Mixture Interpretation Discussion

Florida Statewide Training Meeting

Indian Rocks Beach, FL
May 12-13, 2008

Dr. John M. Butler
National Institute of Standards and Technology

john.butler@nist.gov

NIST and NIJ Disclaimer

Funding: Interagency Agreement 2003-IJ-R-029 between the National Institute of Justice and NIST Office of Law Enforcement Standards

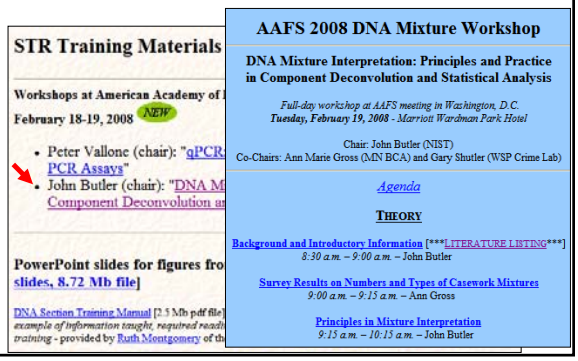
Points of view are mine and do not necessarily represent the official position or policies of the US Department of Justice or the National Institute of Standards and Technology.

Certain commercial equipment, instruments and materials are identified in order to specify experimental procedures as completely as possible. In no case does such identification imply a recommendation or endorsement by the National Institute of Standards and Technology nor does it imply that any of the materials, instruments or equipment identified are necessarily the best available for the purpose.

SWGDAM Disclaimer…

Training Information Available on STRBase

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



AAFS 2008 Workshop Presenters

Ann Marie Gross
MN BCA

John M. Butler
NIST

George Carmody
Carleton University/ Statistical Consultant

Gary Shutler
Wash State Police Crime Lab

Angie Dolph
Marshall University (NIST Summer Intern)

Joanne B. Sgueglia
Mass State Police Crime Lab

Tim Kalafut
US Army Crime Lab

Purpose for Teaching AAFS Workshop

We hope that participants:

- Gain a better understanding of the current approaches being used throughout the community for mixture interpretation
- See worked examples of mixture component deconvolution and statistical analysis
- Come away with ideas to improve your laboratory's interpretation guidelines and training regarding mixtures in forensic casework

AAFS Workshop Morning Agenda - Theory

Background and Introductory Information
8:30 a.m. – 9:00 a.m. – John Butler

Survey Results on Numbers and Types of Casework Mixtures
9:00 a.m. – 9:15 a.m. – Ann Gross

Principles in Mixture Interpretation
9:15 a.m. – 10:15 a.m. – John Butler

***10:15 a.m. – 10:30 a.m.* BREAK**

Strategies for Mixture Deconvolution with Worked Examples
10:30 a.m. – 11:30 a.m. – John Butler

Different Approaches to Statistical Analysis of Mixtures
11:30 a.m. – 12:00 p.m. – George Carmody

***12:00 p.m. – 1:15 p.m.* LUNCH**

Afternoon Agenda – **Practical Application**

Real Case Example – Importance of Properly Stating Your Conclusions
1:15 p.m. – 1:30 p.m. – Gary Shutler

Variability between Labs in Approaches & Mixture Interlaboratory Studies
1:30 p.m. – 2:15 p.m. – John Butler

Validation Studies and Preparing Mixture Interpretation Guidelines
2:15 p.m. – 2:45 p.m. – Joanne Sgueglia

***2:45 p.m. – 3:00 p.m.* BREAK**

Testing of Mixture Software Programs
3:00 p.m. – 3:15 p.m. – Angela Dolph

DNA_DataAnalysis Software Demonstration
3:15 p.m. – 4:00 p.m. – Tim Kalafut

Training Your Staff to Consistently Interpret Mixtures
4:00 p.m. – 4:45 p.m. – Panel Discussion with Ann Gross, Gary Shutler, Joanne Sgueglia

***4:45 p.m. – 5:00 p.m.* – Questions and Answers as needed**

Recent Mixture Workshops
Conducted by John Butler

Helpful feedback obtained from workshop participants

Southern Association of Forensic Scientists (SAFS)
September 11, 2007 (Atlanta, GA)
- **Mixture Interpretation (theory)**
- Along with Software discussion (Rhonda Roby) and demonstration (Tom Overson/Tim Kalafut)
- **33 attendees from 13 different labs**

Northeastern Association of Forensic Scientists (NEAFS)
November 2-3, 2007 (Bolton Landing, NY)
- **The Cutting Edge of DNA Testing: Mixture Interpretation, miniSTRs, and Low Level DNA**
- **42 attendees from 13 different labs**

NEAFS Workshop materials (70 pages) available on STRBase:
http://www.cstl.nist.gov/biotech/strbase/pub_pres/NEAFS2007_CuttingEdgeDNA.pdf

Mixture Basics

From J.M. Butler (2005) *Forensic DNA Typing, 2nd Edition*, p. 154

- Mixtures arise when two or more individuals contribute to the sample being tested.
- Mixtures can be challenging to detect and interpret without extensive experience and careful training. **Even more challenging with poor quality data when degraded DNA is present…**
- Differential extraction can help distinguish male and female components of many sexual assault mixtures. **Y-chromosome markers can help here in some cases…**

More on Mixtures...

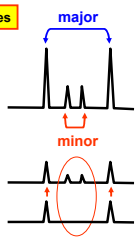
Most mixtures encountered in casework are 2-component mixtures arising from a combination of victim and perpetrator DNA profiles

Ann Gross will discuss some recent collected casework summaries

Torres *et al.* (2003) *Forensic Sci. Int.* 134:180-186 examined 1,547 cases from 1997-2000 containing **2,424 typed samples** of which **163 (6.7%) contained a mixed profile** with **only 8 (0.3%) coming from more than two contributors**

95.1% (155/163) were 2-component mixtures

Ratios of the various mixture components stay fairly constant between multiple loci enabling deduction of the profiles for the major and minor components

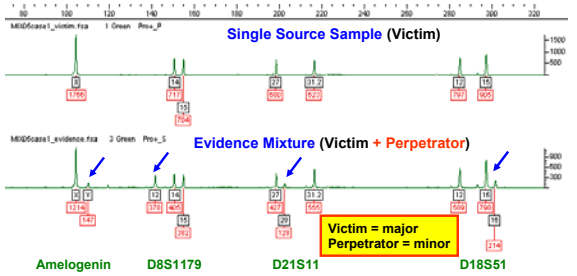


Some mixture interpretation strategies involve using victim (or other reference) alleles to help isolate obligate alleles coming from the unknown portion of the mixture

http://www.cstl.nist.gov/biotech/strbase/interlab/MIX05.htm

MIX05 Case #1; Profiler Plus green loci

Example Mixture Data (MIX05 Study-Profiler Plus)



Obligate Alleles (not present in the victim reference)

Y	12	28	16

True "Perpetrator" Profile

X,Y	12,12	28,31.2	15,16

Sources of DNA Mixtures

- **Two (or more) individuals** contribute to the biological evidence examined in a forensic case (e.g., sexual assault with victim and perpetrator or victim, consensual sexual partner, and perp)

  Victim Reference and Spouse or Boyfriend Reference
- **Contamination** of a single source sample from
  - evidence collection staff
  - laboratory staff handling the sample
  - Low-level DNA in reagents or PCR tubes or pipet tips

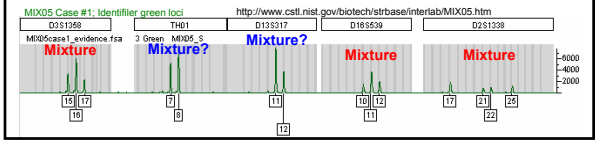
  Examine Staff Profiles (Elimination Database), etc.

Reference elimination samples are useful in deciphering both situations due to possibility of intimate sample profile subtraction

Mixtures: Issues and Challenges

From J.M. Butler (2005) *Forensic DNA Typing, 2nd Edition*, p. 155

- The probability that a mixture will be detected improves with the use of more loci and genetic markers that have a high incidence of heterozygotes.
- The detectability of multiple DNA sources in a single sample relates to the ratio of DNA present from each source, the specific combinations of genotypes, and the total amount of DNA amplified.
- Some mixtures will not be as easily detectable as other mixtures.

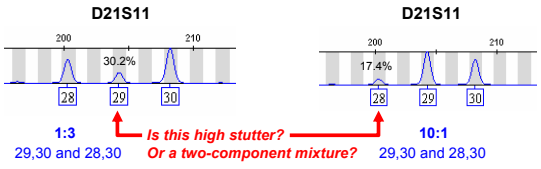


Detecting Mixtures

- Review and compile information from the entire profile – **don't just focus on a single locus!**
- **Tri-allelic patterns exist** in single source samples
  - **145 different tri-alleles recorded for the 13 core CODIS loci** on STRBase as of Jan 22, 2008
  - CSF1PO (5), FGA (22), TH01 (1), TPOX (15), VWA (18), D3S1358 (6), D5S818 (4), D7S820 (7), D8S1179 (11), D13S317 (8), D16S539 (8), D18S51 (21), D21S11 (19)
- A mixture often declared when **>2 peaks in ≥2 loci**

Mixtures: Issues and Challenges

- Artifacts of PCR amplification such as stutter products and heterozygote peak imbalance complicate mixture interpretation
- Thus, only a limited range of mixture component ratios can be solved routinely



1

Responses to Questions
from a Previous Mixture Workshop (Fall 2007)

What are the biggest obstacles you face in your lab in terms of mixture interpretation?

- Trying to be consistent in my interpretation and with coworkers
- Consistency between analysts
- No consistency – based on analysts discretion/experience; due to lack of consistent training
- Vague SOP leading to inconsistency between analysts due to differences in how "conservative" or not each analyst is
- There is a lot of "individual interpretation" in our lab
- Varying opinions between interpreting analysts due to lack of uniform guidelines
- Resistance to change from other analysts/supervisors
- Getting management to commit to guidelines that will be followed by everyone

2

Responses to Questions
from a Previous Mixture Workshop (Fall 2007)

What are the biggest obstacles you face in your lab in terms of mixture interpretation?

- Where to draw the line without throwing away valuable data
- Partial minor contributors
- Stochastic effects in minor components
- STATS and presenting them in court so that the jury will understand them
- When to do stats and what stats to do in different cases
- Lack of concrete/uniform guidelines from statisticians

DNA Mixture Interpretation:
Principles and Practice in Component Deconvolution and Statistical Analysis

Numbers and Types of Casework Mixtures

Handouts available on STRBase at
http://www.cstl.nist.gov/biotech/strbase/training/AAFS2008_MixtureWorkshop.htm

AAFS 2008 Workshop #16
Washington, DC
February 19, 2008

Ann Marie Gross
ann.gross@state.mn.us

Mixtures……

- How often are mixtures obtained
- What types of mixtures are we seeing
  - Where should we focus our attention for training
  - What info can we give to the forensic community regarding mixtures
- What types of samples most often yield mixtures

Torres et al. 4 year Spanish study

- Four year study (1/1997 to 12/2000)
- 2412 samples typed
  - 955 samples from sexual assaults
  - 1408 samples from other offenses
  - 49 samples from human remains identifications
- 163/2412 samples (6.7% showed mixed profile)

Spreadsheet Information Requested

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

Labs requested to also provide info on kit, PCR volume used, etc.

- **Case#**
- **Item#**

This information retained by lab and not returned…

- **Type of sample (biological material if ID'd)**
- **Type of substrate**
- **Quantity amp'd**
- **Minimum # of contributors (1, 2, 3, 4, or >4)**
- **Predominant type (major profile) determined?**
- **Stats reported**
- **Comments**

We would love to have your lab mixture numbers…

Email information to Ann.Gross@state.mn.us

12 Labs Submitted Data
(prior to AAFS meeting)

- **Palm Beach Sheriff's Office Crime Lab, Florida**
- Centre for Forensic Science, Toronto
- Connecticut State Police
- Washington State Police
- New Jersey State Police
- Georgia Bureau of Investigation
- Royal Canadian Mounted Police, Ottawa
- USACIL, Georgia
- Michigan State Police
- Kern County Crime Lab, California
- CAL DOJ
- Minnesota Bureau of Criminal Apprehension

We would still like to collect more case summary data…

All Laboratory Data Combined

Case type	N = 3106	1	2	3	4	>4
			# contributors			
Sexual Assault	N = 1408	51%	40%	8%	--	--
Major Crime	N = 1388	66%	24%	8%	2%	--
High Volume	N = 310	43%	37%	19%	1%	--
		Single source	Mixtures			

Overall Summary – 3106 samples

- 57% of samples from all types of cases are single source
- 43% of samples from all types of cases are mixtures
  - 33% of mixtures of at least two contributors
  - 9% of mixtures of at least three contributors
  - 1% of mixtures of at least four contributors

Focus in training materials will be on two-person mixtures as they presently predominate

Principles of Mixture Interpretation

Topics for Discussion

- SWGDAM Mixture Interpretation Committee progress
- Different statistical approaches: CPE or LR
- ISFG Mixture Interpretation Recommendations
  - UK response
  - German categories for mixtures
- Validation as it relates to mixture interpretation
  - Stochastic threshold vs analytical threshold
- Low-level DNA and mixtures
- Important elements of interpretation guidelines

SWGDAM Mixture Interpretation Subcommittee

- **John Butler** (NIST) - chair
- **Gary Sims** (CA DOJ) - co-chair
- Mike Adamowicz (CT)
- Jack Ballantyne (UCF/NCFS)
- George Carmody (Carleton U)
- Cecelia Crouse (PBSO)
- Allison Eastman (NYSP)
- Roger Frappier (CFS-Toronto)
- Ann Gross (MN BCA)
- Phil Kinsey (MT)
- Jeff Modler (RCMP)
- Gary Shutler (WSP)

Everyone not at every meeting…

Have met 3 times:
Jan 2007
July 2007
Jan 2008

Through the Jan 2008 meeting we have also had to deal with Y-STR issues – which has limited our focus on mixtures

Additional Participants (Jan 2008)
Bruce Heidebrecht (MD)
Steve Lambert (SC)

Started in January 2007

Progress and Plans for Mixture Committee

- **Guidelines in process** of being discussed and written
- Collecting data on number and type of mixture cases observed in various labs
- Plan to create a training workbook with worked examples
- Considering flow charts to aid mixture interpretation
- Have discussed responses to ISFG Recommendations

I invite your input as to what should be included in the guidelines…

Your HOMEWORK…

Elements of DNA Mixture Interpretation

Principles (theory)

↓

Protocols (validation)

↓

Practice (training & experience)

Consistency across analysts

ISFG Recommendations
SWGDAM Guidelines

⋮

Your Laboratory SOPs

↕

Training within Your Laboratory

We discussed and would advocate periodic training to aid accuracy and efficiency within your laboratory.

ISFG Recommendations on Mixture Interpretation

July 13, 2006 issue of *Forensic Science International*

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

ELSEVIER

Forensic Science International 160 (2006) 90–101

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 510-330-3911, USA

Abstract

The DNA commission of the International Society of Forensic Genetics (ISFG) was convened at the 21st congress of the International Society for Forensic Genetics held between 13 and 17 September in the Azores, Portugal. The purpose of the group was to agree on guidelines to encourage best practice that can be universally applied to assist with mixture interpretation. In addition the commission was tasked to provide guidance on low copy number (LCN) reporting. Our discussions have highlighted a significant need for continuing education and research into this area. We have attempted to present a consensus from experts but to be practical we do not claim to have conveyed a clear vision in every respect in this difficult subject. For this reason, we propose to allow a period of time for feedback and reflection by the scientific community. Then the DNA commission will meet again to consider further recommendations.

© 2006 Elsevier Ireland Ltd. All rights reserved.

Keywords: STR typing; Biostatistical analysis; Likelihood ratio; Probability of exclusion; Mixtures; ISFG DNA commission

Who is the ISFG
and why do their recommendations matter?

International Society of Forensic Genetics
http://www.isfg.org/

- An international organization responsible for the promotion of scientific knowledge in the field of genetic markers analyzed with forensic purposes.
- Founded in 1968 and represents more than 1100 members from over 60 countries.
- **A DNA Commission regularly offers recommendations on forensic genetic analysis.**

DNA Commission of the ISFG

- DNA polymorphisms (1989)
- PCR based polymorphisms (1992)
- Naming variant alleles (1994)
- Repeat nomenclature (1997)
- Mitochondrial DNA (2000)
- Y-STR use in forensic analysis (2001)
- Additional Y-STRs - nomenclature (2006)
- **Mixture Interpretation (2006)**
- Disaster Victim Identification (2007)

http://www.isfg.org/Publications/DNA+Commission

ISFG Executive Committee

President
Niels Morling
(Copenhagen, Denmark)

Vice-President
Peter Schneider
(Köln, Germany)

Working Party Representative
Mecki Prinz
(New York City, USA)

Treasurer
Leonor Gusmão
(Porto, Portugal)

Secretary
Wolfgang Mayr
(Vienna, Austria)

Angel Carracedo
***FSI Genetics* Editor-in-Chief**
(former ISFG President, VP)
(Santiago de Compostela, Spain)

Authors of ISFG Mixture Article

Peter Gill
Pioneer of forensic DNA techniques and applications
UK's Forensic Science Service (1978-2008)
University of Strathclyde (Apr 2008 – present)

The Statisticians

Charles Brenner
DNA-View,
Berkeley, CA, USA

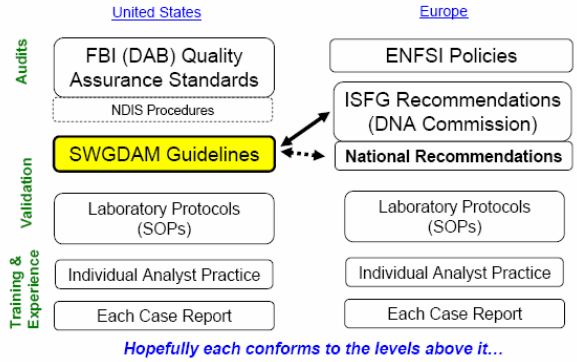
John Buckleton
ESR,
Auckland, New Zealand

Michael Krawczak
Christian-Albrechts-University,
Kiel, Germany

Bruce Weir
U. Washington,
Seattle, USA

My perspective…

Hierarchy of Rules for Forensic DNA Labs



Hopefully each conforms to the levels above it…

UK Response to ISFG Mixture Recommendations

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

Using the published UK response as a model, let us review the nine ISFG Recommendations on mixture interpretation…

From Report to the **Virginia Scientific Advisory Committee** by the DNA Subcommittee – Addendum January 8, 2008

(authored by Dr. Norah Rudin and Dr. Artie Eisenberg)

- "Among the many reasons that Forensic DNA analysis has become the gold standard for forensic science is the relatively discrete nature of the data. For strong, single source samples, a profile can readily be determined, and is subject to little or no analyst judgment. **However, ambiguity may arise when interpreting more complex samples, such as those containing multiple contributors, of poor quality (e.g. degraded or inhibited DNA), of low quantity (e.g. contact samples), or various combinations of these challenging situations**…"

http://www.dfs.virginia.gov/about/minutes/saCommittee/20080108.pdf

From Report to the **Virginia Scientific Advisory Committee** by the DNA Subcommittee – Addendum January 8, 2008

(authored by Dr. Norah Rudin and Dr. Artie Eisenberg)

- "…**These kinds of samples are encountered with increasing frequency, as the sensitivity of the technology has increased, and as law enforcement has become more sophisticated about the kinds of samples they submit for analysis**. Difficult samples are also frequently encountered when reanalyzing historical cases, in which samples were not collected and preserved using the precautions necessary for DNA analysis…"

"Cold cases" or Innocence Project samples…

http://www.dfs.virginia.gov/about/minutes/saCommittee/20080108.pdf

From Report to the **Virginia Scientific Advisory Committee** by the DNA Subcommittee – Addendum January 8, 2008

(authored by Dr. Norah Rudin and Dr. Artie Eisenberg)

- "It is for these types of challenging samples, where the evidence profile may not exactly "match" a reference profile, that confirmation bias becomes a concern. **The interpretation of an evidentiary DNA profile should not be influenced by information about a subject's DNA profile.** Each item of evidence must be interpreted independently of other items of evidence or reference samples. Yet forensic analysts are commonly aware of submitted reference profiles when interpreting DNA test results, creating the opportunity for confirmatory bias, despite the best intentions of the analyst…"

http://www.dfs.virginia.gov/about/minutes/saCommittee/20080108.pdf

DNA Mixture Interpretation:

Principles and Practice in Component Deconvolution and Statistical Analysis

Principles in Mixture Interpretation

Handouts available on STRBase at
http://www.cstl.nist.gov/biotech/strbase/training/AAFS2008_MixtureWorkshop.htm

AAFS 2008 Workshop #16
Washington, DC
February 19, 2008

John M. Butler

john.butler@nist.gov

Two Parts to Mixture Interpretation

- Determination of alleles present in the evidence and **deconvolution of mixture components** where possible
  - Many times through comparison to victim and suspect profiles
- **Providing some kind of statistical answer** regarding the weight of the evidence
  - There are multiple approaches and philosophies

Software tools can help with one or both of these…

Status of Software for Mixture Interpretation

- NIJ Expert System Testbed (**NEST**) Project
  - Evaluating software programs for DNA analysis of single-source (Phase I) and mixtures (Phase II)
  - **http://forensics.marshall.edu/NEST/NEST-Intro.html**
- US Army Crime Laboratory (**USACIL**)
  - Commonly deal with complex sexual assaults
  - Developed software for aiding mixture interpretation and statistical analysis

Steps in the Interpretation of Mixtures

(Clayton *et al.* 1998)

Step #1 Identify the Presence of a Mixture

Step #2 Designate Allele Peaks

Step #3 Identify the Number of Potential Contributors

Step #4 Estimate the Relative Ratio of the Individuals Contributing to the Mixture

Step #5 Consider All Possible Genotype Combinations

Step #6 Compare Reference Samples

Analysis and interpretation of mixed forensic stains using DNA STR profiling

Clayton *et al.* (1998) *Forensic Sci. Int.* 91:55-70

Adapted from Peter Schneider slide (presented at EDNAP meeting in Krakow in April 2007)

Mixture Classification Scheme

Schneider *et al.* (2006) *Rechtsmedizin* 16:401-404

(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 **Type B** **Type C**

Type of mixture and interpretation

- **Type A:** Mixed profile without stochastic effects, a biostatistical analysis has to be performed
- **Type B:** Profile of a major contributor can be unambiguously described and interpreted as a profile from an unmixed stain
- **Type C:** due to the complexity of the mixture, the occurrence of stochastic effects such as allele and locus drop-outs have to be expected:
  - a clear decision to include or exclude a suspect may be difficult to reach, thus a biostatistical interpretation is not appropriate.

Slide from Peter Schneider (presented at EDNAP meeting in Krakow in April 2007)

Biostatistical approaches

- Calculation of the **probability of exclusion** for a randomly selected stain donor* [P(E)]
  (*RMNE - "random man not excluded")
- Calculation of the **likelihood ratio** [LR] based on defined hypotheses for the origin of the mixed stain

Slide from Peter Schneider (presented at EDNAP meeting in Krakow in April 2007)

Which approach should be used?

- If the basis for clearly defined and mutually exclusive hypotheses is given, i.e.:
  - the number of contributors to the stain can be determined,
  - unambiguous DNA profiles across all loci are observed (type A mixtures, or type B, if the person considered as "unknown" contributor is part of the minor component of the mixture),

  then the calculation of a likelihood ratio is appropriate.

Slide from Peter Schneider (presented at EDNAP meeting in Krakow in April 2007)

Which approach should be used?

- If major/minor contributors cannot be identified based on unambiguous DNA profiles, or if the the number of contributors cannot be determined, then the calculation of the probability of exclusion is appropriate.
- The calculation of P(E) is always possible for type A and type B mixtures.

Slide from Peter Schneider (presented at EDNAP meeting in Krakow in April 2007)

Not acceptable …

- … is the inclusion of a genotype frequency of a non-excluded suspect into the report, if the given mixed stain does not allow a meaningful biostatistical interpretation.
  - this would lead to the wrongful impression that this genotype frequency has any evidentiary value regarding the role of the suspect as a contributor to the mixed stain in question.

Slide from Peter Schneider (presented at EDNAP meeting in Krakow in April 2007)

Conclusions

- The likelihood ratio has a significant weight of evidence, as it relates directly to the role of the suspect in the context of the origin of the stain.
- The exclusion probability makes a general statement without relevance to the role of the suspect.
- However, this does not imply that P(E) is always more "conservative" in the sense that the weight of evidence is not as strong compared to the LR.

Slide from Peter Schneider (presented at EDNAP meeting in Krakow in April 2007)

GEDNAP 32

Mixture interpretation exercise:

- 3 person mixture without major contributor
- Person A from group of reference samples was not excluded
- Allele frequencies for eight German database systems provided for exercise
- German-speaking GEDNAP participants invited to participate based on published recommendations

Slide from Peter Schneider (presented at EDNAP meeting in Krakow in April 2007)

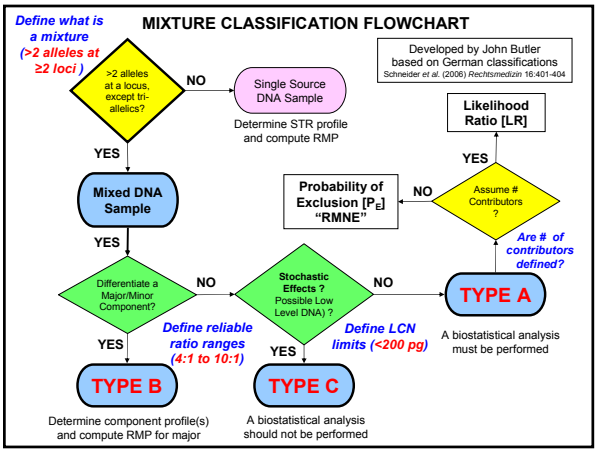
GEDNAP 32

Results:

- 22 labs submitted results (from approx. 80 German-speaking GEDNAP participants)
- Calculations submitted were all correct and consistent:
  - 15x LR approach:
    - Person A + 2 unknown vs. 3 unknown contributors
  - 11x RMNE calculation
- Will be offered again next time

Training and Specific Guidelines/Classification Schemes yielded consistent results among laboratories

Slide from Peter Schneider (presented at EDNAP meeting in Krakow in April 2007)

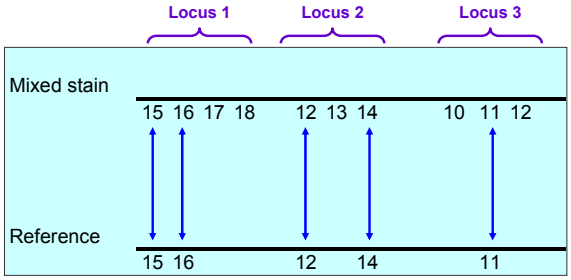


German Type A,B, and C mixture classifications

- **Type A**, where major/minor contributors cannot be deduced, require stats
  - LR
  - RMNE
- **Type B** enables major contributor to be deduced
  - RMP (which is 1/LR)
- **Type C** no stats should be attempted because of the possibility of failure to account for allele dropout due to stochastic effects with low level DNA samples

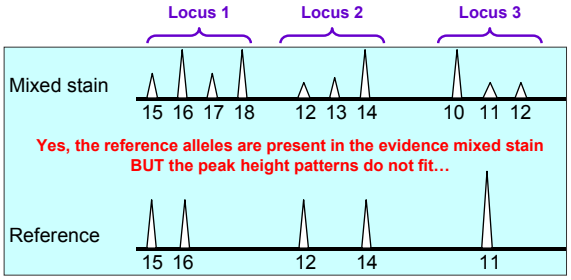
Mixture Example

Comparing Alleles Only



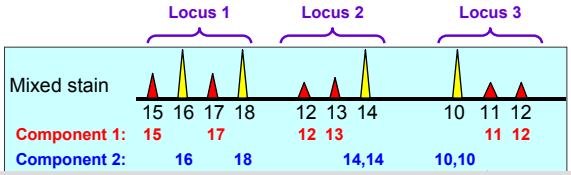
Mixture Example

Showing Importance of Using Peak Height Information



Mixture Example

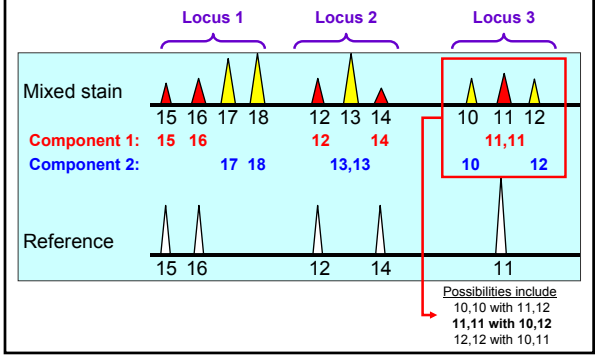
Solving Components Prior to Comparison to Suspect Reference



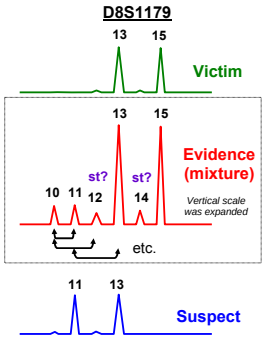
Reference (suspect) does not match either component of the mixed stain and therefore could not have contributed to the evidence sample

Mixture Example

Different Evidence Sample…



Another Mixture Example



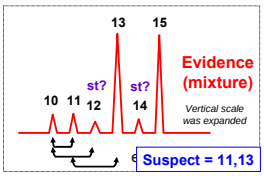
Conclusions from the evidence:

1. **Major contributor = 13,15 (victim) –** *to be expected with an intimate sample like a fingernail or vaginal swab*
2. Alleles 12 and 14 are **likely stutter products** of the major contributor's 13 and 15 alleles but could also be masking minor contributor alleles
3. A number of minor contributor combinations are possible (e.g., 10,11 or 10,12 or 10,13 or 11,13, etc.)
4. Could have more than two contributors present in this mixture

"Suspect cannot be excluded" BUT statement needs to be qualified by statistics because a large percentage of the population might also not be able to be excluded…

Probability of Exclusion Calculation
for a Single STR Locus

The case may grow stronger against a suspect with information from additional STR loci…



The fact that in this case a suspect is included is not very informative because ~9 out of 10 people examined from any population could potentially be included in the evidence mixture…

From VA DFS STR Allele Frequencies
http://www.dfs.virginia.gov/manuals/manuals.cfm?id=5

D8S1179 alleles	AA (n=384)	C (n=346)	H (n=366)
10	0.0287	0.1069	0.0820
11	0.0495	0.0925	0.0465
12	0.1094	0.1416	0.1093
13	0.2422	0.3093	0.3224
14	0.2969	0.1965	0.2623
15	0.1849	0.0896	0.1202
SUM	0.9115	0.9364	0.9426
Sq SUM = PI	0.8308	0.8769	0.8886
PE = 1-PI	0.1692	0.1231	0.1114
PE (%)	**16.9%**	**12.3%**	**11.1%**
	African Am.	Caucasians	Hispanics

"Suspect cannot be excluded" BUT we would expect to see, for example, only 11.1% of Hispanics excluded (or 88.9% cannot be excluded) based on results at this one locus

The Statistic (Determining the Weight of the Evidence) Should Be **Calculated from the Evidence**

Evidence (partial profile):

	Type	Statistic
Locus 1	16,17	1 in 9
Locus 2	17,18	1 in 9
Locus 3	21,22	1 in 12
Locus 4	12,14	1 in 16
Locus 5	28,30	1 in 11

		Product = 1 in 171,000

Match Observed at All Loci that May Be Compared

Reference (full profile):

	Type	Statistic
Locus 1	16,17	1 in 9
Locus 2	17,18	1 in 9
Locus 3	21,22	1 in 12
Locus 4	12,14	1 in 16
Locus 5	28,30	1 in 11
Locus 6	14,16	1 in 26
Locus 7	12,13	1 in 9
Locus 8	11,14	1 in 31
Locus 9	9,9	1 in 32
Locus 10	9,11	1 in 14
Locus 11	6,6	1 in 19
Locus 12	8,8	1 in 3
Locus 13	10,10	1 in 21

		Product = 1 in 665 trillion

The reference sample is still a "match" – just not as much information is available from the evidence for comparison

Statistical Approaches with Mixtures

See Ladd *et al.* (2001) *Croat Med J.* 42:244-246

- **Inferring Genotypes of Contributors** - Separate major and minor components into individual profiles and compute the random match probability estimate as if a component was from a single source
- **Calculation of Exclusion Probabilities** - CPE/CPI (RMNE) – The probability that a random person (unrelated individual) would be excluded as a contributor to the observed DNA mixture
- **Calculation of Likelihood Ratio Estimates** – Comparing the probability of observing the mixture data under two (or more) alternative hypotheses; in its simplest form LR = 1/RMP

RMNE = Random Man Not Excluded (same as CPE)
CPE = Combined Probability of Exclusion (CPE = 1 – CPI)
CPI = Combined Probability of Inclusion (CPI = 1 – CPE)

Advantages and Disadvantages

RMNE (CPE/CPI)

- **Advantages**
  - Does not require an assumption of the number of contributors to a mixture
  - Easier to explain in court
- **Disadvantages**
  - Weaker use of the available information (robs the evidence of its true probative power because this approach does not consider the suspect's genotype)
  - Likelihood ratio approaches are developed within a consistent logical framework

Likelihood Ratios (LR)

- **Advantages**
  - Enables full use of the data including different suspects
- **Disadvantages**
  - More difficult to calculate

John Buckleton, *Forensic DNA Evidence Interpretation*, p. 223

Assumptions for CPE/CPI Approach

- There is no allele dropout (i.e., all alleles are above stochastic threshold) – low-level mixtures can not reliably be treated with CPE
- All contributors are from the same racial group (i.e., you use the same allele frequencies for the calculations)
- All contributors are unrelated
- Peak height differences between various components are irrelevant (i.e., component deconvolution not needed) – this may not convey all information from the available sample data…

Likelihood Ratio (LR)

- Provides ability to express and evaluate both the prosecution hypothesis, H_p (the suspect is the perpetrator) and the defense hypothesis, H_d (an unknown individual with a matching profile is the perpetrator)

$$LR = \frac{H_p}{H_d}$$

- **The numerator, H_p, is usually 1** – since in theory the prosecution would only prosecute the suspect if they are 100% certain he/she is the perpetrator
- The denominator, $\mathbf{H_d}$, is typically the profile frequency in a particular population (based on individual allele frequencies and assuming HWE) – i.e., **the random match probability**

LR is not a probability but a ratio of probabilities

DAB Recommendations on Statistics

February 23, 2000
Forensic Sci. Comm. 2(3); available on-line at
http://www.fbi.gov/hq/lab/fsc/backissu/july2000/dnastat.htm

"The DAB finds either one or both PE or LR calculations acceptable and strongly recommends that one or both calculations be carried out whenever feasible and a mixture is indicated"

- Probability of exclusion (PE)
  - Devlin, B. (1993) Forensic inference from genetic markers. *Statistical Methods in Medical Research*, 2, 241–262.
- Likelihood ratios (LR)
  - Evett, I. W. and Weir, B. S. (1998) *Interpreting DNA Evidence*. Sinauer, Sunderland, Massachusetts.

ISFG DNA Commission on Mixture Interpretation

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

Available online at www.sciencedirect.com
SCIENCE DIRECT
ELSEVIER
Forensic Science International

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

P. Gill, C.H. Brenner, J.S. Buckleton, A. Carracedo, M. Krawczak, W.R. Mayr, N. Morling, M. Prinz, P.M. Schneider, B.S. Weir

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

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

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

Thoughts by Peter Gill on Recommendation #5

(ENFSI meeting, Krakow, Poland, April 19, 2007)

- Prosecution and defense each want to maximize their respective probabilities
- Recommendation 5 places ownership for each hypothesis.
- In order to perform the LR calculation(s), the forensic scientist decides on both the prosecution and defense hypotheses.
- Since the forensic scientists usually cannot discover the defense hypothesis before the trial (as they are typically working with the prosecution if the DNA matches…), assumptions must be clearly stated with the important caveat that you cannot perform calculations on the stand! (For example, you need three weeks warning to make and check calculations.)
- By anchoring the respective hypotheses to each side, the defense can change their hypothesis but the prosecution does not need to change theirs…
- It is worth noting that the likelihood ratio always goes up if the defense lowers their hypothesis (H_d gets lower with more possible combinations)

ISFG (2006) Recommendations

- Recommendation 6: If the crime profile is a major/minor mixture, **where minor alleles are the same size (height or area) as stutters of major alleles, then stutters and minor alleles are indistinguishable.** Under these circumstances alleles in stutter positions that do not support H_p should be included in the assessment.
- In general, stutter percentage is <15%

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

Consideration of Peak in Stutter Position

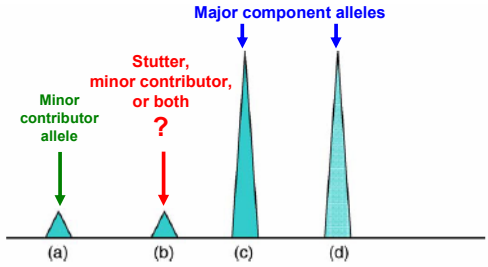


Fig. 4. *c* and *d* are unambiguous alleles, *b* is a minor allele in a stutter position and *a* is an unambiguous minor allele.

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

UK Response

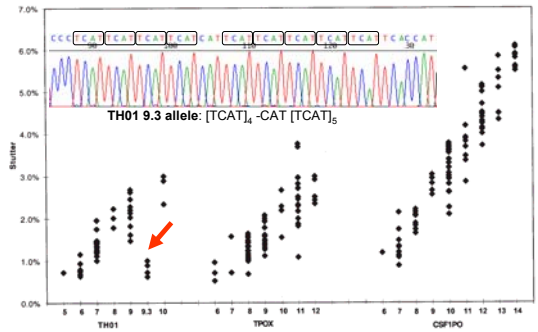
Gill *et al.* (2008) *FSI Genetics* 2(1): 76–82

Recommendation 6:

- Stutters are locus-dependent…
- It is recommended that laboratories make their own maximum experimentally observed stutter sizes per locus determinations since the effects may be technique dependent.
- It is recommended that [maximum stutter percentages be] evaluated per locus.

Measured Stutter Percentages

Variable by Allele Length and Composition



Holt CL, Buoncristiani M, Wallin JM, Nguyen T, Lazaruk KD, Walsh PS. TWGDAM validation of AmpFISTR PCR amplification kits for forensic DNA casework. *J Forensic Sci* 2002; 47(1): 66-96.

UK Response

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

- ***Characterization of +4 base stutters***

We agreed to review +4 bp stutters, however, we note that their presence often relates to over-amplified samples. Preliminary experimental work suggests that they are low level and **generally less then 4% the size of the progenitor allele** (Rosalind Brown, personal communication). Note that 4 bp and +4 bp stutter cannot be distinguished from genetic somatic mutation without experimental work—furthermore, somatic mutations may give rise to peaks that are larger than those caused by stutter artifacts.

ISFG (2006) Recommendations

- Recommendation 7: If drop-out of an allele is required to explain the evidence under H_p: (S = ab; E = a), then the allele should be small enough (height/area) to justify this. Conversely, if a full crime stain profile is obtained where alleles are well above the background level, and the probability of drop-out approaches Pr(D) ≈ 0, then H_p is not supported.

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

UK Response

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

Recommendation 7:

- We recommend slight rewording…[with mention of companion allele]
- If a full crime-stain profile is obtained where alleles are well above the background level, and the probability of dropout Pr(D) approaches zero, then H_p is not supported (Figure 6).

Hypothetical Examples

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

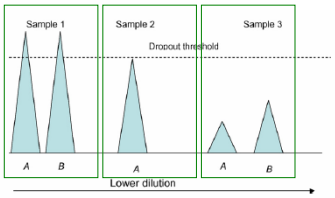


Fig. 4. Results from serial dilutions of the same sample genotype *AB*. The first result (sample 1) shows a locus where both alleles are represented in the profile. One or both of these alleles are above the dropout threshold and consequently are always present in the epg. The second result shows a result where dropout has occurred – the survivor allele is just below the dropout threshold hence this is a rare event, but not impossible. If *A* was just above the dropout threshold we would determine it to be a homozygote *AA* genotype. In the third sample, both alleles are well below the dropout threshold – it is an unambiguous, albeit unbalanced heterozygote. If only one allele was present, then we would have to consider the possibility of dropout of the partner. The same rationale can be applied to any analytical regime, e.g. 28 and 34 PCR cycles.

If Below Dropout Threshold…

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

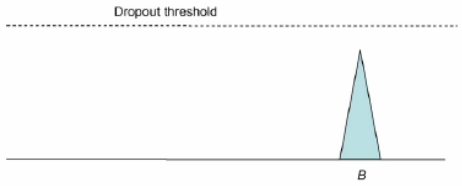


Fig. 5. In this example allele *B* is below the dropout threshold, hence we cannot be confident that it is from a homozygote *BB* individual. It could also be from an individual who is heterozygote, where the missing allele is any other allele. The probability *B*|unknown, H_d is 2Pr(*BF*), where the '*F*' designation is assigned a probability of 1 to take account of the possibility that any allele could have dropped out.

If Above Dropout Threshold…

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

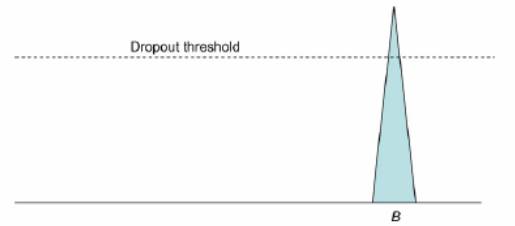


Fig. 6. In this example allele *B* is above the dropout threshold, hence we can be confident that it is from a homozygote *BB* individual. The probability of B|unknown, H_d is $Pr(B)^2$.

Setting Thresholds

- **Detection (analytical) threshold**
  - Dependent on instrument sensitivity
  ~50 RFU
  - Impacted by instrument baseline noise
- **Dropout (stochastic) threshold**
  - Dependent on biological sensitivity
  ~150-200 RFU
  - Impacted by assay and injection parameters

Determining the Dropout (Stochastic) Threshold

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

- The dropout threshold can be determined experimentally for a given analytical technique from a series of pre-PCR dilutions of extracts of known genotype technique (it will probably vary between analytical methods). These samples can be used to determine the point where allelic dropout of a heterozygote is observed relative to the size of the survivor companion allele. The threshold is the maximum size of the companion allele observed. This is also the point where Pr(D) approaches zero (Fig. 4).

Dropout threshold will change depending on instrument and assay conditions (e.g., longer CE injection will raise dropout threshold)

ISFG (2006) Recommendations

- Recommendation 8: If the alleles of certain loci in the DNA profile are at a level that is dominated by background noise, then a biostatistical interpretation for these alleles should not be attempted.

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

UK Response

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

Recommendation 8:

- If there is a band below the experimental threshold where background noise might be prevalent, and it is distinct and clear from the background, then it should be recorded and available on the case file.

ISFG (2006) Recommendations

- Recommendation 9: In relation to low copy number, stochastic effects limit the usefulness of heterozygous balance and mixture proportion estimates. In addition, allelic drop-out and allelic drop-in (contamination) should be taken into consideration of any assessment.

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

UK Response

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

Recommendation 9:

- Case pre-assessment is necessary in order to determine the best scientific method to process a sample. To facilitate this, it is recommended that wherever possible, this should include quantification. Quantification is used to determine the optimum method to process—if low-level DNA, a sample would benefit from procedures to enhance sensitivity of detection. There may be reasons where quantification is not practicable, especially if low levels of DNA are expected, since the result itself may be compromised if a portion of the sample is sacrificed. At low DNA levels, the accuracy of the quantification test itself may be inefficient.

UK Response

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

Recommendation 9 (cont):

- It is possible that a given DNA profile may simultaneously comprise both 'conventional' and 'low-level' loci: for example, if degradation has occurred then low molecular weight loci may be above the dropout threshold, whereas high molecular weight loci may be below the dropout threshold.
- Similarly, if the sample is a mixture, then at a given locus there may be some alleles that are above the dropout threshold (from a major contributor) and others that are below the dropout threshold (from a minor contributor), i.e. different interpretation rationale may be simultaneously applied to different contributors within a locus.

Thank you for your attention…

Questions
or Comments?

http://www.cstl.nist.gov/biotech/strbase
john.butler@nist.gov
301-975-4049

Our team publications and presentations are available at:
http://www.cstl.nist.gov/biotech/strbase/NISTpub.htm