that are contained within (i.e. that are a subset of) the mixed pattern” – RMNE

- 1996
  in referring to the previous (1992) calculation, “this calculation is hard to justify, because it does not make use of some of the information available, namely, the genotype of the suspect. The correct procedure, we believe was described by Evett et al. (1991)” - LR
DNA Advisory Board
DNA Advisory Board Standards (1998)

- DAB created by the DNA Identification Act 1994
  - Staffed and implemented 2005

- QAS Standards for DNA Testing Laboratories
  - Implemented October 1 2008

- No substantive changes from TWGDAM Guidelines for mixtures
  - 8. Validation
    - 8.1.2.2 Species specificity, sensitivity, stability and mixture studies are conducted
  - 9. Analytical Procedures
    - 9.1.3 The laboratory shall have a procedure for differential extraction of stains that potentially contain semen
Statistical and Population Genetics Issues Affecting the Evaluation of the Frequency of Occurrence of DNA Profiles Calculated From Pertinent Population Database(s)

DNA Advisory Board
February 23, 2000
Mixtures are DNA samples derived from two or more contributors. Evidenced typically by the presence of three or more peaks, bands, dots, and/or notable differences in intensities of the alleles for at least one locus in the profile. In some situations, elucidation of a contributor profile is straightforward (e.g., DNA from an intimate swab revealing a mixture consistent with the composition of the perpetrator and the victim). When intensity differences are sufficient to identify the major contributor in the mixed profile, it can be treated statistically as a single source sample. At times, when alleles are not masked, a minor contributor to the mixed profile may be elucidated. Almost always in a mixture interpretation, certain possible genotypes can be excluded. It may be difficult to be confident regarding the number of contributors in some complex mixtures of more than two individuals; however, the number of contributors often can be inferred by reviewing the data at all loci in a profile.
When the contributors of a DNA mixture profile cannot be distinguished, two calculations convey the probative value of the evidence:

- The probability of exclusion (PE) provides an estimate of the portion of the population that has a genotype composed of at least one allele not observed in the mixed profile.
  - Knowledge of the accused and/or victim profiles is not used (or needed) in the calculation.
  - Useful in complex mixtures, because it requires no assumptions about the identity or number of contributors to a mixture.
  - The probabilities derived are valid and for all practical purposes are conservative. However, the PE does not make use of all of the available genetic data.
The Likelihood Ratio (LR) provides the odds ratio of two competing hypotheses, given the evidence.

A case of sexual assault for which the victim reported there were two assailants. A mixture of two profiles is observed in the "male fraction," and the victim is excluded as a contributor of the observed mixed profile. Two men are arrested, and their combined profiles are consistent with the mixture evidence.

A LR calculation logically might compare the probability that the two accused individuals are the source of the DNA in the evidence versus two unknown (random men) are the source of the evidence. Various alternate hypotheses can be entertained as deemed appropriate, given the evidence.

LR considers the identity and actual number of contributors to the observed DNA mixture.

LR makes better use of the available genetic data than does PE.
Short Tandem Repeat (STR) Interpretation Guidelines

Scientific Working Group on DNA Analysis Methods (SWGDAM)

Read about...

- Introduction
- 1. Preliminary Evaluation of Data
- 2. Designation
- 3. Interpretation of Results
- 4. Conclusions
STR Interpretation Guidelines-SWGDAM 2000

3. Interpretation of Results

- **3.1.1. Single Contributor**
  - when the observed number of alleles at each locus and the signal intensity ratios of alleles at a locus are consistent with a profile from a single contributor
  - all loci should be evaluated in making this determination

- **3.1.2. Mixtures With Major/Minor Contributors**
  - if there is a distinct contrast in signal intensities among the alleles. The difference is evaluated on a case-by-case context. All loci should be evaluated in making this determination

- **3.1.3. Mixtures With a Known Contributor(s)**
  - when one of the contributors (e.g., the victim) is known, the genetic profile of the unknown contributor may be inferred.
  - This can be accomplished by subtracting the contribution of the known donor from the mixed profile

- **3.1.4. Mixtures With Indistinguishable Contributors**
  - When major or minor contributors cannot be distinguished because of similarity in signal intensities or the presence of shared or masked alleles, individuals may still be included or excluded as possible contributors
5. Statistical Interpretation

5.2. The formulas used in calculating the frequency of a DNA profile should be defined for the following:

- 5.2.5. Mixture calculations

• BUT HOW DO WE PERFORM THE CALCULATIONS?
  - SILENCE IS GOLDEN?
DNA commission of the International Society of Forensic Genetics: Recommendations on the interpretation of mixtures

P. Gill\textsuperscript{a,*}, C.H. Brenner\textsuperscript{b}, J.S. Buckleton\textsuperscript{c}, A. Carracedo\textsuperscript{d}, M. Krawczak\textsuperscript{e}, W.R. Mayr\textsuperscript{f}, N. Morling\textsuperscript{g}, M. Prinz\textsuperscript{h}, P.M. Schneider\textsuperscript{i}, B.S. Weir\textsuperscript{j}

\textsuperscript{a}Forensic Science Service, Trident Court, 2960 Solihull Parkway, Birmingham, UK
\textsuperscript{b}Forensic Science Group, School of Public Health, University of California, Berkeley, CA 510-339-1911, USA
\textsuperscript{c}ESR, Private Bag 92021, Auckland, New Zealand
\textsuperscript{d}Institute of Legal Medicine, Faculty of Medicine, University of Santiago de Compostela, 15705 Santiago de Compostela, Spain
\textsuperscript{e}Institute of Medical Informatics and Statistics, Kiel, Germany
\textsuperscript{f}Division of Blood Group Serology, Medical University of Vienna, Austria
\textsuperscript{g}Department of Forensic Genetics, Institute of Forensic Medicine, University of Copenhagen, Copenhagen, Denmark
\textsuperscript{h}Office of the Chief Medical Examiner, Department of Forensic Biology, 520 First Avenue, New York, NY 10016, USA
\textsuperscript{i}Institute of Legal Medicine University Clinic of Cologne, Melatengürtel, 60-62 D 50823 Köln, Germany
\textsuperscript{j}University of Washington, Department of Biostatistics, Box 357232, Seattle, WA 98195, USA

Received 4 April 2006; accepted 10 April 2006
Available online 5 June 2006

Summary of ISFG Recommendations on Mixture Interpretation

1. The likelihood ratio (LR) is the preferred statistical method for mixtures over RMNE

2. Scientists should be trained in and use LRs

3. Methods to calculate LRs of mixtures are cited

4. Follow Clayton et al. (1998) guidelines when deducing component genotypes

5. Prosecution determines $H_p$ and defense determines $H_d$ and multiple propositions may be evaluated

6. When minor alleles are the same size as stutters of major alleles, then they are indistinguishable

7. Allele dropout to explain evidence can only be used with low signal data

8. No statistical interpretation should be performed on alleles below threshold

9. Stochastic effects limit usefulness of heterozygote balance and mixture proportion estimates with low level DNA
“…These recommendations have been written to serve two purposes: to define a generally acceptable mathematical approach for typical mixture scenarios and to address open questions where practical and generally accepted solutions do not yet exist. This has been done to stimulate the discussion among scientists in this field. The aim is to invite proposals and criticism in the form of comments and letters to the editors of this journal…We are hoping to continue the process to allow the DNA Commission to critically revise or extend these recommendations in due time…”
Responses to ISFG DNA Commission Mixture Recommendations

- **UK Response**
  - Gill et al. (2008) *FSI Genetics* 2(1): 76–82

- **German Stain Commission**
  - Schneider et al. (2006) *Rechtsmedizin* 16:401-404 (German version)

- **ENFSI Policy Statement**

- **Australia/New Zealand Support Statement**
  - Stringer et al. (2009) *FSI Genetics* 3: 144-145
## NIST MIX 05 Study

<table>
<thead>
<tr>
<th>LabID</th>
<th>Kits Used</th>
<th>Case1</th>
<th>Case2</th>
<th>Case3</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Caucasians</td>
<td>African Americans</td>
<td>Hispanics</td>
</tr>
<tr>
<td>90</td>
<td>ProPlus/Cofiler</td>
<td>1.18E+15</td>
<td>2.13E+14</td>
<td>3.09E+15</td>
</tr>
<tr>
<td>34</td>
<td>ProPlus/Cofiler</td>
<td>2.40E+11</td>
<td>7.00E+09</td>
<td>9.80E+10</td>
</tr>
<tr>
<td>33</td>
<td>ProPlus/Cofiler</td>
<td>2.94E+08</td>
<td>1.12E+08</td>
<td>1.74E+09</td>
</tr>
<tr>
<td>6</td>
<td>ProPlus/Cofiler</td>
<td>40,000,000</td>
<td>3,500,000</td>
<td>280,000,000</td>
</tr>
<tr>
<td>9</td>
<td>ProPlus/Cofiler</td>
<td>1.14E+07</td>
<td>1.97E+07</td>
<td>1.54E+08</td>
</tr>
<tr>
<td>79</td>
<td>ProPlus/Cofiler</td>
<td>930,000</td>
<td>47,900</td>
<td>1,350,000</td>
</tr>
<tr>
<td>16</td>
<td>ProPlus/Cofiler</td>
<td>434,600</td>
<td>31,710</td>
<td>399,100</td>
</tr>
</tbody>
</table>

**same epg: tremendous Inter-Lab variance**
July 2009 Rev. Quality Assurance Standards (QAS)

QAS Standard 5.3.2

A casework CODIS administrator shall be or have been a current or previously qualified DNA analyst ... with documented mixture interpretation training.

QAS Standard 8.3.1

Internal validation studies conducted after the date of this revision shall include as applicable: known and non-probative evidence samples or mock evidence samples, reproducibility and precision, sensitivity and stochastic studies, mixture studies, and contamination assessment.

QAS Standard 8.3.2

Internal validation shall define quality assurance parameters and interpretation guidelines, including as applicable, guidelines for mixture interpretation.

QAS Standard 9.6.4

Laboratories analyzing forensic samples shall have and follow a documented procedure for mixture interpretation that addresses major and minor contributors, inclusions and exclusions, and policies for the reporting of results and statistics.
The Scientific Working Group on DNA Analysis Methods (SWGDAM), is a group of approximately 50 scientists representing federal, state, and local forensic DNA laboratories in the United States and Canada. During meetings, which are held twice a year, subcommittees discuss topics of interest to the forensic DNA community and often develop documents to provide direction and guidance for the community. A mixture interpretation subcommittee was formed in January 2007 and worked for several years to provide a guidance document on autosomal short tandem repeat (STR). This document was presented to the full SWGDAM group and received approval in January 2010.

This document provides guidelines for the interpretation of DNA typing results from short tandem repeats (STR) and supersedes the Scientific Working Group on DNA Analysis Methods (SWGDAM) Short Tandem Repeat (STR) Interpretation Guidelines (2000). The revised guidelines are not intended to be applied retroactively. Guidance is provided for forensic casework analyses on the identification and application of thresholds for allele detection and interpretation, and appropriate statistical approaches to the interpretation of autosomal STRs with further guidance on mixture interpretation. Laboratories are encouraged to review their standard operating procedures and validation data in light of these new guidelines.
Contact Information

Jack Ballantyne
University of Central Florida, Orlando, FL
jballant@mail.ucf.edu