Northwest Association of Forensic Scientists (NWAFS)
Mixture Workshop

Thoughts on Mixture Interpretation

John M. Butler, Ph.D.
National Institute of Standards and Technology

Ft. Collins, CO
September 22, 2009

From ISFG Meeting (Sept 18, 2009)

- Max Baur (German statistician): “RMNE is a deficient method and we should not use it!”

NIST and NIJ Disclaimer

Funding: Interagency Agreement 2008-DN-R-121 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…

Mixture Analysis Efforts at NIST

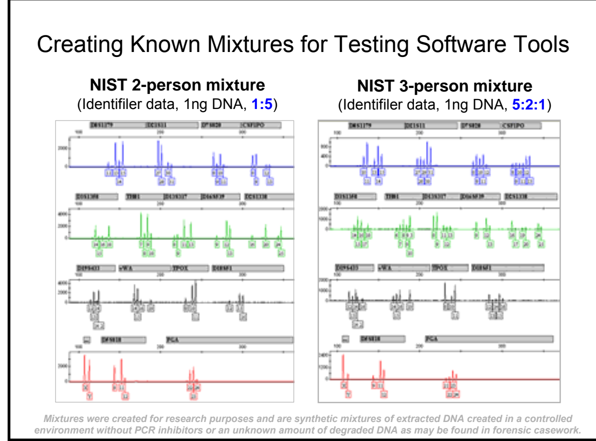
- Interlaboratory Studies: MSS1,2,3 and MIX05
  - **Future ones planned** when software tools and guidelines are available
- Software testing (see posters from AAFS 2008 and Promega 2008)
  - DNA_DataAnalysis (USACIL) – user's manual written
  - FSS-i3 (Promega)
  - Web-LSD (UTenn)
  - GeneMapper *ID-X* v1.1 (ABI)
  - GenoProof Mixture 1.0 (Qualitype)

  Plans to work with TrueAllele 3 software (Mark Perlin, Cybergenetics)
  Some work coordinated with NEST Project (Marshall University)
- Work with SWGDAM Mixture Committee
  - Case summaries
- Training workshops and discussion groups
  - AAFS Feb 2008, MD Apr 2008, FDLE May 2008, MD Dec 2008, AFDIL Jan 2009, Houston Jan 2009, NYC Mar 2009, WI May 2009, UT May 2009, NY TLs June 2009

Creating Known Mixtures for Testing Software Tools

NIST 2-person mixture (Identifiler data, 1ng DNA, 1:5)

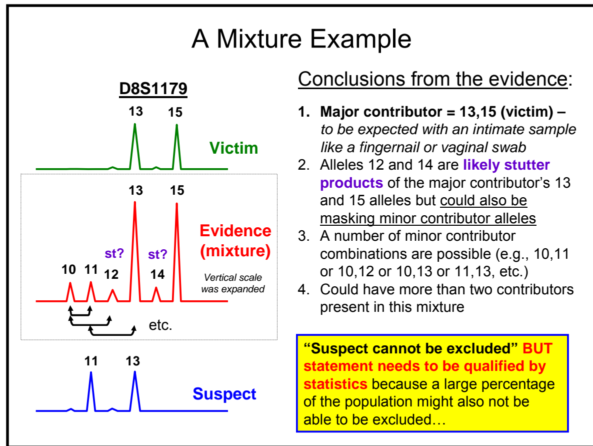
NIST 3-person mixture (Identifiler data, 1ng DNA, 5:2:1)

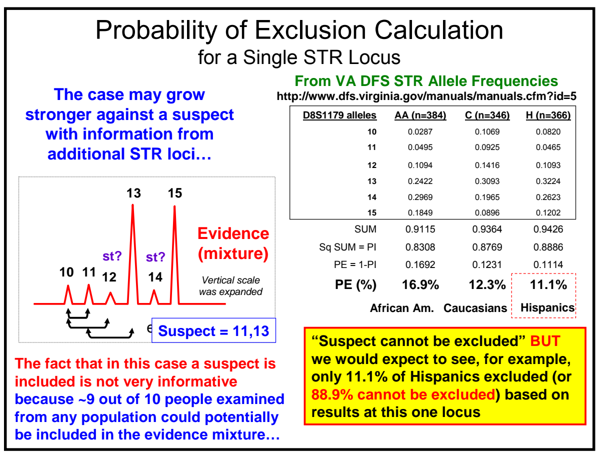


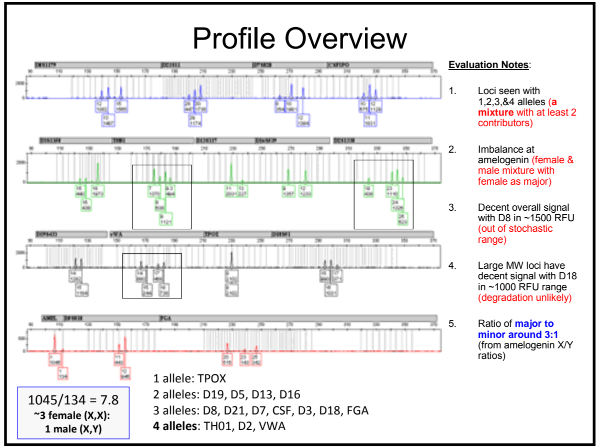
Mixtures were created for research purposes and are synthetic mixtures of extracted DNA created in a controlled environment without PCR inhibitors or an unknown amount of degraded DNA as may be found in forensic casework.

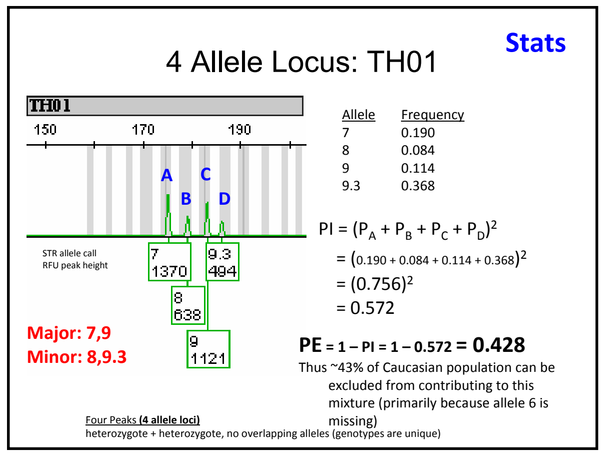
Presentation Plan

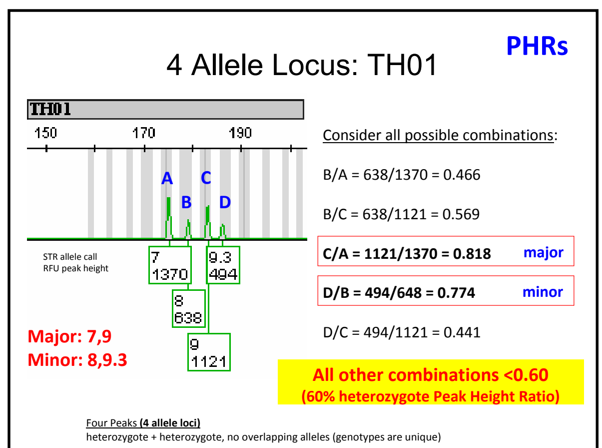
- Discuss and work through a mixture example
- What other labs are doing (or are planning to do)
- Status of SWGDAM mixture guidelines
- **Mixture principles** – what should and should not be done
  - Review of ISFG recommendations
- What to do with low level DNA mixture results
  - Interpretation thresholds
- Other topics?

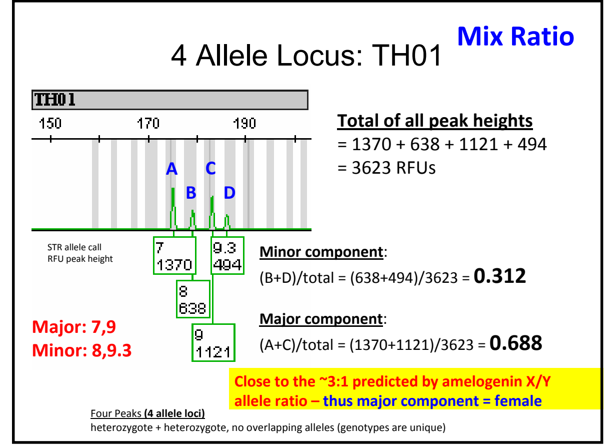

A Mixture Example
D8S1179
13 15
Victim
13 15
Evidence
(mixture)
st? st?
10 11 12 14
Vertical scale was expanded
etc.
11 13
Suspect
Conclusions from the evidence:
1. Major contributor = 13,15 (victim) – to be expected with an intimate sample like a fingernail or vaginal swab
2. Alleles 12 and 14 are likely 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…
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
13 15
Evidence
(mixture)
st? st?
10 11 12 14
Vertical scale was expanded
Suspect = 11,13
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…
"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


Profile Overview
Evaluation Notes:
1. Loci seen with 1,2,3,&4 alleles (a mixture with at least 2 contributors)
2. Imbalance at amelogenin (female & male mixture with female as major)
3. Decent overall signal with D8 in ~1500 RFU (out of stochastic range)
4. Large MW loci have decent signal with D18 in ~1000 RFU range (degradation unlikely)
5. Ratio of major to minor around 3:1 (from amelogenin X/Y ratios)
1045/134 = 7.8
~3 female (X,X):
1 male (X,Y)
1 allele: TPOX
2 alleles: D19, D5, D13, D16
3 alleles: D8, D21, D7, CSF, D3, D18, FGA
4 alleles: TH01, D2, VWA


Stats
4 Allele Locus: TH01
TH01
150 170 190
A C
B D
STR allele call
RFU peak height
7 1370
9.3 494
8 638
9 1121
Major: 7,9
Minor: 8,9.3
Allele Frequency
7 0.190
8 0.084
9 0.114
9.3 0.368
PI = (PA + PB + PC + PD)2
= (0.190 + 0.084 + 0.114 + 0.368)2
= (0.756)2
= 0.572
PE = 1 – PI = 1 – 0.572 = 0.428
Thus ~43% of Caucasian population can be excluded from contributing to this mixture (primarily because allele 6 is missing)
Four Peaks (4 allele loci)
heterozygote + heterozygote, no overlapping alleles (genotypes are unique)


PHRs
4 Allele Locus: TH01
TH01
150 170 190
A C
B D
STR allele call
RFU peak height
7 1370
9.3 494
8 638
9 1121
Major: 7,9
Minor: 8,9.3
Consider all possible combinations:
B/A = 638/1370 = 0.466
B/C = 638/1121 = 0.569
C/A = 1121/1370 = 0.818 major
D/B = 494/648 = 0.774 minor
D/C = 494/1121 = 0.441
All other combinations <0.60
(60% heterozygote Peak Height Ratio)
Four Peaks (4 allele loci)
heterozygote + heterozygote, no overlapping alleles (genotypes are unique)


Mix Ratio
4 Allele Locus: TH01
TH01
150 170 190
A C
B D
STR allele call
RFU peak height
7 1370
9.3 494
8 638
9 1121
Major: 7,9
Minor: 8,9.3
Total of all peak heights
= 1370 + 638 + 1121 + 494
= 3623 RFUs
Minor component:
(B+D)/total = (638+494)/3623 = 0.312
Major component:
(A+C)/total = (1370+1121)/3623 = 0.688
Close to the ~3:1 predicted by amelogenin X/Y allele ratio – thus major component = female
Four Peaks (4 allele loci)
heterozygote + heterozygote, no overlapping alleles (genotypes are unique)

Things to Avoid

- Post PCR clean-up (without threshold changes)
  - **No "enhanced interrogation techniques"**
  - "Water boarding" your DNA will lead to unreliable results
- Casework performed without documentation of assumptions
  - **No "illegal immigrants"**
- Using multiple stats on the same sample
  - **No "mixed marriages" of RMP and CPI**

Presentation Plan

- Discuss and work through a mixture example
- What other labs are doing (or are planning to do)
- Status of SWGDAM mixture guidelines
- **Mixture principles** – what should and should not be done
- Value of and difficulties with standardization (strategy, software, thresholds, etc.)
- What to do with low level DNA mixture results
- Other topics?

What Other Labs Are Doing or Planning…

- CFS Toronto
  - 240 pg cut-off with validated in-house qPCR assay
  - Single threshold of 50 RFU
- FBI and ATF
  - Two thresholds (detection: 50 RFU; stochastic: 200 RFU)
- Wisconsin
  - Setting DNA threshold and variable thresholds with different PCR cycle numbers and injection time
    - **30 cycles and 5 sec. injection = 100 RFU (Milwaukee Lab)**
    - **30 cycles and 10 sec. injection = 150 RFU (Milwaukee Lab)**
    - **32 cycles and 5 sec. injection = 300 RFU (Madison Lab)**
    - **32 cycles and 10 sec. injection = 600 RFU (Madison Lab)**
- New York State
  - Eight labs (NYC, state, and six county labs) looking into uniform protocols

CE User's Group (Primarily Maryland Labs)

- Gathering of all DNA analysts across ~16 labs
- Meets twice a year for detailed discussion
- Mixture principles discussion (April 10, 2008)
- **Mixture exercises discussion** (Dec 5, 2008)
  - 4 examples sent to all participants beforehand
  - Spent 1-2 hours discussing each one
- Threshold evaluations (June 4, 2009)

Mixture Questionnaire Summaries

…strbase\training\FL-May2008-Workshop.htm

2007 Workshop

STRs, CE, and Mixtures

42 participants from 13 different labs

28 responses

>80 analysts from 10 different FL labs

48 responses

76 responses representing >20 different laboratories

20 Questions Asked

Interpretation Guidelines
11 questions including:
What would you like to see in national guidelines on how to perform DNA mixture interpretation and statistical analysis?

Validation and Training
4 questions including:
What kind of training materials would be beneficial to help your laboratory more effectively solve mixtures?

Other Topics
5 questions including:
What are the biggest obstacles you face in your lab in terms of mixture interpretation?

A Few of the Responses…

from the Mixture Workshop Questionnaires (Nov 2007 and May 2008)

What would you like to see in national guidelines on how to perform DNA mixture interpretation and statistical analysis?

- General guidelines for how profiles should be interpreted & when profiles are inconclusive
- What stat calculations to use in various situations; when to use single source stats or mixture stats
- Loose guidelines that provide direction but don't overly limit subjectivity
- Mixture classification scheme
- More detail with examples
- Standard for using RFU and peak height ratios to determine major/minor across loci

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)
- Terry Coons (OR)
- Roger Frappier (CFS-Toronto)
- Ann Gross (MN BCA)
- **Bruce Heidebrecht (MD)**
- Phil Kinsey (MT)
- Jeff Modler (RCMP)
- Tamyra Moretti (FBI DNA Unit I)
- **Steven Myers (CA DOJ)**
- Joanne Sgueglia (MA)
- Gary Shutler (WA)

Everyone not at every meeting…

Have met 6 times:
Jan 2007
July 2007
Jan 2008
July 2008
Nov 2008
July 2009

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)
Cecelia Crouse (PBSO)
Allison Eastman (NYSP)
Steve Lambert (SC)

Also at Gaithersburg mtg (Nov 2008)
Todd Bille (ATF)
Hiron Poon (RCMP)

SWGDAM Efforts Thus Far

- Gathering mixture case summary information
- Developing some training materials
  - **AAFS 2008 mixture workshop** (available on STRBase Training section)
- Working to complete interpretation guidelines (hopefully by Jan 2010)

Will be discussed in more detail tomorrow…

Mixture Interpretation Workshop

http://www.cstl.nist.gov/biotech/strbase/training/AAFS2008_MixtureWorkshop.htm

AAFS (February 19, 2008)

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

- **John Butler (NIST)**
- Ann Gross (MN)
- George Carmody (Carleton U.)
- Gary Shutler (WA)
- Joanne Sgueglia (MA)
- Angela Dolph (Marshall U./NIST)
- Tim Kalafut (USACIL)

196 page handout

Budowle/FBI Mixture Paper

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

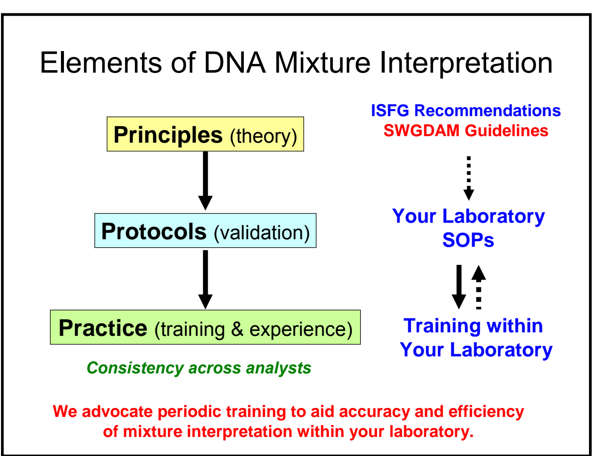
Mixture Interpretation: Defining the Relevant Features for Guidelines for the Assessment of Mixed DNA Profiles in Forensic Casework*

Does NOT represent the opinion of the SWGDAM Mixture Committee as guidelines are still in discussion and development ***(will likely be Jan 2010 before anything will be voted on by SWGDAM)***

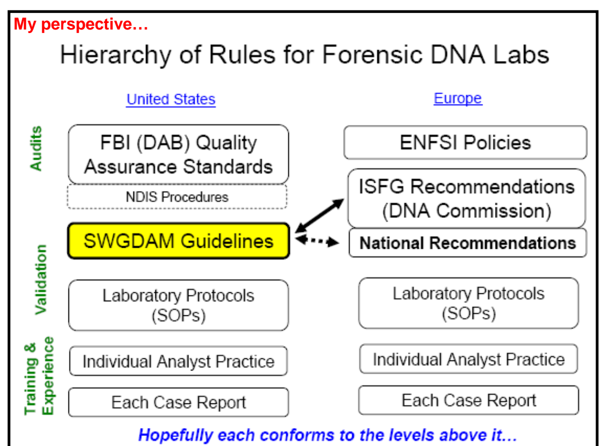
Mention of Mixtures in the July 2009 Revised Quality Assurance Standards (QAS)

- 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.**

Elements of DNA Mixture Interpretation



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



Statistical Approaches

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)

Statistics

- **Required for inclusion** (cannot be excluded) statements
- Choose one approach for the type of mixture and be consistent in application
- Recognize that differences exist with approaches
- Foley case (Apr 2009) statistical differences
  - CPE: 1 in 13 thousand
  - Subtraction: 1 in 23 million
  - TrueAllele Casework: 1 in 189 billion

TrueAllele is more powerful (has great potential), but can you explain the results in court?

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

Restricted vs Unrestricted LR

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…

Combined Probability of Exclusion (CPE)

Each locus is calculated separately and then combined for CPE

$CPE = 1 – (1 – PE_1)(1 – PE_2)(1 – PE_3)…(1-PE_N)$

Probability of exclusion at a single locus:

- ***The combined frequency of alleles detected (P)***
  P = frequency of allele 1 + frequency of allele 2 + frequency of allele 3, … N
- ***The combined frequency of alleles not detected (Q)***
  Q = 1 – P
- **$PE = Q^2 + 2Q(1-Q)$**

US Caucasian Data

Allele	Frequency
8	0.151
10	0.243
12	0.166

P = 0.151 + 0.243 + 0.166 = **0.56**

Q = 1 – 0.56 = **0.44**

PE = $(0.44)^2 + 2(0.44)(1-0.44)$ = 0.1936 + 0.4928

PE = 0.686

CPI = 1 – CPE

Calculation from CPI Perspective

Each locus is calculated separately and then combined for CPE

CPI or $P_{profile} = (P_{locus1}) (P_{locus2}) (P_{locus3}) … (P_{locus(N)})$

Probability of inclusion at a single locus:

- ***Individual frequencies are summed and then squared***
  PI or $P_{locus} = (p_1 + p_2 + p_3 + … + p_N)^2$
- $PE = 1 – P_{locus} = 1 – PI$
- **$PE = Q^2 + 2Q(1-Q)$**

Alleles present in the mixture / **Remaining possible alleles in the population**

Essentially $P^2 + 2 PQ + Q^2 = 1$ (PI: P^2; PE: $2PQ + Q^2$)

P + Q = 1 so P = 1 – Q and Q = 1 – P

Provides probability of an unrelated individual in the population is a contributor to the mixture at the loci examined

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 often 1** – the prosecution is 100% certain that the suspect is the perpetrator
- The denominator, 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.

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

Available online at www.sciencedirect.com

SCIENCE @ DIRECT

Forensic Science International 160 (2006) 90-101

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

Received 4 April 2006; accepted 10 April 2006

Available online 5 June 2006

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

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

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

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

Available online at www.sciencedirect.com

Forensic Science International 160 (2006) 90–101

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

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

ELSEVIER

Available online at www.sciencedirect.com

Forensic Science International 160 (2006) 89

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…"**

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

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

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
© Springer Medizin Verlag 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)

Danksagung. Für die konstruktive Mitarbeit an den hier vorgestellten Empfehlungen zur Bewertung von DNA-Mischspuren sei K. Anslinger (München), P. Berschick (Düsseldorf), M. Eckert (Wiesbaden), C. Hohoff (Münster), S. Jung (Würzburg) und J. Schnee-Griese (Stuttgart) herzlich gedankt.

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)

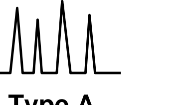
Adapted from Peter Schneider slide (presented at EDNAP meeting in Krakow in April 2007)

Mixture Classification Scheme

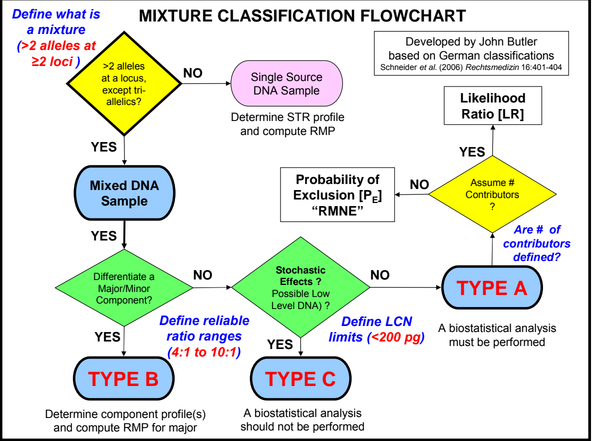
Schneider *et al.* (2006) *Rechtsmedizin* 16:401-404

(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**



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*

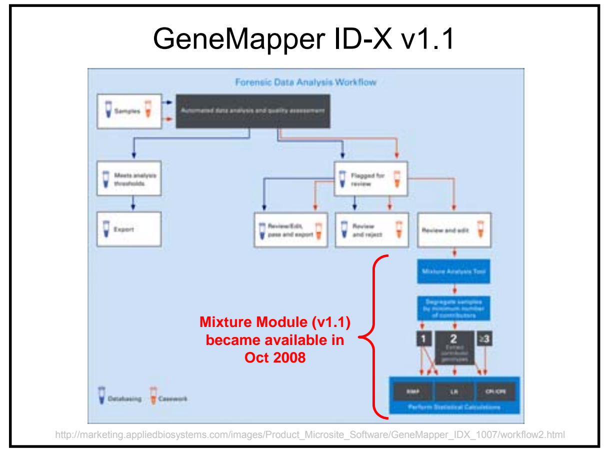
In general we agree with the recommendations of Gill et al. that are: (i) when possible peak height/area should be included in mixture interpretation; (ii) stutter position peaks at similar peak height/area as that of obligate minor contributor alleles should be considered as potential alleles in the interpretation and statistics calculation; and (iii) a stochastic threshold (termed ''dropout threshold'') should be defined.

Challenge of Standardization

- DNA quantitation
  - Need to use same method (assay or kit) and universal standard (e.g., NIST SRM 2372 or something traceable)
- STR kits
  - Need to use same kits and PCR conditions
- **Instrument sensitivities** (interlabs show 10 fold difference!)
  - Need to be consistent within labs and across labs
- Thresholds
  - Dependent on instrument sensitivities and PCR conditions
- Interpretation/Statistics
  - Using the same software (e.g., TrueAllele) will help
- Conclusions
  - Need to have same approach to drawing the line between inclusion/exclusion and inconclusive

Validation experiments: to confirm consistency within and between labs

GeneMapper ID-X v1.1



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

What is a true peak (allele)?

Peak detection threshold

Signal (S)
Noise (N)

Signal > 3x sd of noise

Peak height ratio (PHR)

Allele 1
Allele 2
Heterozygote peak balance

PHR consistent with single source Typically above 60%

Stutter percentage

True allele
Stutter product

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

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

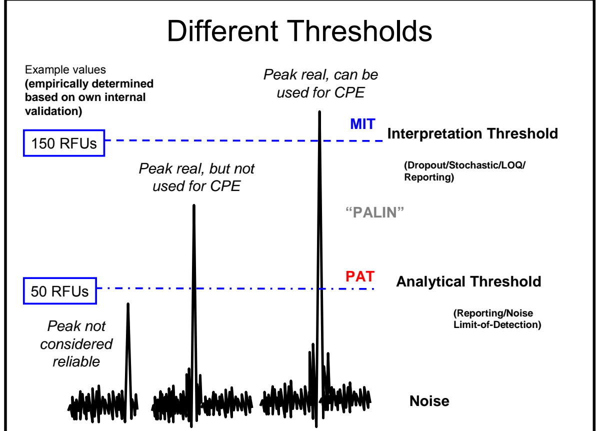
- Peak Amplitude Threshold (**PAT**)
- Match Interpretation Threshold (**MIT**)

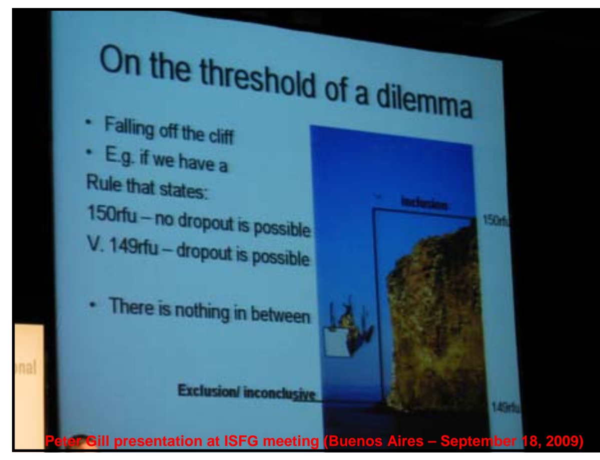
Pat Buchanan

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

Mitt Romney

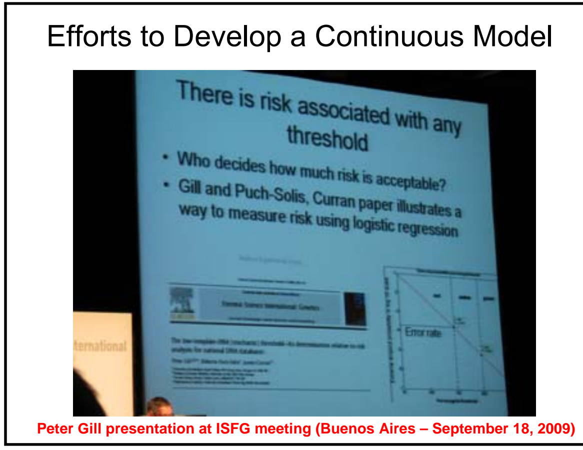
Different Thresholds





Peter Gill presentation at ISFG meeting (Buenos Aires – September 18, 2009)

Efforts to Develop a Continuous Model



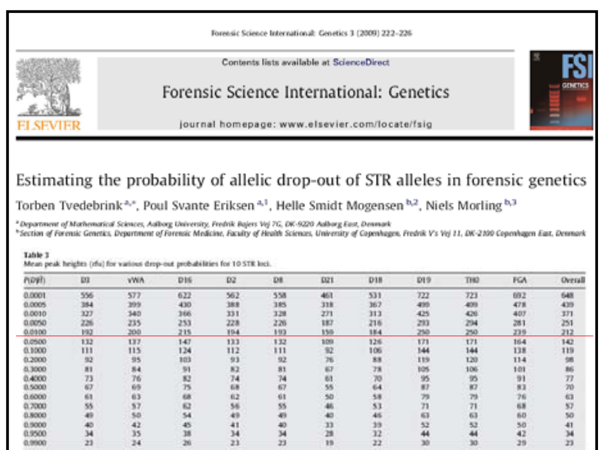
Peter Gill presentation at ISFG meeting (Buenos Aires – September 18, 2009)

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)



Forensic Science International: Genetics 3 (2009) 222–226

Contents lists available at ScienceDirect

Forensic Science International: Genetics

journal homepage: www.elsevier.com/locate/fsig

Estimating the probability of allelic drop-out of STR alleles in forensic genetics

Torben Tvedebrink[a,*], Poul Svante Eriksen[a,1], Helle Smidt Mogensen[b,2], Niels Morling[b,3]

[a] Department of Mathematical Sciences, Aalborg University, Fredrik Bajers Vej 7G, DK-9220 Aalborg East, Denmark

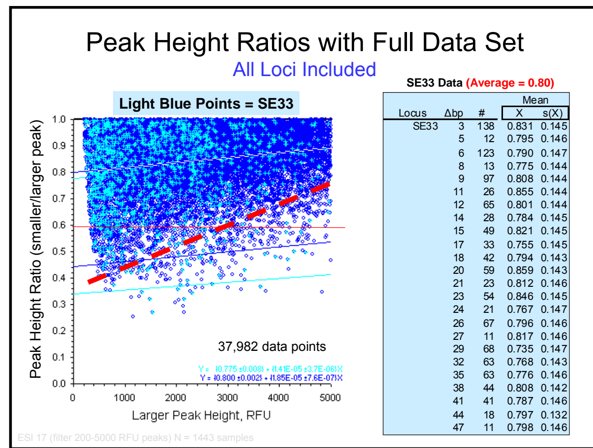
[b] Section of Forensic Genetics, Department of Forensic Medicine, Faculty of Health Sciences, University of Copenhagen, Frederik V's Vej 11, DK-2100 Copenhagen East, Denmark

Table 3
Mean peak heights (rfu) for various drop-out probabilities for 10 STR loci.

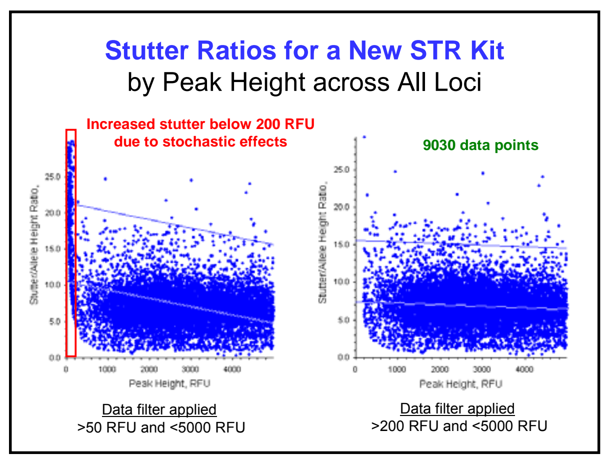
P(D|H)	D3	vWA	D16	D2	D8	D21	D18	D19	TH0	FGA	Overall
0.0001	556	577	622	562	558	461	531	722	723	692	648
0.0005	384	399	430	388	385	318	367	499	499	478	439
0.0010	327	340	366	331	328	271	313	425	426	407	371
0.0050	226	235	253	228	226	187	216	293	294	281	251
0.0100	192	200	215	194	193	159	184	250	250	239	212
0.0500	132	137	147	133	132	109	126	171	171	164	142
0.1000	111	115	124	112	111	92	106	144	144	138	119
0.2000	92	95	103	93	92	76	88	119	120	114	98
0.3000	81	84	91	82	81	67	78	105	106	101	86
0.4000	73	76	82	74	74	61	70	95	95	91	77
0.5000	67	69	75	68	67	55	64	87	87	83	70
0.6000	61	63	68	62	61	50	58	79	79	76	63
0.7000	55	57	62	56	55	46	53	71	71	68	57
0.8000	49	50	54	49	49	40	46	63	63	60	50
0.9000	40	42	45	41	40	33	39	52	52	50	41
0.9500	34	35	38	34	34	28	32	44	44	42	34
0.9900	23	24	26	23	23	19	22	30	30	29	23

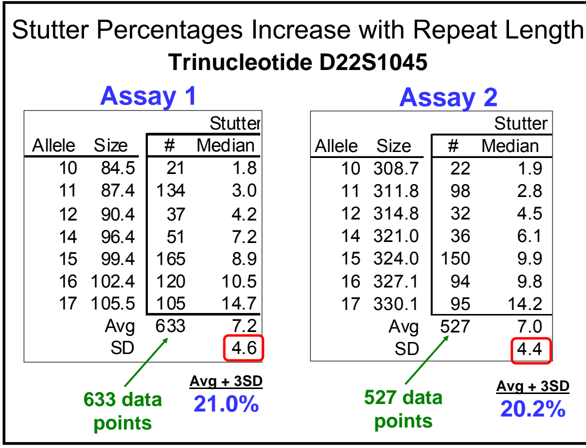
Reliable Mixture Interpretation Cannot Usually Be Performed with Low Level DNA

- 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



Locus	Δbp	#	Mean X	Mean s(X)
SE33	3	138	0.831	0.145
	5	12	0.795	0.146
	6	123	0.790	0.147
	8	13	0.775	0.144
	9	97	0.808	0.144
	11	26	0.855	0.144
	12	65	0.801	0.144
	14	28	0.784	0.145
	15	49	0.821	0.145
	17	33	0.755	0.145
	18	42	0.794	0.143
	20	59	0.859	0.143
	21	23	0.812	0.146
	23	54	0.846	0.145
	24	21	0.767	0.147
	26	67	0.796	0.146
	27	11	0.817	0.146
	29	68	0.735	0.147
	32	63	0.768	0.143
	35	63	0.776	0.146
	38	44	0.808	0.142
	41	41	0.787	0.146
	44	18	0.797	0.132
	47	11	0.798	0.146





Assay 1

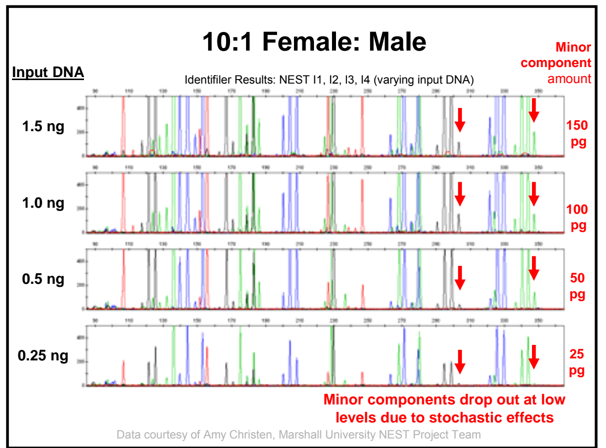
Allele	Size	Stutter #	Stutter Median
10	84.5	21	1.8
11	87.4	134	3.0
12	90.4	37	4.2
14	96.4	51	7.2
15	99.4	165	8.9
16	102.4	120	10.5
17	105.5	105	14.7
	Avg	633	7.2
	SD		4.6

Assay 2

Allele	Size	Stutter #	Stutter Median
10	308.7	22	1.9
11	311.8	98	2.8
12	314.8	32	4.5
14	321.0	36	6.1
15	324.0	150	9.9
16	327.1	94	9.8
17	330.1	95	14.2
	Avg	527	7.0
	SD		4.4

What to do with low level DNA mixtures?

- German Stain Commission "Category C"
  - Cannot perform stats because stochastic effects make it uncertain that all alleles are accounted for
- ISFG DNA Commission Recommendation #9
  - Stochastic effects limit usefulness
- Forensic DNA Typing, 3rd edition
  - Don't go outside the box without supporting validation



Further Thoughts

- Creating universal mixture interpretation protocols will be challenging but will have great value
- Gather mixture case summaries to understand what types of mixtures your labs are seeing
- Establish categories of mixtures (e.g., German A, B, C)
- Draw a line that can be applied consistently (based on your validation of stochastic thresholds) to where your labs will stop in their DNA mixture interpretation

Literature Worth Reviewing

- Clayton, T.M., Whitaker, J.P., Sparkes, R., Gill, P. (1998) Analysis and interpretation of mixed forensic stains using DNA STR profiling. *Forensic Sci. Int.* 91: 55-70.
- Gill, P., Brenner, C.H., Buckleton, J.S., Carracedo, A., Krawczak, M., Mayr, W.R., Morling, N., Prinz, M., Schneider, P.M., Weir, B.S. (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., Fimmers, R., Keil, W., Molsberger, G., Patzelt, D., Pflug, W., Rothämel, T., Schmitter, H., Schneider, H., Brinkman, B. (2009) The German Stain Commission: recommendations for the interpretation of mixed stains. *Int. J. Legal Med.* 123: 1-5.

Claims Have Been Made of No Consensus Regarding Mixture Interpretation

http://www.promega.com/geneticidproc/ussymp17proc/oralpresentations/Perlin.pdf

Scientific Validation of Mixture Interpretation Methods

Mark W. Perlin
Cybergenetics, Pittsburgh, PA

December 5, 2006

In the Proceedings of Promega's Seventeenth International Symposium on Human Identification

- **Different laboratories follow different mixture interpretation guidelines.** Moreover, different examiners within the same laboratory who are following the same guidelines often infer different STR profiles.
- **Therefore, there is no concordance in current forensic practice on what constitutes a "correct" mixture solution.** Thus, it is not possible to conduct a mixture interpretation concordance study in order to validate a mixture interpretation method.
- **DNA mixture evidence currently fails the general acceptance test of both Frye and Daubert, since there are no generally accepted methods for interpreting mixed stains.**

Interpretation of DNA Mixtures – European Consensus on Principles

Morling *et al.* (2007) *FSI Genetics* 1(3):291–292

“We propose that the German paper and the UK response can provide a model for other countries to follow in formulating their local national recommendations.”

“We consider this [support by a formal network of European and national forensic genetics, scientific organizations] to be **sufficient evidence of a scientific consensus** (or general agreement) to support the basic principles concerning the interpretation and formulation of the strength of evidence of DNA [mixture] results.”

Interpretation of DNA Mixtures – European Consensus on Principles

Morling *et al.* (2007) *FSI Genetics* 1(3):291–292

“We would like to draw the attention to…the need for:

(1) clarification of working practices for the interpretation of DNA profiles based on accreditation according to recognized laboratory standards such as ISO 17025,

(2) education in the interpretation of the weight of the evidence of complicated DNA profiles, and

(3) development of computer based expert systems that can assist in the interpretation of complicated DNA profiles.”

Software Programs to Aid Mixture Interpretation and Statistical Calculations

- FSS-i3
- GeneMapperID-X v1.1
- TrueAllele Casework
- DNA_DataAnalysis (USACIL)

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

Likelihood Ratios

Basic Math Terms

- When ‘+’ is used, this means ‘OR’
- When ‘x’ is used, this means ‘AND’
- Pr. is shorthand for probability
- Therefore…
  - the probability of *a* ‘AND’ *b* happening together is **Pr(*a* and *b*) = *a* x *b***
  - the probability of *a* ‘OR’ *b* happening together is **Pr(*a* or *b*) = *a* + *b***

Slide information from Peter Gill (ISFG 2007 workshop, Copenhagen, August 20-21, 2007)

Conditioning

- **Probabilities are conditional**, which means that the probability of something is based on a hypothesis
- In math terms, conditioning is denoted by a vertical bar
  - Hence, **Pr(a|b)** means ‘the probability of *a* ***given*** that *b* is true”
- The probability of an event *a* is dependent upon various assumptions—and these assumptions or hypotheses can change…

Slide information from Peter Gill (ISFG 2007 workshop, Copenhagen, August 20-21, 2007)

Probability Example – Will It Rain? (1)

Defining the Event and Assumptions/Hypotheses

- Let’s suppose that *a* is the probability of an event (e.g., **will it rain?**)
- What is the probability that it will rain in the afternoon – **Pr(*a*)**?
- This probability is dependent upon assumptions
  - We can look at the window in the morning and observe if it is sunny (s) or cloudy (c)
  - Pr(a) **if** it is sunny (s) is less than Pr(a) **if** it is cloudy (c)
- We can write this as **Pr(*a*|*s*)** and **Pr(*a*|c)**
  - Since sunny or cloudy are the only possibilities, Pr(s) + Pr(c) = 1
  - or **Pr(s) = 1 – Pr(c)**

Slide information from Peter Gill (ISFG 2007 workshop, Copenhagen, August 20-21, 2007)

Probability Example – Will It Rain? (2)

Examining Available Data

- Pr(a|s) and Pr(a|c) can be calculated from data
- How often does it rain in the afternoon when its sunny in the morning?
  - 20 out of 100 observations so **Pr(*a*|*s*) = 0.2**
- How often does it rain in the afternoon when it is cloudy in the morning?
  - 80 out of 100 observations so **Pr(*a*|*c*) = 0.8**

Slide information from Peter Gill (ISFG 2007 workshop, Copenhagen, August 20-21, 2007)

Probability Example – Will It Rain? (3)

Formation of the Likelihood Ratio (LR)

- The LR compares two probabilities to find out which of the two probabilities is the most likely

The probability that it will rain in the afternoon when it is cloudy in the morning or **Pr(*a*|*c*)** is divided by the probability that it will rain in the afternoon when it is sunny in the morning or **Pr(*a*|s)**

$$LR = \frac{\Pr(a\,|\,c)}{\Pr(a\,|\,s)} = \frac{0.8}{0.2} = 4$$

Slide information from Peter Gill (ISFG 2007 workshop, Copenhagen, August 20-21, 2007)

Probability Example – Will It Rain? (4)

Explanation of the Likelihood Ratio

$$LR = \frac{\Pr(a\,|\,c)}{\Pr(a\,|\,s)} = \frac{0.8}{0.2} = 4$$

- The probability that it will rain is 4 times more likely **if** it is cloudy in the morning than **if** it is sunny in the morning.
- The word **if** is very important here. It must always be used when explaining a likelihood ratio otherwise the explanation could be misleading.

Slide information from Peter Gill (ISFG 2007 workshop, Copenhagen, August 20-21, 2007)

Likelihood Ratios in Forensic DNA Work

- We evaluate the evidence (*E*) relative to alternative pairs of hypotheses
- Usually these hypotheses are formulated as follows:
  - The probability of the evidence if the crime stain originated with the suspect or Pr(*E*|*S*)
  - The probability of the evidence if the crime stain originated from an unknown, unrelated individual or Pr(*E*|*U*)

$$LR = \frac{\Pr(E\,|\,S)}{\Pr(E\,|\,U)}$$

The numerator

The denominator

Slide information from Peter Gill (ISFG 2007 workshop, Copenhagen, August 20-21, 2007)

The Likelihood Ratio Must Be Stated Carefully

- The probability of the evidence is *x* times more likely **if** the stain came from the suspect Mr. Smith than **if** it came from an unknown, unrelated individual.
- It is not appropriate to say: "The probability that the stain came from Mr. Smith." because we must always include the conditioning statement – i.e., **always make the hypothesis clear in the statement.**
- Always use the word '**if**' when using a likelihood ratio to avoid this trap

Slide information from Peter Gill (ISFG 2007 workshop, Copenhagen, August 20-21, 2007)

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**

Relationship between Likelihood Ratio (LR) and Random Match Probability (RMP)

- For single source samples or deduced major component profiles in a mixture…

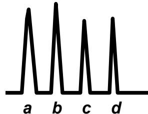
$$LR = \frac{1}{RMP} \quad \text{or} \quad RMP = \frac{1}{LR}$$

Example #1

A Single Locus from a 2-Person Mixture

- Consider a simple **two person mixture** with one locus consisting of two heterozygotes with non-overlapping alleles
- If the suspect is *ab*, then there must be another (unknown person) who is *cd*

Suspect = a,b



Forget peak heights for the time being

Adapted from Peter Gill (ISFG 2007 workshop, Copenhagen, August 20-21, 2007)

Example #1

The Two Hypotheses Are Formed…

- **Prosecution (H_p)**: The DNA result has come from the suspect and one unknown person, or **Pr(*E*|*S,U*)**
- **Defense (H_d)**: The DNA result has come from two unknown people, or **Pr(*E*|*U₁,U₂*)**

$$LR = \frac{\Pr(E\,|\,S,U)}{\Pr(E\,|\,U_1,U_2)}$$

Suspect = a,b

Forget peak heights for the time being

Adapted from Peter Gill (ISFG 2007 workshop, Copenhagen, August 20-21, 2007)

Example #1

Formulating the Numerator
(Prosecution Hypothesis)

- If the prosecution hypothesis is true, then we would expect genotype *ab* to be present with 100% probability or Pr=1.
- The chance of seeing an unknown person of type *cd* is the frequency of that type in the population or $2p_cp_d$, where p_c is the allele frequency for allele c.
- **Pr(*E*|*S,U*)** = 1 x $2p_cp_d$ **=** $\mathbf{2p_cp_d}$

Suspect = a,b

Forget peak heights for the time being

Adapted from Peter Gill (ISFG 2007 workshop, Copenhagen, August 20-21, 2007)

Example #1

Formulating the Denominator
(Defense Hypothesis)

- The defense claims that the evidence could come from any two random individuals
- We must work out **all possible pairwise combinations** from alleles *abcd* and their probabilities (genotype frequencies)

Individual #1	Individual #2	Products
ab $2p_ap_b$	cd $2p_cp_d$	$2p_ap_b$ x $2p_cp_d$ $4p_ap_bp_cp_d$
ac	bd	$2p_ap_c$ x $2p_bp_d$ $4p_ap_bp_cp_d$
ad	bc	$2p_ap_d$ x $2p_bp_c$ $4p_ap_bp_cp_d$
cd	ab	$2p_cp_d$ x $2p_ap_b$ $4p_ap_bp_cp_d$
bd	ac	$2p_bp_d$ x $2p_ap_c$ $4p_ap_bp_cp_d$
bc	ad	$2p_bp_c$ x $2p_ap_d$ $4p_ap_bp_cp_d$
Sum of products		$\mathbf{24p_ap_bp_cp_d}$

Reverse combinations (cd/ab, bd/ac, bc/ad)

Multiplied because you are considering individual #1 **AND** individual #2

Added because you are considering any of the possibilities (combination 1,2,3,4,5, **OR** 6)

Pr(*E*|*U₁,U₂*) = $\mathbf{24p_ap_bp_cp_d}$

Adapted from Peter Gill (ISFG 2007 workshop, Copenhagen, August 20-21, 2007)

Example #1

Formulating the Likelihood Ratio

- The numerator and denominator are combined to form the LR
- And common elements in both numerator and denominator are eliminated to simplify the algebraic equation…

$$LR = \frac{\Pr(E\,|\,S,U)}{\Pr(E\,|\,U_1,U_2)} = \frac{\cancel{2}\cancel{p_c}\cancel{p_d}}{2\cancel{4}p_ap_b\cancel{p_c}\cancel{p_d}} = \frac{1}{12p_ap_b}$$

12

Adapted from Peter Gill (ISFG 2007 workshop, Copenhagen, August 20-21, 2007)

All LR Calculations Follow the Same Basic Rules Just Shown

- Form hypotheses
  - Keep in mind what you are conditioning on
- The LR numerator belongs to the prosecution
- The LR denominator belongs to the defense
- Numerator and denominator are combined and equation is simplified
- Allele frequency values are placed into the equation for each locus
- **The LR from each locus is combined through multiplication if the loci are independently inherited (i.e., the product rule) to form a LR for the entire profile**

Example #2

Another Example…

- The evidentiary mixture profile is from a semen stained vaginal swab and possesses alleles a, b, c, and d.
- The suspect is a,b and the victim is c,d.
- Because it is reasonable to assume that the victim's alleles would be present on the swab (i.e., an intimate sample), we can condition on this…

Suspect = a,b
Victim = c,d

Adapted from Peter Gill (ISFG 2007 workshop, Copenhagen, August 20-21, 2007)

Example #2

With an Intimate Sample, the Hypothesis Changes…

- **Prosecution (H_p)**: The DNA result has come from the suspect and the victim, or **Pr(*E*|*S,V*)**
- **Defense (H_d)**: The DNA result has come from the victim and one unknown person, or **Pr(*E*|*U,V*)**

$$LR = \frac{\Pr(E \mid S, V)}{\Pr(E \mid U, V)}$$

Suspect = a,b
Victim = c,d

Adapted from Peter Gill (ISFG 2007 workshop, Copenhagen, August 20-21, 2007)

Example #2

Formulating the Numerator
(Prosecution Hypothesis)

- The prosecution hypothesis (S+V) is completely explains the evidence. Hence, the probability is Pr=1
- **Pr(*E*|*S,V*) = 1 x 1 = 1**

Suspect = a,b
Victim = c,d

Adapted from Peter Gill (ISFG 2007 workshop, Copenhagen, August 20-21, 2007)

Example #2

Formulating the Denominator
(Defense Hypothesis)

- The defense hypothesis is that the presence of alleles *a* and *b* are the result of an unknown person – and they concede that alleles *c* and *d* come from the victim
- Since the frequency of an unknown, unrelated individual possessing alleles *a* and *b* in the population is $2p_ap_b$, where p_a is the allele frequency for allele *a* and p_b is the allele frequency for allele *b*, then
- **Pr(*E*|*U,V*)** = $2p_ap_b$ x 1 = $\mathbf{2p_ap_b}$

Suspect = a,b
Victim = c,d

Adapted from Peter Gill (ISFG 2007 workshop, Copenhagen, August 20-21, 2007)

Example #2

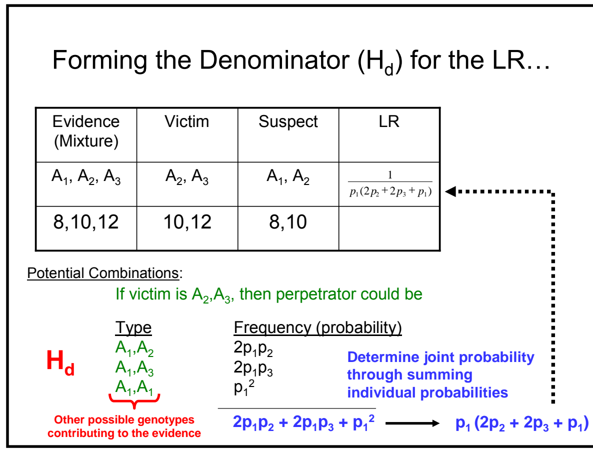
Formulating the Likelihood Ratio

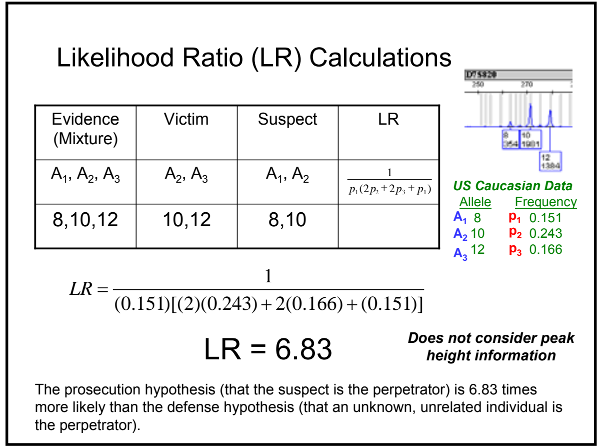
- The numerator and denominator are combined to form the LR

$$LR = \frac{\Pr(E \mid S, V)}{\Pr(E \mid U, V)} = \frac{1}{2 p_a p_b}$$

- **Note that this LR is the same as for a non-mixed sample comprising the suspect alone.**
- This example then is an illustration of simplification by "subtraction" (victim's alleles are being removed from mathematical consideration…).

Adapted from Peter Gill (ISFG 2007 workshop, Copenhagen, August 20-21, 2007)





Likelihood Ratios for the Following Hypotheses

H_p: The mixture contains the DNA of the victim and the suspect

H_d: The mixture contains the DNA of the victim and an unknown, unrelated individual

Evidence (Mixture)	Victim	Suspect	LR
A_1, A_2, A_3, A_4	A_1, A_2	A_3, A_4	$\frac{1}{2p_3p_4}$
A_1, A_2, A_3	A_1, A_2	A_1, A_3 or A_2, A_3 or A_3, A_3	$\frac{1}{p_3(2p_1+2p_2+p_3)}$
A_1, A_2, A_3	A_1, A_1	A_2, A_3	$\frac{1}{2p_2p_3}$
A_1, A_2	A_1, A_2	A_1, A_1 or A_1, A_2 or A_2, A_2	$\frac{1}{(p_1+p_2)^2}$
A_1, A_2	A_1, A_1	A_1, A_2 or A_2, A_2	$\frac{1}{p_2(2p_1+p_2)}$
A_1, A_1	A_1, A_1	A_1, A_1	$\frac{1}{p_1^2}$

Adapted from Buckleton (2005) *Forensic DNA Evidence Interpretation*, Table 7.1, p. 229