Topics and Techniques for Forensic DNA Analysis

Mixture Interpretation

Forensic Statistics Course

Towson, MD
April 7, 2009

Dr. John M. Butler
National Institute of Standards and Technology
john.butler@nist.gov

Presentation Plans

- Background and Resources
- Thresholds
- Statistical Approaches
- ISFG Recommendations & Responses
- German Categorization of Mixtures
- Examples

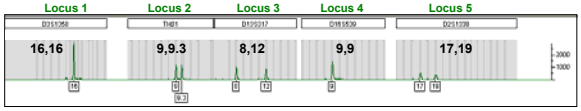
- *Please ask questions!*

Mention of Mixtures in the July 2009 Revised Quality Assurance Standards (QAS)

- QAS Standard 5.3.2
  - **A casework CODIS administrator** shall be or have been a current or previously qualified DNA analyst … with **documented mixture interpretation training**.
- QAS Standard 8.3.1
  - **Internal validation studies** conducted after the date of this revision shall include as applicable: known and non-probative evidence samples or mock evidence samples, reproducibility and precision, sensitivity and stochastic studies, **mixture studies**, and contamination assessment. Internal validation studies shall be documented and summarized…
- QAS Standard 8.3.2
  - **Internal validation shall define quality assurance parameters and interpretation guidelines, including as applicable, guidelines for mixture interpretation.**
- QAS Standard 9.6.4
  - **Laboratories analyzing forensic samples shall have and follow a documented procedure for mixture interpretation that addresses major and minor contributors, inclusions and exclusions, and policies for the reporting of results and statistics**.

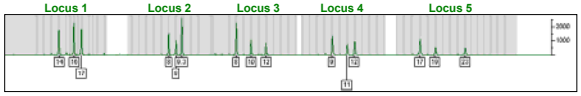
Single Source vs. Mixture Samples

Single Source Sample



One or two peaks observed at each locus (tested DNA region)

Mixture Sample



More than two peaks observed at more than two loci (tested DNA regions)

Different possible combinations could have given rise to the particular mixture observed

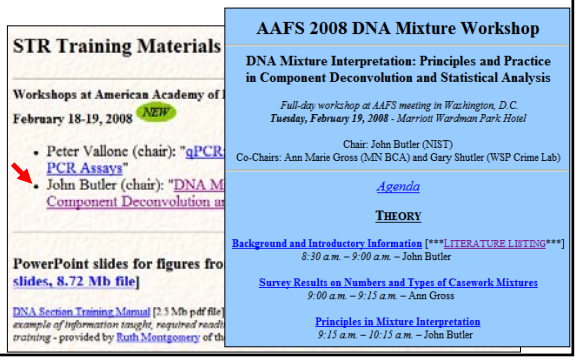
Did anyone here attend this workshop?

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

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

John M. Butler
Ann Marie Gross
Gary G. Shutler

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

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**

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.
- Differential extraction can help distinguish male and female components of many sexual assault mixtures.

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…

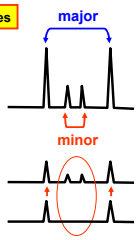
More on Mixtures...

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

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**

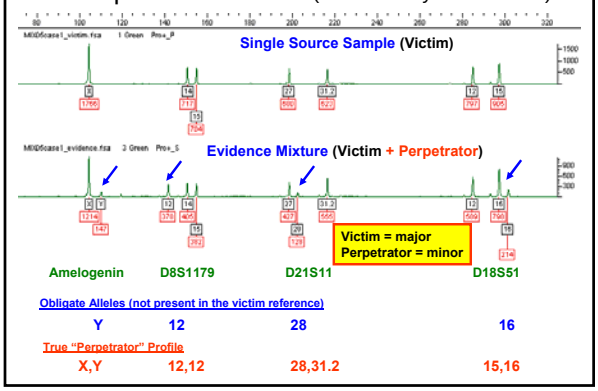
95.1% (155/163) were 2-component mixtures

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

Example Mixture Data (MIX05 Study-Profiler Plus)



Sources of DNA Mixtures

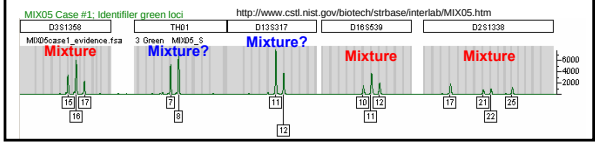
- **Two (or more) individuals** contribute to the biological evidence examined in a forensic case (e.g., sexual assault with victim and perpetrator or victim, consensual sexual partner, and perp)
  - **Victim Reference and Spouse or Boyfriend Reference**
- **Contamination** of a single source sample 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

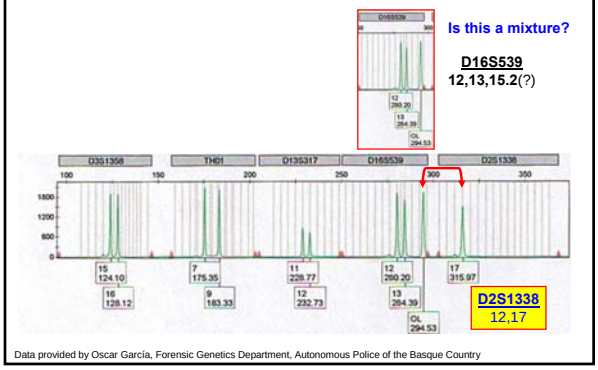
- The probability that a mixture will be detected improves with the use of more loci and genetic markers that have a high incidence of heterozygotes.
- The detectability of multiple DNA sources in a single sample relates to the ratio of DNA present from each source, the specific combinations of genotypes,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



Detecting Mixtures

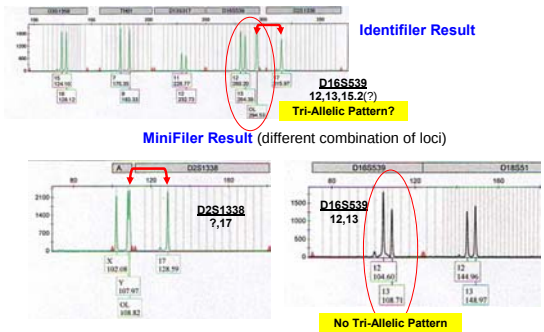
- Review and compile information from the entire profile – **don't just focus on a single locus!**
- **Tri-allelic patterns exist** in single source samples
  - **145 different tri-alleles recorded for the 13 core CODIS loci** on STRBase as of Jan 22, 2008
  - CSF1PO (5), FGA (22), TH01 (1), TPOX (15), 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**

Importance of Considering Entire Profile and Not Just a Single Locus



Data provided by Oscar García, Forensic Genetics Department, Autonomous Police of the Basque Country

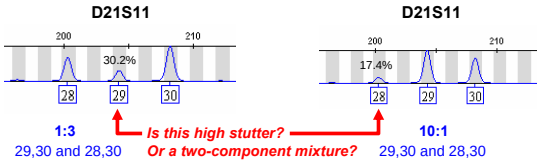
Value of Testing with Another STR Kit to Confirm Off-Ladder Alleles



Data provided by Oscar García, Forensic Genetics Department, Autonomous Police of the Basque Country

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



Gathered Case Summary Data

During 2007 and early 2008, **Ann Gross** (MN BCA) from the SWGDAM Mixture Interpretation Committee **coordinated the collection of case summary data** from **14 different forensic labs** who collectively reported on **4780 samples**.

A preliminary summary of this information is divided by crime classifications: sexual assault, major crime (homicide), and high volume (burglary). **Over half of the samples examined were single source and ~75% of all reported mixtures were 2-person.**

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

CFS Toronto Case Summary Data

Case type	N = 276	1	2	3	4	>4
			# contributors			
Sexual Assault	N = 152	42%	52%	7%	1%	--
High Volume	N = 56	69%	16%	16%	--	--
Major Crime	N = 68	59%	34%	7%	--	--
		Single source	Mixtures			

Mixture Case Summaries

	minimum # of contributors					
Crime Class	**1**	**2**	**3**	**4**	**>4**	**N**
Sexual Assault	884	787	145	11	0	1827
Major Crime	1261	519	182	32	0	1994
High Volume	344	220	140	11	5	720
Total	2489	1526	467	54	5	4541
Single source	54.8%	33.6%	10.3%	1.2%	0.1%	mixtures

http://www.cstl.nist.gov/biotech/strbase/pub_pres/Promega2008poster.pdf

"Final" Data Set from 14 Different Labs

Plan to conduct further data analysis and publish results

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

1

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

Elements of DNA Mixture Interpretation

Principles (theory)

Protocols (validation)

Practice (training & experience)

Consistency across analysts

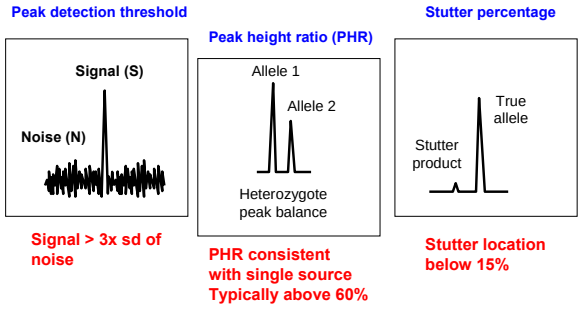
ISFG Recommendations
SWGDAM Guidelines

Your Laboratory SOPs

Training within Your Laboratory

We advocate periodic training to aid accuracy and efficiency of mixture interpretation within your laboratory.

What is a true peak (allele)?



Signal > 3x sd of noise

PHR consistent with single source Typically above 60%

Stutter location below 15%

Setting Thresholds

- **Detection (analytical) threshold**
  - Dependent on instrument sensitivity ~50 RFU
  - Impacted by instrument baseline noise

what is a peak?

- **Dropout (stochastic) threshold**
  - Dependent on biological sensitivity ~150-200 RFU
  - Impacted by assay and injection parameters

what is reliable PCR data?

Validation studies should be performed in each laboratory

Validation Studies

- Information from validation studies should be used to set laboratory-specific
  - Stutter %
  - Peak Height Ratios
  - Minimum Peak Heights (detection thresholds)
  - Relative balance across loci
- These values are all dependent on amount of input DNA
  - If low-level DNA is amplified, stutter % may be higher and peak height ratios may be lower

Threshold Values

- Critical for proper interpretation of STR data
- Establish minimum RFU that a PCR product must display for quantitative and/or qualitative evaluation
- Signal-to-noise ratio is really irrelevant as PCR variability is the bigger issue (stochastic effects with low levels of DNA template)

Bruce Budowle, "Guidelines for the Interpretation of Mixtures", Promega 2008 meeting breakout session on mixture interpretation (Hollywood, CA) – Oct 15, 2008

Threshold 1

- A **Peak Amplitude Threshold (PAT)** must be established that operationally defines the minimum peak height in RFUs for confidently ascribing a true PCR amplicon peak
- Defines when confidence is high for peak assignment
- Quantitative threshold based on a signal-to-noise ratio (and may be slightly higher – i.e., 50 RFUs)
- May also be called "Detection Threshold"

Bruce Budowle, "Guidelines for the Interpretation of Mixtures", Promega 2008 meeting breakout session on mixture interpretation (Hollywood, CA) – Oct 15, 2008

Threshold 2

- A **Match Interpretation Threshold (MIT)** must be established based on empirical studies performed in your laboratory
  - FBI's MIT was 200 RFU and has now been lowered to 150 RFUs based on instruments getting better
- The minimum peak height in RFUs that all amplicon peaks at a given locus must display to confidently conclude that no genetic components of the sample failed to be detected due to stochastic affects (such as might occur with low copy number template)
  - Can exclude but not use statistics if alleles fall between PAT and MIT
- Necessary for avoiding standard interpretation where potential stochastic affects may result in allele drop out, peak height ratio variation, or non-reproducible results
  - This threshold does not apply to LCN
- May be called "Interpretation Threshold"

Bruce Budowle, "Guidelines for the Interpretation of Mixtures", Promega 2008 meeting breakout session on mixture interpretation (Hollywood, CA) – Oct 15, 2008

Two Thresholds

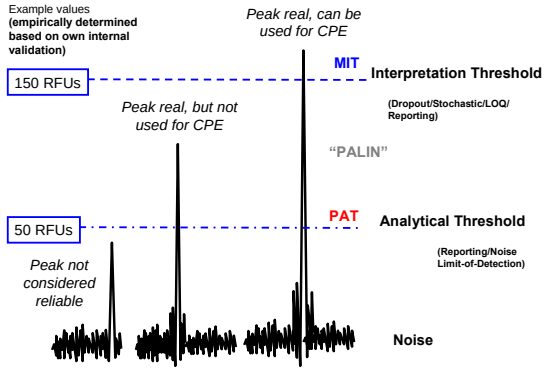
- Peak Amplitude Threshold (**PAT**)
- Match Interpretation Threshold (**MIT**)

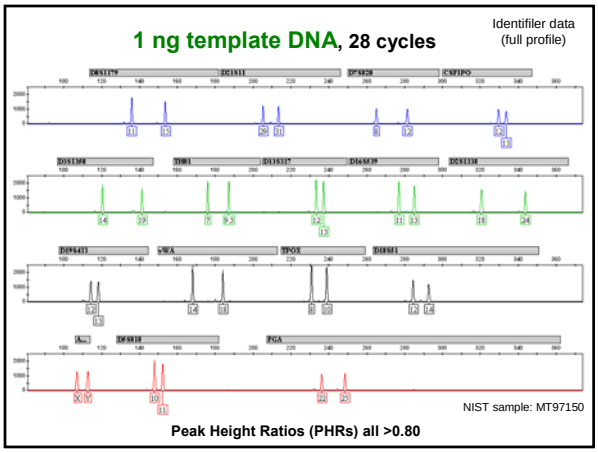
Pat Buchanan

If between PAT and MIT, **can exclude** but not use statistics

Mitt Romney

Different Thresholds





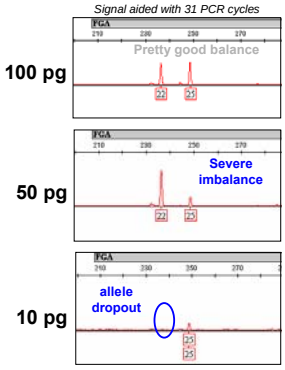
Reliable Mixture Interpretation Cannot Usually Be Performed with Low Level DNA

- Intra-locus peak height ratios vary significantly
- Stutter products can be artificially high
- Allele dropout occurs
- Allele drop-in confuses results
  - can only be caught with replicate amplifications and analyses

Peak Height Ratio Measurements

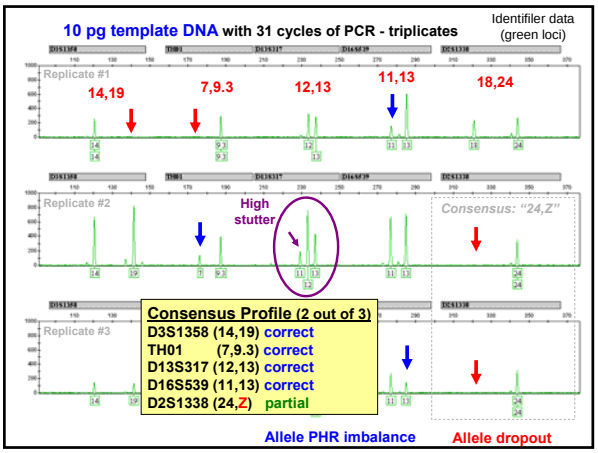
Identifiler STR Kit – only FGA shown

Signal aided with 31 PCR cycles

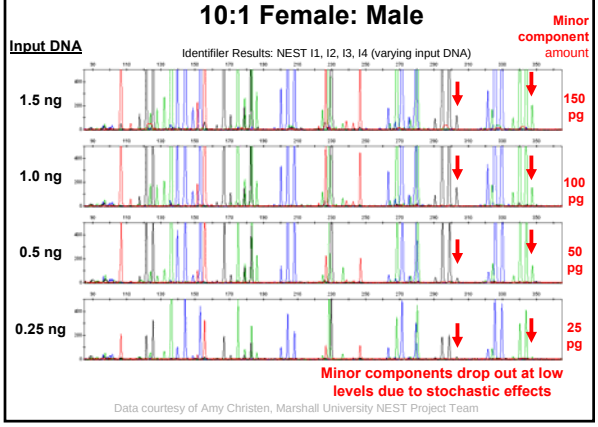


	Peak Heights (RFUs) FGA-22	FGA-25	PHR	Average PHR
100 pg (1)	1692	1517	0.90	0.69 (±0.23)
100 pg (2)	1915	864	0.45	
100 pg (3)	1239	909	0.73	
50 pg (1)	992	260	0.26	0.49 (±0.36)
50 pg (2)	1422	419	0.29	
50 pg (3)	895	805	0.90	
10 pg (1)	--	66	0	0.37 (±0.32)
10 pg (2)	54	107	0.50	
10 pg (3)	130	219	0.59	

All levels performed in triplicate...



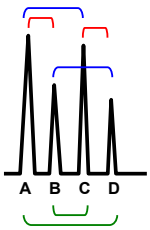
10:1 Female: Male



Data courtesy of Amy Christen, Marshall University NEST Project Team

Statistical Approaches

With Some Mixtures, Multiple Genotype Combinations Are Possible



AC
BD
AB
CD
BC
AD

Depends on PHR and proportion of mixture components from the various contributors

Peak Height Ratios (PHR)
Minimum Peak Height (mPH)
Proportion (p) or mixture proportion (M_x)

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)

Calculating Statistics for Mixtures

There are various statistical approaches that can be used for reporting mixture results:

- Probability of exclusion (PE)/Probability of inclusion (PI)
- Random Match Probability (RMP)
- Likelihood Ratio (LR)

Probability of Exclusion/Inclusion

Also known as the Combined Probability of Exclusion/Inclusion (CPE/CPI)

Prob. of Inclusion (PI) is the combined frequency of all combinations of genotypes that CANNOT BE EXCLUDED from contributing to the mixture

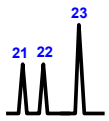
-Makes no assumptions about # of contributors

- aka random man not excluded (RMNE)

Prob. of Exclusion (PE) is the probability of EXCLUDING a randomly selected person

Probability of Exclusion/Inclusion

Suppose the following scenario (from AFDIL guidelines):



FGA

From X population database

Allele	Frequency
21	0.187
22	0.182
23	0.156

$PI = (P_A + P_B + P_C)^2$

$= (0.187 + 0.182 + 0.156)^2$

$= (0.525)^2$

$= 0.276$

Thus it is expected that 28% of a group of randomly selected persons will not be excluded as contributors or 1 out of 4 randomly selected persons (RMNE)

PE = 1 – PI = 1 – 0.276 = 0.724

Thus it is expected that 72% of a group of randomly selected persons will be excluded as contributors

Random Match Probability

Random Match Probability (RMP) is the probability of obtaining a match between two distinct and unrelated individuals

RMP is calculated by taking the inverse of the genotype frequency for a marker or a full profile

For example,

Locus	Allele 1	Allele 2	Allele 1 Freq (p)	Allele 2 Freq (q)	Calc	Genotype Freq
D13S317	11	12	0.3394	0.24834	2pq	0.1686
D16S539	12	11	0.32616	0.32119	2pq	0.2095
				Combined Freq		0.0353217
				1 in		28.311208

RMP = 1 in 28

Random Match Probability

What does this mean?

RMP = 1 in 28

This is the theoretical chance that if one person is pulled at random from a population, they will have this particular profile. Obviously in this case, there are only 2 loci therefore the chance is relatively high.

-It does NOT mean the chance that someone else is guilty

-It is NOT the chance that the defendant is not guilty

Locus	Allele 1	Allele 2	Allele 1 Freq (p)	Allele 2 Freq (q)	Calc	Genotype Freq
D13S317	11	12	0.3394	0.24834	2pq	0.1686
D16S539	12	11	0.32616	0.32119	2pq	0.2095
				Combined Freq		0.0353217
				1 in		28.311208

Likelihood Ratio

Likelihood Ratio is based on defined hypotheses as to the origin of the mixture. This calculation compares the probabilities of the evidence as 2 alternatives.

-The prosecution hypothesis (H_p)

-The defense hypothesis (H_d)

Typically, the prosecution's hypothesis is that DNA profile generated from the crime scene originates from the victim and the suspect. The defense's hypothesis is that the evidence originates from the victim and an unknown person.

LR provides a numerical value that indicates how many more times likely the observed DNA profile originated from H_p than H_d

Likelihood Ratio

- Likelihood Ratio requires a description of the scenario
- Hypotheses must clearly state who contributed to the stain
- Hypotheses must state how many **unknown** contributors are assumed

This allows the evidential value of a stain to be calculated with reference to a specific person involved in the case (i.e. the accused stain donor)

Note: This information is from Schneider PM, Fimmers R, Keil W, Molsberger G, Patzelt D, Pflug W, Rothämel T, Schmitter H, Schneider H, Brinkmann B.The German Stain Commission: recommendations for the interpretation of mixed stains. Int J Legal Med. 2009 Jan;123(1):1-5.

Likelihood Ratio

We must define the scenario:

Two contributors, unambiguous DNA profile

The hypothesis of the prosecution is that the victim and the defendant contributed to the mixture; $H_p = 1$ (or 100% probability)

However, the defense claims the victim and an unknown person contributed to the mixture.

Example

The victim's genotype is 21,23. The suspect's genotype is 22,23. The defense hypothesis must explain the 22 allele and would include the following possible combinations: (21,22) (22,22) (22,23)

From X population database

Allele	Frequency
21 (a)	0.187
22 (b)	0.182
23 (c)	0.156

LR = 1/(2ab + b² + 2bc) LR = 1/[(2(0.187)(0.182) + 0.182² + 2(0.182)(0.156)] **LR = 6.33**

Likelihood Ratio

LR = 6.33

The result can be described as follows:

It is 6.33 times more likely to observe the DNA profile if the mixed stain originated from the victim and the suspect than if it originated from the victim and an unknown person in X population.

Note: This information is from Schneider PM, Fimmers R, Keil W, Molsberger G, Patzelt D, Pflug W, Rothämel T, Schmitter H, Schneider H, Brinkmann B.The German Stain Commission: recommendations for the interpretation of mixed stains. Int J Legal Med. 2009 Jan;123(1):1-5.

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

Summarized from 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

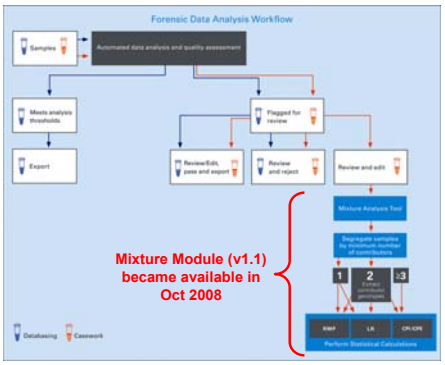
DAB Recommendations on Statistics

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

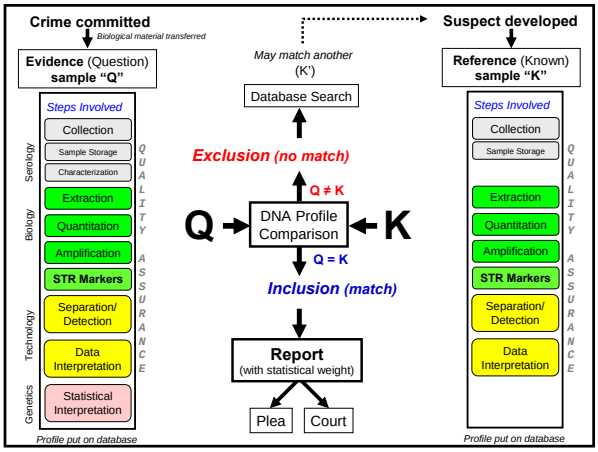
"The DAB finds either one or both PE or LR calculations acceptable and strongly recommends that one or both calculations be carried out whenever feasible and a mixture is indicated"

- 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.

GeneMapper ID-X v1.1



http://marketing.appliedbiosystems.com/images/Product_Microsite_Software/GeneMapper_IDX_1007/workflow2.html



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

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

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
ELSEVIER
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** – nothing yet…
  - a Mixture Interpretation subcommittee was started Jan 2007

Recent Article from 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.

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)
- Biostatistics for Parentage Analysis (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)

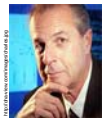

Secretary
Wolfgang Mayr
(Vienna, Austria)


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

	United States	Europe
Audits	FBI (DAB) Quality Assurance Standards	ENFSI Policies
	NDIS Procedures	
Validation	SWGDAM Guidelines	ISFG Recommendations (DNA Commission)
		National Recommendations
	Laboratory Protocols (SOPs)	Laboratory Protocols (SOPs)
Training & Experience	Individual Analyst Practice	Individual Analyst Practice
	Each Case Report	Each Case Report

Hopefully each conforms to the levels above it…

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

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**

German Stain Commission on DNA Mixtures

Rechtsmedizin 2006, 16 : 401 - 404

Rechtsmedizin 2006 · 16:401–404
DOI 10.1007/s00194-006-0411-1
Online publiziert: 16. November 2006

P. M. Schneider[1] · R. Fimmers[4] · W. Keil[2] · G. Molsberger[5] · D. Patzelt[6] · W. Pflug[7] · T. Rothämel[8] · H. Schmitter[3] · H. Schneider[9] · B. Brinkmann[10]
[1] Institut für Rechtsmedizin, Universitätsklinikum, Köln
[2] Institut für Rechtsmedizin, Ludwig-Maximilians-Universität, München
[3] Bundeskriminalamt, Wiesbaden
[4] Institut für Medizinische Biometrie, Informatik und Epidemiologie, Bonn
[5] Landeskriminalamt Nordrhein-Westfalen, Düsseldorf
[6] Institut für Rechtsmedizin, Bayr. Julius-Maximilians-Universität, Würzburg
[7] Landeskriminalamt Baden-Württemberg, Stuttgart
[8] Institut für Rechtsmedizin der Medizinischen Hochschule, Hannover
[9] Hessisches Landeskriminalamt, Wiesbaden
[10] Vorsitzender der Spurenkommission, Institut für Rechtsmedizin, Münster

Article in German
(English version published in Jan 2009)

General recommendations of the stain commission on the interpretation of DNA results from mixed stains

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

English Version

Int J Legal Med (2009) 123:1–5
DOI 10.1007/s00414-008-0244-4

REVIEW ARTICLE

The German Stain Commission: recommendations for the interpretation of mixed stains

P. M. Schneider · R. Fimmers · W. Keil · G. Molsberger · D. Patzelt · W. Pflug · T. Rothämel · H. Schmitter · H. Schneider · B. Brinkmann

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)

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)

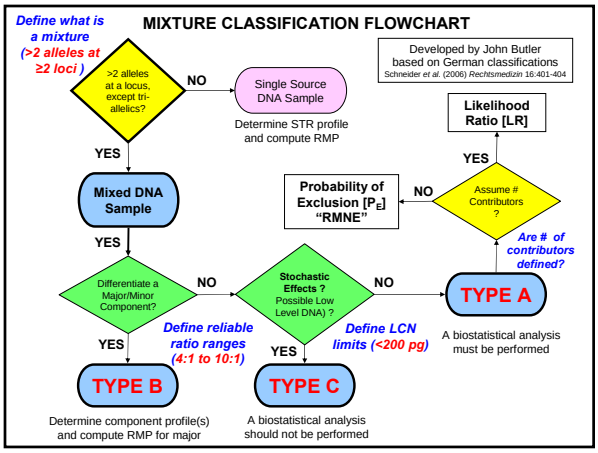
GEDNAP 32

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

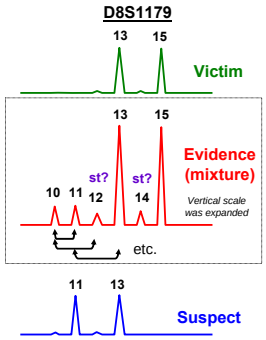
Slide from Peter Schneider (presented at EDNAP meeting in Krakow in April 2007)



German Type A,B, and C mixture classifications

- **Type A**, where major/minor contributors cannot be deduced, require stats
  - LR
  - RMNE (CPE/CPI)
- **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

Another Mixture Example



Conclusions from the evidence:

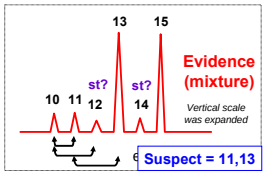
1. **Major contributor = 13,15 (victim) –** *to be expected with an intimate sample like a fingernail or vaginal swab*
2. Alleles 12 and 14 are **likely stutter products** of the major contributor's 13 and 15 alleles but could also be masking minor contributor alleles
3. A number of minor contributor 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

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