DNA Mixture Interpretation:
Where did we come from? What are we doing?
Where are we going?

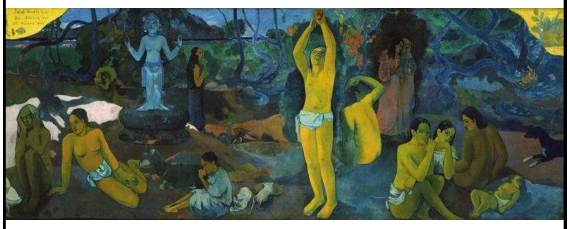
Michael Coble, PhD
NIST

Official Disclaimer

The opinions and assertions contained herein are solely those of the author and are not to be construed as official or as views of the U.S. Department of Commerce, U.S. Department of Justice, or the U.S. Department of Defense.

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 U.S. Department of Commerce, U.S. Department of Justice, or the U.S. Department of Defense nor does it imply that any of the materials, instruments or equipment identified are necessarily the best available for the purpose.

Where Do We Come From? What Are We? Where Are We Going?



Paul Gauguin, 1897

http://en.wikipedia.org/wiki/File:Woher_kommen_wir_Wer_sind_wir_Wohin_gehen_wir.jpg

What are we doing?
(2013 - 2015)

Interlaboratory Studies

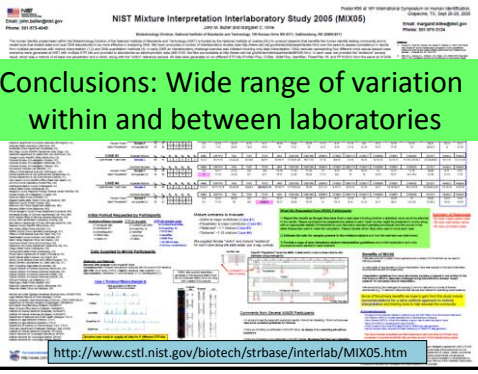
- The method by which multiple laboratories compare results and demonstrate that the methods used in one's own laboratory are reproducible in another laboratory.
- These tests are essential to demonstrate consistency in results from multiple laboratories.

(J.M. Butler, *Forensic DNA Typing*, p. 216)

Previous Interlaboratory Studies

- MSS 1 (1997) – 22 labs participated
- MSS 2 (1999) – 45 labs participated
- MSS 3 (2000-2001) – 74 labs participated
- MIX05 (2005) – 69 labs participated

MIX05 Poster Presentation at ISHI



How MIX13 differs from MIX05 study

	MIX13 (2013)	MIX05 (2005)
Response	108 labs	69 labs
Number of cases provided	5 cases	4 cases
Case types being mimicked	Sexual assault & touch evidence	Sexual assault evidence
Mixture complexity	2, 3, >3-person (potentially related, low-template, inclusion/exclusion)	all 2-person (all unrelated, male/female; various major/minor ratios)
Scenarios provided	Yes	No

MIX 13 – NIST Interlaboratory Study on Mixture Interpretation - **Purpose**

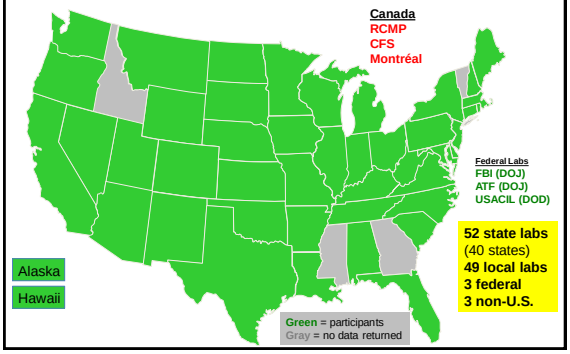
- MIX05 – conducted in 2005. Since then a great deal of effort has been focused on improvements in DNA mixture interpretation.
- 2010 SWGDAM Guidelines approved in January 2010 – many labs have changed their protocols recently.
- MIX13 – Interpretation challenge – no samples to run.

MIX 13 – NIST Interlaboratory Study on Mixture Interpretation - **Goals**

- (1) To evaluate the current "lay of the land" regarding STR mixture interpretation across the community.
- (2) To measure consistency in mixture interpretation across the U.S. after the publication of the 2010 SWGDAM guidelines.
- (3) To learn where future training and research could help improve mixture interpretation and reporting.

MIX13 Participants from **108 Laboratories**

46 states had at least one lab participate



Due to the number of laboratories responding and the federal, state, and local coverage obtained, this MIX13 interlaboratory study can be assumed to provide **a reasonable representation of current U.S. forensic DNA lab procedures across the community**

MIX13 was also used an intra-lab study

Comments from TL of a MIX13 Lab

- Thank you for the opportunity to participate in this exercise! Some of these were very challenging and provoked a lot of conversation.
- I had a majority of the analysts in our Forensic Biology Unit interpret these profiles ***independently in an effort to identify inconsistencies and areas where we need to improve.*** **It was very interesting how much the results varied!** I've included two spreadsheets that demonstrate this – "MIX13 summary of allele calls" and "MIX13 summary of stats and conclusions."

16 different analysts examined the data in this particular lab

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 = ?

Case 01 – Sexual Assault Evidence

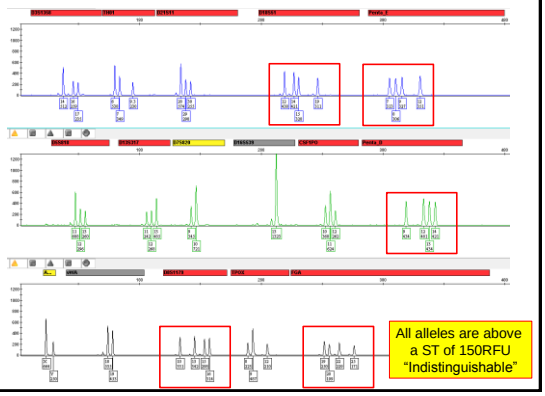
~1:1 mixture (2-person)

"German Type A"

Scenario

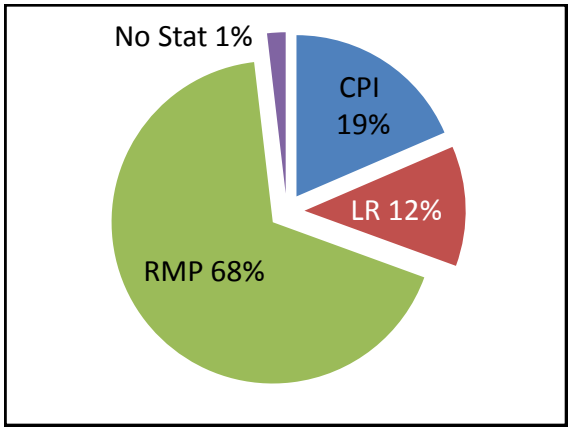
- Evidence: sperm fraction from a vaginal swab.
- A female meets a male acquaintance at a bar after work and they return to her apartment for a nightcap. She recalls the drink tasting funny and then wakes up 14 hours later after a co-worker has her landlord to open her apartment. She is confident that she did not have consensual sex and was probably drugged. She reports the incident to the police and goes to the hospital for an examination.
- The accused male gives a buccal swab for comparison.

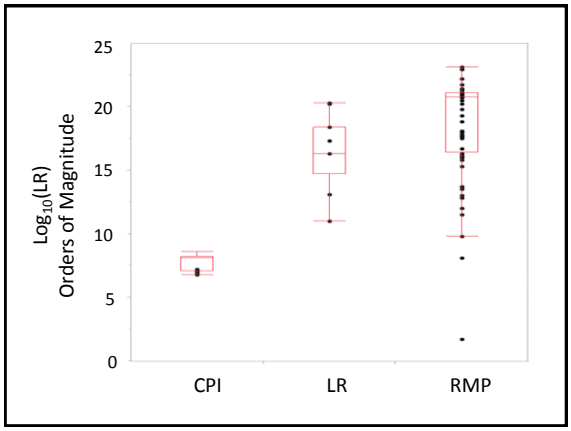
Case 01 – PP16HS

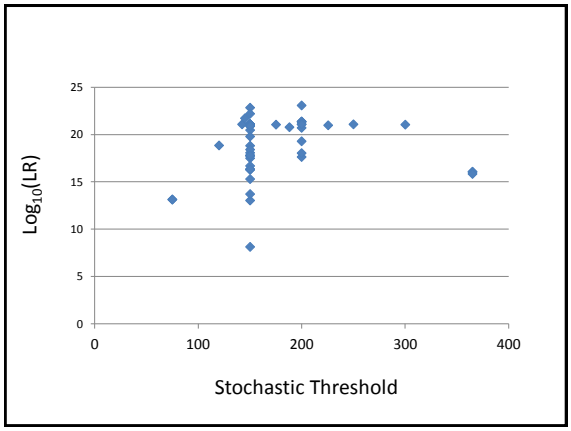


Primary Goals

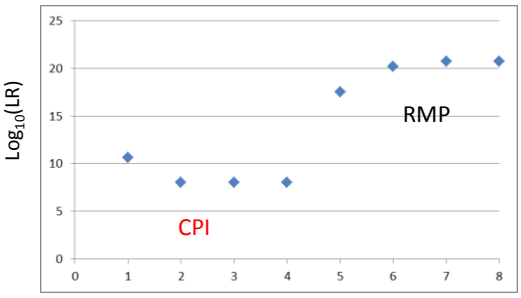
- Primary purpose – will labs attempt to subtract the victim from the evidence and generate a mRMP/LR stat **or** simply use CPI.


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


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


25
20
15
10
5
0
Log10(LR)
0
100
200
300
400
Stochastic Threshold

Intra-Laboratory Results (n = 8)



Case 02 – Handgun (Touch) Evidence

Low template profile with potential dropout (3-person)

"German Type C"

Case 02 – IDFiler

Case 02 – Four Suspects

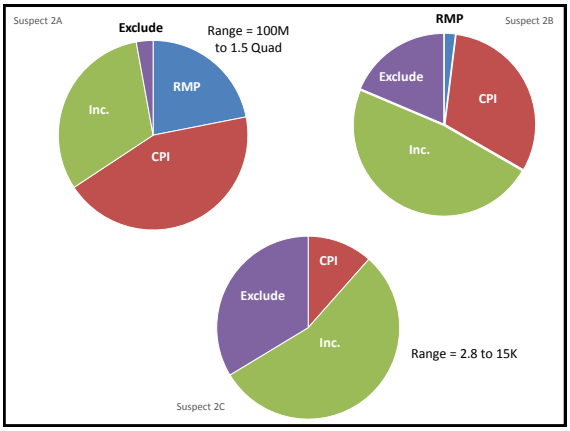
	Individual	Included?	Ratio
212 pg	Suspect A	Yes	6
53 pg	Suspect B	Yes	1.5
35 pg	Suspect C	Yes	1
	Suspect D	No	--

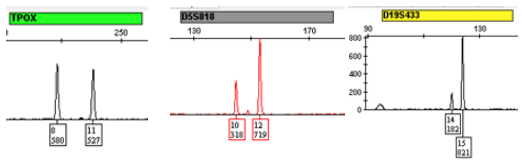
Drop Out Is Possible

Total Input DNA = 300 pg

Primary Goals

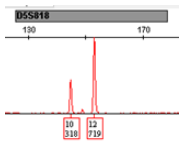
- Primary purpose – is this mixture too complex for interpretation due to the potential of drop-out?
- Several labs – CPI for Suspects A, B and C using a limited number of loci.
- One lab has included Suspect D (Not in the mixture).





~ 1 in 35

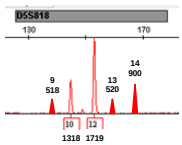
A cautionary note on using CPI when drop-out is possible



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

PI = $(0.051 + 0.361)^2$

PI = 0.169 or **1 in 5.92**



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**

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

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

Case 03 – Sexual Assault Evidence

Potential relative involved
(3-person)

Case 03 – Two Suspects

Individual	Inclusion?	Ratio
Victim	Included	7
Boyfriend (CP)	Included	2
Suspect 3A (Brother)	Included	1
Suspect 3B (Friend)	Excluded	--

Drop-out Possible

The Brothers

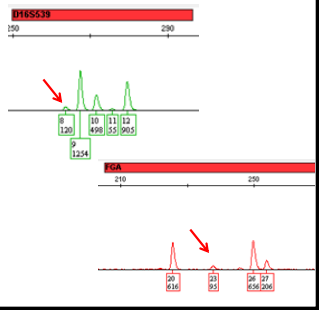
Markers	D8S1179	D21S11	D7S820	CSF1PO	D3S1358	TH01	D13S317	vWA	TPOX	D18S51	D5S818
Victim 03A	12,15	31.2,31.2	10,10	10,11	14,14	9.3,9.3	11,12	15,15	9,11	12,13	11,12
Cons Partner	14,14	28,35	10,11	10,12	14,18	7,8	12,13	17,21	6,8	13,16	10,12
Suspect 03A	14,15	28,35	10,11	12,12	14,18	7,8	12,13	17,21	8,9	13,16	10,12

For 11 of the 13 CODIS loci – all of suspect 03A's alleles are masked by either his brother or the victim

The Brothers

For D16 and FGA – two alleles of the suspected brother are present in the epg

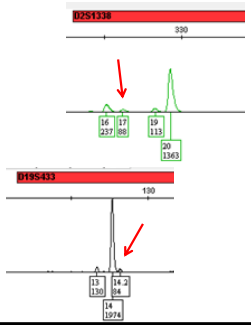
Markers	D16S539	FGA
Victim 03A	9,12	20,26
Cons Partner	10,10	26,27
Suspect 03A	8,9	23,27



The Brothers

The kit-specific markers give some additional information

Markers	D2S1338	D19S433
Victim 03A	20,20	14,14
Cons Partner	16,20	14,14
Suspect 03A	16,17	14,14.2



Primary Goals

- Primary purpose – is this mixture too complex for interpretation due to the potential of drop-out in the low level suspect?
- Only one lab included Suspect B (Not in the mixture)
- Most labs are using CPI stats for this case…

RMNE

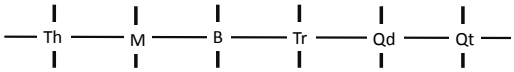
- Random Man Not Excluded (CPE/CPI) – The probability that a ***random person*** (unrelated individual) would be excluded as a contributor to the observed DNA mixture.
- Only a few labs have stated this – "Due to the relatedness of the exemplars submitted for comparison, a statistical analysis cannot be provided at this time."

Intra-Laboratory Results (n = 8)

CPI

13.1 Mill

1.4 Mill



Inconclusive
Inconclusive
Inconclusive
Inconclusive

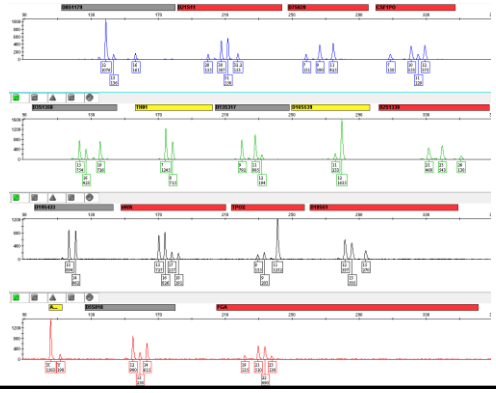
Excluded
Excluded

Case 04 – Bite Mark Evidence

Minor component (2-person)

"German Type B"

Case 04 – IDPlus



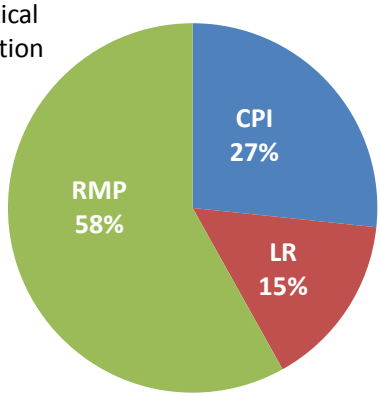
Case 04 – One Suspect

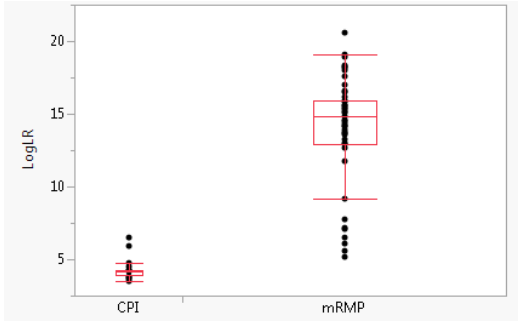
Individual	Inclusion?	Ratio
Victim	Included	~3.5
Suspect 4A	Included	1

Primary Goals

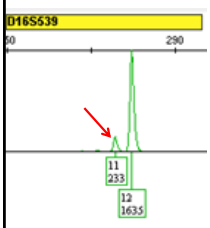
- Primary purpose – will labs choose to deconvolve this mixture since the mixture ratio is close to the limit of deconvolution for many labs?
- All labs have included the suspect in the mixture.

Statistical Evaluation





Focus on Uncertainty with D16 (stutter)

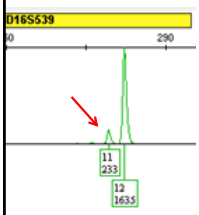


If 10% stutter from the 12 allele (163 RFU) is part of the 11 allele, then the remaining peak (70 RFU) is below the ST

No CPI labs excluded D16 from the stat

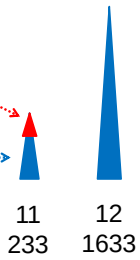
POI = 11, 12

Uncertainty with D16



70 RFU
(POI)

163 RFU
(Stutter)



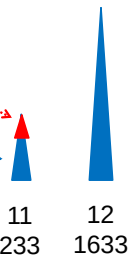
11 12
233 1633

POI = 11, 12

Uncertainty with D16
Observed Approaches

70 RFU
(POI)

163 RFU
(Stutter)



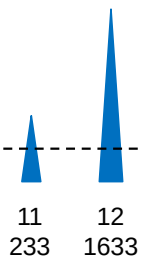
11 12
233 1633

Drop the locus
(Below ST)

POI = 11, 12

Uncertainty with D16
Observed Approaches

150 RFU
(ST)



11 12
233 1633

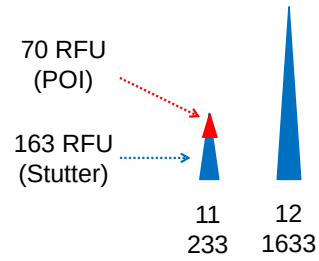
Use 2pq or p^2

(to infer genotypes)

11,12 or 11,11

POI = 11, 12

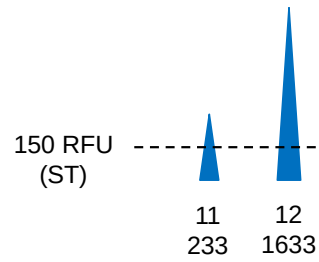
Uncertainty with D16
Observed Approaches



Use 2p
(Below ST)

POI = 11, 12

Uncertainty with D16
Observed Approaches

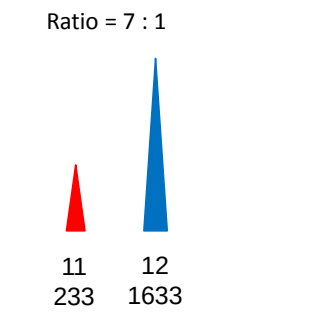


Use 2pq
(to infer genotype)

11,12

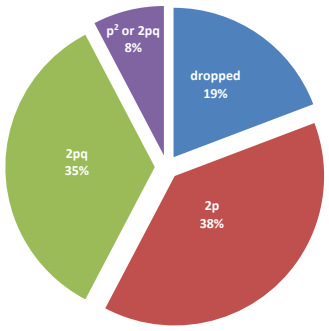
(matches POI)
"Certainty"

POI = 11, 12



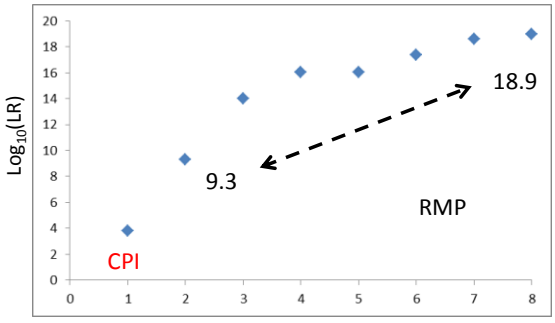
Assumes 0% Stutter

Statistics with D16 (mRMP/LR)



Probabilistic Software can help with consistency

Intra-Laboratory Results (n = 8)



Summary

- Most labs have validated and implemented AT and STs since MIX05. However, there is still a great deal of variation in interpretation across the U.S.
- An Idea – if everyone uses the same AT/ST, then one would expect to see similar results.
- Reality – the results were all over the place, Some of this was to be expected since each lab's protocol is different (e.g. dropping a locus vs, 2p).

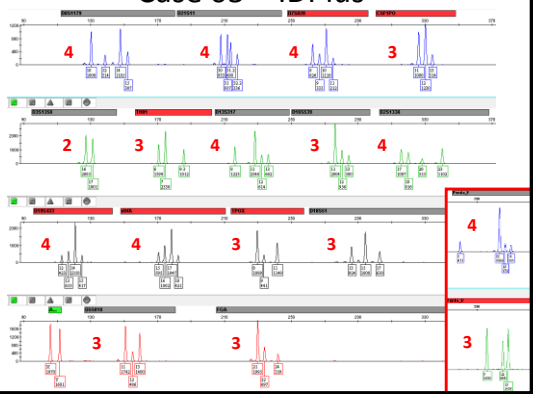
Case 05 – Ski Mask
(Robbery Evidence)

Complex mixture (>3-person) with # of contributors; inclusion/exclusion issues

Scenario

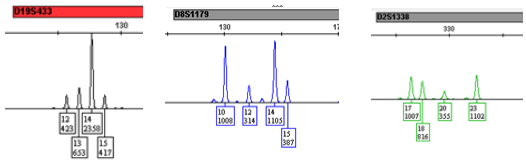
- Evidence: Ski mask recovered at a bank robbery.
- A number of gang-related robberies have targeted several banks in the city. The robberies have typically involved 2-3 perpetrators. A ski mask was recovered in a trash can one block away from the latest bank robbery and is submitted for DNA testing.
- A confidential informant has implicated two suspects in at least three of the armed robberies. Police have obtained buccal swab references from the two suspects identified from the CI, and another known accomplice of the suspects.

Case 05 – IDPlus



No more than 4 alleles at a locus

- Suggests a 2 person mixture

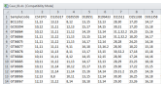


- Peak Height information does not agree

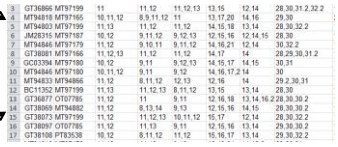
Case 05

- Is actually a 4 person mixture with no more than 4 alleles at any locus.
- Created with Virtual Mixture Maker (David Duewer, NIST) using 259 Caucasian samples from the NIST population data.

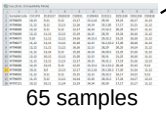
An overview of the mixture creation



65 samples
(ID and Pentas)

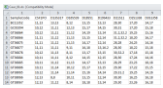


Groups 1 and 2

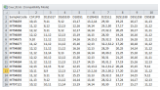


65 samples
(ID and Pentas)

An overview of the mixture creation



65 samples
(ID and Pentas)



64 samples
(ID and Pentas)

3	GT36866 MT97199	11	11,12	11,12,13	13,15	12,14	28,30,31.2,32.2
4	MT94818 MT97165	10,11,12	8,9,11,12	11	13,17,20	14,16	29,30
5	MT94803 MT97199	11,13	11,12	11,12	14,15,18	13,14	28,30,32.2
6	JM28315 MT97187	10,12	9,11,12	9,12,13	12,15,16	12,14,15	28,30
7	MT94846 MT97179	11,12	9,10,11	9,11,12	14,16,21	12,14	30,32.2
8	GT38081 MT97166	11,12,13	11,12	11,12	14,17	14	28,29,30,31.2
9	GC03394 MT97180	10,12	9,11	9,12,13	14,15,17	14,15	30,31
10	MT94846 MT97180	10,11,12	9,11	9,12	14,16,17	2,14	30
11	MT94833 MT94866	11,12	8,11,12	12,13	12,16	14	29.2,30,31
12	BC11352 MT97199	11,13	11,12,13	8,11,12	13,15	13,14	28,30
13	GT36877 OT07785	11,12	11	9,11	12,15,16	13,14,16	2,28,30,32.2
14	GT38069 MT94882	11,12	8,13,14	9,13	12,15,16	14,15	28,30,30.2
15	GT38073 MT97199	11,12	11,12,13	10,11,12	15,17	12,14	28,30,32.2
16	GT38097 OT07785	11,12	11,13	9,11	12,15,16	13,14	29,30,30.2
17	GT38108 PT83538	10,12	8,11,12	11,12	15,16,17	13,14	29,30,32.2

Groups 3 and 4

Groups 1 and 2 (only the top 25 "least # of shares")

Select combinations with no more than 4 alleles at any locus

Groups 3 and 4 (only the top 25 "least # of shares")

		#1	#2	#3	#4	#5	#6
GT36866_MT97199	UT58299_Y12	0	3	9	4	0	0
MT94803_MT97199	UT57301_TT50705	0	3	7	6	0	0
GT36866_OT07776	UT58299_TT50705	0	2	8	6	0	0
GT36885_MT97192	WT51386_TT50705	0	2	8	6	0	0
GT38069_GT38119	UT58299_MT94884	0	2	8	6	0	0
GT38098_MT97199	UT57301_TT50705	0	2	8	6	0	0
MT94803_MT97199	UT58299_Y12	0	3	6	7	0	0
MT97126_MT97173	UT58318_UT57299	0	3	6	7	0	0
MT97126_MT97173	UT58318_TT50705	0	3	6	7	0	0

Note: All samples are unrelated
(relative testing, mtDNA, Y-STRs, X-STRs, etc…)

Impact of New Loci on Mixture Assumptions

Probability of incorrectly assigning the specific number of contributors based on observed alleles
(not considering peak height imbalances)

True # of contributors	Using NIST Caucasians (Hill et al. 2013)	1	2	3	4	5
6	CODIS13	1.75E-40	6.34E-09	0.161242	0.945657	0.999873
	CODIS22	0 (< E-99)	9.59E-21	5.32E-05	0.188138	0.859901
5	CODIS13	9.78E-33	2.10E-06	0.41432	0.989651	
	CODIS22	6.36E-61	7.01E-15	0.004837	0.610149	
4	CODIS13	7.02E-25	0.000515	0.785495		
	CODIS22	3.50E-46	3.49E-09	0.16523		
3	CODIS13	8.42E-17	0.059486			
	CODIS22	5.77E-31	0.000433			
2	CODIS13	1.70E-08				
	CODIS22	2.05E-15				

0.05%

With 13 CODIS loci, **5.9% of 3-person contributors could falsely be considered a 2-person mixture** based on observed alleles (using NIST Caucasian allele frequencies)

With expanded CODIS loci, this drops to **0.04%**

Data from Coble, Bright, Buckleton and Curran (article in review)

Case 05 – 3 Suspects

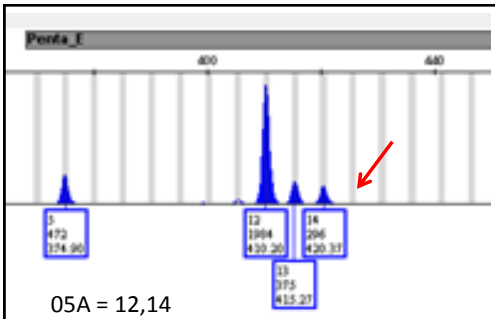
Individual	
Suspect 5A	Included
Suspect 5B	Included
Suspect 5C	Not in the mixture

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)



05A = 12,14
05B = 5,12
05C = 12,**15**

Summary of Issues

- **Use of CPI has significant limitations when it comes to complex mixtures** because this approach delivers information regarding the presence of alleles rather than specific suspect genotypes
- A CPI approach **has the potential to falsely include innocent suspects** as demonstrated in MIX13 Case 5
- The U.S. forensic DNA community adopted CPI for simplicity in 1990s and early 2000s when 2-person mixtures were common and have now inappropriately extrapolated the approach to more complex mixtures

What Needs to Be Done

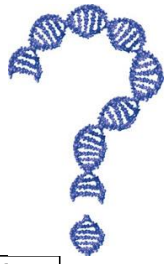
- **STOP! using CPI on complex mixtures**, such as MIX13 Case 5
  - **it is better to declare a result inconclusive** than to potentially falsely include an innocent person (you are more likely to have false inclusions with a low stat as the power to exclude has been reduced when loci are removed)
- **Set a complexity threshold** to aid in determining when to not interpret a mixture
- **Adopt a probabilistic genotyping approach** (will involve software) after validation studies with complex mixtures

Acknowledgments

National Institute of Justice and NIST Law Enforcement Standards Office

Dr. Charlotte Word
Dr. Robin Cotton
Dr. John Butler

Dan Katz



Contact info:
mcoble@nist.gov
+1-301-975-4330