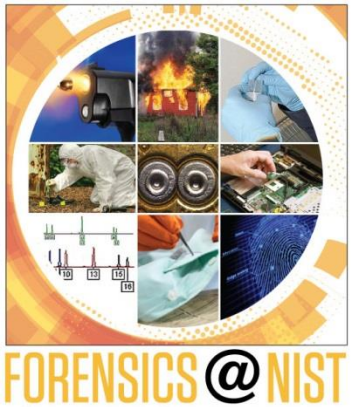


DNA Mixture Interpretation Webcast

April 12, 2013

http://www.nist.gov/oles/forensics/dna-analyst-training-on-mixture-interpretation.cfm

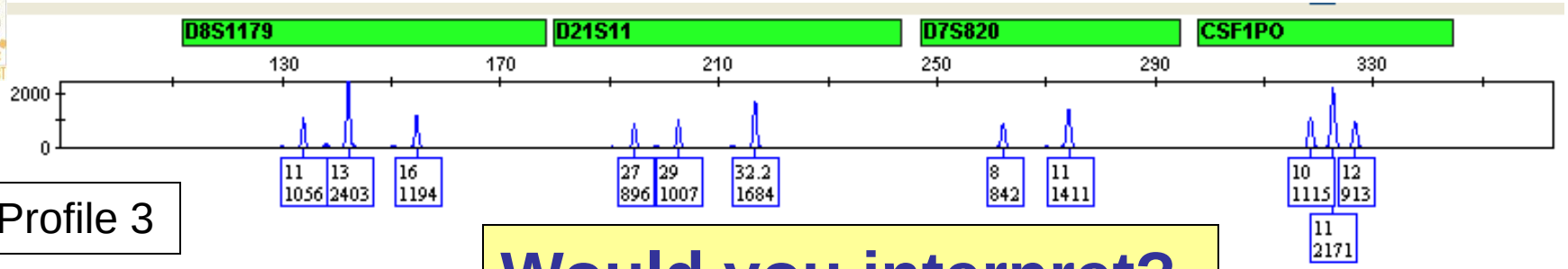
http://www.cstl.nist.gov/strbase/mixture.htm

Complex Mixtures

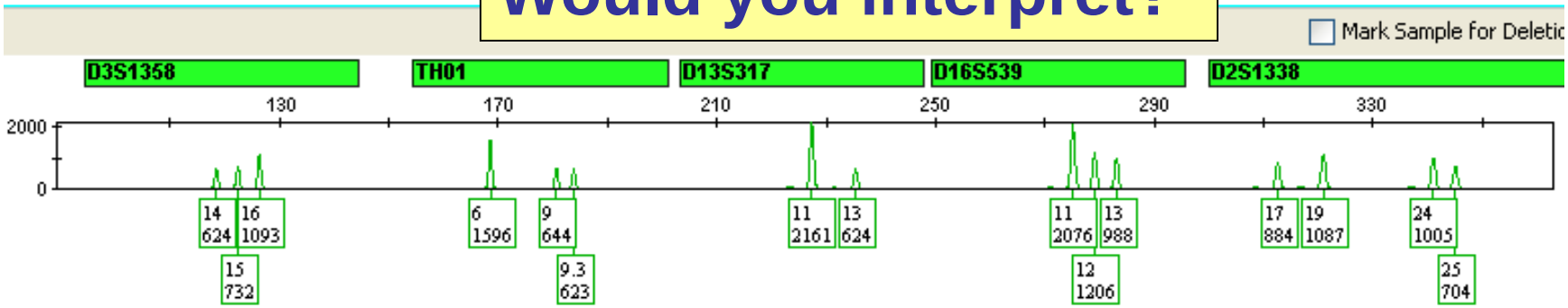
Charlotte J. Word

Consultant

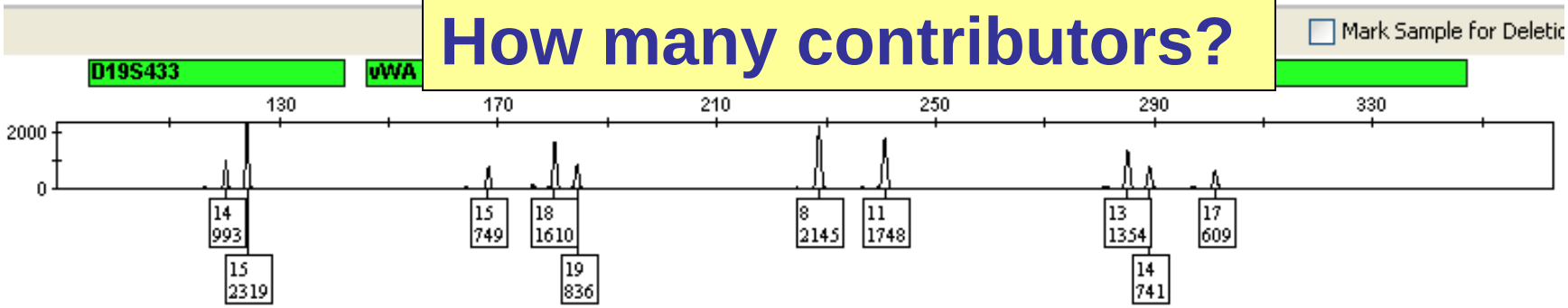
Profile 3

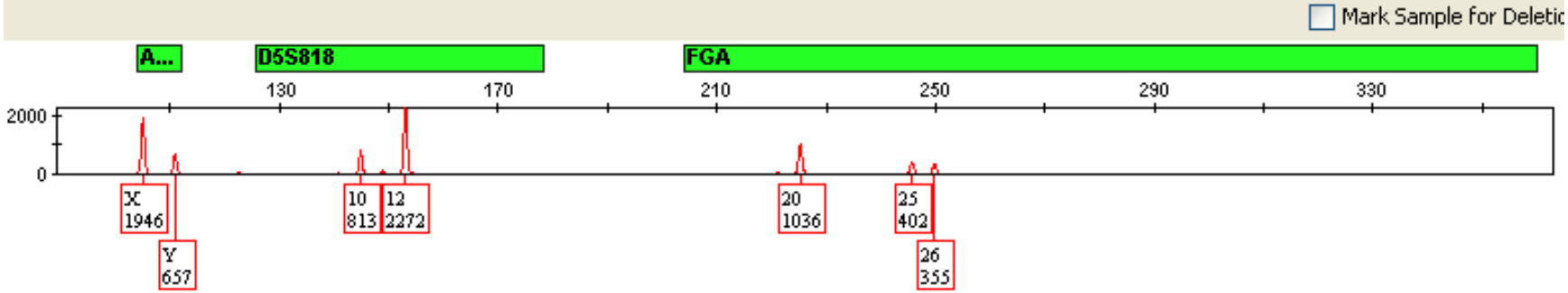


Would you interpret?



How many contributors?





Two-Person Mixtures

- **Lots of experience and familiarity with two-person mixtures, literature, validation studies, training samples**
- **Published guidelines for interpretation**
- **Well developed SOPs for interpretation**
- **Routine amount of input DNA in amplification generally leads to nice profiles**

Two-Person Mixtures

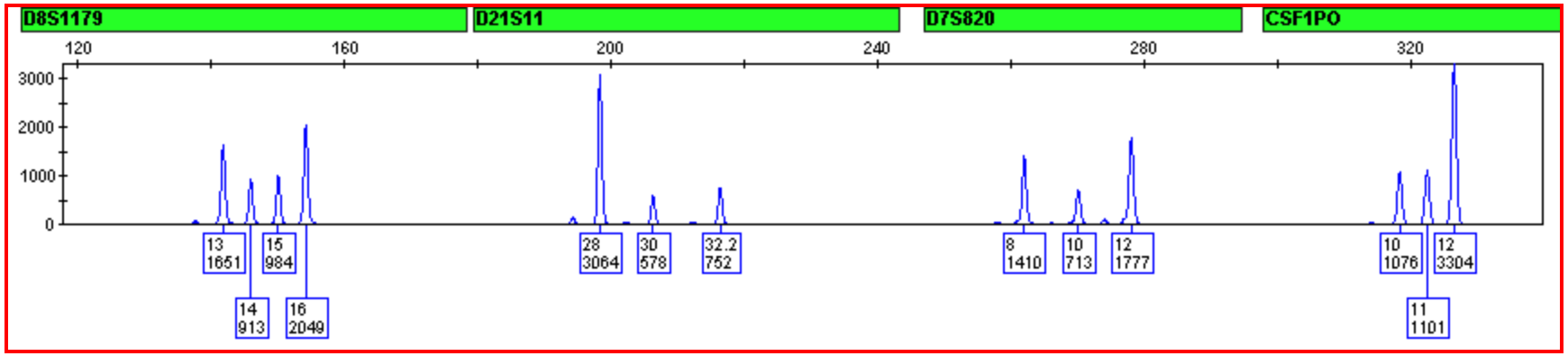
High Certainty Leads to High Confidence

- Only two contributors present
- Distinguishing stutter/artifacts from true alleles
- Use stochastic threshold to assess if all alleles are likely present vs. LT DNA with stochastic effects
- Assessing mixture ratio (distinguishable/ major:minor or indistinguishable mixture)
- Deducing second contributor if one contributor is known

Two-Person Mixtures

Assume number of contributors is two:

- **Aids in allele association at each locus based on peak height ratios**
- **May aid in genotype association for full profile based on mixture ratio**
- **Statistics calculations often straight forward**



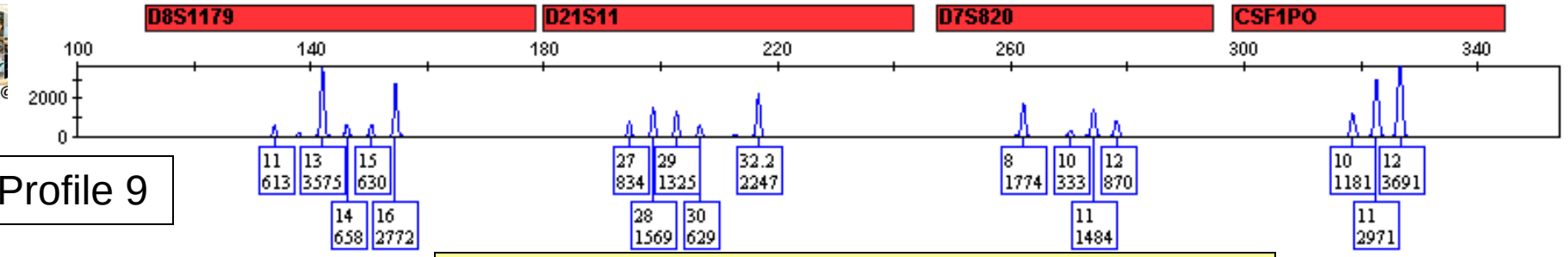
Complex Mixtures

- **Multiple contributors**
  - **3- & 4- person (or more!)**
- **Relatives in the Mixtures**

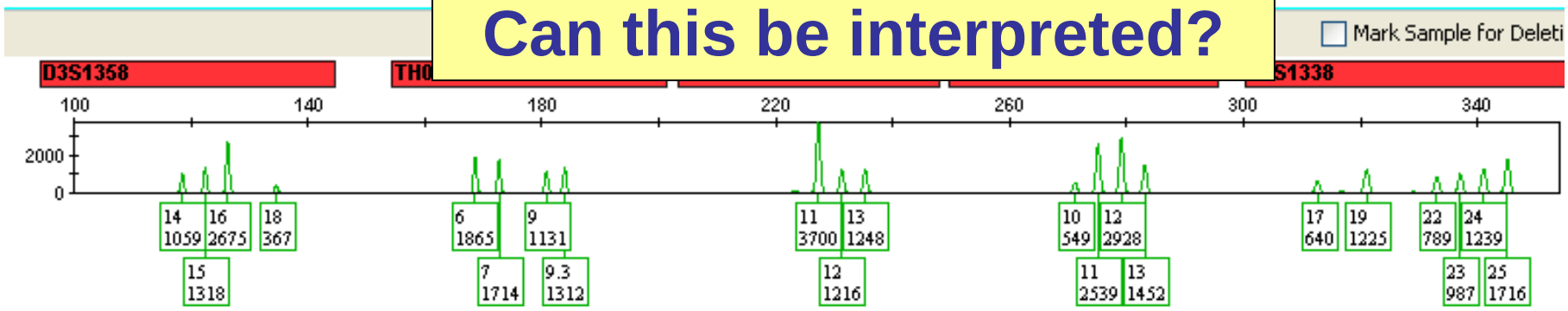
MYTH

It is easy to determine the number of contributors to a DNA profile.

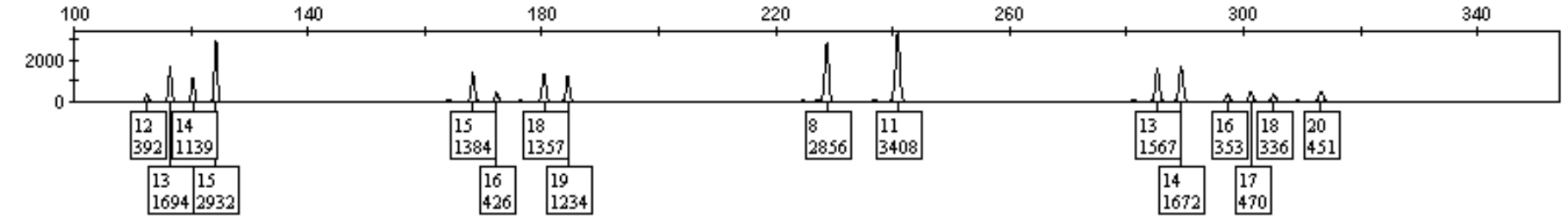
Profile 9



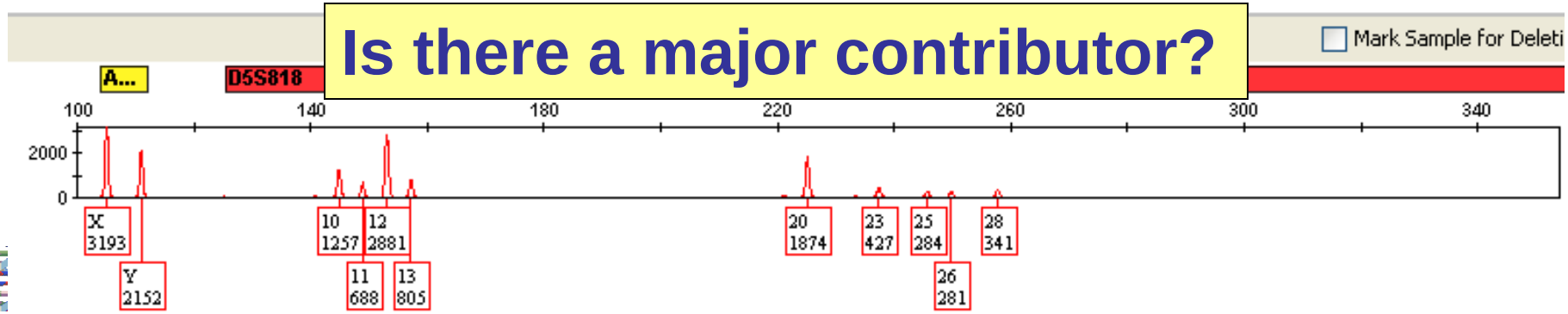
Can this be interpreted?



How many contributors assumed for interpretation?

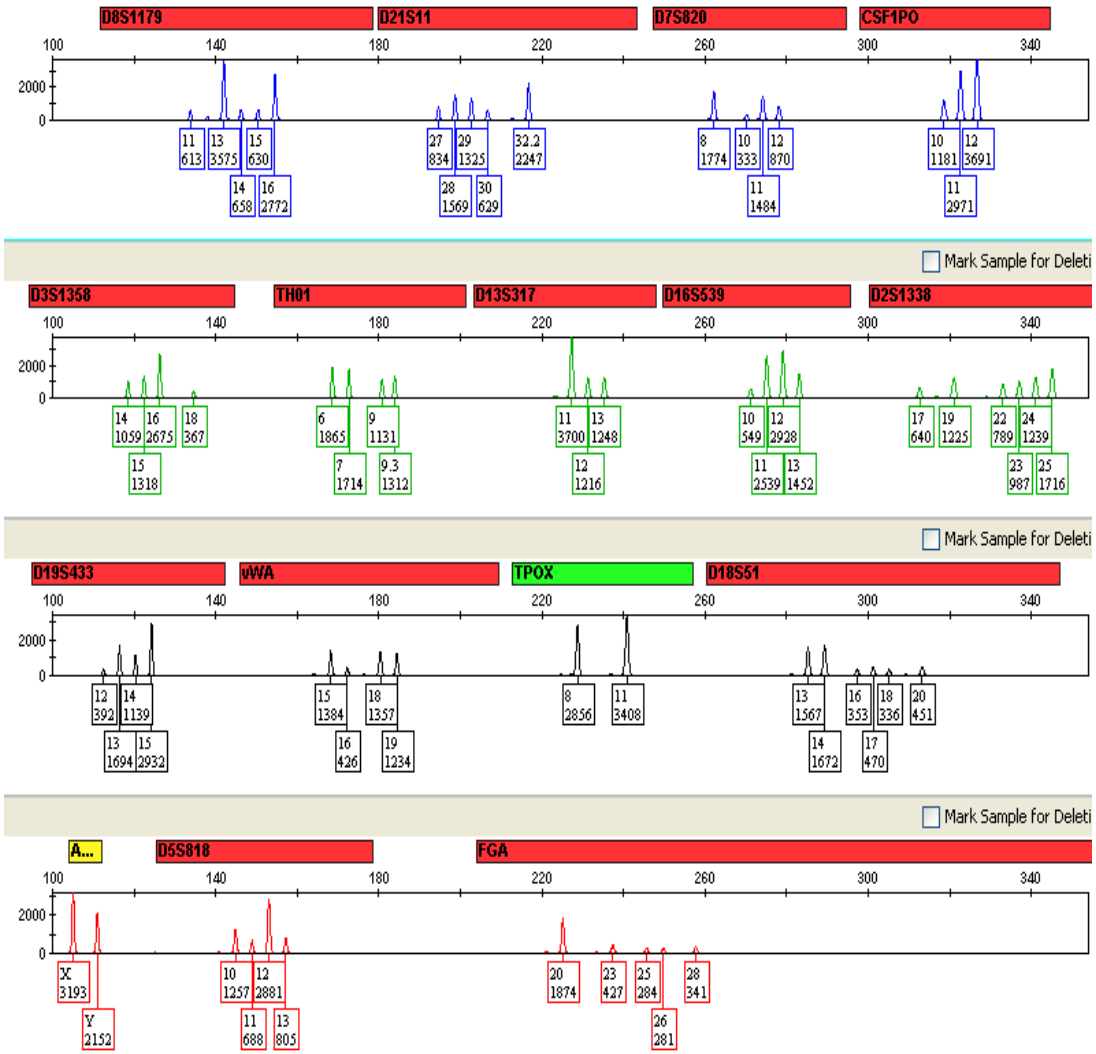


Is there a major contributor?



Complex Mixture – Allele Summary

- 6 alleles at 2 loci
- 5 alleles at 3 loci
- 4 alleles at 7 loci
- 3 alleles at 2 loci
- 2 alleles at 1 locus
- 1 allele at 0 loci
- 63 total alleles

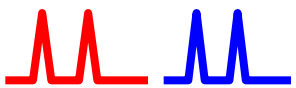


Two-Person Mixtures

14 total combinations

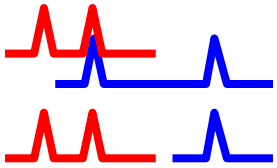
Observed profile

A B



4 alleles

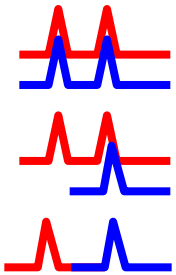
All heterozygotes and non-overlapping alleles



3 alleles

Heterozygote + heterozygote, one overlapping allele

Heterozygote + homozygote, no overlapping alleles

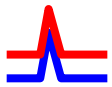


2 alleles

Heterozygote + heterozygote, two overlapping alleles

Heterozygote + homozygote, one overlapping allele

Homozygote + homozygote, no overlapping alleles



1 allele

Homozygote + homozygote, overlapping allele

Three-Person Mixtures

Observed profile

150 total combinations

6 alleles

All heterozygotes and non-overlapping alleles

5 alleles

Two heterozygotes and one homozygote
Three heterozygotes, one overlapping allele

4 alleles

Six combinations of heterozygotes, homozygotes and overlapping alleles

3 alleles

Eight combinations of heterozygotes, homozygotes, and overlapping alleles

2 alleles

Five combinations of heterozygotes, homozygotes, and overlapping alleles

1 allele

All homozygotes, overlapping allele

Four-Person Mixtures

Observed profile

MANY combinations

8 alleles

All heterozygotes and non-overlapping alleles

7 alleles

Several combinations of heterozygotes, homozygotes, and overlapping alleles

6 alleles

Many combinations

5 alleles

Many combinations

4 alleles

Many combinations

3 alleles

Many combinations

2 alleles

Many combinations

1 allele

All homozygotes, overlapping allele

Available online at www.sciencedirect.com
ScienceDirect
Forensic Science International: Genetics 1 (2007) 20–28
FSI GENETICS
www.elsevier.com/locate/fsig

Towards understanding the effect of uncertainty in the number of contributors to DNA stains

John S. Buckleton [a], James M. Curran [b,*], Peter Gill [c]

[a] The Institute of Environmental Science and Research Ltd., Private Bag 92021, Auckland, New Zealand
[b] Department of Statistics, University of Auckland, Private Bag 92019, Auckland, New Zealand
[c] The Forensic Science Service, Trident Court, Solihull Parkway, Birmingham Business Park, Solihull B37 7YN, UK

Received 31 May 2006; received in revised form 12 September 2006; accepted 13 September 2006

Abstract

DNA evidence recovered from a scene or collected in relation to a case is generally declared as a mixture when more than two alleles are observed at several loci. However, in principle, all DNA profiles may be considered to be potentially mixtures, even those that show not more than two alleles at any locus. When using a likelihood ratio approach to the interpretation of mixed DNA profiles it is necessary to postulate the number of potential contributors. However, this number is never known with certainty. The possibility of a, say three-person mixture, presenting four or fewer peaks at each locus of the CODIS set was explored by Paoletti et al. [D.R. Paoletti, T.E. Doom, C.M. Krane, M.L. Raymer, D.E. Krane, Empirical analysis of the STR profiles resulting from conceptual mixtures, J. Forensic Sci. 50 (2005) 1361–1366]. In this work we extend this analysis to consider the profiler plus and SGM plus multiplices. We begin the assessment of the risk associated with current practice in the calculation of LR's. We open the discussion of possible ways to surmount this ambiguity.

Forensic Science International: Genetics 1 (2007) 20–28

Two-Person Simulated Mixtures – SGM+ Number of Alleles at each Locus

Table 1

The probability of observing a given number of alleles in a two-person mixtures for simulated profiles at the SGM^{+TM} loci

Loci	No. of alleles			
	1	2	3	4
D3	0.011	0.240	0.559	0.190
vWA	0.008	0.194	0.548	0.250
D16	0.016	0.287	0.533	0.164
D2	0.003	0.094	0.462	0.441
D8	0.011	0.194	0.521	0.274
D21	0.007	0.147	0.505	0.341
D18	0.003	0.095	0.472	0.430
D19	0.020	0.261	0.516	0.203
THO	0.016	0.271	0.547	0.166
FGA	0.003	0.116	0.500	0.381

Buckleton et al. Forensic Science International: Genetics 1 (2007) 20–28

Three-Person Simulated Mixtures – SGM+ Number of Alleles at each Locus

Table 2

The probability of observing a given number of alleles in a three-person mixtures for simulated profiles at the SGM^{+TM} loci

Loci	No. of alleles showing					
	1	2	3	4	5	6
D3	0.000	0.053	0.366	0.463	0.115	0.002
vWA	0.000	0.037	0.285	0.468	0.194	0.016
D16	0.001	0.086	0.397	0.411	0.100	0.005
D2	0.000	0.008	0.104	0.385	0.393	0.110
D8	0.001	0.041	0.258	0.436	0.236	0.029
D21	0.000	0.023	0.192	0.428	0.302	0.055
D18	0.000	0.007	0.109	0.392	0.396	0.096
D19	0.003	0.078	0.352	0.401	0.152	0.014
THO	0.001	0.074	0.395	0.439	0.088	0.002
FGA	0.000	0.012	0.144	0.424	0.346	0.074

Buckleton et al. Forensic Science International: Genetics 1 (2007) 20–28

2, 3, 4-Person Simulated Mixtures – CODIS Loci Number of Alleles at each Locus

J Forensic Sci, Nov. 2005, Vol. 50, No. 6
Paper ID JFS2004475
Available online at: www.astm.org

David R. Paoletti,[1] M.S.; Travis E. Doom,[1,2] Ph.D.; Carissa M. Krane,[3] Ph.D.; Michael L. Raymer,[1,2] Ph.D.; and Dan E. Krane,[4] Ph.D.

Empirical Analysis of the STR Profiles Resulting from Conceptual Mixtures

ABSTRACT: Samples containing DNA from two or more individuals can be difficult to interpret. Even ascertaining the number of contributors can be challenging and associated uncertainties can have dramatic effects on the interpretation of testing results. Using an FBI genotypes dataset, containing complete genotype information from the 13 Combined DNA Index System (CODIS) loci for 959 individuals, all possible mixtures of three individuals were exhaustively and empirically computed. Allele sharing between pairs of individuals in the original dataset, a randomized dataset and datasets of generated cousins and siblings was evaluated as were the number of loci that were necessary to reliably deduce the number of contributors present in simulated mixtures of four or less contributors. The relatively small number of alleles detectable at most CODIS loci and the fact that some alleles are likely to be shared between individuals within a population can make the maximum number of different alleles observed at any tested loci an unreliable indicator of the maximum number of contributors to a mixed DNA sample. This analysis does not use other data available from the electropherograms (such as peak height or peak area) to estimate the number of contributors to each mixture. As a result, the study represents a worst case analysis of mixture characterization. Within this dataset, approximately 3% of three-person mixtures would be mischaracterized as two-person mixtures and more than 70% of four-person mixtures would be mischaracterized as two- or three-person mixtures using only the maximum number of alleles observed at any tested locus.

Paoletti et al. J Forensic Sci, Nov. 2005, Vol. 50, No. 6

2- to 5-Person Simulated Mixtures – Identifiler Number of Alleles vs. Likelihood Estimator

J Forensic Sci, January 2011, Vol. 56, No. 1
doi: 10.1111/j.1556-4029.2010.01550.x
Available online at: interscience.wiley.com

PAPER

CRIMINALISTICS

Hinda Haned,[1] M.S.; Laurent Pène,[2] M.S.; Jean R. Lobry,[1] Ph.D.; Anne B. Dufour,[1] Ph.D.; and Dominique Pontier,[1] Ph.D.

Estimating the Number of Contributors to Forensic DNA Mixtures: Does Maximum Likelihood Perform Better Than Maximum Allele Count?

Haned et al. J Forensic Sci, January 2011, Vol. 56, No. 1

Number of Contributors – Total Number of Alleles

314 FORENSIC SCIENCE

doi: 10.3325/cmj.2011.52.314

Estimating the number of contributors to two-, three-, and four-person mixtures containing DNA in high template and low template amounts

Jaheida Perez, Adele A. Mitchell, Nubia Ducasse, Jeannie Tamariz, Theresa Caragine

Office of Chief Medical Examiner of the City of New York, The Department of Forensic Biology, New York, NY, USA

Perez et al., Croat Med J. 2011; 52:314-26

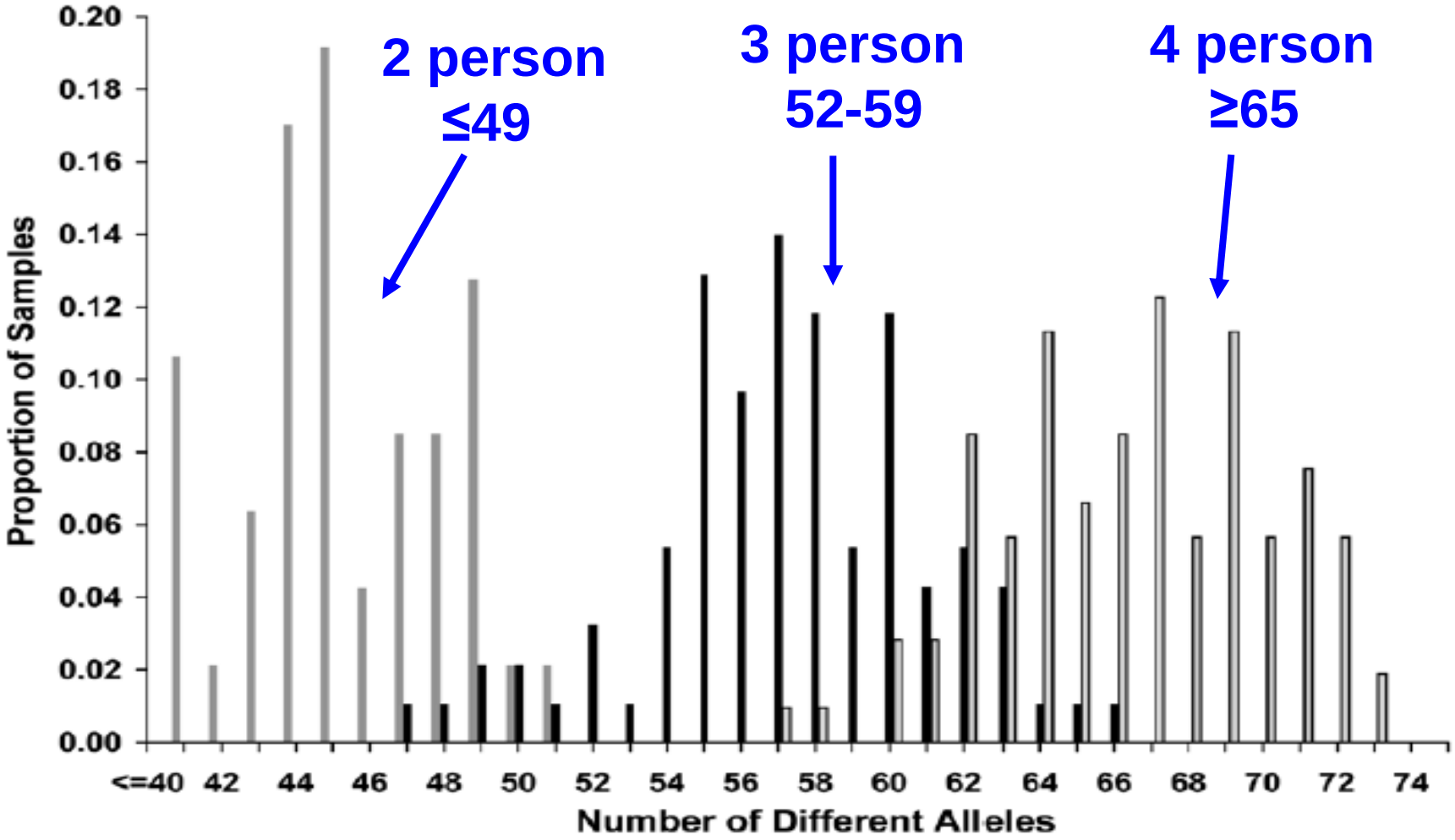


Figure 1. Expected # of different alleles from mixtures.

Estimating the number of contributors to two-, three-, and four-person mixtures containing DNA in high template and low template amounts
Perez et al., Croat Med J. 2011; 52:314-26

Two-Person Mixture Studies Summary

Based on Allele Counts Alone:

- **Always** recognized as a mixture – no risk of confusing as a single-source
  - Loci with 3 or 4 alleles
  - Peak height ratio imbalance at loci with 2 alleles
- Observe more loci with 2 or 3 alleles than 4 alleles – even when DNA from two heterozygous individuals were mixed
- 49 or fewer total alleles

Buckleton et al. Forensic Science International: Genetics 1 (2007) 20–28; Paoletti et al. J Forensic Sci, Nov. 2005, Vol. 50, No. 6; Haned et al. J Forensic Sci, January 2011, Vol. 56, No. 1; Perez et al., Croat Med J. 2011; 52:314-26

Three-Person Mixture Studies Summary

- No risk of confusing as a single-source
- Small risk of confusing with two-person mixture
  - Observe at least one locus with 5 or 6 alleles in ~97% of profiles (3% have ≤4 alleles)
  - Maximum allele count works most of time
  - 3% profiles look like 2-person mixture
  - Risk if LT-DNA, degradation, inhibition, primer mutation to look like 2-person mixture
- Most loci have 3 or 4 alleles
- 52-59 total alleles

Buckleton et al. Forensic Science International: Genetics 1 (2007) 20–28; Paoletti et al. J Forensic Sci, Nov. 2005, Vol. 50, No. 6; Haned et al. J Forensic Sci, January 2011, Vol. 56, No. 1; Perez et al., Croat Med J. 2011; 52:314-26

Four-Person Mixture Studies Summary

- No risk of confusing as a single-source
- Very small risk of confusing with two-person mixture
  - Likely to have peak height imbalance
- **Very small number of loci with 8 alleles and very few with 7 alleles**
  - High risk of confusing with three-person mixture
  - Risk if LT-DNA, degradation, inhibition, primer mutation
- ≥65 total alleles

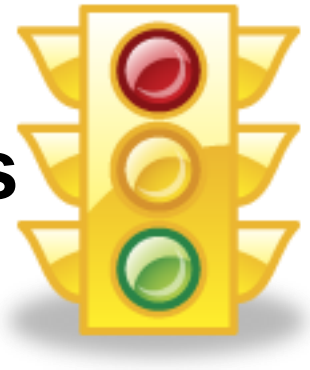
Buckleton et al. Forensic Science International: Genetics 1 (2007) 20–28; Paoletti et al. J Forensic Sci, Nov. 2005, Vol. 50, No. 6; Haned et al. J Forensic Sci, January 2011, Vol. 56, No. 1; Perez et al., Croat Med J. 2011; 52:314-26

Four-Person Mixture Studies Summary

>70% of 4-person mixtures would NOT be recognized as 4-person mixtures based on maximum number allele count at a locus

Buckleton et al. Forensic Science International: Genetics 1 (2007) 20–28; Paoletti et al. J Forensic Sci, Nov. 2005, Vol. 50, No. 6; Haned et al. J Forensic Sci, January 2011, Vol. 56, No. 1; Perez et al., Croat Med J. 2011; 52:314-26

Five-, Six- Person Mixture Studies Summary

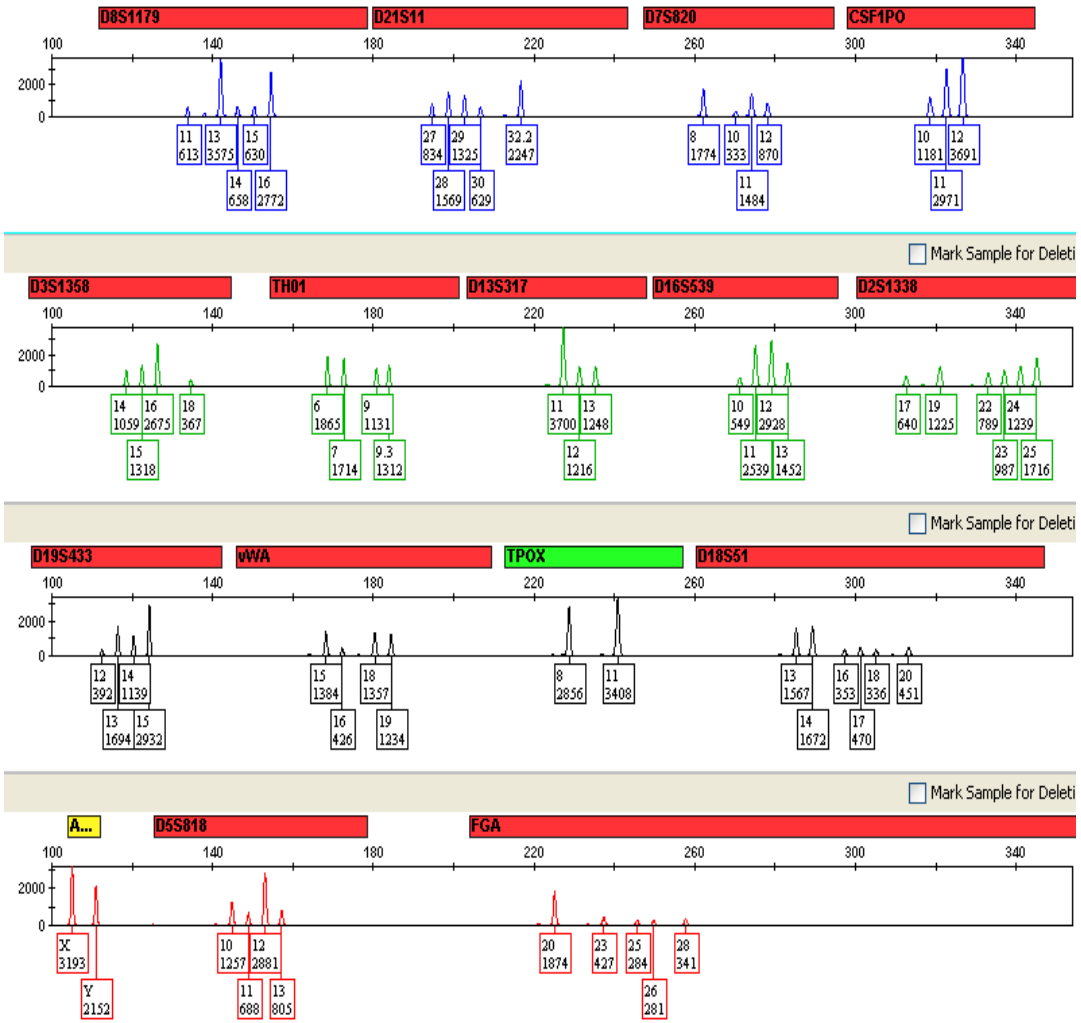


- >99% of 5 person mixtures would look like 4 person mixtures (~60%) or 3-person mixtures (~40%)
- Most 6 person mixtures would look like 5 person mixture (6%), 4-person mixtures (80%) or 3-person mixtures (14%)

Wang, T.W., Kalet, P., Pendleton, J., Gilbert, K., Lucas, L. and Birdwell, J.D. 2005 The probable number of contributors to a STR DNA mixture. http://www.promega.com/products/pm/genetic-identity/ishi-conference-proceedings/16th-ishi-poster-abstracts/; Haned et al. J Forensic Sci, January 2011, Vol. 56,(1), 23-28

Complex Mixture – Allele Summary

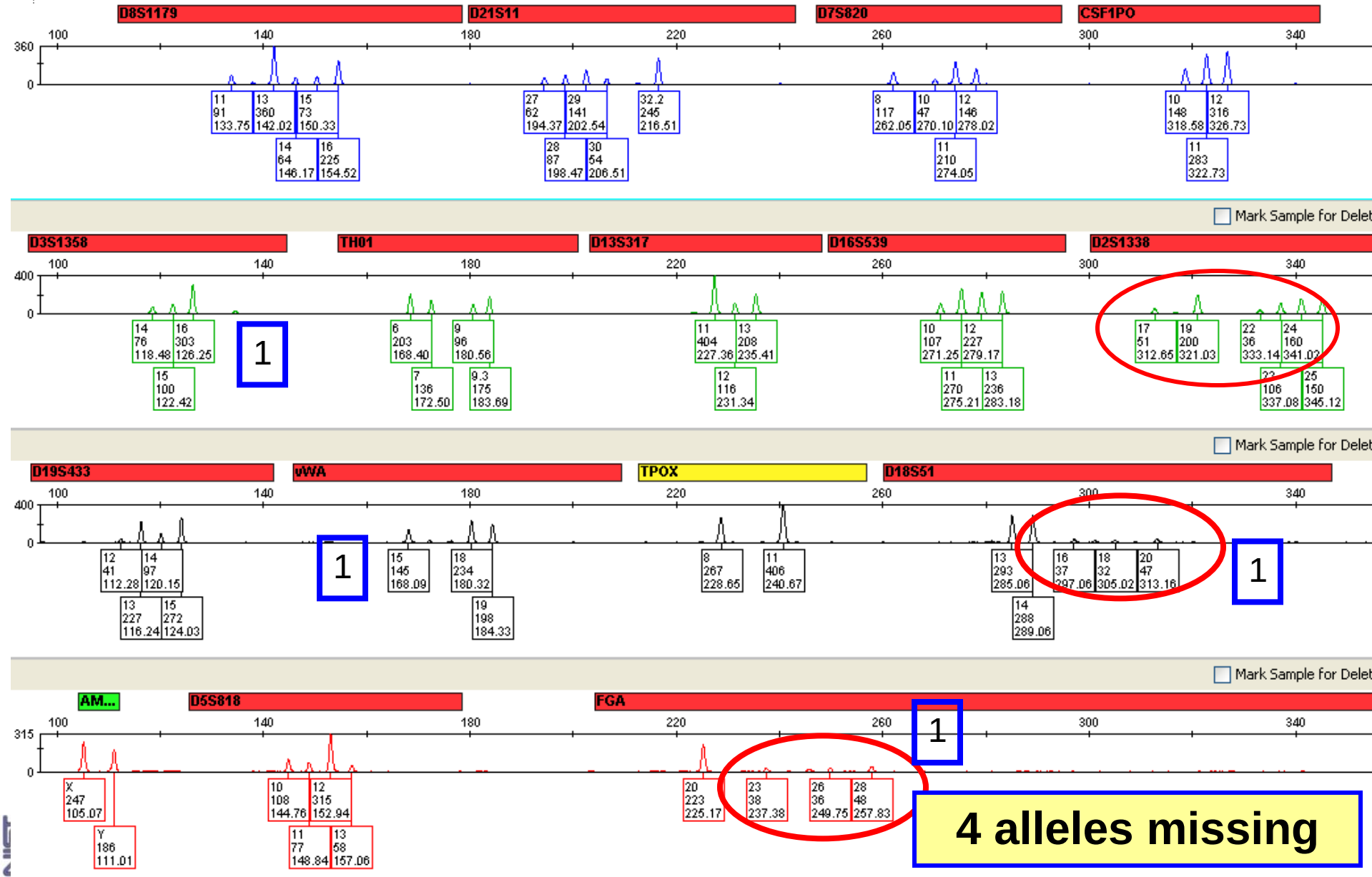
- 6 alleles at 2 loci
- 5 alleles at 3 loci
- 4 alleles at 7 loci
- 3 alleles at 2 loci
- 2 alleles at 1 locus
- 1 allele at 0 loci
- 63 total alleles



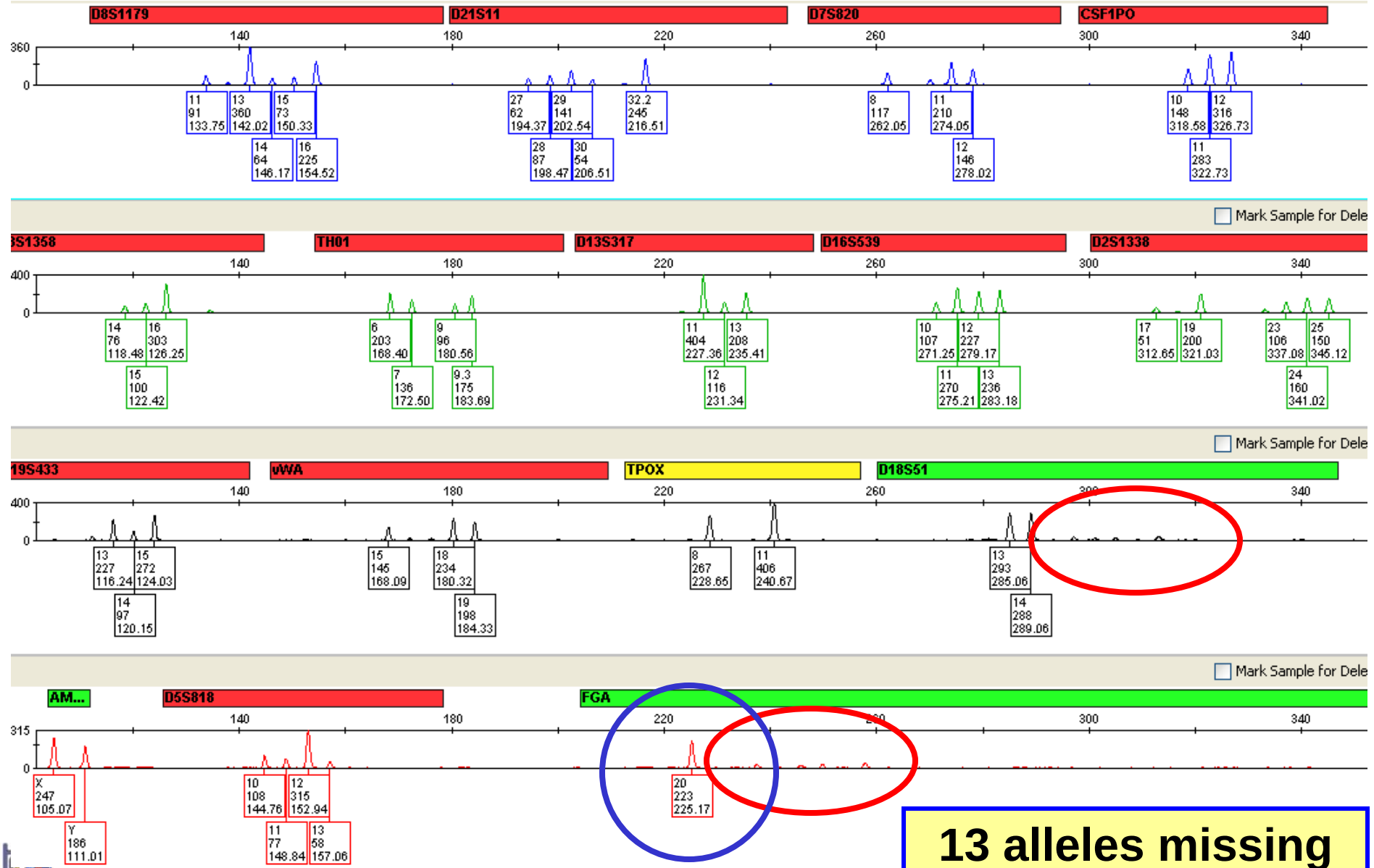
No Major Contributor!

A 4-person mixture @ 3:2:1.6:1 ratio!!

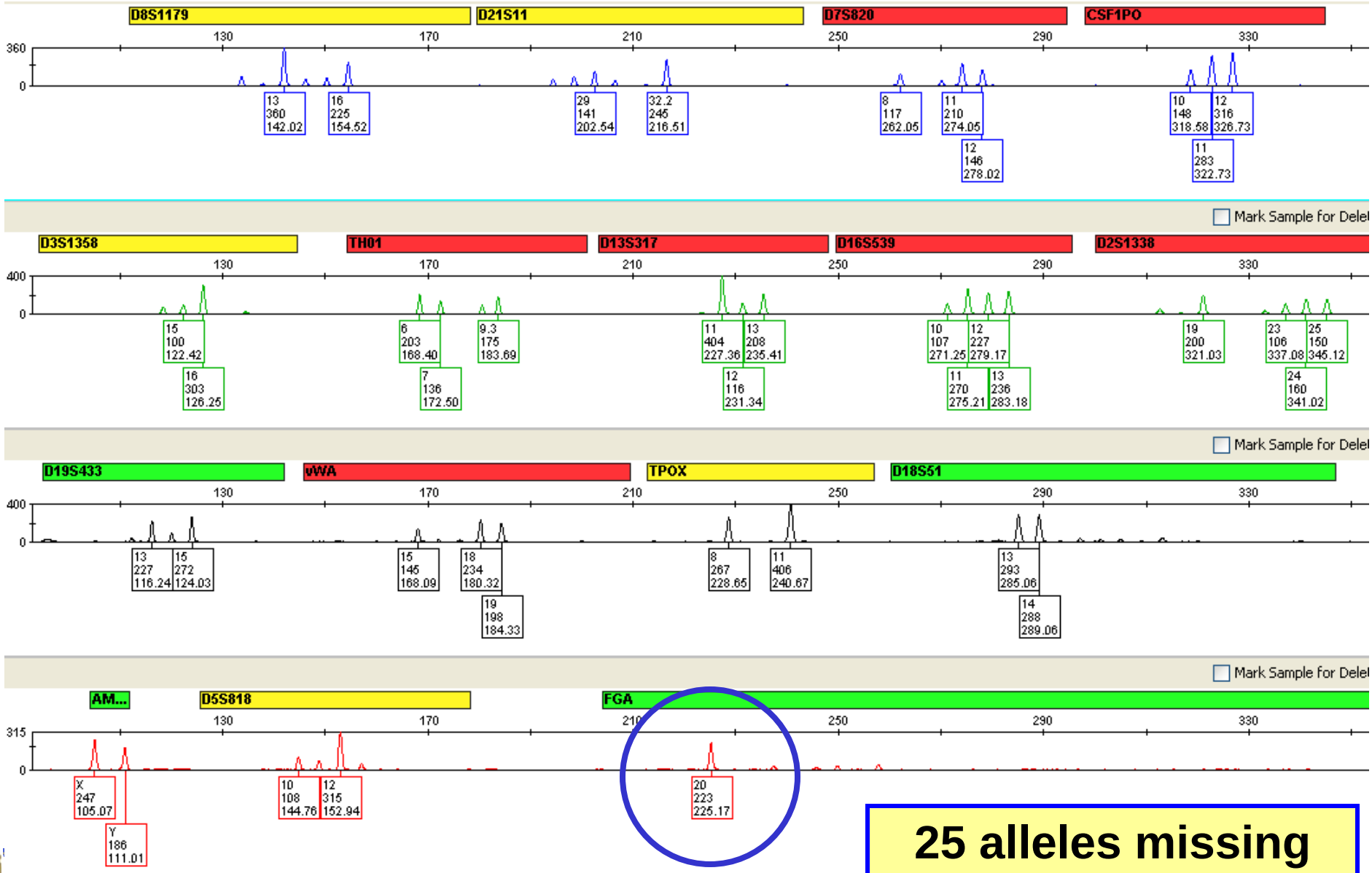
Mixture with 30 RFU Analytical Threshold



Mixture with 50 RFU Analytical Threshold



Mixture with 100 RFU Analytical Threshold



Mixture with 100 RFU Analytical Threshold

- Looks like it could be a two-person mixture
- Looks like it may have a major contributor at some loci, but not all → indistinguishable mixture?
- Many alleles near or above 150-200 RFU

Good to interpret?

Mixture with 100 RFU Analytical Threshold

- If compare this profile to the known contributors:
  - The highest peak or peaks are not always from the person with the most DNA (3:2:1.6:1)
  - The highest peaks are not consistent with any of the known contributors over the profile
  - Cannot correctly “pull out” any one or two of the correct contributors at all loci
  - The “major” contributor is missing an allele from this profile
- Allele shares complicate mixture interpretation
- Allele shares can cause high peaks that are suggestive of major contributor profiles
- Stochastic effects lead to loss of data

CPI Statistical Frequencies with Different Analytical Thresholds

	Frequency of 1 in __ unrelated individuals			
	Full Profile	**30 RFU**	**50 RFU**	**100 RFU**
Caucasian	5,300	45,000	2,400,000*	5.7 billion*
African American	25,000	250,000	290,000,000*	870 billion*
SW Hispanic	4,400	75,000	10,000,000*	20 billion*

*Single allele at one locus; p^2 in calculation rather than 2p

Total # of Alleles	63	59	50	38
# of Alleles Missing	--	4	13	25

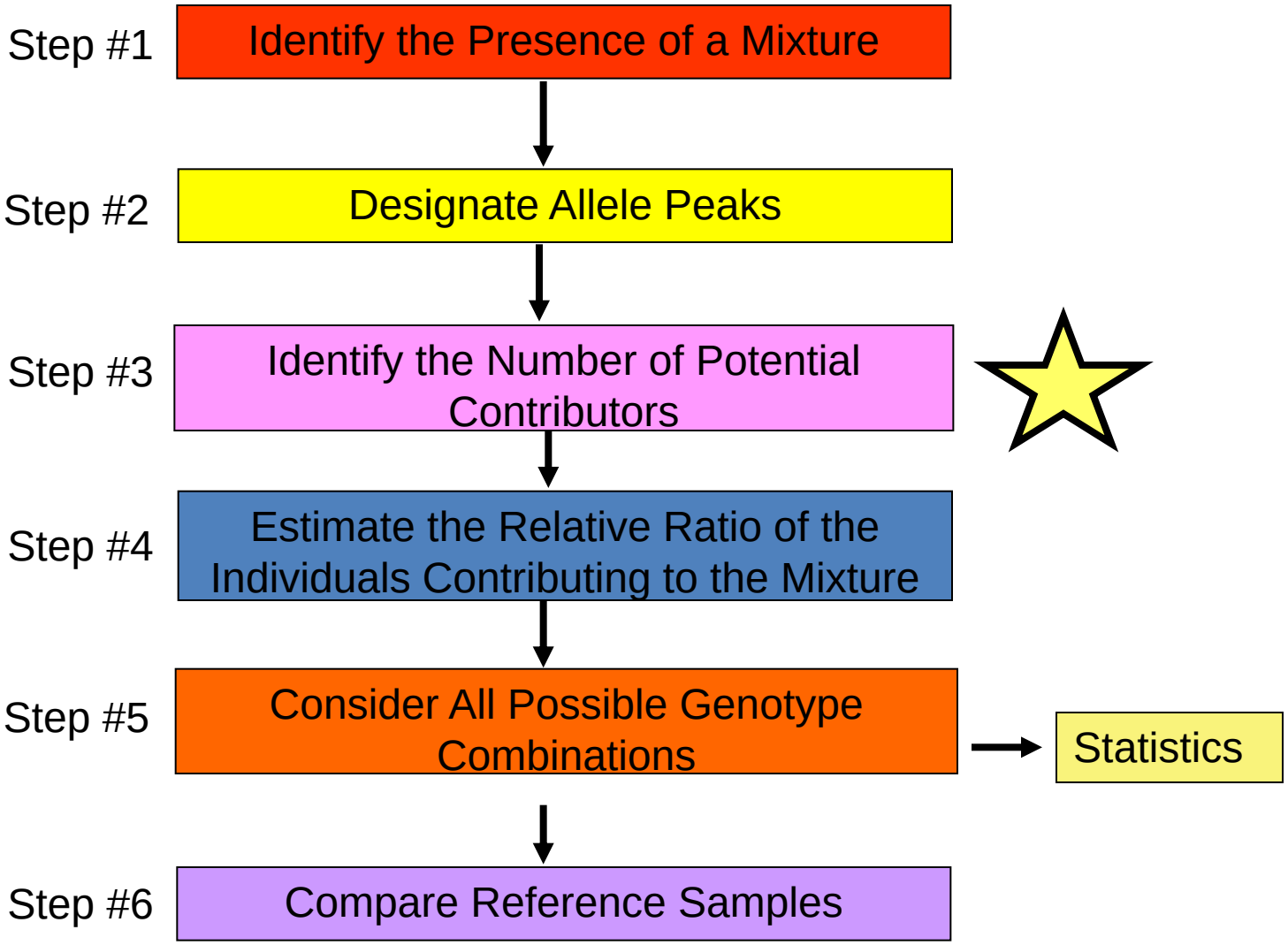
Thanks to Liz Benzinger and Kristen Slaper for the PopStats Calculations!

MYTH

It is easy to determine the number of contributors to a DNA profile.

Steps in the interpretation of mixtures

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



Modified from slide from Dr. John Butler

INCREASED COMPLEXITY

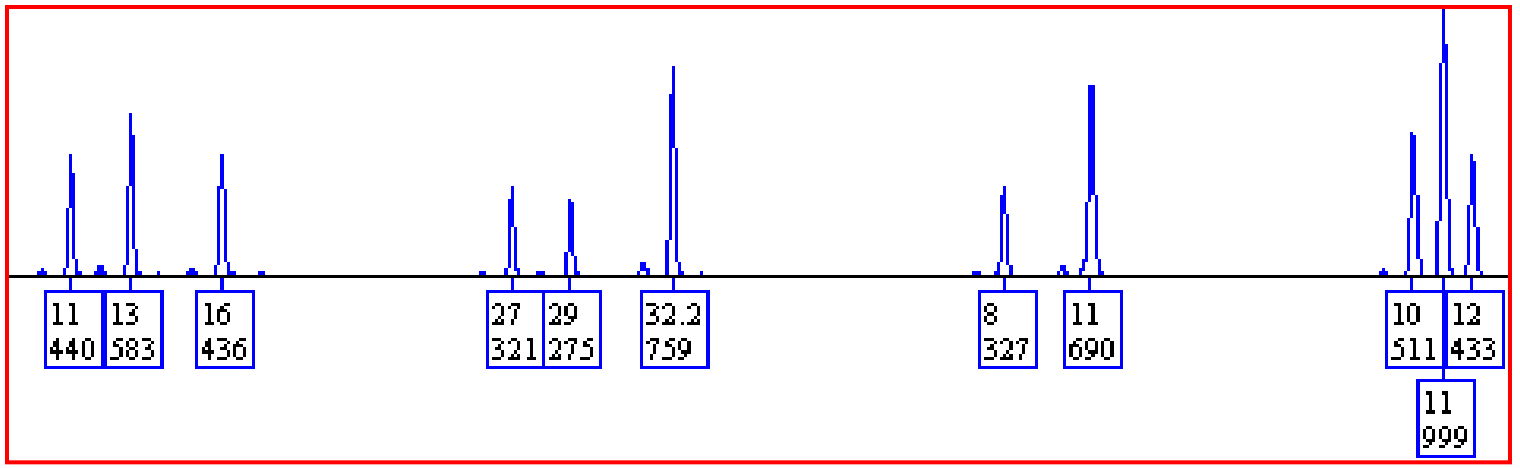
HIGH UNCERTAINTY

LACK OF CONFIDENCE

Complex Mixtures

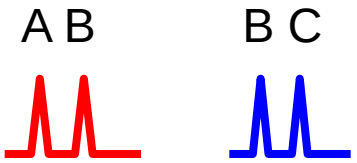
Mixtures with Relatives

Parent-Child
Sibling-Sibling



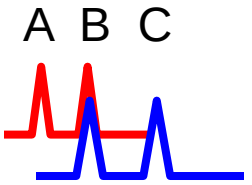
Parent + Child

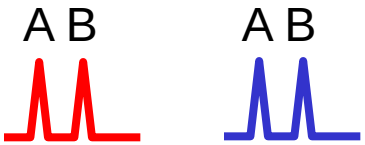
Mixture DNA Profile Pattern



Maximum: 3 alleles

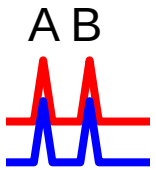
Both heterozygote, one shared allele

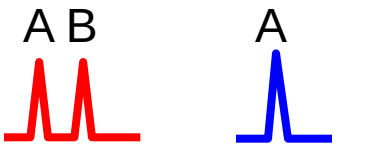




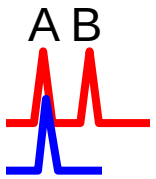
2 alleles

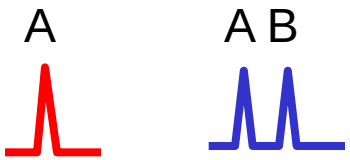
Heterozygote + heterozygote, two shared alleles



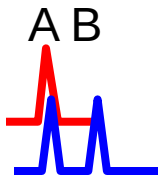


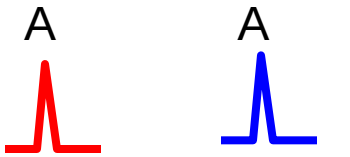
Heterozygote + homozygote, one shared allele





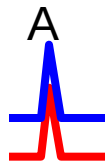
Homozygote + heterozygote, one shared allele



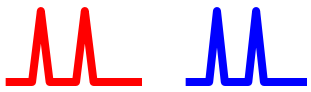
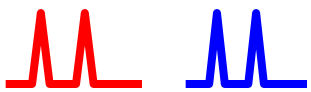
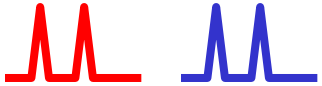
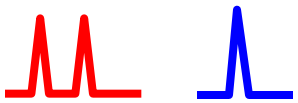
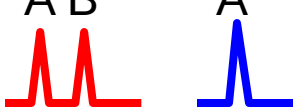
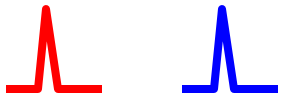
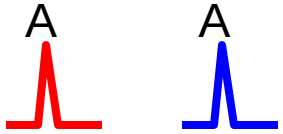


1 allele

Homozygote + homozygote, one shared allele



ALLELE SHARE AT EACH LOCUS

P1 + P2	Genotypes of Children	% Sibling Allele Sharing
A B C D 	AC or AD or BC or BD	0%, 50% or 100%
A B B C 	AB or AC or BB or BC	0%, 50% or 100%
A B A B 	AB/BA or AA or BB	0%, 50% or 100%
A B C 	AC or BC	50% or 100%
A B A 	AA or BA	50% or 100%
A B 	AB	100%
A A 	AA	100%

P1 = Parent 1; P2 = Parent 2

Allele Sharing in Relatives

ELSEVIER

Forensic Science International 131 (2003) 85–89

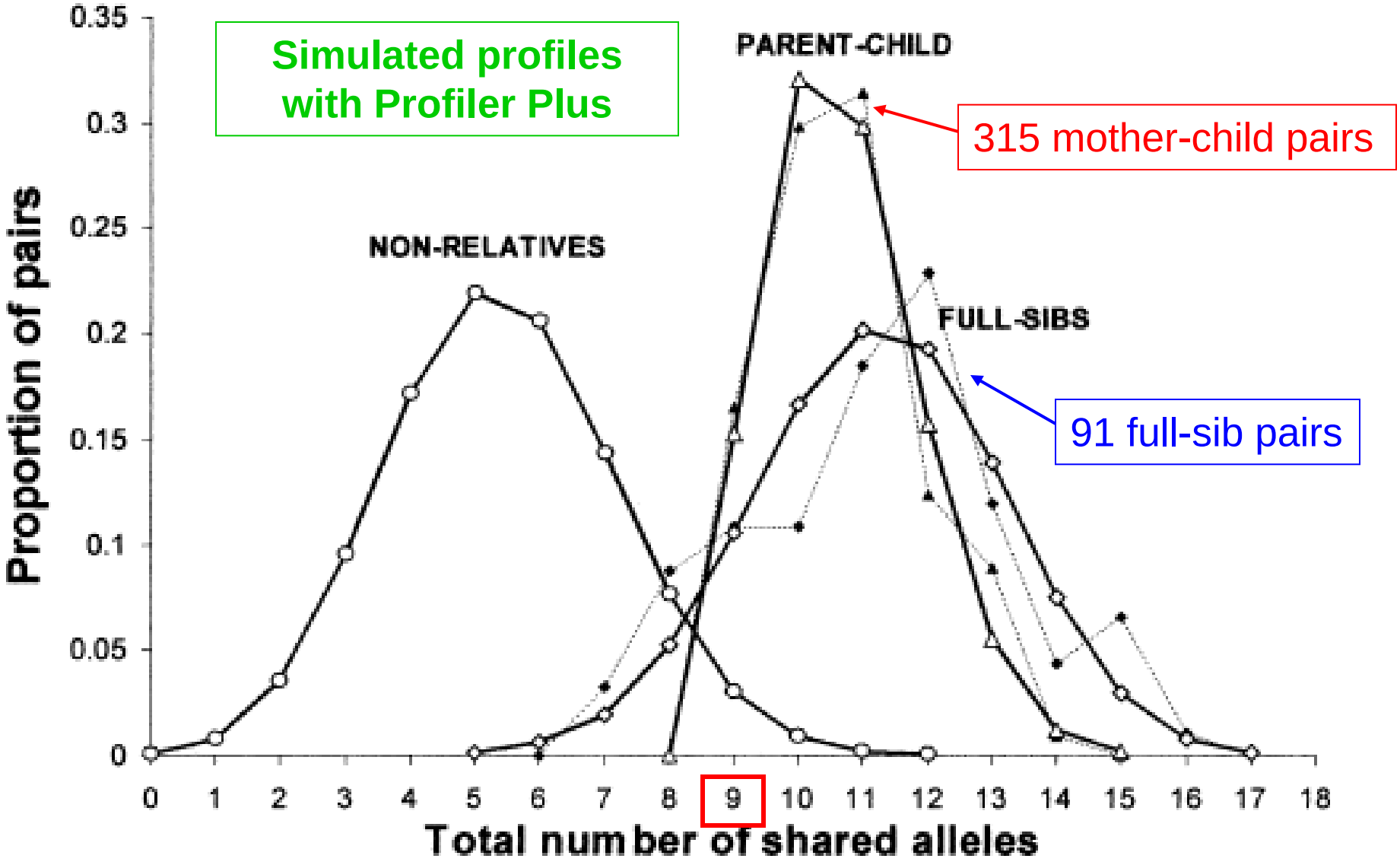
Forensic Science International

www.elsevier.com/locate/forsciint

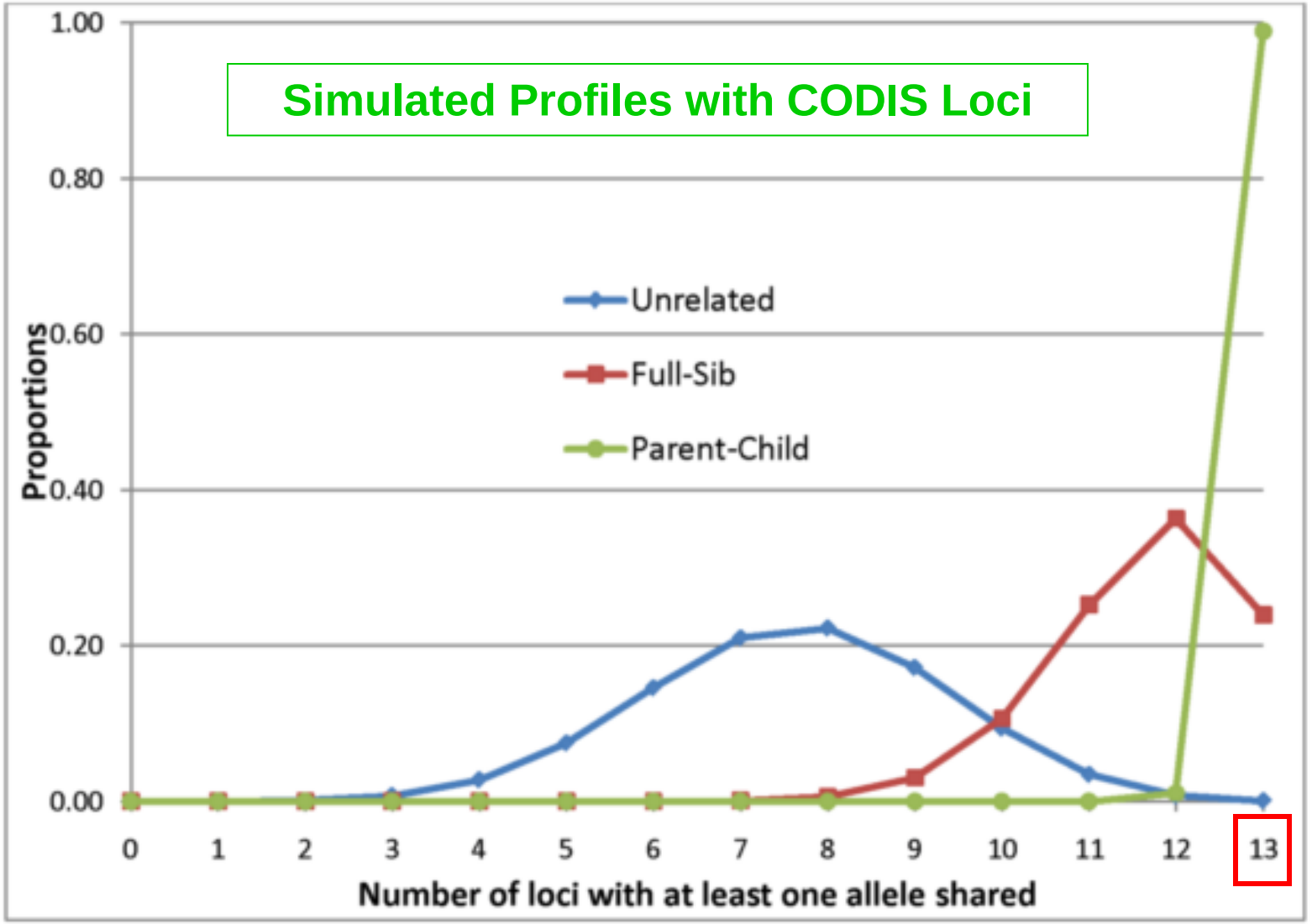
Allele sharing in first-degree and unrelated pairs of individuals in the Ge.F.I. AmpFlSTR® Profiler Plus™ database

Silvano Presciuttini[a,l,*], Francesca Ciampini[a], Milena Alù[b], Nicoletta Cerri[c], Marina Dobosz[d], Ranieri Domenici[e], Gabriella Peloso[f], Susi Pelotti[g], Andrea Piccinini[h], Elena Ponzano[i], Ugo Ricci[j], Adriano Tagliabracci[k], J.E. Baley-Wilson[l], Francesco De Stefano[m], Vincenzo Pascali[d,l]

Presciuttini et al. Forensic Science International 131 (2003) 85-89



Forensic Science International 131 (2003) 85-89



Ge et al. Comparisons of the familial DNA databases searching policies. J. Forensic Sci. 2011;56(6):1448-56.

Mixtures with Relatives – Summary

Parent-Child

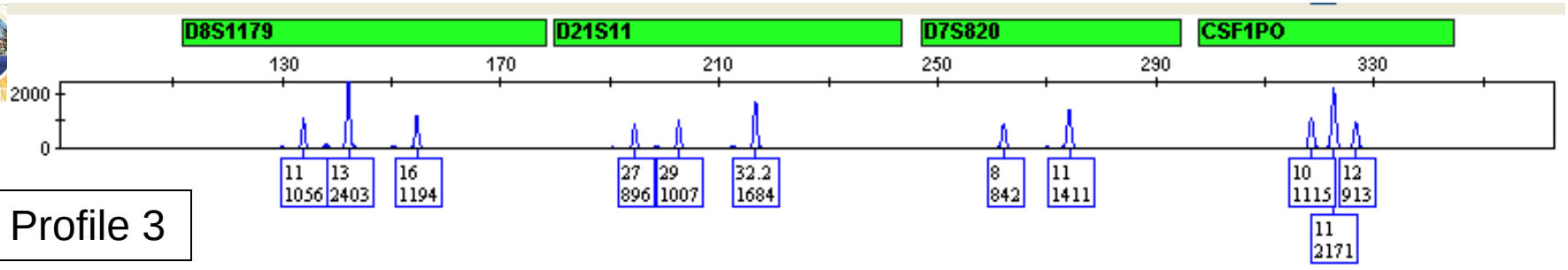
- Expect at least 50% allele share
- Expect at least one shared allele at each locus
- Maximum 3 alleles per locus (in absence of mutation)
- If test X loci, expect >X allele shares (9-14 Profiler Plus; 13-20 CODIS)

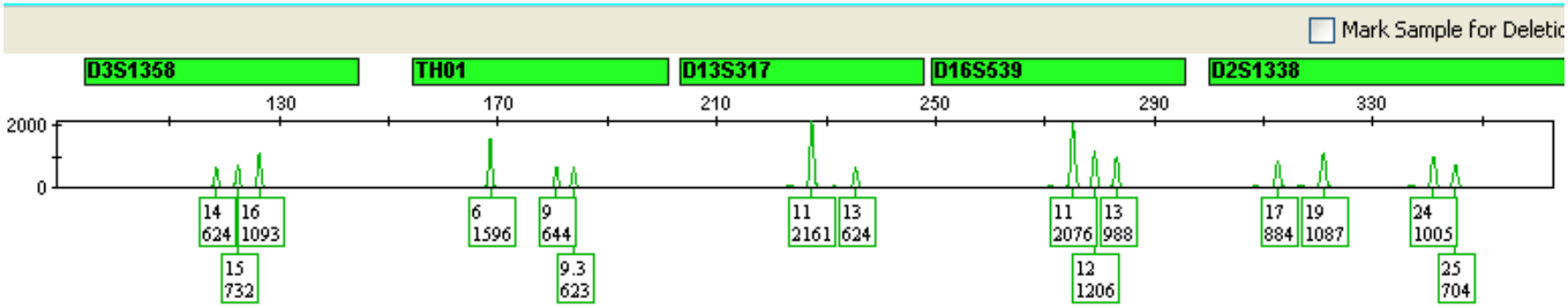
Mixtures with Relatives – Summary

Sibling-Sibling

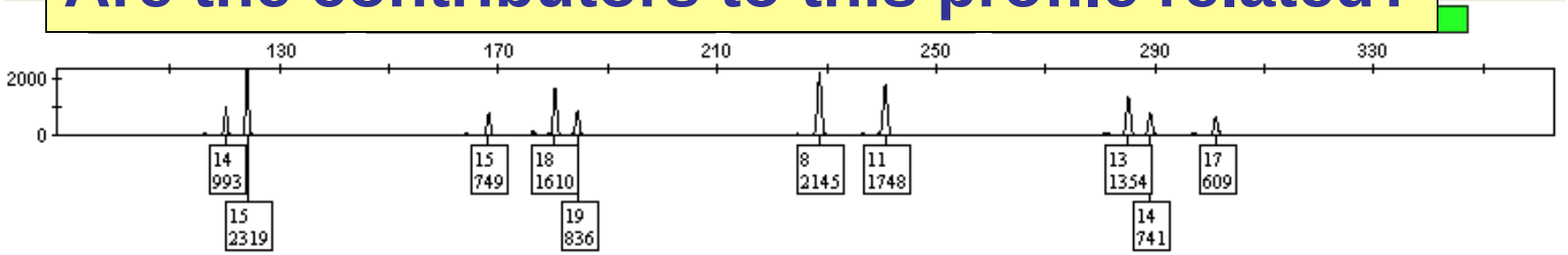
- Expect at least 50% allele share overall, but variable: 7-16 Profiler Plus; 12-22 CODIS (≥X-1)
- Expect 0, 50 or 100% allele share at each locus
- Expect at least one allele share at 9-13 loci (CODIS data)

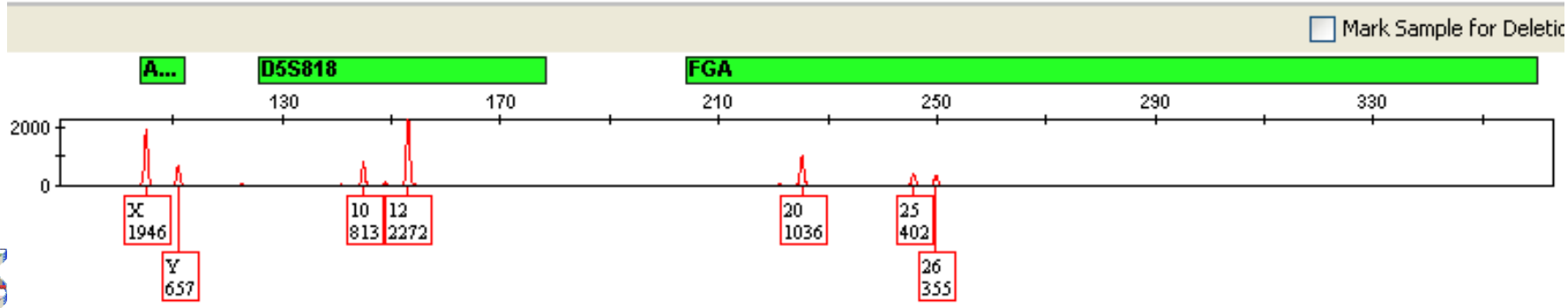
Profile 3





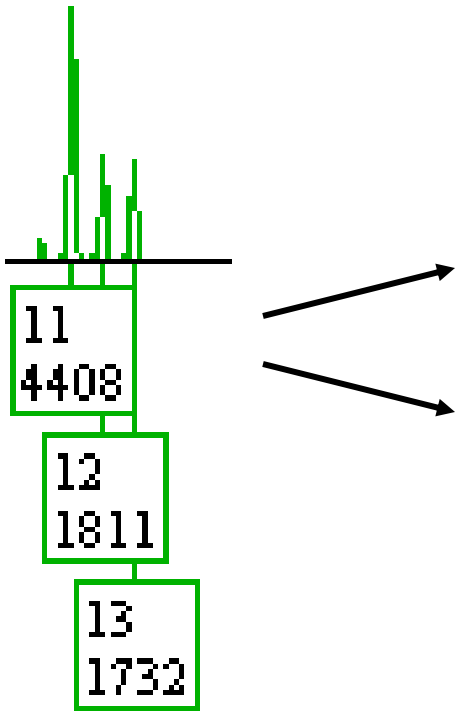
Are the contributors to this profile related?





Mixtures with Relatives –
Working Backwards from Mixed DNA Profile

- With mixed DNA profile from unknowns, may not know if alleles are shared
- Data in the graphs are not helpful



11,12 + 11,13

or

11,11 + 12,13

Unrelated?

Relative?
Parent-Child?
Sibs?

Relative?
Sibs?

True Known Contributors to Previous Profile

- Share 14 alleles over 15 Identifiler loci
  - 8 alleles at 9 Profiler Plus loci
  - 13 alleles at 13 CODIS loci
  - 15 alleles 17 loci (Identifiler + PowerPlex 16 HS)
- One allele in common at each locus, except D2, FGA and Penta E
- Likely not parent, unless mutations occurred
- Sibs?
  - Using known contributors' profiles : Inconclusive from allele #; Ge locus data suggests sibs

Relatives in DNA Testing

What if the true contributors **in a mixed DNA sample** are closely related?

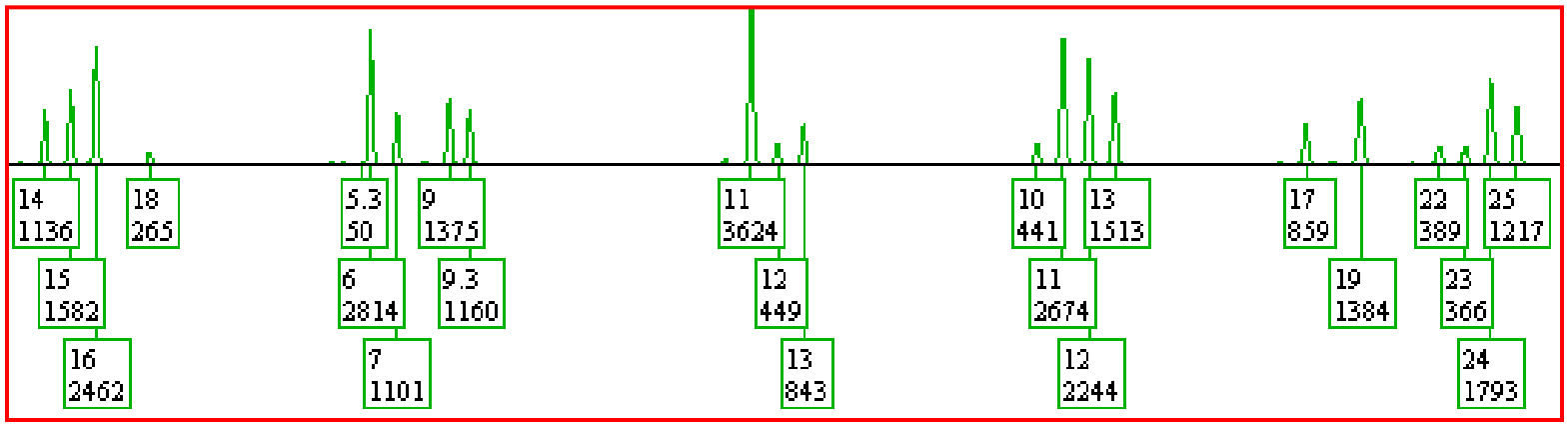
- Significant issue with the types of samples being tested today (e.g., “Touch” DNA)
  - Any item likely to routinely be used/shared by related individuals (e.g., living in same household, driving same car, sharing clothing)
  - Relatives committing crimes together (e.g., shared clothing, weapons)
  - Multiple homicides involving family members
- NO statistical method to address this
- Statistics reported for “random individuals”
  - However, a relative is more likely to be included

Complex Mixture Interpretation

- We have limited experience with known complex mixtures (training, validation, or proficiency tests)
- No or limited published guidelines for interpretation
- Limited interpretation SOPs available
- Routine amount of DNA amplified → poor quality profiles, LT DNA likely for 1 or more contributors
- How do you do the statistical calculations?

Complex Mixture Interpretation

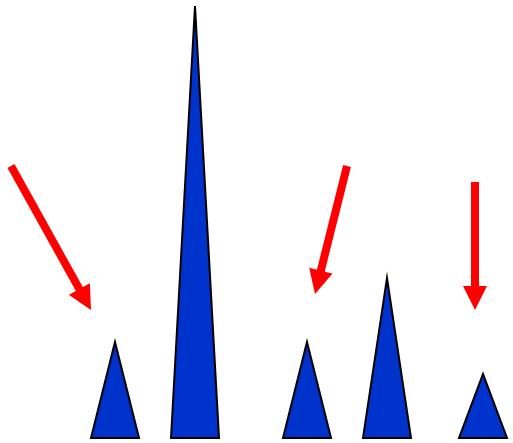
Is hard because the parameters used to interpret two-person mixtures often may not be directly applicable to complex mixtures



Complex Mixtures

More Uncertainty and Lack of Confidence

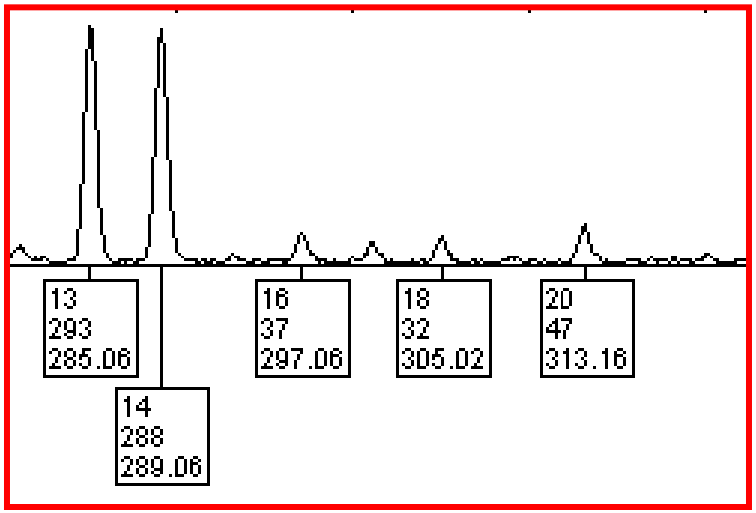
- Peak vs. Artifacts
  - Stutter?
  - Pull-up?
  - True Allelle?



Complex Mixtures

More Uncertainty and Lack of Confidence

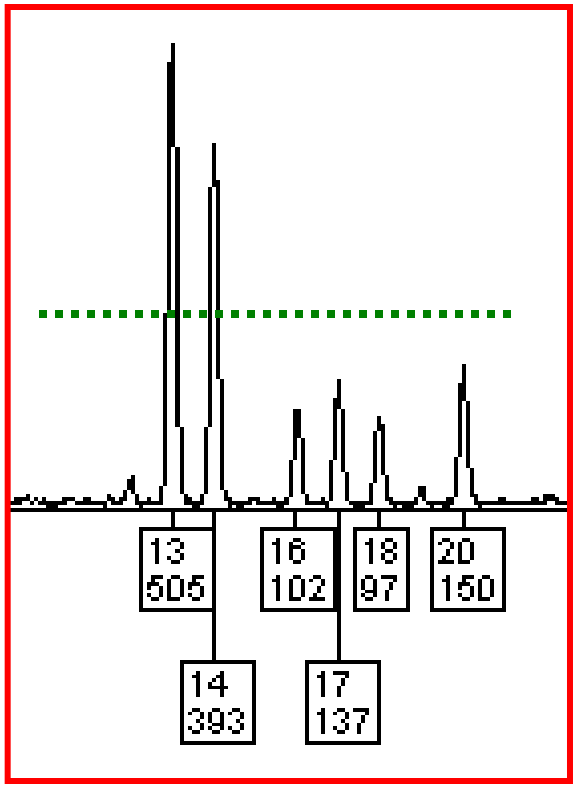
- High likelihood that DNA from one or more contributors is below optimal range
  - LT DNA = stochastic effects
  - Missing alleles? (allele drop out)
  - Elevated Stutter? True allele vs. Stutter?
  - Allele drop-in?



Complex Mixtures

More Uncertainty and Lack of Confidence

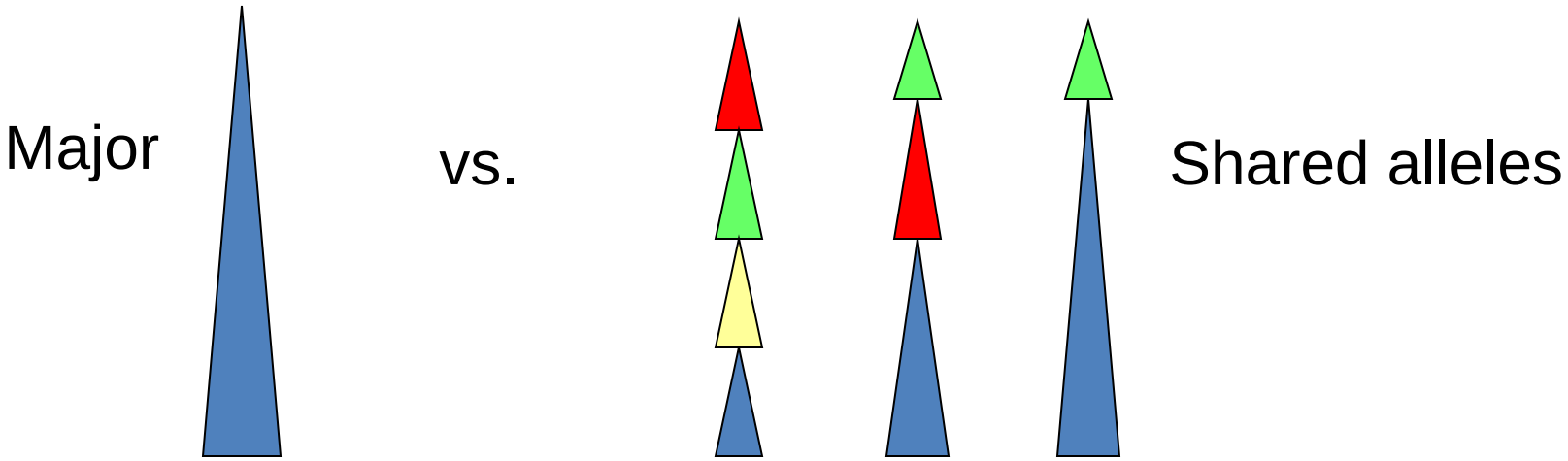
- **Stochastic threshold**
  - **Only meaningful for the peaks below the value – may be missing sister allele**
  - **Only helps with assessing if ALL alleles are likely present**



Complex Mixtures

More Uncertainty and Lack of Confidence

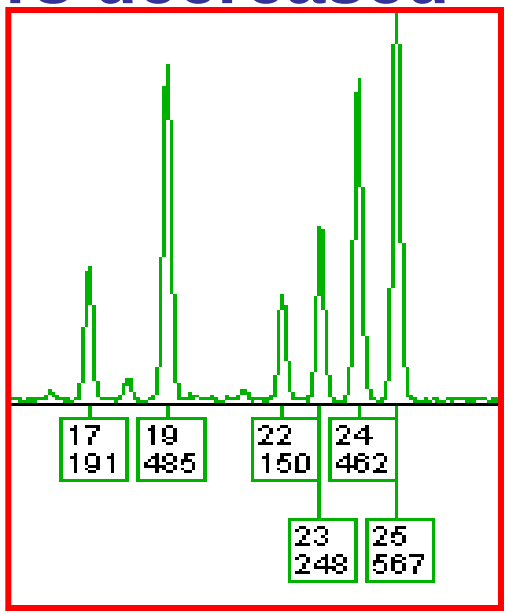
- **Stochastic threshold**
  - **NO meaning for peaks above the value –**
    - **Major contributor?**
    - **Shared alleles? How many shares? Relatives or unrelated?**



Complex Mixtures

More Uncertainty and Lack of Confidence

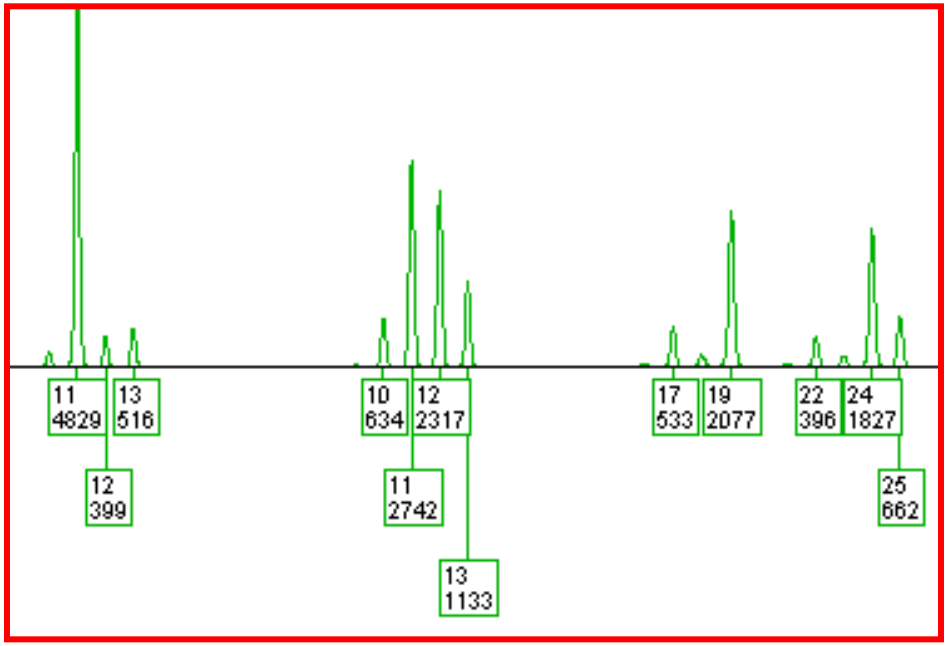
- Peak height ratios have no meaning at most or all loci
  - Cannot use to associate alleles into genotypes
  - Ability to deduce other contributors decreased even if you know one contributor



Complex Mixtures

More Uncertainty and Lack of Confidence

- Mixture ratio cannot be calculated
  - Different amount from each contributor likely with no way to determine
  - Cannot use to associate genotypes into profiles



Complex Mixtures

More Uncertainty and Lack of Confidence

- Number of contributors – maximum allele count/minimum number often an underestimate
  - What number to assume?
  - May need to interpret under multiple assumptions (especially if the conclusion changes)

Complex Mixtures

More Uncertainty and Lack of Confidence

- **"Inclusion" based on alleles NOT based on genotypes → may not be correct inclusion**
- **False Inclusions**
  - Increased risk as # of alleles increase
- **How calculate statistical frequency?**

Complex Mixtures

Exclusions less likely/ Exclusion criteria difficult to develop

- Can anyone be excluded if LT DNA present?
- Partial “inclusions”

Estimate frequency of included individuals can be quite common – can become meaningless (1 in 2 individuals)

Inconclusive reporting increased

What can we do?

- Amplify more DNA?
- Test another portion of the sample?
- Test another sample in the case?
- Probabilistic approaches to interpretation? (stay tuned)

Conclusions

- Criteria routinely used in crime laboratories for the interpretation of two-person mixtures may not apply for most complex mixtures
- LT-DNA, degradation, inhibition play more significant role
- Additional complex mixtures need to be generated and evaluated for establishment of scientifically supported interpretation guidelines

THANK YOU!!

John Butler

Mike Coble

Robin Cotton

Catherine Grgicak

Bruce Heidebrecht

& Workshop attendees

For many hours of discussions!

Catherine Grgicak

Robin Cotton

NIJ Grant to Boston University

For all of the profiles!