DNA Mixture Analysis:

Principles and Practice of Mixture Interpretation and Statistical Analysis Using the SWGDAM STR Interpretation Guidelines

---

SWGDAM Guidelines and Mixture Literature

John M. Butler

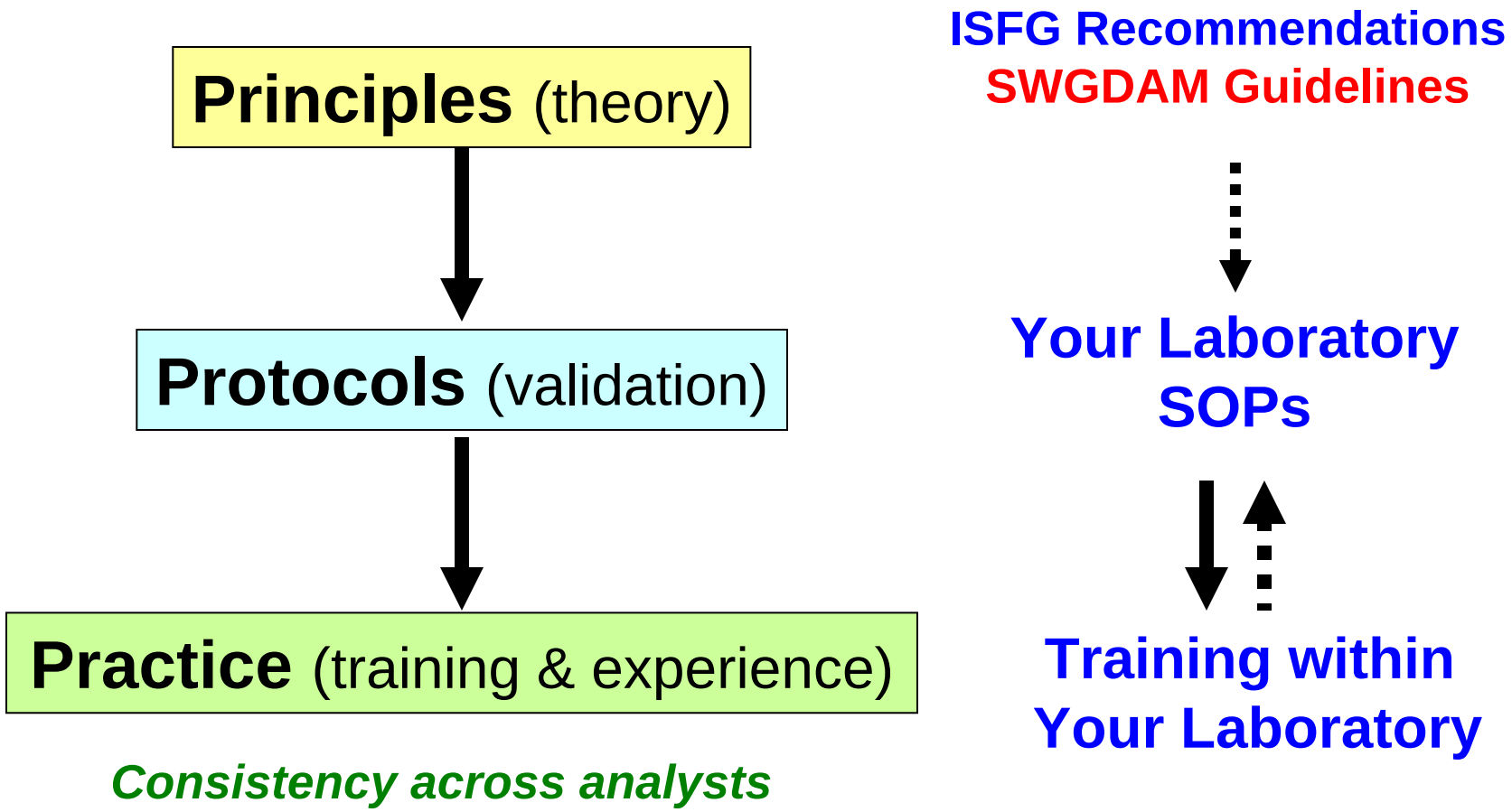
NIST

AAFS 2011 Workshop #17

Chicago, IL

February 22, 2011

Elements of DNA Mixture Interpretation



Periodic training will aid accuracy and efficiency 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

http://www.fbi.gov/about-us/lab/codis/swgdam-interpretation-guidelines

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)

Purpose and Scope of Document

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

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**.

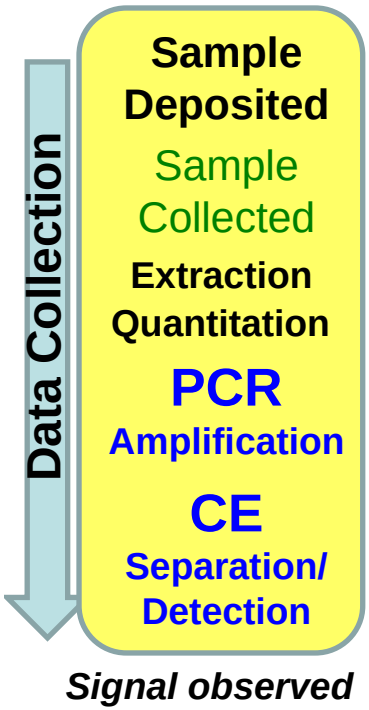
http://www.fbi.gov/about-us/lab/codis/swgdam-interpretation-guidelines

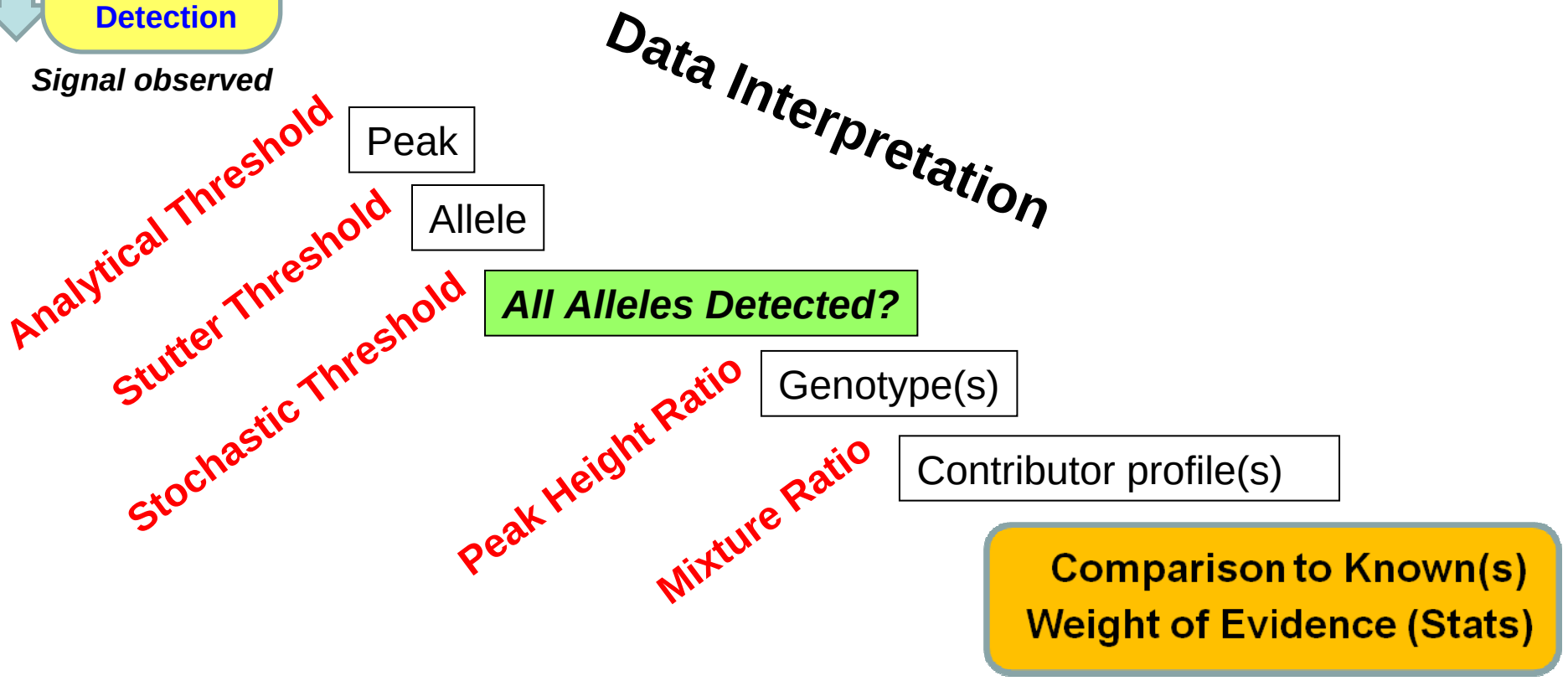
Overview of these SWGDAM Guidelines

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





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**."

Buckleton, J. and Curran, J. (2008) A discussion of the merits of random man not excluded and likelihood ratios. *Forensic Sci. Int. Genet.* 2: 343-348.

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).

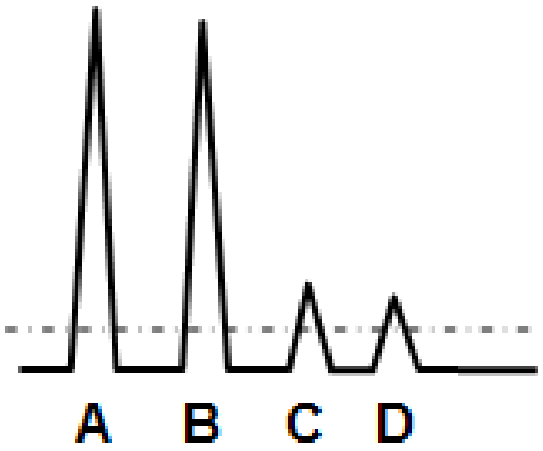
Category of DNA Typing Result	RMP	CPE/CPI	LR (1)
Single Source	✔		✔
Single Major Contributor to a Mixture	✔		✔
Multiple Major Contributors to a Mixture	✔ (2)	✔ (2)	✔
Single Minor Contributor to a Mixture	✔	✔ (3)	✔
Multiple Minor Contributors to a Mixture	✔ (2)	✔ (3)	✔
Indistinguishable Mixture	✔ (1)	✔	✔

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

http://www.fbi.gov/about-us/lab/codis/swgdam-interpretation-guidelines

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

http://www.fbi.gov/hq/lab/html/codis_swgdam.pdf

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.

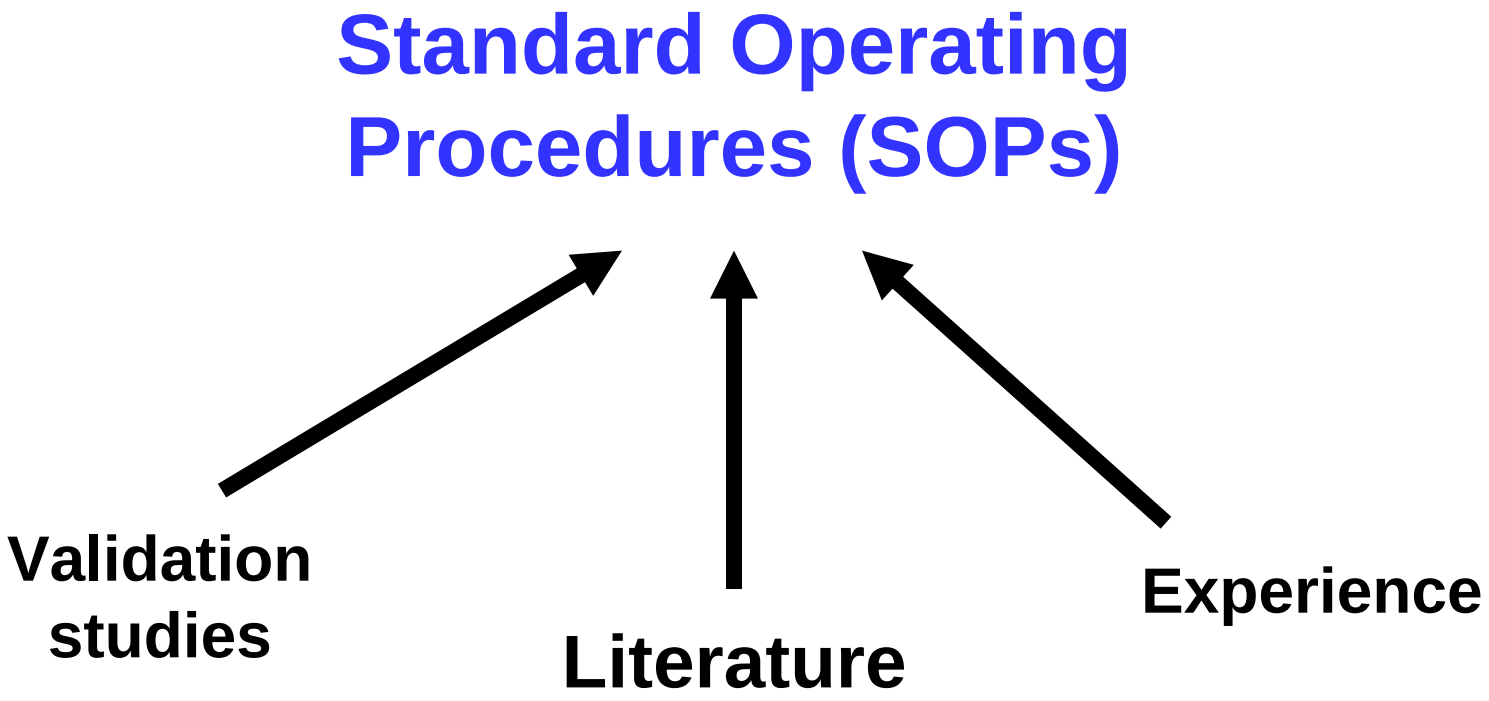
Artifact: a non-allelic product of the amplification process (e.g., stutter, non-templated nucleotide addition, or other non-specific product), an anomaly of the detection process (e.g., pull-up or spike), or a by-product of primer synthesis (e.g., "dye blob").

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.

http://www.fbi.gov/about-us/lab/codis/swgdam-interpretation-guidelines

Your Laboratory Interpretation Protocols



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.*

The Mixture Literature

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

Quality Assurance Standards for Forensic DNA Testing Laboratories
(effective July 1, 2009)

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**.

http://www.fbi.gov/hq/lab/fsc/backissu/oct2008/standards/2008_10_standards01b.htm

Useful Articles on DNA Mixture Interpretation

- **Buckleton, J.S. and Curran, J.M. (2008) A discussion of the merits of random man not excluded and likelihood ratios. *Forensic Sci. Int. Genet.* 2: 343-348.**

- Budowle, B., *et al.* (2009) Mixture interpretation: defining the relevant features for guidelines for the assessment of mixed DNA profiles in forensic casework. *J. Forensic Sci.* 54: 810-821.

- Clayton, T.M., *et al.* (1998) Analysis and interpretation of mixed forensic stains using DNA STR profiling. *Forensic Sci. Int.* 91: 55-70.

- **Gill, P., *et al.* (2006) DNA commission of the International Society of Forensic Genetics: Recommendations on the interpretation of mixtures. *Forensic Sci. Int.* 160: 90-101.**

- Gill, P., *et al.* (2008) National recommendations of the technical UK DNA working group on mixture interpretation for the NDNAD and for court going purposes. *FSI Genetics* 2(1): 76–82.

- **Schneider, P.M., *et al.* (2009) The German Stain Commission: recommendations for the interpretation of mixed stains. *Int. J. Legal Med.* 123: 1-5.**

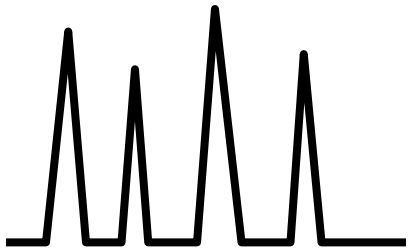
Articles in bold font are included in the workshop handouts

German Mixture Classification Scheme

Schneider *et al*. (2009) *Int. J. Legal Med*. 123: 1-5

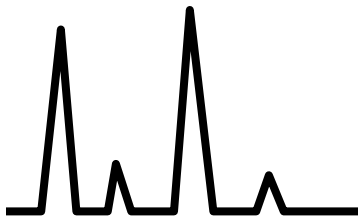
(German Stain Commission, 2006):

- **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



Type A

"Indistinguishable"



Type B

"Distinguishable"

Type C

"Uninterpretable"

SWGDAM

Available for download from the ISFG Website:
http://www.isfg.org/Publication;Gill2006

ELSEVIER

Available online at www.sciencedirect.com
SCIENCE DIRECT®

Forensic Science International 160 (2006) 90–101

Forensic Science International
www.elsevier.com/locate/forsciint

DNA commission of the International Society of Forensic Genetics: Recommendations on the interpretation of mixtures

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

[a] *Forensic Science Service, Trident Court, 2960 Solihull Parkway, Birmingham, UK*
[b] *Forensic Science Group, School of Public Health, University of California, Berkeley, CA 510-339-1911, USA*
[c] *ESR, Private Bag 92021, Auckland, New Zealand*
[j] *University of Washington, Department of Biostatistics, Box 357232, Seattle, WA 98195, USA*

Our discussions have highlighted a significant need for continuing education and research into this area.

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

Gill *et al.* (2006) DNA Commission of the International Society of Forensic Genetics: Recommendations on the interpretation of mixtures. *Forensic Sci. Int.* 160: 90-101

Budowle et al. (2009) Article from the FBI Mixture Committee

J Forensic Sci, May 2009, Vol. 54, No. 3
doi: 10.1111/j.1556-4029.2009.01046.x
Available online at: www.blackwell-synergy.com

Bruce Budowle,[1] Ph.D.; Anthony J. Onorato,[1] M.S.F.S., M.C.I.M.; Thomas F. Callaghan,[1] Ph.D.; Angelo Della Manna,[2] M.S.; Ann M. Gross,[3] M.S.; Richard A. Guerrieri,[1] M.S.; Jennifer C. Luttman,[1] M.F.S.; and David Lee McClure,[4] B.S.

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

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.

Budowle, B., *et al.* (2009) Mixture interpretation: defining the relevant features for guidelines for the assessment of mixed DNA profiles in forensic casework. *J. Forensic Sci.* 54: 810-821.

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

Gill *et al.* (2006) DNA Commission of the International Society of Forensic Genetics: Recommendations on the interpretation of mixtures. *Forensic Sci. Int.* 160: 90-101

Available online at www.sciencedirect.com
SCIENCE DIRECT®
Forensic Science International 160 (2006) 89
Forensic Science International
www.elsevier.com/locate/forsciint

Editorial

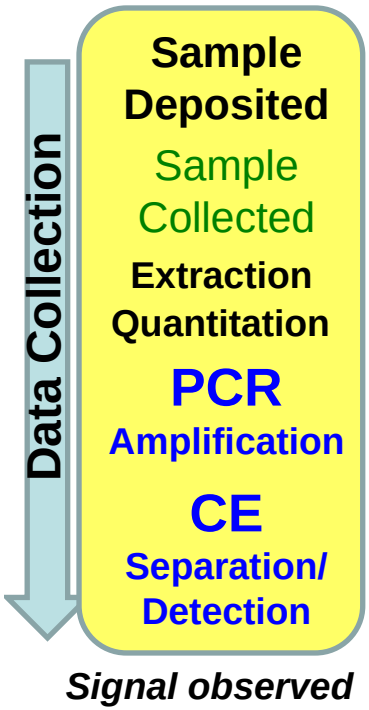
Editorial on the recommendations of the DNA commission of the ISFG on the interpretation of mixtures

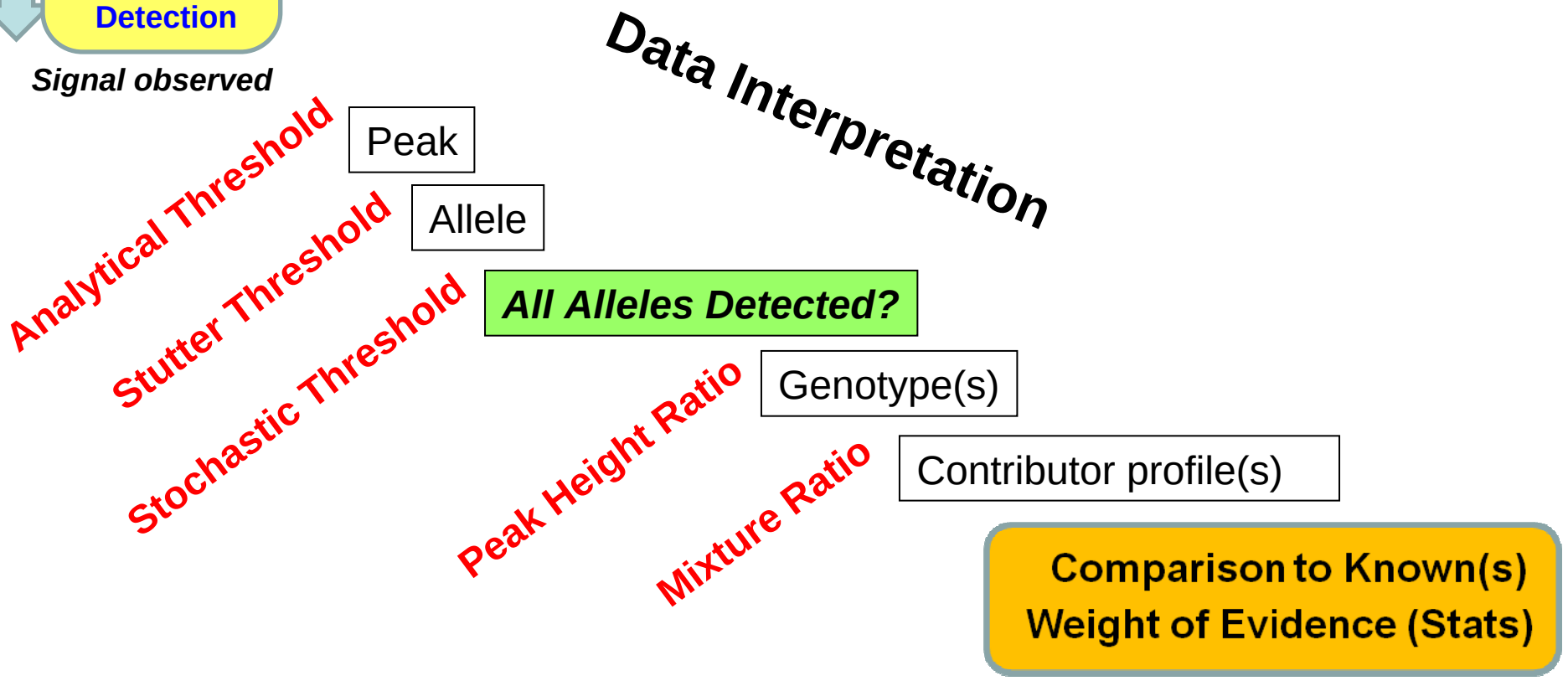
"…**These recommendations have been written** to serve two purposes: to define a generally acceptable mathematical approach for typical mixture scenarios and to address open questions where practical and generally accepted solutions do not yet exist. This has been done **to stimulate the discussion among 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)
  - Schneider *et al.* (2009) *Int. J. Legal Med.* 123: 1-5 (English version)
- ENFSI Policy Statement
  - Morling *et al.* (2007) *FSI Genetics* 1(3):291–292
- 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





Principles Behind Thresholds

Thresholds (example values)	**Principles Behind** *(if properly set based on lab- & kit-specific empirical data)*
Analytical Threshold (e.g., 50 RFU)	Below this value, observed peaks cannot be reliably distinguished from instrument noise (baseline signal)
Limit of Linearity (e.g., 5000 RFU)	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
Stochastic Threshold (e.g., 250 RFU)	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
Stutter Threshold (e.g., 15%)	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)
Peak Height Ratio (e.g., 60%)	Above this value, two heterozygous alleles can be grouped as a possible genotype (often lower with lower DNA amounts)
Major/Minor Ratio (e.g., 4:1)	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

Threshold Decisions

Thresholds to Determine	Decisions to Make (lab & kit specific)	Useful Validation Data
Analytical = _____ RFU	Single overall value or color specific	Noise levels in negative controls or non-peak areas of positive controls
Stochastic = _____ RFU	Minimum peak height RFU value or alternative criteria such as quantitation values or use of a probabilitistic genotype approach	Level where dropout occurs in low level single-source heterozygous samples under conditions used (e.g., different injection times, post-PCR cleanup)
Stutter filter = ___%	Profile, locus, or allele-specific	Stutter in single-source samples (helpful if examined at multiple DNA quantities)
Peak Height Ratio = ___%	Profile, locus, or signal height (quantity) specific	Heterozygote peak height ratios in single-source samples (helpful if examined at multiple DNA quantities)
Major/Minor Ratio = _____	When will you attempt to separate components of a mixture into major and minor contributors for profile deductions?	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

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