Invited Presentation to the American Society of Crime Laboratory Directors/
Laboratory Accreditation Board (ASCLD/LAB)
January 8, 2015 – Raleigh, NC

DNA Mixture Interpretation: *State-of-the-Art*

Michael D. Coble, PhD

John M. Butler, PhD

National Institute of Standards and Technology

Presentation Outline

Mike
- Challenges that exist with DNA mixtures
  - Stacking, number of contributors (relatives), drop-out with low level

Mike
- NIST interlaboratory studies (MIX05 & MIX13)

Mike
- Concerns raised with MIX13 study
  - Why does this range of results exist?

John
- Comparison to other fields (e.g., cell line authentication)

John
- Accreditation issues
  - The QAS is not specific enough to provide much help here

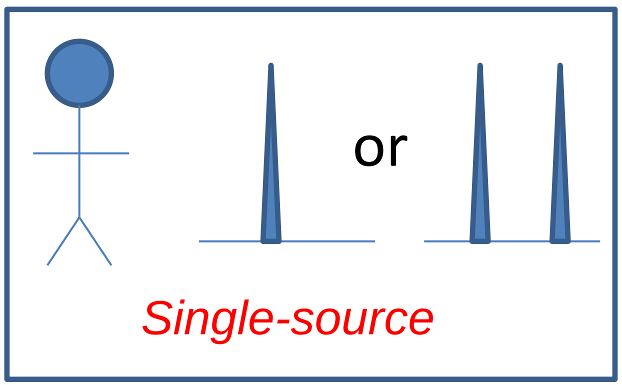
John
- Validation needs for the community
  - Can we reliably extrapolate from 2-person mixture studies?
- Closing thoughts: *where do we go from here?*

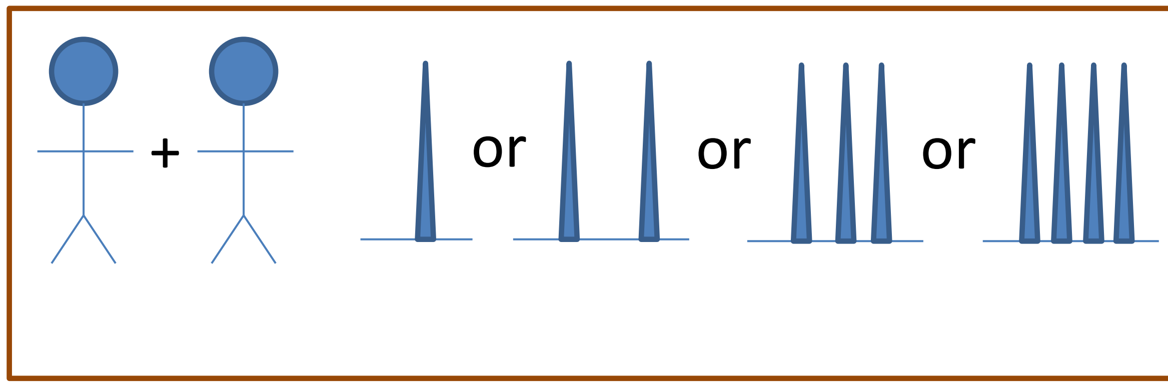
Challenges that Exist with DNA Mixtures

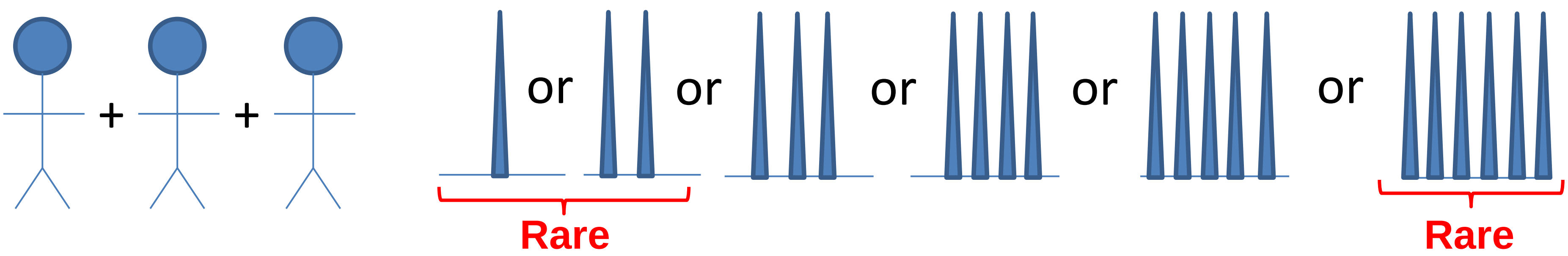
Challenges with Complex DNA Mixtures

Complex mixture = more than two individuals contribute to a biological sample; often with low amounts of DNA

- **Interpretation <u>uncertainty increases</u> (errors are more likely)**
  - Allele stacking because mixture contributors have peaks in common







Challenges with Complex DNA Mixtures

Complex mixture = more than two individuals contribute to a biological sample; often with low amounts of DNA

- **Interpretation <u>uncertainty increases</u> (errors are more likely)**
  - Allele stacking because mixture contributors have peaks in common
  - Number of contributors becomes more uncertain

		Appearing as				
Configuration	*N* contributor mixture	1	2	3	4	5
Existing CODIS	6	1.93E-41	1.81E-09	8.21E-02	0.8739	0.9993
Proposed CODIS with SE33		0	4.58E-21	1.18E-04	0.3303	0.9384
Proposed CODIS without SE33		1.87E-69	7.71E-18	2.24E-03	0.6963	0.9981
Fusion		1.50E-81	3.76E-22	1.59E-04	0.5220	0.9937

Challenges with Complex DNA Mixtures

Complex mixture = more than two individuals contribute to a biological sample; often with low amounts of DNA

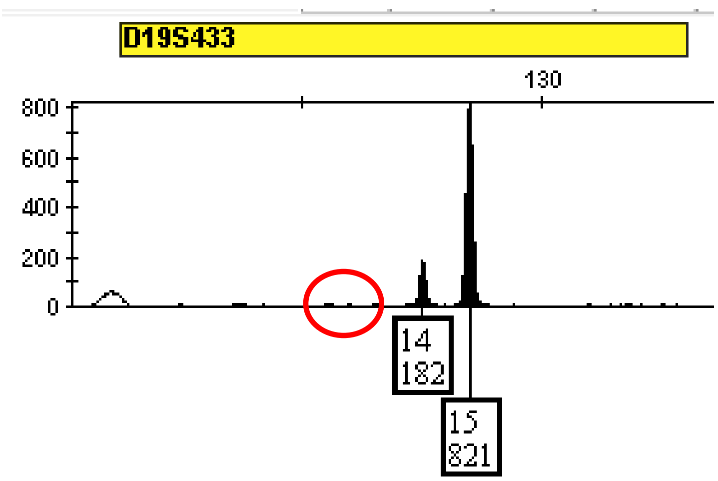
- **Interpretation <u>uncertainty increases</u> (errors are more likely)**
  - Allele stacking because mixture contributors have peaks in common
  - Number of contributors becomes more uncertain
  - Drop-out issues where allele information is lost when you attempt to work with low amounts of DNA

<u>Contributors</u>

A = 15, 15

B = 14, 15

C = **12**, 14

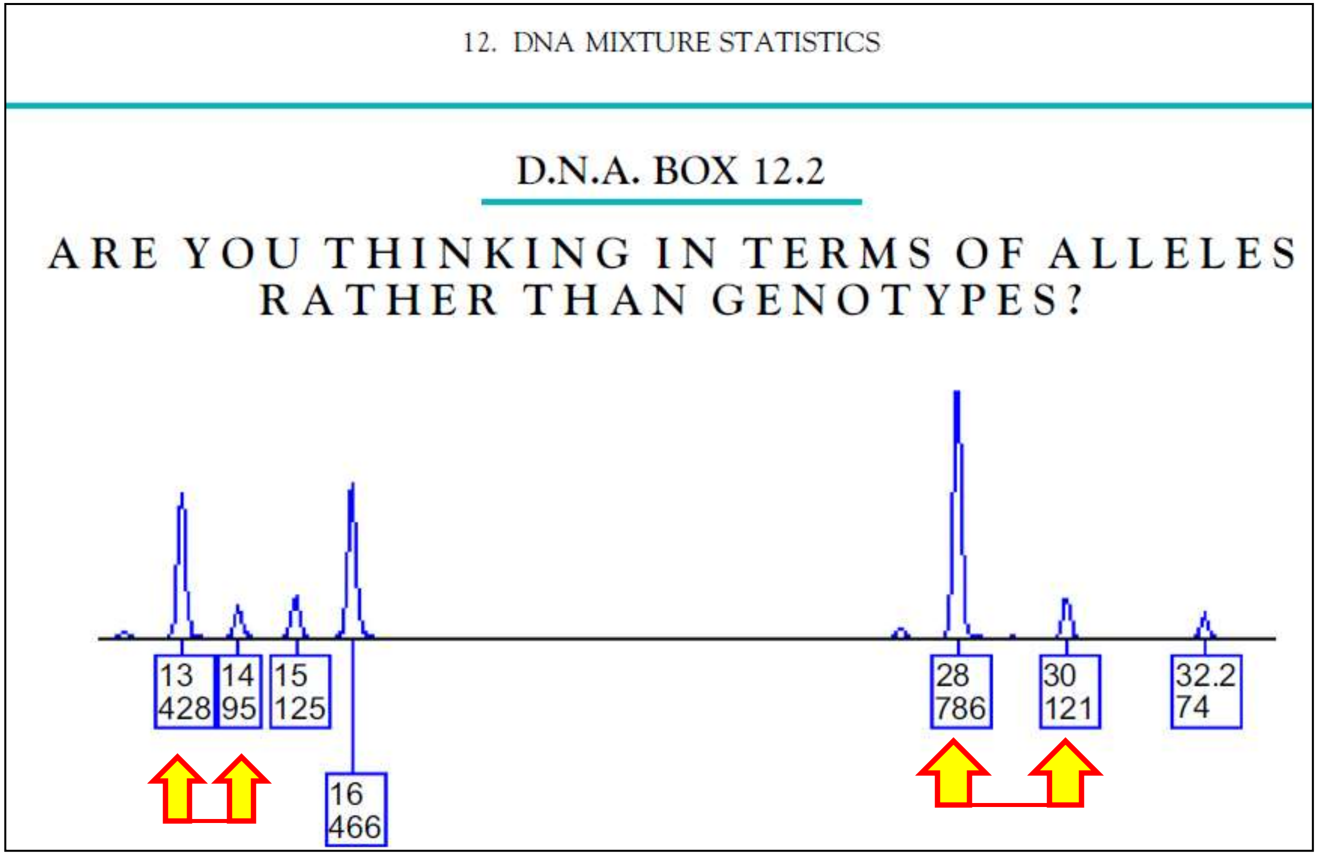


Thinking in Terms of Genotypes vs Alleles

- CPI – The probability that a random, unrelated person would be included as a contributor to the observed DNA mixture.
- CPI provides a simple statistical model for equating observed alleles with all possible genotypes.
- The focus is on ALLELES and not GENOTYPES

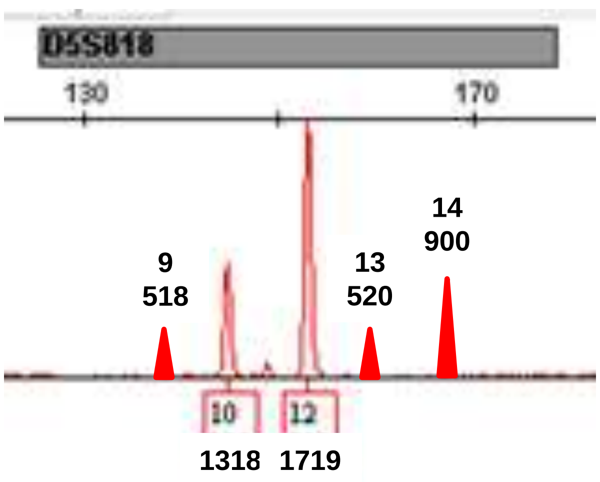
Thinking in Terms of Genotypes vs Alleles

Would you include or exclude a reference sample that is **13,14** and **28,30** at these two loci?



Butler, J.M. (2015) *Advanced Topics in Forensic DNA Typing: Interpretation*. Elsevier Academic Press: San Diego, p. 312

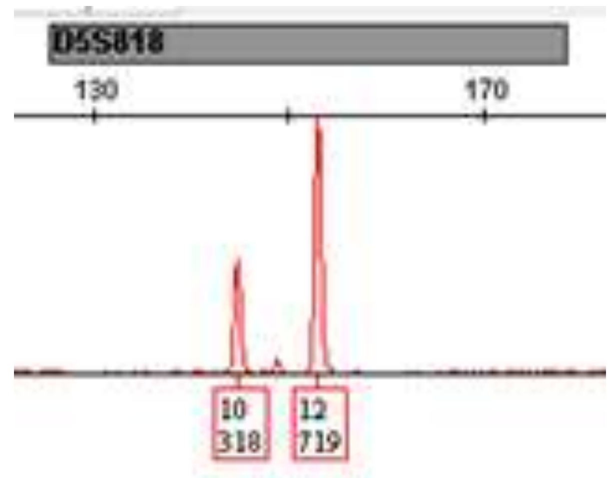
A cautionary note on using CPI when drop-out is possible



$$PI = (f_9 + f_{10} + f_{12} + f_{13} + f_{14})^2$$

$$PI = (0.049+0.051+0.361+0.384+0.141)^2$$

PI = 0.986 or **1 in 1.01**



$$PI = (f_{10} + f_{12})^2$$

$$PI = (0.051 + 0.361)^2$$

PI = 0.169 or **1 in 5.92**

Drop-out ***inflates*** your statistics for CPI (not conservative!)

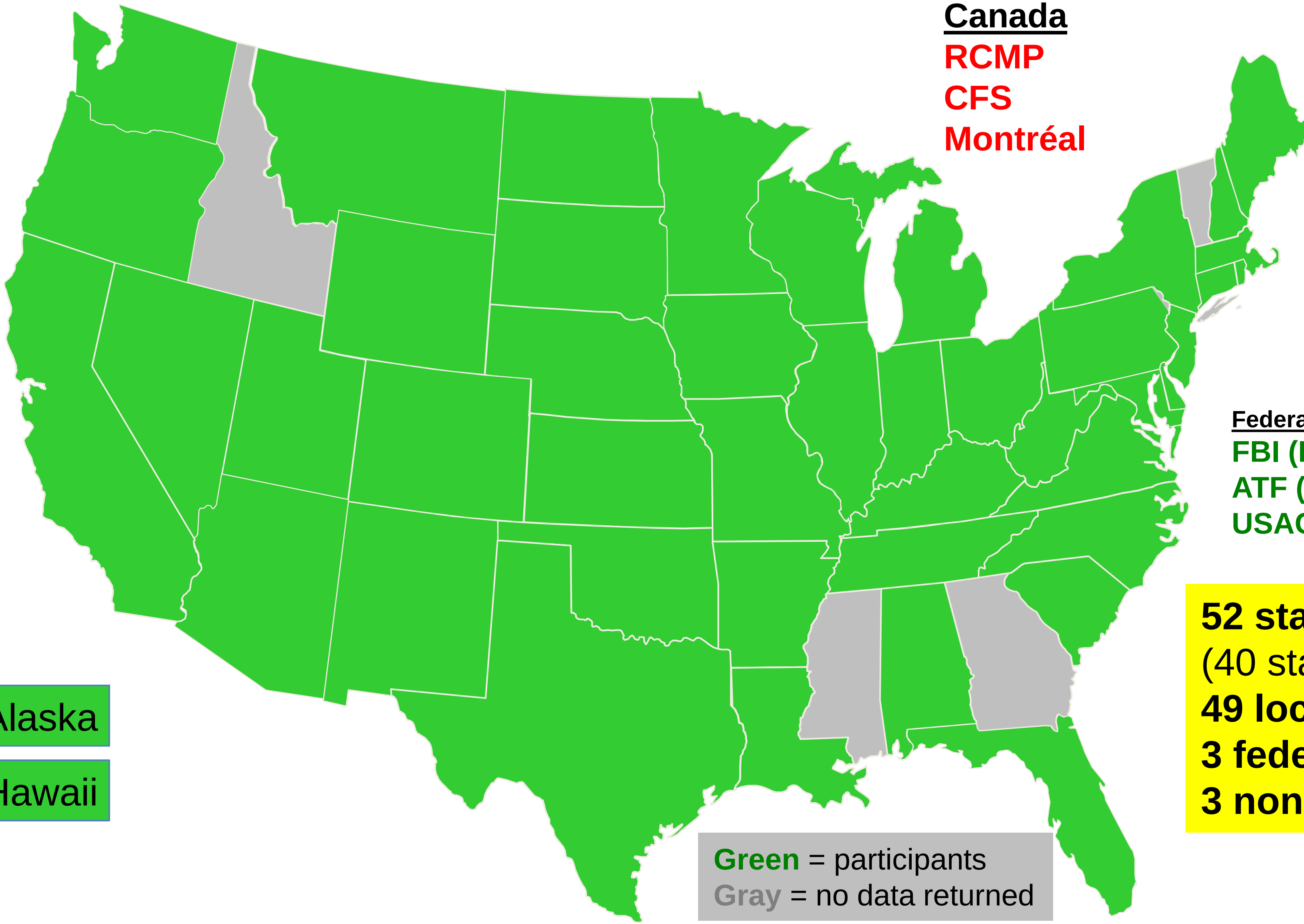
Summary of DNA Mixture Interlaboratory Studies

Study	Year	# Labs	# Samples	Mixture Types
MSS 1	1997	22	11 stains	ss, 2p, 3p
MSS 2	1999	45	11 stains	ss, 2p, 3p
MSS 3	2000-01	74	7 extracts	ss, 2p, 3p
MIX05	2005	69	4 cases (.fsa)	only 2p
MIX13	2013	108	5 cases (.fsa)	2p, 3p, 4p

- Other recent studies
  - UK Regulator
  - USACIL

MIX13 Participants from **108 Laboratories**

46 states had at least one lab participate



Canada
RCMP
CFS
Montréal

Federal Labs
FBI (DOJ)
ATF (DOJ)
USACIL (DOD)

Alaska

Hawaii

52 state labs
(40 states)
49 local labs
3 federal
3 non-U.S.

Green = participants
Gray = no data returned

Purpose of MIX13 Cases

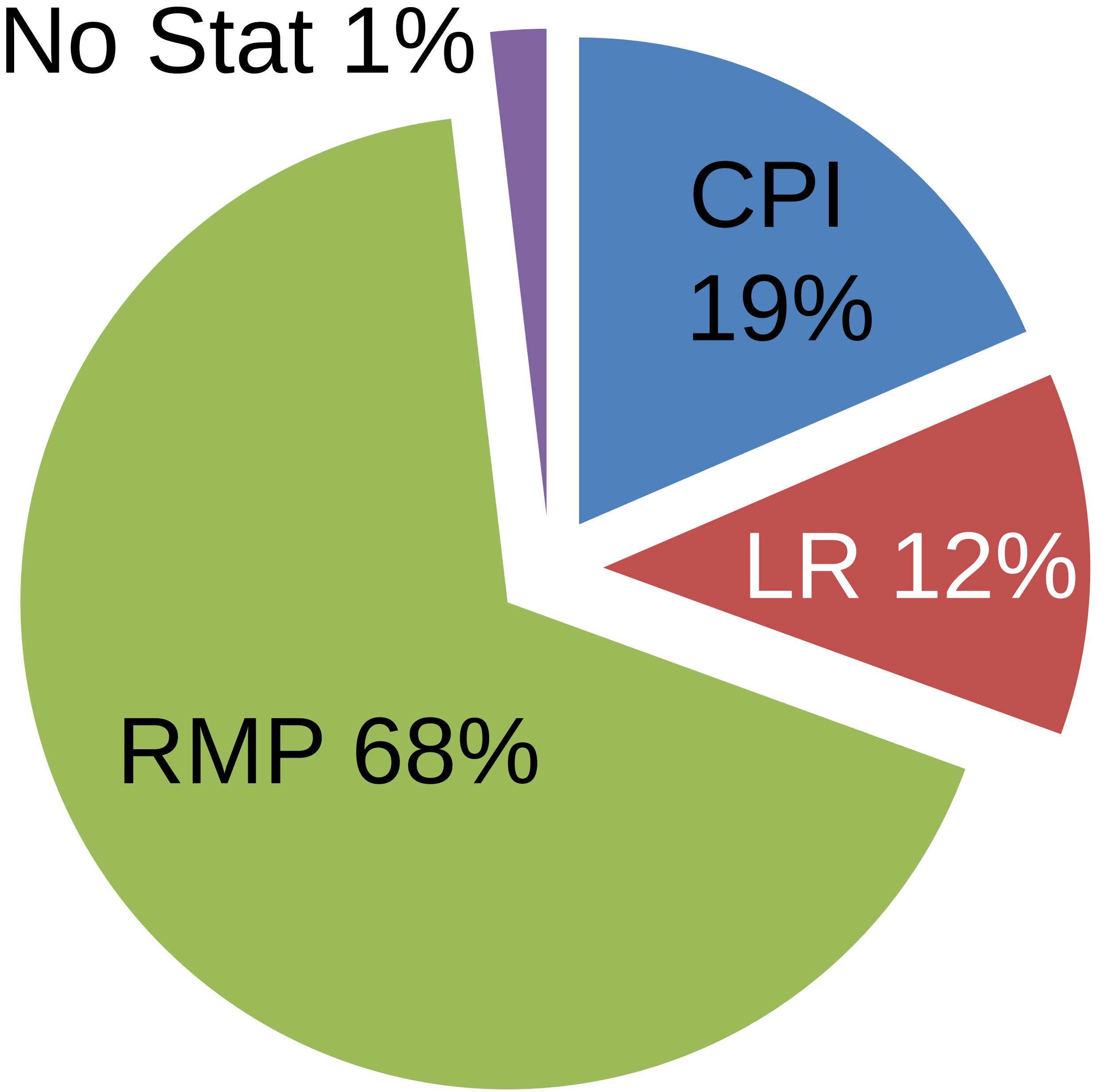
	Challenge provided to study responses
Case 1	~1:1 mixture (2-person)
Case 2	**Low template** profile with potential dropout (3-person)
Case 3	Potential **relative** involved (3-person)
Case 4	Minor component (2-person)
Case 5	Complex mixture (>3-person) with **# of contributors**; inclusion/exclusion issues

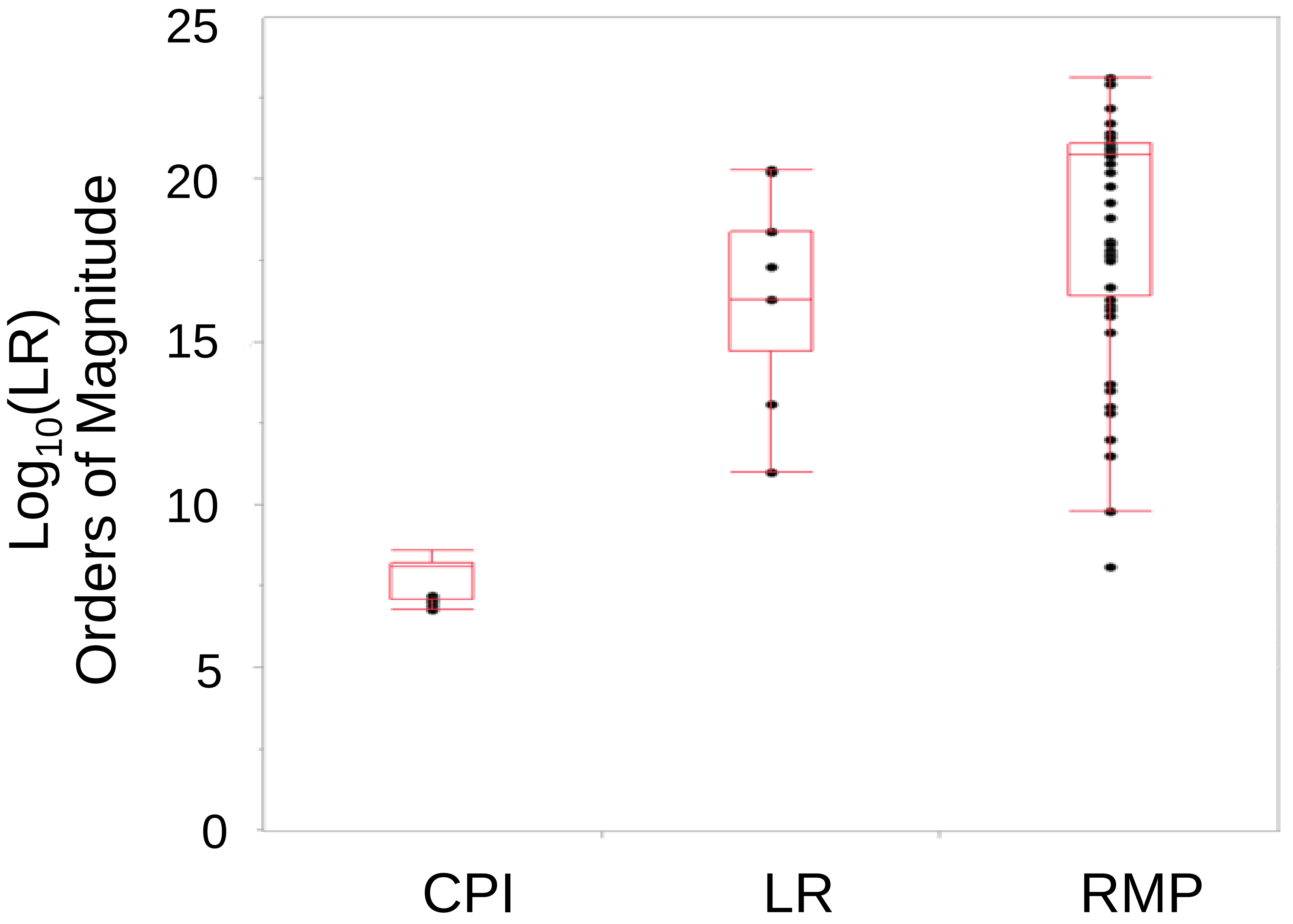
According to German Stain Commission (2009) mixture types: 1 = A, 2 = C, 3 = ?, 4 = B, 5 = ?

MIX13 Study (Case 01)

- Summary – Mock sexual assault, 2 person 50:50 mixture, all alleles above a ST of 150 RFU.

- Purpose – How many labs would consider the victim's profile and determine genotypes (deconvolution) for a mRMP statistic?


No Stat 1%
CPI 19%
LR 12%
RMP 68%



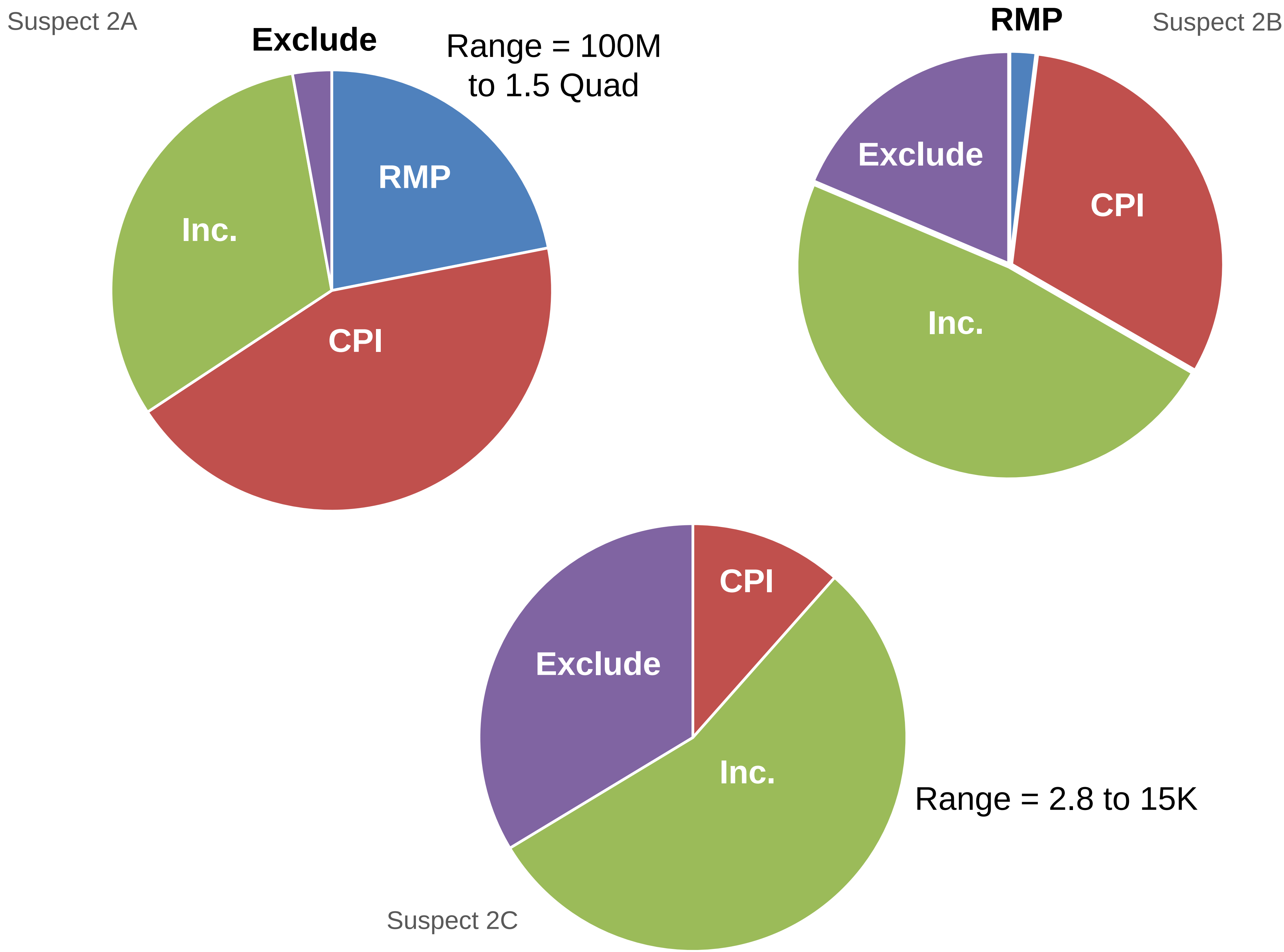
Log10(LR)
Orders of Magnitude
25
20
15
10
5
0
CPI
LR
RMP

MIX13 Study (Case 02)

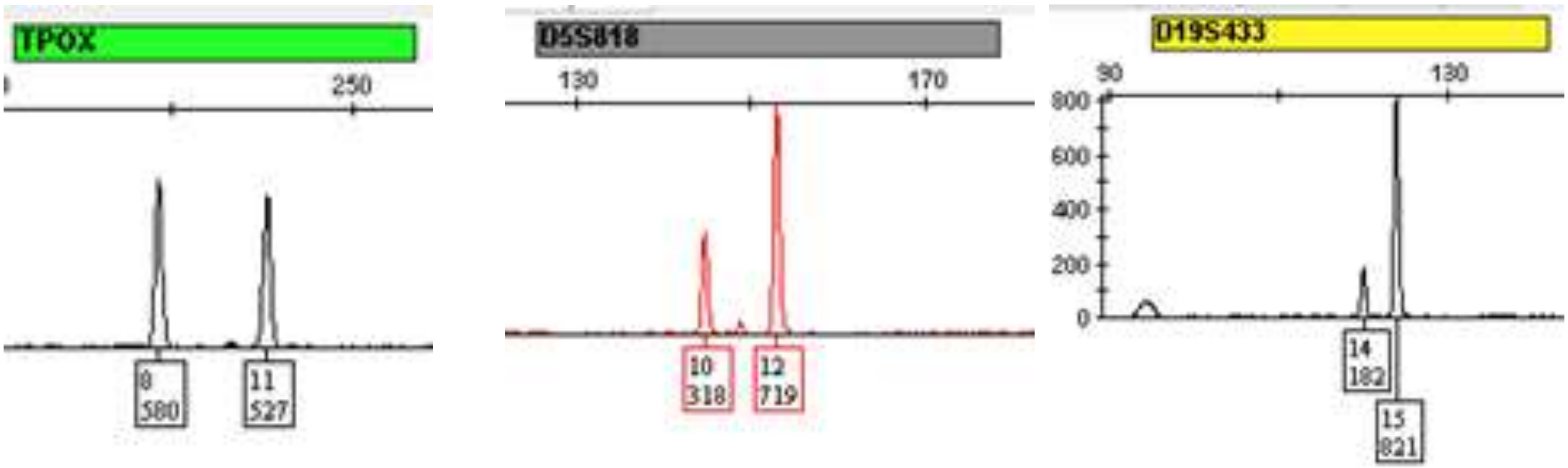
- Summary – Mock handgun (touch DNA), 3 person **6**:**1.5**:**1** mixture, total DNA amplified was 300 pg, potential for drop-out with the 2 low-level contributors. An additional contributor profile (suspect D) was provided, but is not in the mixture.
- Purpose – How many labs would consider this mixture as too complex to interpret?

Primary Goals

- Most labs – CPI for some combination of Suspects A, B and C using a limited number of loci.
- One lab included Suspect D (Not in the mixture).



Suspect 2A
Exclude
Range = 100M to 1.5 Quad
RMP
Inc.
CPI
RMP
Suspect 2B
Exclude
CPI
Inc.
CPI
Exclude
Inc.
Range = 2.8 to 15K
Suspect 2C


TPOX
250
8
580
11
527
D5S818
130
170
10
318
12
719
D19S433
90
130
800
600
400
200
0
14
182
15
821

~ 1 in 35

Intra-Laboratory Results (n = 8)

Analyst	Suspect A	Suspect B	Suspect C	Suspect D
1	Inconclusive - A, B, C			Excluded
2	6.74 Quad	**23.6**	Excluded	Excluded
3	Inconclusive - A, B, C			Excluded
4	**9.4** for A, B, C			Excluded
5	4.1 Quint	**37**	Excluded	Excluded
6	**230** for A, B		Inconclusive	Excluded
7	**9.4** for A, B		Excluded	Excluded
8	**37.3** for A, B		Excluded	Excluded

MIX13 Study (Case 03)

- Summary – Mock sexual assault, 3 person **7**:**2**:**1** mixture, The two minor contributors are brothers, An additional contributor profile (suspect 3B) was provided, but is not in the mixture.

- Most of the suspected brother’s alleles are masked in the mixture

- Purpose – Given the relatedness of the individuals in the mixture, is this too complex for interpretation?

Primary Goals

- Only one lab included Suspect B (Not in the mixture)

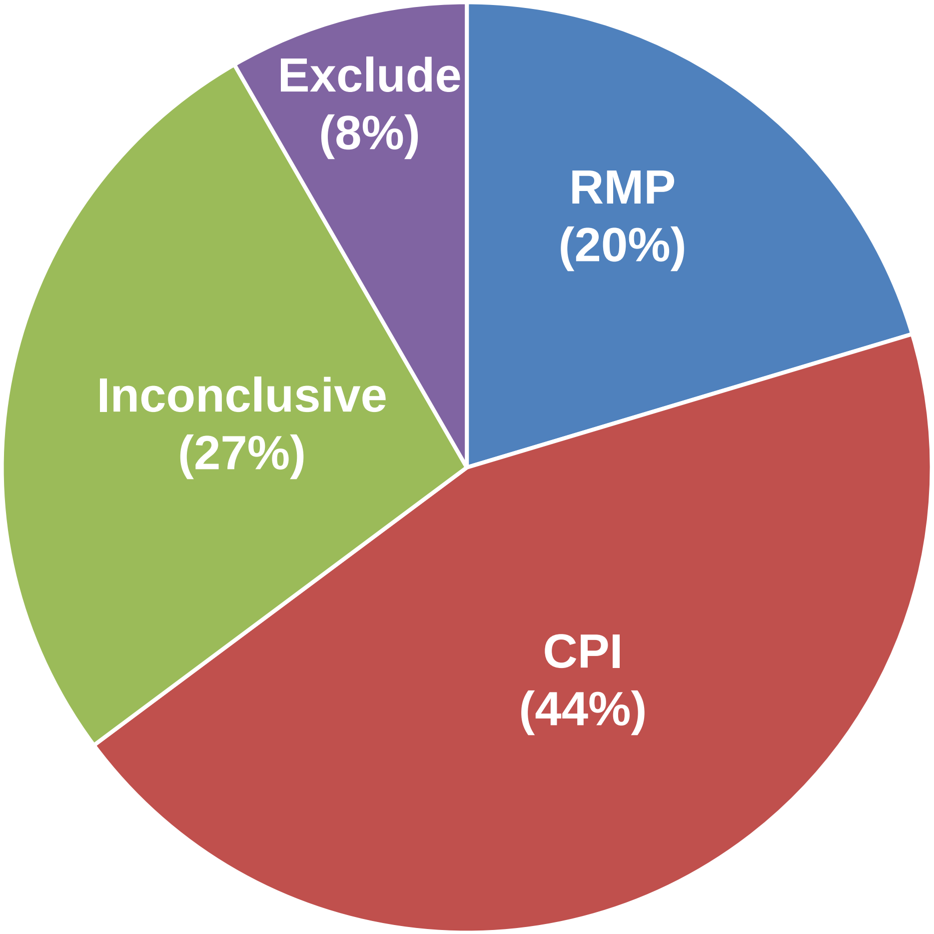
- Most labs are using CPI stats for this case…

RMNE

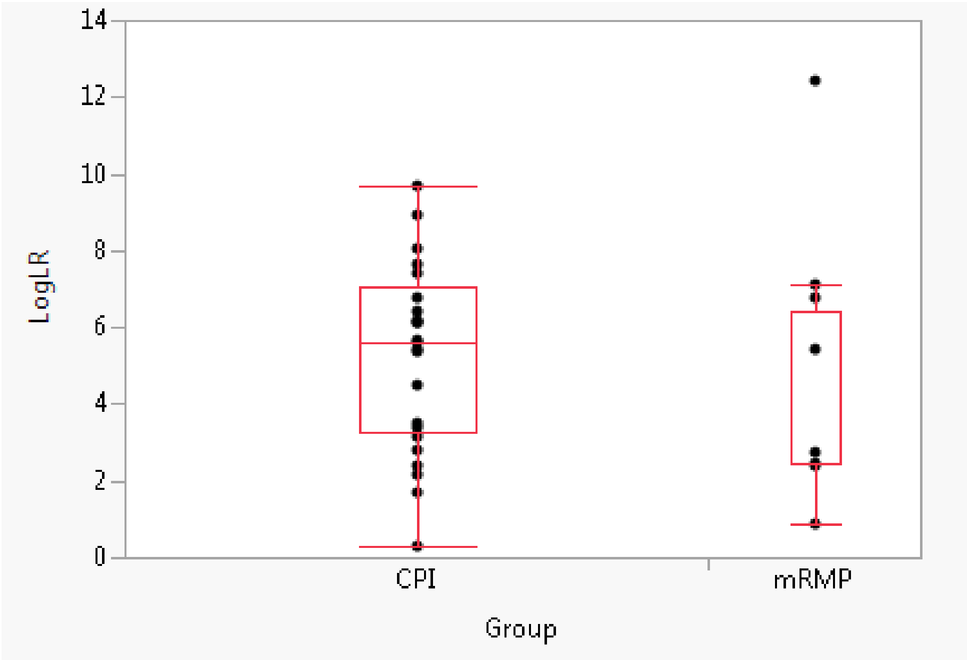
- Random Man Not Excluded (CPE/CPI) – The probability that a ***random person*** (unrelated individual) would be excluded as a contributor to the observed DNA mixture.

- Only a few labs have stated this – “Due to the relatedness of the exemplars submitted for comparison, a statistical analysis cannot be provided at this time.”

Case 03



Exclude
(8%)
RMP
(20%)
Inconclusive
(27%)
CPI
(44%)



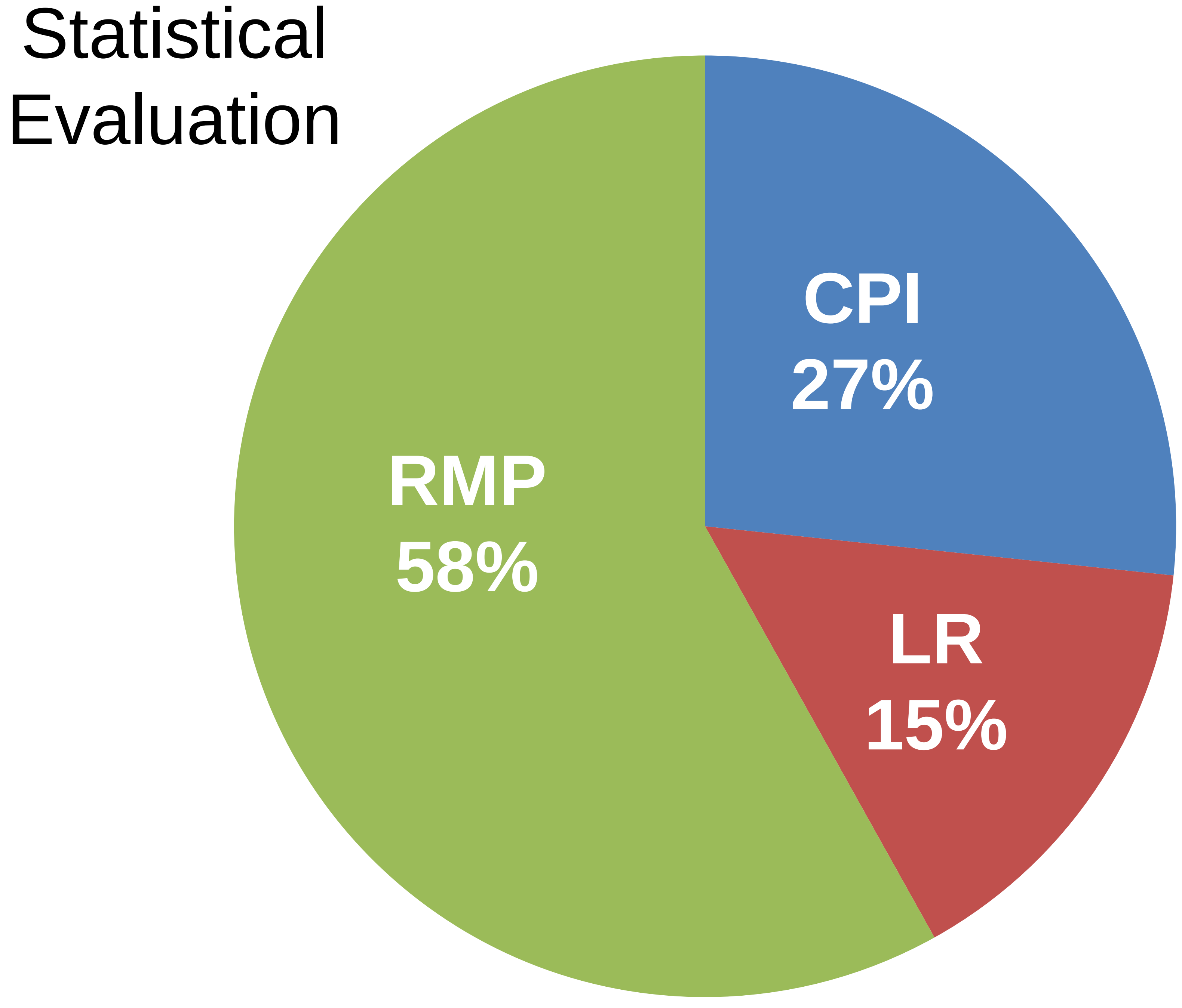
LogLR
0
2
4
6
8
10
12
14
CPI
mRMP
Group

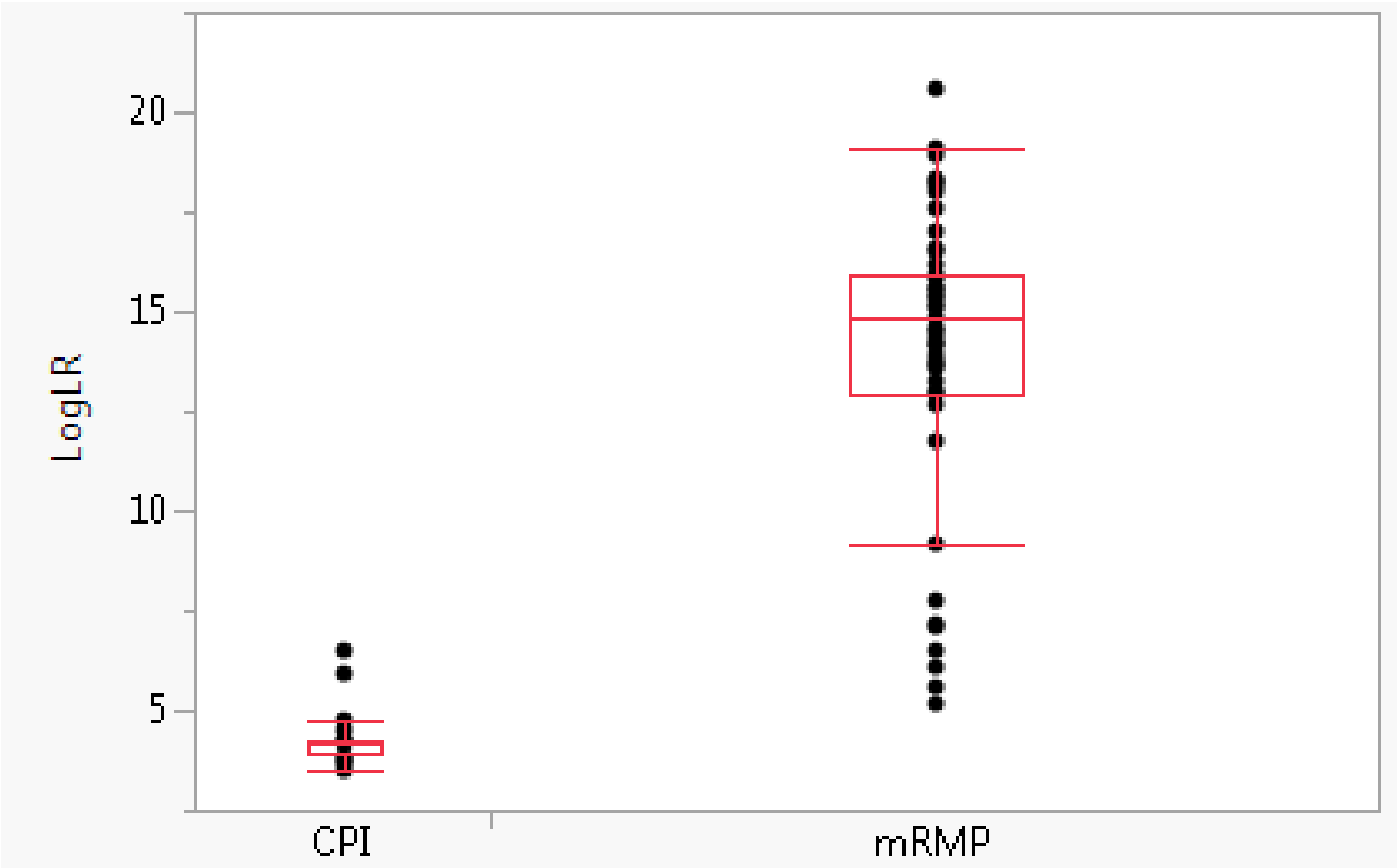
MIX13 Study (Case 04)

- Summary – Mock sexual assault, 2 person **3.5**:**1** mixture, minor component has alleles below the ST of 150 (required by all labs!)

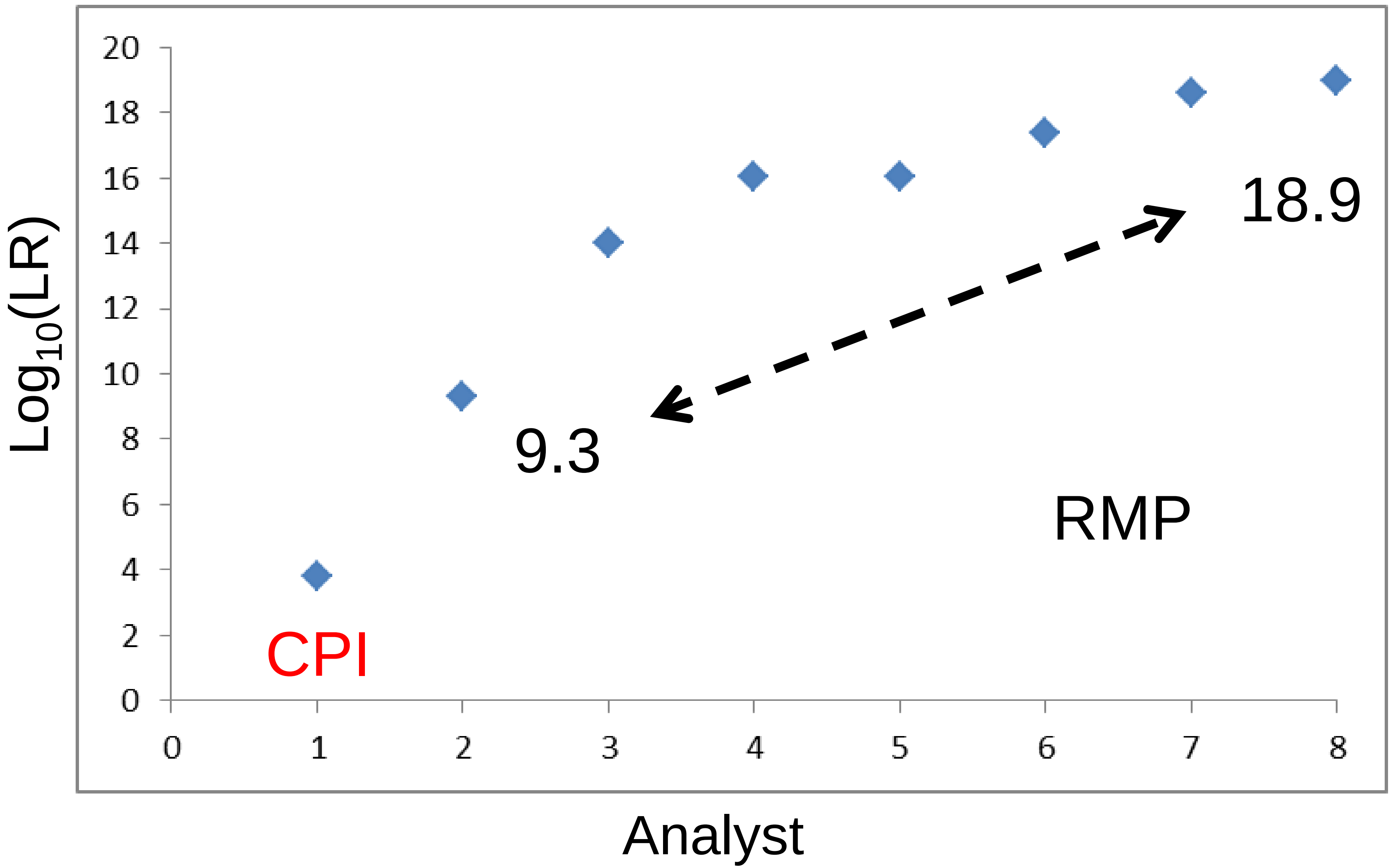
- Purpose – How many labs would attempt to separate the two components?

- With all labs using the AT/ST – how much variation is expected?


Statistical
Evaluation
CPI
27%
RMP
58%
LR
15%


20
15
10
5
LogLR
CPI
mRMP

Intra-Laboratory Results (n = 8)



MIX13 Study (Case 05)

- Summary – Mock bank robbery with ski mask evidence (touch DNA), 4 person **1:1:1:1** mixture.
- However – this mixture had no more than 4 alleles at any locus (appears as a 2p mixture). 2 of the 4 contributors were provided along with a non-contributor.
- Purpose – How many labs would consider this mixture as too complex to interpret?

MIX13 Case 5 Outcomes with Suspect C

(whose genotypes were not present in the mixture)

# Labs	Report Conclusions	Reasons given
7	Exclude Suspect C	detailed genotype checks (ID+); TrueAllele negative LR (ID+); assumed major/minor and suspects did not fit (ID+); 4 of 18 labs noted Penta E missing allele 15 (PP16HS)
3	Inconclusive with C only (A & B included)	All these labs used PP16HS
22	Inconclusive for A, B, and C	
76	**Include & provide CPI statistics**	*All over the road…*

Range of CPI stats for Caucasian population:

FBI allele frequencies: **1 in 9** (Labs 12 & 54) to **1 in 344,000** (Lab 107)

Concerns raised with MIX13 study

Concerns raised with MIX13 (1)

- From my perspective – labs are substituting **CPI** for **Interpretation** when it comes to mixtures – even for simple mixtures.

Case 04 – D16 locus
If 10% stutter from the 12 allele (163 RFU) is part of the 11 allele, then the remaining peak (70 RFU) would be below the ST

No CPI labs excluded D16 from the stat (N = 22)

POI = 11, 12

Concerns raised with MIX13 (2)

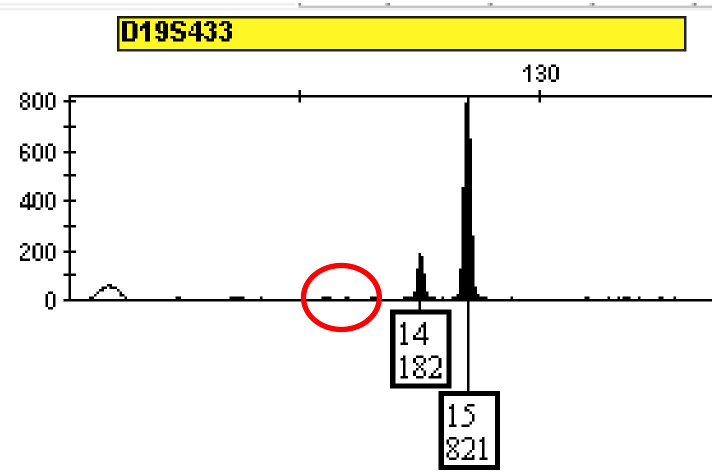
- Another example – Case 02, D19 locus

Contributors

A = 15, 15

B = 14, 15

C = **12**, 14



15 of 108 labs used CPI to include Suspect C (13.8%)

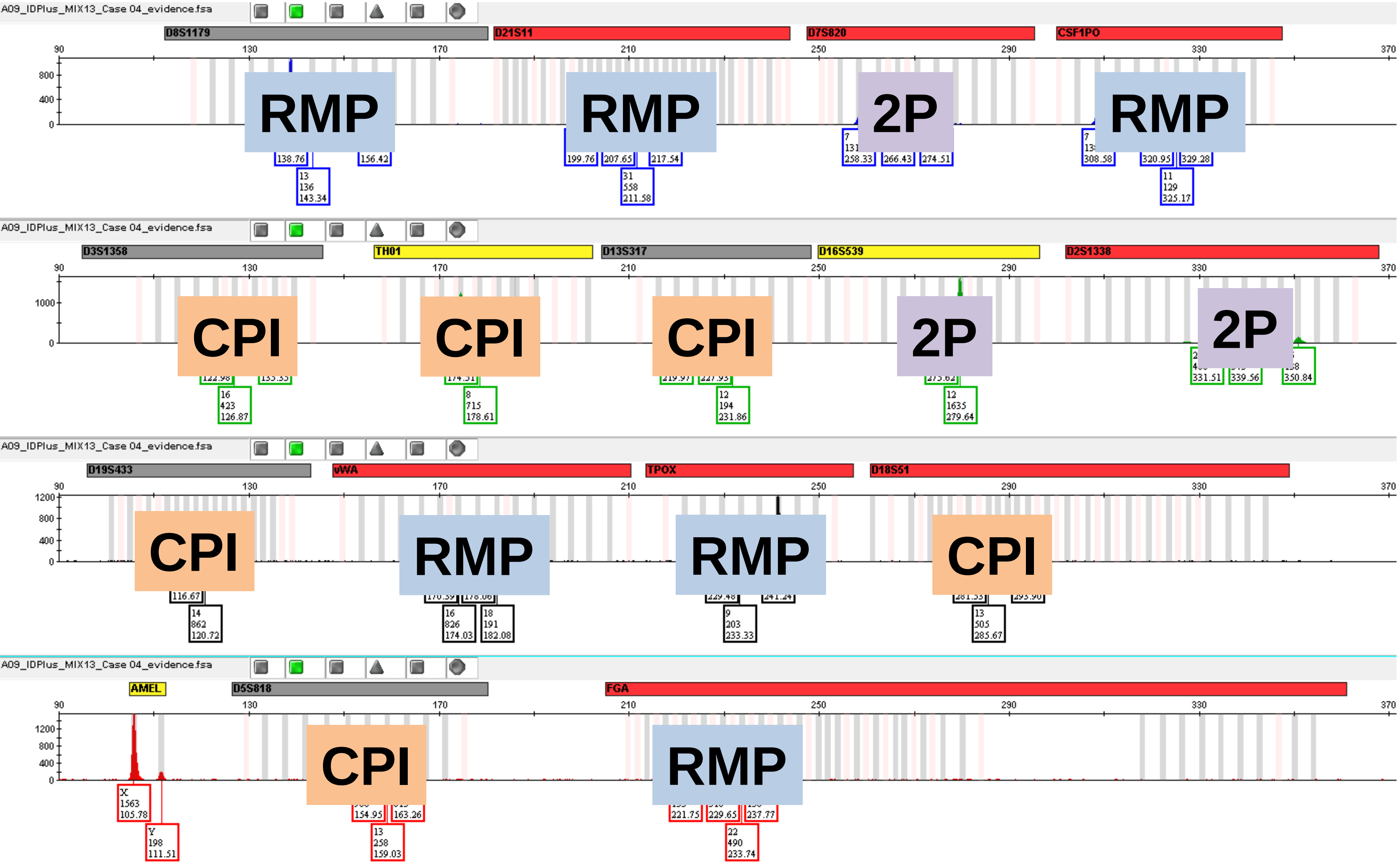
4 of these 15 (26.6%) used D19 as a locus for CPI

Concerns raised with MIX13 (3)

- Labs using RMP, LR – all over the place

- 4.6.2. It is not appropriate to calculate a composite statistic using multiple formulae for a multi-locus profile. For example, the CPI and RMP cannot be multiplied across loci in the statistical analysis of an individual DNA profile because they rely upon different fundamental assumptions about the number of contributors to the mixture.

One Lab's Interpretation



Switch to JOHN

Confronting Irreproducibility

Cell line authentication example:

- This field now has a **standard** (ANSI/ATCC ASN-0002-2011, Authentication of Human Cell Lines: Standardization of STR Profiling), which NIST scientists Margaret Kline and John Butler helped write
  - Yet many laboratories are not following it and some labs even deny there is a problem!
  - **Compliance is estimated to be only 15-30%**

See C&E News article (Dec 15, 2014) Vol. 92 (50) pp. 28-30

Arnaud, C.H. (2014) Confronting irreproducibility in life sciences research. *Chemical & Engineering News* 92(50): 28-30

Some Differences in Reporting Statistics

		Case1		
LabID	Kits Used	Caucasians	African Americans	Hispanics
90	ProPlus/Cofiler	1.18E+15	2.13E+14	3.09E+15
34	ProPlus/Cofiler	2.40E+11	7.00E+09	9.80E+10
33	ProPlus/Cofiler	2.94E+08	1.12E+08	1.74E+09
6	ProPlus/Cofiler	40,000,000	3,500,000	280,000,000
9	ProPlus/Cofiler	4.14E+07	1.97E+07	1.54E+08
79	ProPlus/Cofiler	930,000	47,900	1,350,000
16	ProPlus/Cofiler	434,600	31,710	399,100

~10 orders of magnitude difference (10^5 to 10^{15}) based on which alleles were deduced and reported

Remember that these labs are interpreting the same MIX05 electropherograms

Further Examination of These 7 Labs

		Case 1		
LabID	Kits Used	Caucasians	ASCLD-LAB accredited?	Solved loci listed?
90	ProPlus/Cofiler	1.18E+15	Yes	Yes
34	ProPlus/Cofiler	2.40E+11	Yes	Yes
33	ProPlus/Cofiler	2.94E+08	Yes	No
6	ProPlus/Cofiler	40,000,000	Yes	Yes
9	ProPlus/Cofiler	4.14E+07	No	No (CPE)
79	ProPlus/Cofiler	930,000	Yes	Yes
16	ProPlus/Cofiler	434,600	Yes	No

Possible Reasons for Variability in Reported Statistics:

- Different types of calculations (CPE vs RMP)
- Different loci included in calculations (due to different thresholds used)
- Different allele frequency population databases (most use PopStats)
- Use of victim (e.g., major component in Case 1) profile stats

DNA Mixture Workshop Attendees

50 states and >25 other countries

Federal Labs

FBI

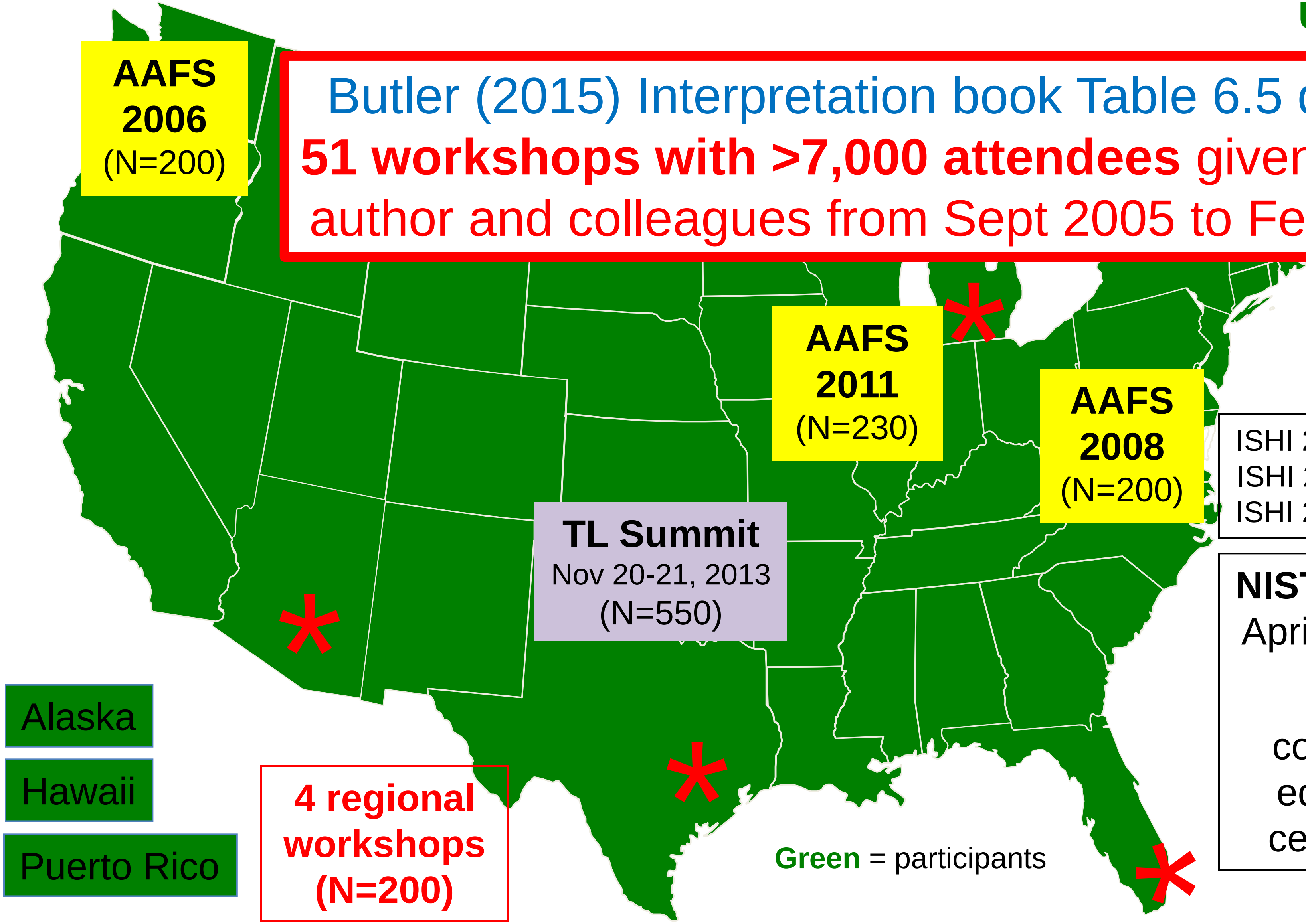
ATF

AFDIL

USACIL

Butler (2015) Interpretation book Table 6.5 details

51 workshops with >7,000 attendees given by the author and colleagues from Sept 2005 to Feb 2014



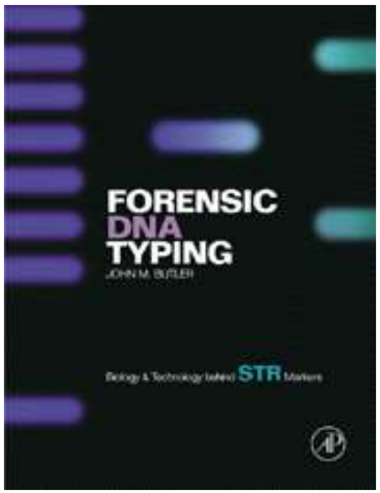
ISHI 2010 (N=200)
ISHI 2011 (N=160)
ISHI 2012 (N=145)

NIST Webinar
April 12, 2013

>1000 continuing education certificates

DNA Mixture Information Coverage
in *Forensic DNA Typing* Textbooks

1st Edition

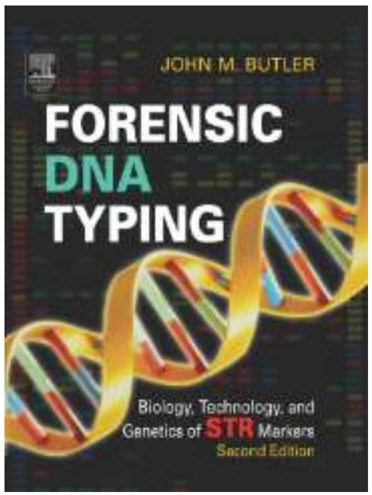


Jan 2001

335 pages

13 pages

2nd Edition

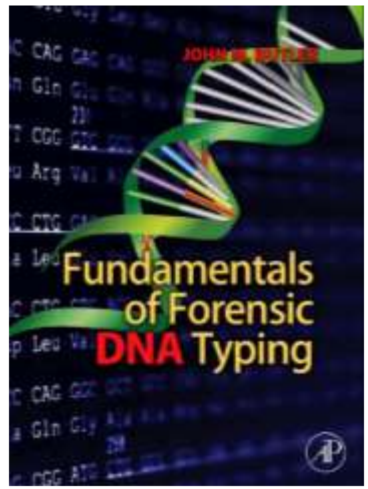


Feb 2005

688 pages

25 pages

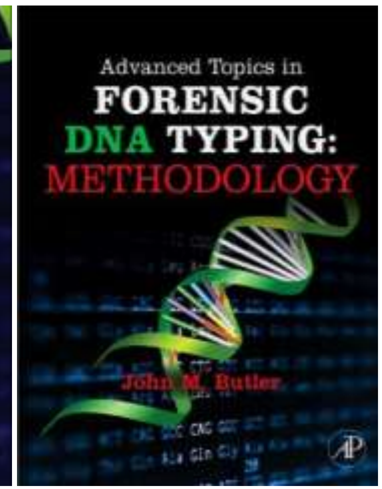
3rd Edition (3 volumes)



Sept 2009

520 pages

10 pages

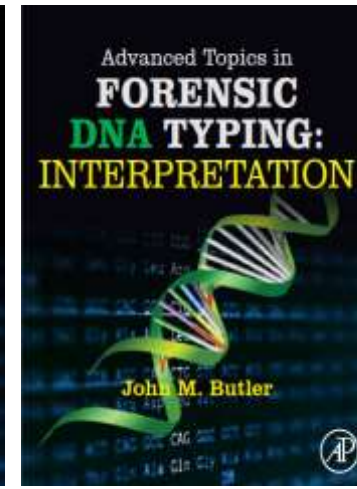


Aug 2011

704 pages

1 page

p. 235



Oct 2014

608 pages

126 pages

Chapters 6, 7, 12, 13
Appendix 4 (low-level, 2-person example)

People Know They Can Pass Audits and Still Be Doing Something Scientifically Invalid

- Within the past two years, I received a phone call ***during an ASCLD/LAB audit*** from a laboratory director (and the next day from DNA TL) sharing a concern raised by an auditor about how their laboratory was performing DNA mixture interpretation (they were doing what is known as "suspect-driven CPI")
- When I explained that what their laboratory was doing was not scientifically sound, **the response was "I don't care because I can still pass my audit! We have a protocol and we are following it."**

Audit Document
with Current FBI Quality Assurance Standards (2011)

						Yes	No	N/A
9.6	Does the laboratory **have and follow written guidelines for the interpretation of data**?							
9.6.1	Does the laboratory **verify that all control results meet the laboratory's interpretation guidelines for all reported results**?							
9.6.2	Has the 1996 National Research Council report and/or a court-directed method been used for the statistical interpretation of a DNA profile for a given population and/or hypothesis or relatedness, and are these calculations derived from an established population database(s) appropriate for the calculation?							
9.6.3	Does the laboratory **have and follow specific documented statistical interpretation guidelines** if genetic analyses that are not addressed by Standard 9.6.2 are being performed?							
9.6.4	Does the laboratory **have and follow documented procedures for mixture interpretation** to include the following:							
	a. Major and minor contributors?	Yes		No				
	b. Inclusions and exclusions?	Yes		No				
	c. Policies for reporting results and statistics?	Yes		No				

NRC II (1996), p. 130 emphasizes LR over CPI as a better approach for DNA mixtures

http://www.fbi.gov/about-us/lab/biometric-analysis/codis/forensic-qas-audit-9-1-11

From FBI QAS Audit Document

Discussion (following Section 9.6)

A laboratory shall have and follow written guidelines for the interpretation of data that are supported through its validation. A laboratory shall verify that all control results meet the laboratory's interpretation guidelines for all reported results. A documented method must exist to demonstrate that control values are verified when used (e.g., check-off, technical review).

The statistical interpretation of autosomal loci shall be made following recommendations 4.1, 4.2, or 4.3, as deemed applicable, of the National Research Council report titled "The Evaluation of Forensic DNA Evidence" (1996) and/or a court-directed method. **The laboratory shall provide documentation for the interpretation method being used.** These calculations shall be derived from a documented population database(s) appropriate for the calculation.

If a laboratory is performing genetic analyses not addressed by Standard 9.6.2, (e.g., Y-chromosome, mtDNA), the laboratory shall have and follow documented statistical interpretation guidelines for that testing.

A laboratory shall have and follow a documented procedure for mixture interpretation supported by its validation. Based upon a laboratory's validation, it shall have and follow procedures to discern major and minor contributors, inclusions and exclusions, and policies for reporting results and applicable statistics.

http://www.fbi.gov/about-us/lab/biometric-analysis/codis/forensic-qas-audit-9-1-11

SWGDAM Guidelines are Always Not Followed

- How many labs ignore the SWGDAM **guidelines** because they **are not auditable standards**?
  - Example: does a laboratory have and use a stochastic threshold (ST) that is necessary when using a CPI approach to mixtures?
- In June 2013 I received a phone call from a Los Angeles prosecutor who expressed concern that the two labs she used were getting different results from DNA mixtures (LASD uses a ST while LAPD does not → both are ASCLD/LAB accredited for DNA testing)

Los Angeles County Sheriff's Department (**LASD**)

http://www.ascld-lab.org/cert/ALI-131-T.pdf

American Society of Crime Laboratory Directors / Laboratory Accreditation Board

ASCLD/LAB-*International* Program

SCOPE of ACCREDITATION

Corresponds to Certificate Number ALI-131-T

Name and Address of Accredited Laboratory

Los Angeles County Sheriff's Department
Scientific Services Bureau
1800 Paseo Rancho Castilla
Los Angeles, California 90032

Laboratory Contact Information

Captain Clay Porlier, Laboratory Director
Phone: 323-260-8501
Fax: 323-415-4544
E-Mail: csporlie@lasd.org

Los Angeles Police Department (**LAPD**)

http://www.ascld-lab.org/cert/ALI-346-T.pdf

American Society of Crime Laboratory Directors / Laboratory Accreditation Board

ASCLD/LAB-*International* Program

SCOPE of ACCREDITATION

Corresponds to Certificate Number ALI-346-T

Name and Address of Accredited Laboratory

Los Angeles Police Department
Scientific Investigation Division
Criminalistics Laboratory
1800 Paseo Rancho Castilla
Los Angeles, California 90032

Laboratory Contact Information

Doreen Hudson, Laboratory Director
Phone: 323-415-8112
Fax: 323-276-1942
E-Mail: a9990@lapd.lacity.org

ISO/IEC 17025 validation requirement

5.4.5 Validation of methods

- 5.4.5.2 **The laboratory shall validate** … **standard methods used outside their intended scope** [e.g., CPI stats on complex DNA mixtures] … **to confirm that the methods are fit for the intended use**. The validation shall be as extensive as is necessary to meet the needs of the given application…

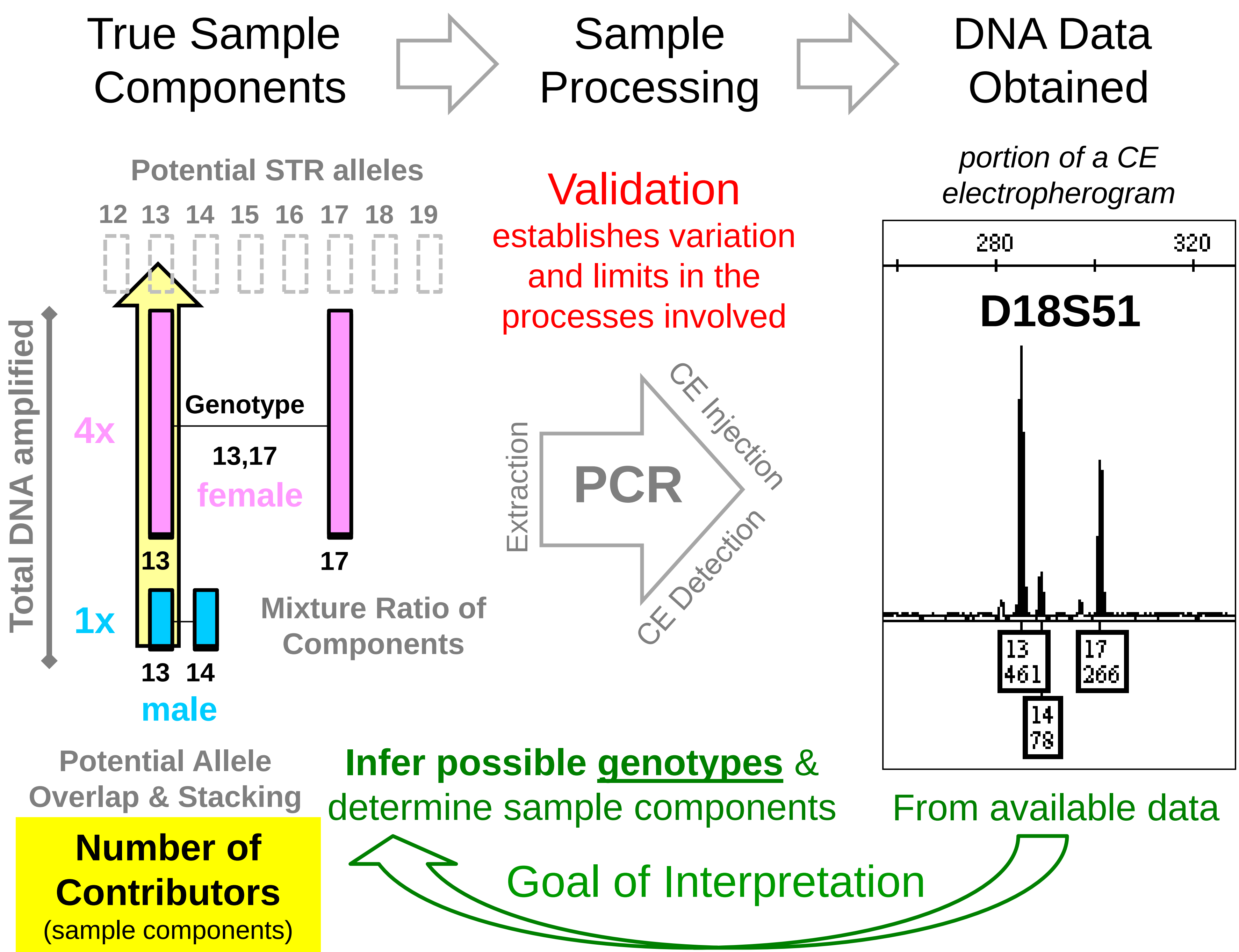
  Note 2 **The techniques used for the determination of the performance of a method should be one of, or a combination of, the following**:
  - (a) calibration using reference standards or materials;
  - (b) comparison of results achieved with other methods;
  - (c) **interlaboratory comparisons**;
  - (d) systematic assessment of the factors influencing the result;
  - (e) assessment of the uncertainty of the results based on **scientific understanding of the theoretical principles of the method** and **practical experience**.

The Perils of “Validation Extrapolation”

“Laboratories cannot adequately understand performance characteristics of low-template, complex DNA mixtures from having run a few high-template, simple DNA mixtures such as a few mixtures of 9947A and 9948. **Attempts at *validation extrapolation*, where a simple two-person mixture study is expected to provide guidance for proper interpretation of less optimal mixtures, <u>will not enable creation of robust protocols</u> that provide consistent, reliable results**. Every DNA interpretation protocol should be based on validation data, the scientific literature, and experience (SWGDAM 2010). **<u>Empirical data are always needed to establish limitations for a technique</u>**.”

SWGDAM 2010: SWGDAM interpretation guidelines for autosomal STR typing by forensic DNA testing laboratories. Available at http://www.swgdam.org.

Butler, J.M. (2015) *Advanced Topics in Forensic DNA Typing: Interpretation*. Elsevier Academic Press: San Diego, p. 166



True Sample Components
Sample Processing
DNA Data Obtained
Potential STR alleles
12 13 14 15 16 17 18 19
Total DNA amplified
4x
1x
Genotype
13,17
female
13
17
13 14
male
Mixture Ratio of Components
Potential Allele Overlap & Stacking
Validation
establishes variation and limits in the processes involved
Extraction
PCR
CE Injection
CE Detection
portion of a CE electropherogram
280
320
D18S51
13
461
14
78
17
266
Infer possible genotypes & determine sample components
From available data
Goal of Interpretation
Number of Contributors
(sample components)

Validation Needs to Match Sample Types

in a complex mixture. Of course, if peaks are observed below a stochastic threshold in a complex mixture, then allele drop-out of a sister allele is possible as with simple two-person mixtures and the detected allele could be a false homozygote.

Validation Needs to Match Sample Types

If a laboratory desires to develop appropriate protocols that will enable reliable interpretation of DNA from low-level DNA or mixtures involving three or more contributors, then validation studies need to be performed with known samples that mimic the amounts of DNA and complexity of profiles where stochastic effects and allele dropout are expected. In short, three- or four-person mixtures of known genotypes should be mixed at specific ratios and amplified multiple times. Then these complex mixture profiles should be subjected to interpretation approaches to see if a true contributor is appropriately associated with the mixture and if non-contributors are appropriately excluded.

In my opinion, a laboratory cannot run a single two-person mixture series (e.g. 9:1, 5:1, 3:1, 1:1, 1:3, 1:5, and 1:9) and feel confident that minimum requirements for "mixture validation" have been met. This type of a limited validation may simply be able to help determine that a minor contributor can be detected down to a certain level. Determining that a mixture exists is not the same as fully interpreting a mixture. Developing robust interpretation protocols will require considering more samples – especially ones that go beyond a cursory combination of control samples 9947A and 9948 (D.N.A. Box 7.1).

Butler, J.M. (2015) *Advanced Topics in Forensic DNA Typing: Interpretation*. Elsevier Academic Press: San Diego, p. 164

Validation Needs to Match Sample Types (1)

"**If a laboratory desires to develop appropriate protocols that will enable reliable interpretation of DNA from low-level DNA or mixtures involving three or more contributors, then validation studies need to be performed with known samples that mimic the amounts of DNA and complexity of profiles where stochastic effects and allele dropout are expected.** In short, three- or four-person mixtures of known genotypes should be mixed at specific ratios and amplified multiple times. Then these complex mixture profiles should be subjected to interpretation approaches to see if a true contributor is appropriately associated with the mixture and if non-contributors are appropriately excluded."

Butler, J.M. (2015) *Advanced Topics in Forensic DNA Typing: Interpretation*. Elsevier Academic Press: San Diego, p. 164

Validation Needs to Match Sample Types (2)

In my opinion, a laboratory cannot run a single two-person mixture series (e.g. 9:1, 5:1, 3:1, 1:1, 1:3, 1:5, and 1:9) and feel confident that minimum requirements for "mixture validation" have been met. This type of a limited validation may simply be able to help determine that a minor contributor can be detected down to a certain level. Determining that a mixture exists is not the same as fully interpreting a mixture. Developing robust interpretation protocols will require considering more samples – especially ones that go beyond a cursory combination of control samples 9947A and 9948 (D.N.A. Box 7.1).

Butler, J.M. (2015) *Advanced Topics in Forensic DNA Typing: Interpretation*. Elsevier Academic Press: San Diego, p. 164

Complexity Increases with More Contributors

D.N.A. Box 7.2 Number of Possible Genotype Combinations

	1 contributor *either hom or het*	# 2	2 contributors *7 classes*	# 14	3 contributors *23 classes*	# 150	4 contributors *41 classes*	# >600
1 allele (1 hom, 0 het)	homozygote (hom)	1	(a) 2 hom, 1 shared	1	(a) 3 hom, 1 shared	1	(a) 4 hom, 1 shared	1
2 alleles (2 hom, 1 het)	heterozygote (het)	1	(a) hom + hom, 0 shared	1	(a) 2 hom + hom, 0 shared	2	(a) 3 hom + hom, 1 shared	2
			(b) hom + het, 1 shared	2	(b) 2 hom + het, 1 shared	2	(b) 2 hom + 2 hom, 2 shared	1
			(c) hom + het, 2 shared	1	(c) hom + hom + het, 2 shared	1	(c) 3 hom + het, 1 shared	2
					(d) hom + 2 het, 2 shared	2	(d) 2 hom + hom + het, 2 shared	2
					(e) 3 het, 2 shared	1	(e) 2 hom + 2 het, 2 shared	1
							(f) hom + hom + 2 het, 2 shared	1
							(g) hom + 3 het, 2 shared	2
							(h) 4 het, 2 shared	1
3 alleles (3 hom, 3 het)	– *possible tri-allele?*		(a) hom + het, 1 shared	3	(a) hom + hom + hom, 0 shared	1	(a) 2 hom + hom + hom, 1 shared	3
			(b) het + het, 1 shared	3	(b) hom + hom + het, 0 shared	3	(b) 3 hom + het, 1 shared	3
					(c) hom + hom + het, 1 shared	6	(c) hom + hom + hom + het, 2 shared	3
					(d) hom + 2 het, 2 shared (2:2:2)	3	(d) hom + hom + het + het, 2 shared	5
					(e) hom + het + het, 2 shared (4:1:1)	3	(e) hom + 2 het + het, 2 shared	9
					(f) hom + het + het, 2 shared (3:2:1)	6	(f) hom + het + het + het, 3 shared	3
					(g) het + het + het, 2 shared (3:2:1)	6	(g) 3 het + het, 2 shared	6
					(h) het + het + het, 3 shared (2:2:2)	1	(h) 2 het + 2 het, 3 shared	2
							(i) 2 het + het + het, 3 shared	3
4 alleles (4 hom, 6 het)	–		(a) het + het, 0 shared	3	(a) hom + hom + het, 0 shared	6	(a) hom + hom + hom + hom, 0 shared	1
					(b) hom + het + het, 1 shared (3:1:1:1)	12	(b) 2 hom + hom + het, 1 shared (4:2:1:1)	12
					(c) hom + het + het, 1 shared (2:2:1:1)	12	(c) hom + hom + hom + het, 1 shared (3:2:2:1)	12
					(d) 2 het + het, 2 shared (2:2:1:1)	6	(d) hom + hom + 2 het, 2 shared (2:2:2:2)	6
					(e) het + het + het, 1 shared (3:1:1:1)	4	(e) hom + hom + het + het, 2 shared (3:2:2:1)	24
					(f) het + het + het, 2 shared (2:2:1:1)	12	(f) het + het + het + het, 4 shared (2:2:2:2)	2
							(g) het + het + het + het, 3 shared (3:2:2:1)	8
							(h) 2 hom + het + het, 2 shared (4:2:1:1)	12
							(i) hom + het + het + het, 3 shared (3:2:2:1)	24
							(j) 2 het + het + het, 3 shared (3:2:2:1)	24
							(k) 2 het + het + het, 3 shared (4:2:1:1)	12
							(l) 2 het + het + het, 2 shared (3:3:1:1)	24
5 alleles (5 hom, 10 het)	–		–		(a) hom + het + het, 0 shared	15	(a) hom + hom + hom + het, 0 shared	10
					(b) het + het + het, 1 shared	30	(b) hom + hom + het + het, 1 shared	30
							(c) hom + het + het + het, 2 shared	30
							(d) het + het + het + het, 3 shared (2:2:2:1:1)	>4
							(e) 2 het + het + het, 2 shared (3:2:1:1:1)	>11
6 alleles (6 hom, 15 het)	–		–		(a) het + het + het, 0 shared	15	(a) hom + hom + het + het, 0 shared	30
							(b) hom + het + het + het, 1 shared	48
							(c) het + het + het + het, 2 shared	>13
7 alleles (7 hom, 21 het)	–		–		–		(a) hom + het + het + het, 0 shared	70
							(b) het + het + het + het, 1 shared	>35
8 alleles (8 hom, 28 het)	–		–		–		(a) het + het + het + het, 0 shared	105

Butler, J.M. (2015) *Advanced Topics in Forensic DNA Typing: Interpretation*. Elsevier Academic Press: San Diego, p. 167

Some Key Principles

- Everything in **science involves** mapping observed data to **models** (“hypotheses”)
  - Hardy Weinberg Equilibrium (HWE) models expected genotype frequencies (p^2 or 2pq) assuming unrelated individuals
  - Theta corrections (θ=0.01 or θ=0.03) model potential variation from assumptions of unrelated individuals
- All **models require assumptions,** some of which are more reliable than others depending on data obtained
- **Validation studies generate data that inform the model being used** or enable a model to be constructed
  - For example, a test for HWE is comparing population (validation) data to a model to see goodness-of-fit
- **Genotypes—not alleles—matter** in deciphering mixtures
- **Probabilistic genotyping involves modeling observed data** against potential genotype combinations

Why are we where we are today?

- The incredible success of DNA has lead to **more sensitive methods and more "touch-evidence" samples being provided** which has led to more complex mixtures (we are pushing the envelope)
  - Lower template DNA profiles have more uncertainty associated with them in terms of allele peak height variation

- **Statistical interpretation techniques have not kept pace with the methodology improvements**
  - Much of the U.S. forensic DNA community is effectively using a 1992 statistical tool on 21st century data

What does the scientific literature say?

- **ISFG DNA Commission (2006) states that likelihood ratios (LRs) are preferred over use of the combined probability of inclusion (CPI)** [termed the random man not excluded, RMNE] because LRs address the question about whether a specific suspect's profile may be included in a mixture; NRC II (1996) p. 130 also supports ISFG

- Yet **most U.S. labs still use CPI** in large measure because of the DAB 2000 statement that use of either CPI or LR is okay (I would argue though that this was made in the context of simple two-person mixtures, primarily sexual assault evidence being done at the time)

- SWGDAM 2010 Autosomal STR Guidelines provide guidance for use of CPI, LR, and deconvolution with RMP (again primarily with a focus on two-person mixtures that are not low-level touch evidence – see SWGDAM.org FAQ statement)

Historical Perspective on DNA Mixture Approaches

Probabilistic genotyping software in development…

LR commonly used in Europe and other labs around the world

Today

ISFG DNA Commission
LR with drop-out
2012

2013 DNA TL Summit

SWGDAM guidelines
(RMP, CPI, LR)
2010

2008 NIJ burglary report increases touch evidence

ISFG DNA Commission
LR over CPI
2006

complex mixtures increase

2-person mixtures predominate

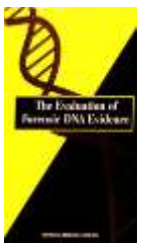
Few mixtures

2000

DAB Stats (Feb 2000)
CPI and LR okay

CPI becomes routine in U.S.

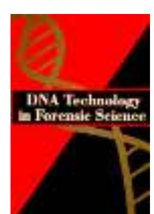
Weir et al. describe LRs for mixtures
1997



NRC II report (p.130) supports LR
1996

Evett et al. describe LRs for mixtures
1991

1992
NRC I report (p.59) supports CPI



1985
RMNE (CPI) used in paternity testing

LR = likelihood ratio
CPI = combined probability of inclusion
RMNE = random man not excluded

Forensic Laboratories Accredited for Biology/DNA

Accrediting Body	# Labs	% total
ASCLD/LAB	**240**	**88.6%**
FQS ANSI-ASQ National Accreditation Board (ANAB)	31	11%
A2LA	1	0.4%
Total	271	Bode Technology is both ASCLD/LAB and FQS accredited

ASCLD/LAB 240 labs: *26 legacy (11%) + 214 international (89%)*
FQS: *about half government labs and half private labs*

Based on searches of the following websites on 3 Jan 2015:
http://www.ascld-lab.org/accredited-laboratory-index/
http://search.anab.org/search-accredited-companies.aspx (and http://houstonforensicscience.org/accreditation.html)
http://www.a2la.org/dirsearchnew/searchbyspec.cfm?fieldpk=18&title=Forensic

Important Questions to Consider

- **Do we have laboratories getting incorrect results in spite of being accredited by ASCLD/LAB (or others) to ISO/IEC 17025?**
  - Most MIX05 and all MIX13 participants were from accredited forensic DNA laboratories…

- If being accredited does not lead to getting consistent and accurate DNA mixture results, will this observation erode confidence in the accreditation process?
  - **Will more specificity in the QAS and other audit documents help in the future?**

What does ASCLD/LAB want in terms of specificity in standards?

- SWGDAM and OSAC can work to produce documents with more specificity to aid interpretation but **these documents must be enforceable by accrediting bodies or they may be ignored by some**

- SWGDAM Autosomal STR Interpretation Committee is current revising the 2010 guidelines (see SWGDAM.org FAQ)
  - Current version including a detailed examples document is almost 100 pages long

Thank you for your attention

STRBase validation information available at:
http://www.cstl.nist.gov/strbase/validation.htm
STRBase mixture information available at:
http://www.cstl.nist.gov/strbase/mixture.htm

Contact Information

Michael D. Coble, Ph.D.
Research Biologist, NIST Applied Genetics Group
michael.coble@nist.gov
301-975-4330

John M. Butler, Ph.D.
NIST Fellow & Special Assistant to the Director for Forensic Science
Vice-Chair, National Commission on Forensic Science
john.butler@nist.gov
301-975-4049

Postscript: This slide was added after the meeting upon further thought about effective preparation to handle complex DNA mixtures

- **Forensic DNA testing** can be equated to different levels of **mathematics**:
  1. **Single-source samples** (e.g., reference samples) are like **basic arithmetic**
  2. **Two-person mixtures** (e.g., sexual assault evidence) are like **algebra**
  3. **Complex mixtures** (e.g., touch DNA) are like **calculus**

- Validation studies can be considered classroom instruction to help understand the topic and prepare for the casework “exams”

- Proficiency tests are like homework – a graded experience where feedback is received to prepare students for the casework exams when the true answers are not known to the test takers

- If homework is not challenging enough or if your classroom instruction is not to the level needed to be prepared, how can you hope to pass the test? Algebraic principles are necessary for calculus (just like two-person mixture principles apply to complex mixtures), but to truly solve calculus problems and complex mixtures a different level of thinking is required and more study is necessary. ***Would you want to go into a calculus final with only instruction in algebra and experience doing only basic math homework problems?***