Foundations for Mixture Interpretation: ISFG and SWGDAM

John M. Butler
National Institute of Standards and Technology

Baltimore, MD
April 12, 2011
Elements of DNA Mixture Interpretation

**Principles** (theory)

**Protocols** (validation)

**Practice** (training & experience)

Consistency across analysts

Periodic training will aid accuracy and efficiency within your laboratory.

ISFG Recommendations

SWGDAM Guidelines

Your Laboratory

SOPs

Training within Your Laboratory
SWGDAM STR Interpretation Guidelines

• SWGDAM approved on January 14, 2010

• Publicly released **April 8, 2010** on the FBI website for the CODIS group

SWGDAM Interpretation Guidelines for Autosomal STR Typing by Forensic DNA Testing Laboratories

- **Guidelines**
  - Not Standards
  - No lab should be audited against this document

- **Autosomal STR Typing**
  - This document does not address Y-STRs, mitochondrial DNA testing, or CODIS entries

- **Forensic DNA Testing Laboratories**
  - Databasing labs may have different issues since they are working with known single source samples
Members of SWGDAM Mixture Committee over the time period of Jan 2007 to Jan 2010

- **John Butler** (NIST) – chair
- **Mike Adamowicz** (CT)
- **Terry Coons** (OR)
- Jeff Modler (RCMP)
- **Phil Kinsey** (MT)
- Todd Bille (ATF)
- Allison Eastman (NYSP)
- **Bruce Heidebrecht** (MD)
- **Tamyra Moretti** (FBI DNA Unit I)
- **George Carmody** (Carleton U)
- Roger Frappier (CFS-Toronto)
- **Jack Ballantyne** (UCF/NCFS)
- **Gary Sims** (CA DOJ) - co-chair
- Joanne Sgueglia (MA)
- **Gary Shutler** (WA)
- Cecelia Crouse (PBSO)
- **Hiron Poon** (RCMP)
- Steve Lambert (SC)
- **Steven Myers** (CA DOJ)
- Ann Gross (MN BCA)

The 15 members in bold font were involved with most of the writing (July-Oct 2009)
Due to the multiplicity of forensic sample types and the potential complexity of DNA typing results, it is impractical and infeasible to cover every aspect of DNA interpretation by a preset rule. However, the laboratory should utilize written procedures for interpretation of analytical results with the understanding that specificity in the standard operating protocols will enable greater consistency and accuracy among analysts within a laboratory.

Overview of these SWGDAM Guidelines

1. Preliminary evaluation of data – *is something a peak and is the analysis method working properly?*
2. Allele designation – *calling peaks as alleles*
3. Interpretation of DNA typing results – *using the allele information to make a determination about the sample*
   1. Non-allelic peaks
   2. Application of peak height thresholds to allelic peaks
   3. Peak height ratio
   4. Number of contributors to a DNA profile
   5. Interpretation of DNA typing results for mixed samples
   6. Comparison of DNA typing results
4. Statistical analysis of DNA typing results – *assessing the meaning (rarity) of a match*

*Other supportive material: statistical formulae, references, and glossary*
Steps in DNA Interpretation

Data Collection
- Sample Deposited
- Sample Collected
- Extraction
- Quantitation
- PCR Amplification
- CE Separation/Detection

Data Interpretation
- Signal observed
- Peak (vs. noise)
- Allele (vs. artifact)
- Genotype (allele pairing)
- Profile (genotype combining)

All Alleles Detected?
- Genotype(s)
- Contributor profile(s)
- Comparison to Known(s)
- Weight of Evidence (Stats)
Interpretation of Evidence Completed before Comparison to Known(s)

• “3.6.1. The laboratory must establish guidelines to ensure that, to the extent possible, DNA typing results from evidentiary samples are interpreted before comparison with any known samples, other than those of assumed contributors.”

– While the FBI QAS do not address this issue, this is an example of an issue felt by the committee members to be of such importance that it warranted a “must.”
Stats Required for Inclusions

SWGDAM Interpretation Guideline 4.1:
“The laboratory must perform statistical analysis in support of any inclusion that is determined to be relevant in the context of a case, irrespective of the number of alleles detected and the quantitative value of the statistical analysis.”

Buckleton & Curran (2008): “There is a considerable aura to DNA evidence. Because of this aura it is vital that weak evidence is correctly represented as weak or not presented at all.”

All Statistical Approaches Are Considered

Table 1 – Suitable Statistical Analyses for DNA Typing Results
The statistical methods listed in the table cannot be combined into one calculation. For example, combining RMP at one locus with a CPI calculation at a second locus is not appropriate. However, an RMP may be calculated for the major component of a mixture and a CPE/CPI for the entire mixture (as referred to in section 4.6.2).

<table>
<thead>
<tr>
<th>Category of DNA Typing Result</th>
<th>RMP</th>
<th>CPE/CPI</th>
<th>LR (1)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Single Source</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Single Major Contributor to a Mixture</td>
<td>✓</td>
<td></td>
<td>✓</td>
</tr>
<tr>
<td>Multiple Major Contributors to a Mixture</td>
<td>✓ (2)</td>
<td>✓ (2)</td>
<td>✓</td>
</tr>
<tr>
<td>Single Minor Contributor to a Mixture</td>
<td>✓</td>
<td>✓ (3)</td>
<td>✓</td>
</tr>
<tr>
<td>Multiple Minor Contributors to a Mixture</td>
<td>✓ (2)</td>
<td>✓ (3)</td>
<td>✓</td>
</tr>
<tr>
<td>Indistinguishable Mixture</td>
<td>✓ (1)</td>
<td>✓</td>
<td>✓</td>
</tr>
</tbody>
</table>

(1) Restricted or unrestricted
(2) Restricted
(3) All potential alleles identified during interpretation are included in the statistical calculation

Unrestricted vs. Restricted

Use of peak height information to select only certain combinations

Unrestricted

All combinations of alleles are deemed possible (relative peak height differences are not utilized)

\[ AB + AC + AD + BC + BD + CD \]

Restricted

Based on relative peak heights, alleles are paired only where specific combinations of alleles are deemed possible

\[ AB + AC + AD + BC + BD + CD \]

Glossary with 46 Defined Terms

Glossary for this document

Allelic dropout: failure to detect an allele within a sample or failure to amplify an allele during PCR.

Analytical threshold: the minimum height requirement at and above which detected peaks can be reliably distinguished from background noise; peaks above this threshold are generally not considered noise and are either artifacts or true alleles.

Artifact: a non-allelic product of the amplification process (e.g., stutter, non-templated nucleotide addition, or other non-specific product), an anomaly of the detection process (e.g., pull-up or spike), or a by-product of primer synthesis (e.g., “dye blob”).

Coincidental match: a match which occurs by chance.

Composite profile: a DNA profile generated by combining typing results from different loci obtained from multiple injections of the same amplified sample and/or multiple amplifications of the same DNA extract. When separate extracts from different locations on a given evidentiary item are combined prior to amplification, the resultant DNA profile is not considered a composite profile.

Your Laboratory Interpretation Protocols

**Standard Operating Procedures (SOPs)**

- Validation studies
- Literature
- Experience

**SWGDAM Guidelines (2010) Introduction:** ...the laboratory should utilize written procedures for interpretation of analytical results with the understanding that specificity in the standard operating protocols will enable greater consistency and accuracy among analysts within a laboratory. It is recommended that standard operating procedures for the interpretation of DNA typing results be sufficiently detailed that other forensic DNA analysts can review, understand in full, and assess the laboratory’s policies and practices. The laboratory’s interpretation guidelines should be based upon validation studies, scientific literature, and experience.
See provided reference list with over 100 relevant references for further information on each topic discussed in today’s workshop (4 articles along with the SWGDAM Guidelines have also been included in the handout)
Revised Quality Assurance Standard
Requirement for Literature Review

5.1.3.2. The laboratory shall have a program approved by the technical leader for the annual review of scientific literature that documents the analysts’ ongoing reading of scientific literature. The laboratory shall maintain or have physical or electronic access to a collection of current books, reviewed journals, or other literature applicable to DNA analysis.
Useful Articles on DNA Mixture Interpretation


*Articles in bold font are included in the workshop handouts*
German Mixture Classification Scheme


(German Stain Commission, 2006):

- **Type A**: no obvious major contributor, no evidence of stochastic effects
- **Type B**: clearly distinguishable major and minor contributors; consistent peak height ratios of *approximately 4:1* (major to minor component) for all heterozygous systems, no stochastic effects
- **Type C**: mixtures without major contributor(s), evidence for stochastic effects

“Indistinguishable”  “Distinguishable”  “Uninterpretable”
Our discussions have highlighted a significant need for continuing education and research into this area.
In general we agree with the recommendations of Gill et al. that are:
(i) when possible peak height/area should be included in mixture interpretation; (ii) stutter position peaks at similar peak height/area as that of obligate minor contributor alleles should be considered as potential alleles in the interpretation and statistics calculation; and (iii) a stochastic threshold (termed “dropout threshold”) should be defined.
ISFG Recommendations on Mixture Interpretation

http://www.isfg.org/Publication;Gill2006

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

• New Zealand/Australia Support Statement
  – Stringer et al. (2009) FSI Genetics 3(2):144-145

• **SWGDAM – Interpretation Guidelines**
  – Approved Jan 2010 and released April 2010 on FBI website
Steps in DNA Interpretation

Data Collection
- Sample Deposited
- Sample Collected
- Extraction
- Quantitation

PCR
- Amplification

CE
- Separation/Detection

Signal observed

Data Interpretation
- Peak (vs. noise)
- Allele (vs. artifact)
- Genotype (allele pairing)
- Profile (genotype combining)

All Alleles Detected?
- Genotype(s)
- Contributor profile(s)

Comparison to Known(s)
- Weight of Evidence (Stats)
# Principles Behind Thresholds

<table>
<thead>
<tr>
<th>Thresholds</th>
<th>Principles Behind</th>
</tr>
</thead>
<tbody>
<tr>
<td>Analytical Threshold (e.g., 50 RFU)</td>
<td>Below this value, observed peaks cannot be reliably distinguished from instrument noise (baseline signal)</td>
</tr>
<tr>
<td>Limit of Linearity (e.g., 5000 RFU)</td>
<td>Above this value, the CCD camera can become saturated and peaks may not accurately reflect relative signal quantities (e.g., flat-topped peaks) and lead to pull-up/bleed-through between dye color channels</td>
</tr>
<tr>
<td>Stochastic Threshold (e.g., 250 RFU)</td>
<td>Above this peak height value, it is reasonable to assume that allelic dropout of a sister allele of a heterozygote has not occurred at that locus; single alleles above this value in single-source samples are assumed to be homozygous</td>
</tr>
<tr>
<td>Stutter Threshold (e.g., 15%)</td>
<td>Below this value, a peak in the reverse (or forward) stutter position can be designated as a stutter artifact with single-source samples or some mixtures (often higher with lower DNA amounts)</td>
</tr>
<tr>
<td>Peak Height Ratio (e.g., 60%)</td>
<td>Above this value, two heterozygous alleles can be grouped as a possible genotype (often lower with lower DNA amounts)</td>
</tr>
<tr>
<td>Major/Minor Ratio (e.g., 4:1)</td>
<td>When the ratio of contributors is closer than this value in a two-person mixture, it becomes challenging and often impossible to correctly associate genotype combinations to either the major or minor contributor</td>
</tr>
</tbody>
</table>
# Threshold Decisions

<table>
<thead>
<tr>
<th>Thresholds to Determine</th>
<th>Decisions to Make (lab &amp; kit specific)</th>
<th>Useful Validation Data</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Analytical = ____ RFU</strong></td>
<td>Single overall value or color specific</td>
<td>Noise levels in negative controls or non-peak areas of positive controls</td>
</tr>
<tr>
<td><strong>Stochastic = ____ RFU</strong></td>
<td>Minimum peak height RFU value or alternative criteria such as quantitation values or use of a probabilistic genotype approach</td>
<td>Level where dropout occurs in low level single-source heterozygous samples under conditions used (e.g., different injection times, post-PCR cleanup)</td>
</tr>
<tr>
<td><strong>Stutter filter = ____%</strong></td>
<td>Profile, locus, or allele-specific</td>
<td>Stutter in single-source samples (helpful if examined at multiple DNA quantities)</td>
</tr>
<tr>
<td><strong>Peak Height Ratio = ____%</strong></td>
<td>Profile, locus, or signal height (quantity) specific</td>
<td>Heterozygote peak height ratios in single-source samples (helpful if examined at multiple DNA quantities)</td>
</tr>
<tr>
<td><strong>Major/Minor Ratio = ____</strong></td>
<td>When will you attempt to separate components of a mixture into major and minor contributors for profile deductions?</td>
<td>Defined mixture ratios (e.g., 1:1, 1:3, 1:9) with known samples to observe consistency across loci and to assess ability to deduce correct contributor profiles</td>
</tr>
</tbody>
</table>
An Example...
Case Example #3

D8S1179  D21S11  D7S820  CSF1PO

D3S1358  TH01  D13S317  D16S539  D2S1338

D19S433  vWA  TPOX  D18S51

Amelogenin  D5S818  FGA

AT = 30 RFU
ST = 150 RFU
Stutter filter off

125 pg total DNA
Impact of Results with Low Level DNA

When amplifying low amounts of DNA (e.g., 125 pg), allele dropout is a likely possibility leading to higher uncertainty in the potential number of contributors and in the possible genotype combinations.
Case Example #3

Identifiler
125 pg total DNA

AT = 30 RFU
ST = 150 RFU
Stutter filter off

Peaks below stochastic threshold

TPOX
D5S818
D18S51

5 alleles

y-axis zoom to 100 RFU
What Can We Say about this Result?

• Low level DNA (only amplified 125 pg total DNA)
  – likely to exhibit stochastic effects and have allele dropout

• Mixture of at least 3 contributors
  – Based on detection of 5 alleles at D18S51
  – If at equal amounts, ~40 pg of each contributor (if not equal, then less for the minor contributors); we expect allele dropout

• At least one of the contributors is male
  – Based on presence of Y allele at amelogenin

• Statistics if using CPI/CPE
  – Would appear that we can only use TPOX and D5S818 results with a stochastic threshold of 150 RFU (will explore this further)

• Due to potential of excessive allele dropout, we are unable to perform any meaningful Q-K comparisons
Uncertainty in the Potential Number of Contributors with this Result

- Several of the peaks are barely above the analytical threshold of 30 RFU. In fact, with an analytical threshold of 50 RFU or even 35 RFU, there would only be three detected alleles at D18S51.

- Stochastic effects could result in a high degree of stutter off of the 17 allele making alleles 16 and 18 potential stutter products.

- No other loci have >4 alleles detected.
All Detected Alleles Are Above the Stochastic Threshold – Or Are They?

Does this result guarantee no allele drop-out?

We have assumed three contributors. If result is from an equal contribution of 3 individuals…

Then some alleles from individual contributors would be below the stochastic threshold and we could not assume that all alleles are being observed!
Assuming Three Contributors…
Some Possible Contributions to This Result

1:1:1

3:1:1

Stochastic alert!
All Loci Are Not Created Equal when it comes to mixture interpretation

- In the case of less polymorphic loci, such as TPOX, there are fewer alleles and these occur at higher frequency. Thus, there is a greater chance of allele sharing (peak height stacking) in mixtures.

- **Higher locus heterozygosity is advantageous for mixture interpretation** – we would expect to see more alleles (within and between contributors) and thus have a better chance of estimating the true number of contributors to the mixture.
Even if you did attempt to calculate a CPI/CPE statistic using loci with all observed alleles above the stochastic threshold on this result...

**TPOX Allele Frequencies** (NIST Caucasian, Butler et al. 2003)
- 8 = 0.53
- 11 = 0.24
- CPI = \((0.53 + 0.24)^2 = 0.59\) or **59%**

Combine loci = 0.59 x 0.18 = 0.11 or **11%**

**D5S818 Allele Frequencies** (NIST Caucasian, Butler et al. 2003)
- 10 = 0.05
- 12 = 0.38
- CPI = \((0.05 + 0.38)^2 = 0.18\) or **18%**

*Approximately 1 in every 9 Caucasians could be included in this mixture*
Impact of Amplifying More DNA

D19S433

125 pg total DNA amplified

Allele 12 is missing

D19S433

500 pg total DNA amplified

True Contributors
3 contributors with a 2:1:1 mixture

15,15 (2x)
14,15 (1x)
12,14 (1x)
How should you handle the suspect comparison(s) with this case result?

• **No suspect comparisons should be made** as the mixture result has too much uncertainty with stochastic effects that may not account for all alleles being detected

• **Declare the result “inconclusive”**
How not to handle this result

• “To heck with the analytical and stochastic thresholds”, I am just going to see if the suspect profile(s) can fit into the mixture allele pattern observed – and then if an allele is not present in the evidentiary sample try to explain it with possible allele dropout due to stochastic effects

• This is what Bill Thompson calls “painting the target around the arrow (matching profile)…”

Value of Using a Profile Interpretation Worksheet

**PROFILE INTERPRETATION WORKSHEET**

**IDENTIFIER**

**PROFILE NAME:** Case Example #3  
**ANALYST:** John Butler  
**DATE:** 11 October 2010  
**MIXTURE:** □ yes □ no □ unsure

**Allele and Locus Assessments**

| ID LOCUS | Alleles called | Alleles above Stochastic Threshold | Stutter or other peaks to consider | Possible allele dropout? | Y/N | Stochastic issues? (e.g., elevated stutter, PHR imbalance, drop-in, etc.) | Y/N | Degradation / Inhibition (obvious)? | Y/N | If mixture, restricted genotypes can be used? | Y/N | Can this locus be interpreted? | Y/N | Additional Comments |
|----------|----------------|-----------------------------------|-----------------------------------|--------------------------|-----|--------------------------------------------------------------------|-----|----------------------------------|-----|------------------------------------------|-----|-------------------------------|-----|
| D8S1179  | 11,13,16       | 13                                | Maybe                             | Y                         | Y/N | Y                                                                  | Y/N | N                                | N   | N                                      | N   | N                             | N   |

**Analytical threshold:** 30 RFU  
**Stutter % used:** 0% (**filter turned-off**)  
**Stochastic threshold:** 150 RFU  
**Peak height ratio:** 60%  
**Comments:** low level DNA (125 pg)

Make decisions on the evidentiary sample and document them prior to looking at the known(s) for comparison purposes.
What to do with low level DNA mixtures?

• **German Stain Commission “Category C”**  
  (Schneider et al. 2006, 2009)  
  – Cannot perform stats because stochastic effects make it uncertain that all alleles are accounted for

• **ISFG Recommendations #8 & #9**  
  (Gill et al. 2006)  
  – Stochastic effects limit usefulness

• **Fundamentals of Forensic DNA Typing** (2010)  
  Butler 3rd edition (volume 1), chapter 18  
  – Don’t go “outside the box” without supporting validation
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

A Complexity/Uncertainty Threshold

New Scientist article (August 2010)

• How DNA evidence creates victims of chance
  – 18 August 2010 by Linda Geddes

• From the last paragraph:
  – In really complex cases, analysts need to be able to draw a line and say "This is just too complex, I can't make the call on it," says Butler. "Part of the challenge now, is that every lab has that line set at a different place. But the honest thing to do as a scientist is to say: I'm not going to try to get something that won't be reliable."

Summary

• Do not blindly use a stochastic threshold with complex mixtures as assumptions regarding the number of contributors can impact interpretation

• Going back to try and get a better sample from the evidence (if available) is wiser than spending a lot of time trying to work with a poor quality DNA result
Future of Complex, Low-level Mixtures

• If you want to work in this area, you need supporting validation data (collecting a few results at high DNA levels and extrapolating to greater complexity and smaller amounts of DNA will not be sufficient)

• Recent efforts in Europe are focused on modeling uncertainty through probabilistic genotype approaches

• Will require software to perform all of the calculations
STRBase Mixture Section

http://www.cstl.nist.gov/biotech/strbase/mixture.htm

Section launched in October 2010 and will continue to develop over time

• Updated literature lists by topic

• Workshop slides and links to other info

• Useful freeware programs (e.g., Excel macros) will be available for download
Acknowledgments

• **SWGDAM Mixture Committee members** for their hard work through many long hours of discussing and writing these guidelines
  
  – Ted Staples for his support as SWGDAM chair
  – Bruce Heidebrecht (for some of the slides)

• NIJ Funding to our NIST Group through NIST OLES interagency agreement 2008-DN-R-121

http://www.cstl.nist.gov/biotech/strbase/training.htm

john.butler@nist.gov

301-975-4049