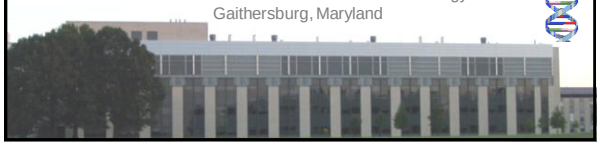
The Copenhagen Forensic Genetic Summer School
Advanced Topics in STR DNA Analysis
June 27-28, 2012

Application of Thresholds for Interpretation

Mike Coble and Becky Hill

NIST Applied Genetics Group
National Institute of Standards and Technology
Gaithersburg, Maryland



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

Available online at www.sciencedirect.com
SCIENCE DIRECT
Forensic Science International 160 (2006) 90–101
www.elsevier.com/locate/forsciint

DNA commission of the International Society of Forensic Genetics: Recommendations on the interpretation of mixtures

P. Gill [a,*], C.H. Brenner [b], J.S. Buckleton [c], A. Carracedo [d], M. Krawczak [e], W.R. Mayr [f], N. Morling [g], M. Prinz [h], P.M. Schneider [i], B.S. Weir [j]

[a] Forensic Science Service, Trident Court, 2960 Solihull Parkway, Birmingham, UK
[b] Forensic Science Group, School of Public Health, University of California, Berkeley, CA 510-339-1911, USA
[c] ESR, Private Bag 92021, Auckland, New Zealand
[j] University of Washington, Department of Biostatistics, Box 357232, Seattle, WA 98195, USA

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

Received 4 April 2006; accepted 10 April 2006
Available online 5 June 2006

Gill *et al.* (2006) DNA Commission of the International Society of Forensic Genetics: Recommendations on the interpretation of mixtures. *Forensic Sci. Int.* 160: 90-101

Available online at www.sciencedirect.com
SCIENCE DIRECT
Forensic Science International 160 (2006) 89
www.elsevier.com/locate/forsciint

Editorial

Editorial on the recommendations of the DNA commission of the ISFG on the interpretation of mixtures

"…**These recommendations have been written** to serve two purposes: to define a generally acceptable mathematical approach for typical mixture scenarios and to address open questions where practical and generally accepted solutions do not yet exist. This has been done **to stimulate the discussion among scientists in this field**. **The aim is to invite proposals and criticism in the form of comments and letters to the editors of this journal**…**We are hoping to continue the process to allow the DNA Commission to critically revise or extend these recommendations in due time…"**

Summary of ISFG Recommendations on Mixture Interpretation

1. The likelihood ratio (LR) is the preferred statistical method for mixtures over RMNE
2. Scientists should be trained in and use LRs
3. Methods to calculate LRs of mixtures are cited
4. Follow Clayton et al. (1998) guidelines when deducing component genotypes
5. Prosecution determines H_p and defense determines H_d and multiple propositions may be evaluated
6. When minor alleles are the same size as stutters of major alleles, then they are indistinguishable
7. Allele dropout to explain evidence can only be used with low signal data
8. No statistical interpretation should be performed on alleles below threshold
9. Stochastic effects limit usefulness of heterozygote balance and mixture proportion estimates with low level DNA

Gill *et al.* (2006) DNA Commission of the International Society of Forensic Genetics: Recommendations on the interpretation of mixtures. *Forensic Sci. Int.* 160: 90-101

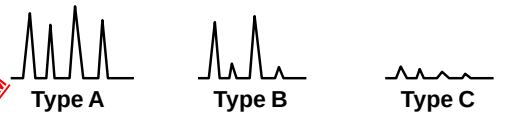
German Mixture Classification Scheme

Schneider *et al*. (2009) *Int. J. Legal Med*. 123: 1-5

(German Stain Commission, 2006):

- **Type A**: no obvious major contributor, no evidence of stochastic effects
- **Type B**: clearly distinguishable major and minor contributors; consistent peak height ratios of ***approximately 4:1*** (major to minor component) for all heterozygous systems, no stochastic effects
- **Type C**: mixtures without major contributor(s), evidence for stochastic effects

SWGDAM



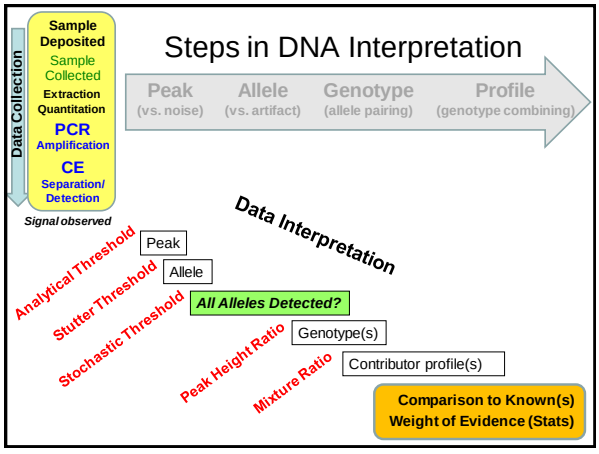
Type A "Indistinguishable"

Type B "Distinguishable"

Type C "Uninterpretable"

Responses to ISFG DNA Commission Mixture Recommendations

- UK Response
  - Gill *et al.* (2008) *FSI Genetics* 2(1): 76–82
- **German Stain Commission**
  - Schneider *et al.* (2006) *Rechtsmedizin* 16:401-404 (German version)
  - Schneider *et al.* (2009) *Int. J. Legal Med.* 123: 1-5 (English version)
- ENFSI Policy Statement
  - Morling *et al.* (2007) *FSI Genetics* 1(3):291–292
- New Zealand/Australia Support Statement
  - Stringer *et al.* (2009) *FSI Genetics* 3(2):144-145
- **SWGDAM** – Interpretation Guidelines
  - Approved Jan 2010 and released April 2010 on FBI website



Principles Behind Thresholds

Thresholds (example values)	Principles Behind *(if properly set based on lab- & kit-specific empirical data)*
Analytical Threshold (e.g., 50 RFU)	Below this value, observed peaks cannot be reliably distinguished from instrument noise (baseline signal)
Limit of Linearity (e.g., 5000 RFU)	Above this value, the CCD camera can become saturated and peaks may not accurately reflect relative signal quantities (e.g., flat-topped peaks) and lead to pull-up/ bleed-through between dye color channels
Stochastic Threshold (e.g., 250 RFU)	Above this peak height value, it is reasonable to assume that allelic dropout of a sister allele of a heterozygote has not occurred at that locus; single alleles above this value in single-source samples are assumed to be homozygous
Stutter Threshold (e.g., 15%)	Below this value, a peak in the reverse (or forward) stutter position can be designated as a stutter artifact with single-source samples or some mixtures (often higher with lower DNA amounts)
Peak Height Ratio (e.g., 60%)	Above this value, two heterozygous alleles can be grouped as a possible genotype (often lower with lower DNA amounts)
Major/Minor Ratio (e.g., 4:1)	When the ratio of contributors is closer than this value in a two-person mixture, it becomes challenging and often impossible to correctly associate genotype combinations to either the major or minor contributor

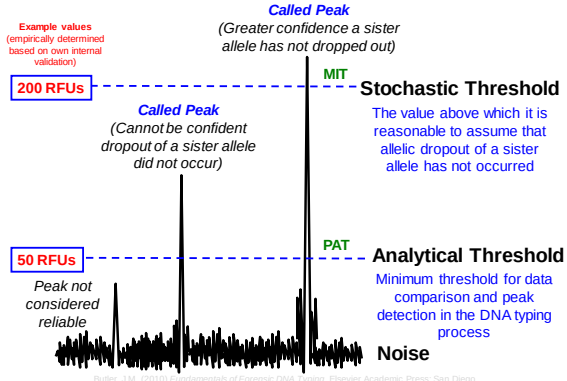
Your Laboratory Interpretation Protocols

Standard Operating Procedures (SOPs)

Validation studies — **Literature** — **Experience**

SWGDAM Guidelines (2010) Introduction: *…the laboratory should utilize written procedures for interpretation of analytical results with the understanding that specificity in the standard operating protocols will enable greater consistency and accuracy among analysts within a laboratory. It is recommended that standard operating procedures for the interpretation of DNA typing results be sufficiently detailed that other forensic DNA analysts can review, understand in full, and assess the laboratory's policies and practices. The laboratory's interpretation guidelines should be based upon validation studies, scientific literature, and experience.*

Overview of Two Thresholds

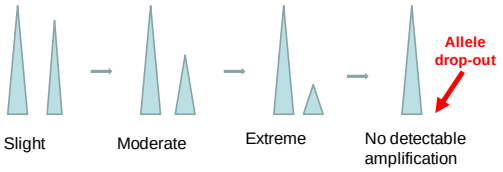


How can we characterize variation?

- Look at total amount of variation at end of process
  - Follow the positive control over time
- Experimentally break process into components and characterize using appropriate statistics
  - e.g., separate amplification variation from injection variation
- Analyze existing or new validation data, training sample data, SRM data, kit QC data
- Use casework data
  - e.g., variation between knowns (victim's DNA profile within an intimate sample) and matching single-source evidence profiles

Problem with Stochastic Effects

- **Allele drop-out** is an extension of the amplification disparity that is observed when heterozygous peaks heights are unequal
  - Occurs in single-source samples and mixtures
  - Analyst is unable to distinguish complete allele drop-out in a true heterozygote from a homozygous state



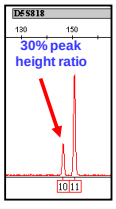
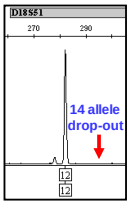
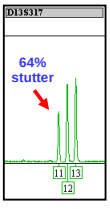
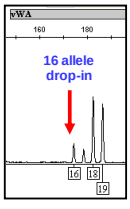
What is Allele Drop Out?

- Scientifically
  - **Failure to detect** an allele within a sample or failure to amplify an allele during PCR. *From SWGDAM Guidelines, 2010*
  - Note that: Failure to detect ≠ failure to amplify
- Operationally
  - Setting a threshold(s) or creating a process, based on validation data and information in the literature, which allows assessment of the likelihood of drop-out of an allele or a locus.

Stochastic Effects with Low Levels of DNA

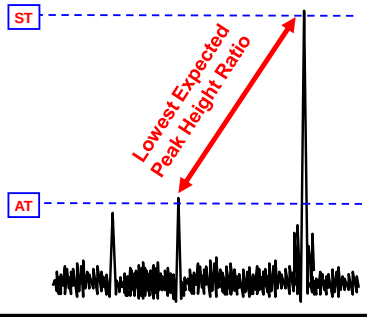
When Combined with Higher Sensitivity Techniques

*Loss of True Signal (**False Negative**)* | *Gain of False Signal (**False Positive**)*

Severe **Peak Imbalance**	Allelic **Drop-out**	**High Stutter**	Allelic **Drop-in**
			
Identifiler, 30 pg DNA, 31 cycles	Identifiler, 30 pg DNA, 31 cycles	Identifiler, 10 pg DNA, 31 cycles	Identifiler, 10 pg DNA, 31 cycles
Correct genotype: **10,11**	**12,14**	**12,13**	**18,19**

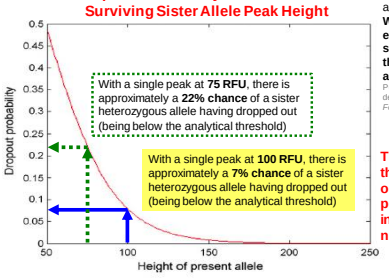
Stochastic and Analytical Thresholds

Impact Lowest Expected Peak Height Ratio



Setting a Stochastic Threshold is Essentially Establishing a Risk Assessment

Drop Out Probability as a Function of Surviving Sister Allele Peak Height



"Currently, most laboratories use an arbitrary stochastic threshold. **When a protocol is changed, especially if it is made more sensitive to low-level DNA, then the stochastic threshold must also change.**"

Puch-Solis R, et al. (2011). Practical determination of the low template DNA threshold. *Forensic Sci. Int. Genet.* 5(5): 422-427.

The position and shape of this curve may change based on anything that can impact peak detection (e.g., CE injection time, PCR cycle number, post-PCR cleanup).

Gill, P., et al. (2009). The *low-template* (stochastic) threshold-Its determination relative to risk analysis for national DNA databases. *FSI Genetics*, 3, 104-111.

Keep in Mind…

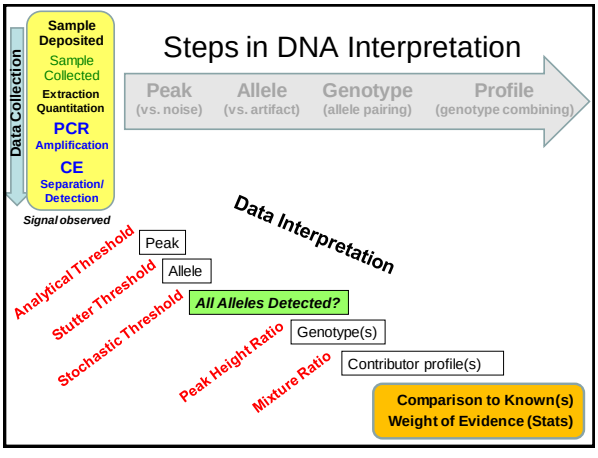
"The use of bounds **applied to data that show continuous variation** is common in forensic science and is often a pragmatic decision. However it should be borne in mind that applying such bounds has arbitrary elements to it and that **there will be cases where the data lie outside these bounds**."

Bright, J.A., et al. (2010). Examination of the variability in mixed DNA profile parameters for the Identifiler multiplex. *Forensic Science International: Genetics, 4,* 111-114.

Appropriately Applying a Stochastic Threshold

Limitations of Stochastic Thresholds

- The possibility of allele sharing with a complex mixture containing many contributors may make a stochastic threshold meaningless
- "Enhanced interrogation techniques" to increase sensitivity (e.g., increased PCR cycles) may yield false homozygotes with >1000 RFU
- **New turbo-charged kits with higher sensitivity will need to be carefully evaluated to avoid allele drop-out and false homozygotes**

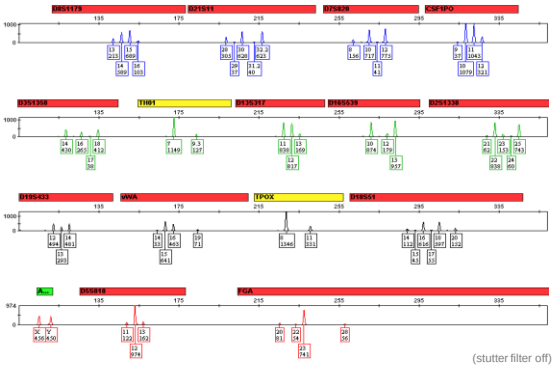


Interpretation of Potential Stutter Peaks in a Mixed Sample

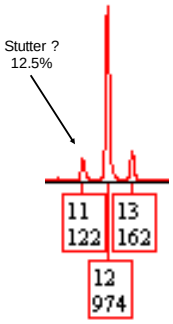
- 3.5.8.1. For mixtures in which minor contributors are determined to be present, a peak in stutter position (generally n-4) may be determined to be 1) a stutter peak, 2) an allelic peak, or 3) indistinguishable as being either an allelic or stutter peak.

Profile 1

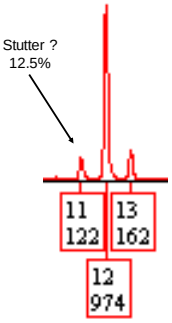
AT = 30 RFU
ST = 200 RFU



Profile 1 - D5S818



Profile 1 - D5S818



Possible Genotypes

13, 13
12, 13
11, 13
13, F

ISFG Recommendation #6 Example

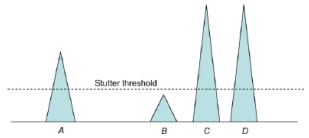


Fig. 2. A two person mixture with major peaks *C, D* and minor peaks *A*. There is an additional peak present in a stutter position (*B*).

Likely a AA

(homozygote)

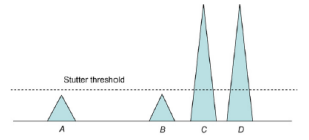


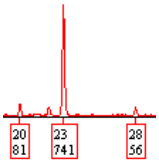
Fig. 3. A two person mixture with major peaks *C, D* and minor peaks *A, B*, where *B* is in a stutter position.

Possibly AB

(heterozygote)

Could also be AC, AD, AA, or A,? (dropout)

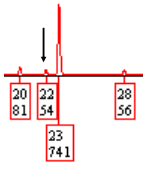
Profile 1 - FGA



Major	Minor

If Assume 2 Contributors….

23,23	20,28

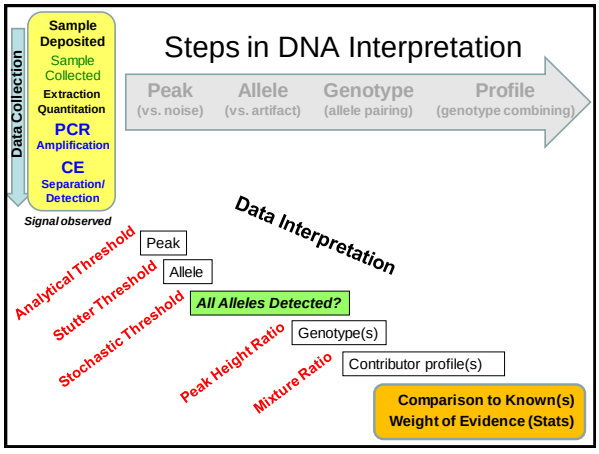


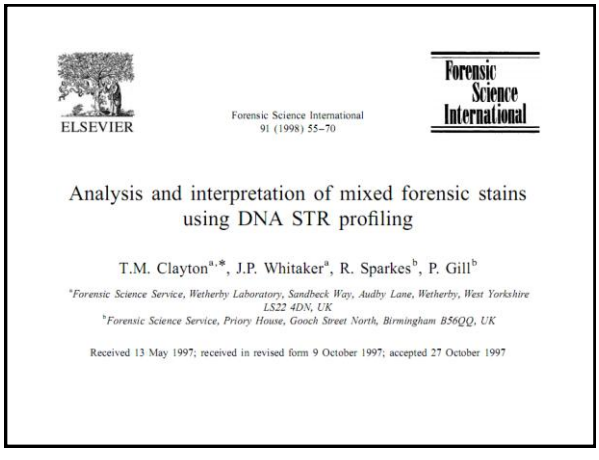
If Assume >2 Contributors…

23,23	20,?; 28,?; 22,?; ?,?

Summary

- Stutter can vary across profiles, loci, or alleles.
- Stutter becomes especially problematic for mixtures when samples are at low [DNA] levels.
- Labs should decide when is it appropriate to turn off stutter filters, especially when the minor component alleles are nearly the same height as stutter peaks.




ELSEVIER

Forensic Science International
91 (1998) 55–70

Forensic Science International

Analysis and interpretation of mixed forensic stains using DNA STR profiling

T.M. Clayton[a,*], J.P. Whitaker[a], R. Sparkes[b], P. Gill[b]

[a]Forensic Science Service, Wetherby Laboratory, Sandbeck Way, Audby Lane, Wetherby, West Yorkshire LS22 4DN, UK

[b]Forensic Science Service, Priory House, Gooch Street North, Birmingham B56QQ, UK

Received 13 May 1997; received in revised form 9 October 1997; accepted 27 October 1997

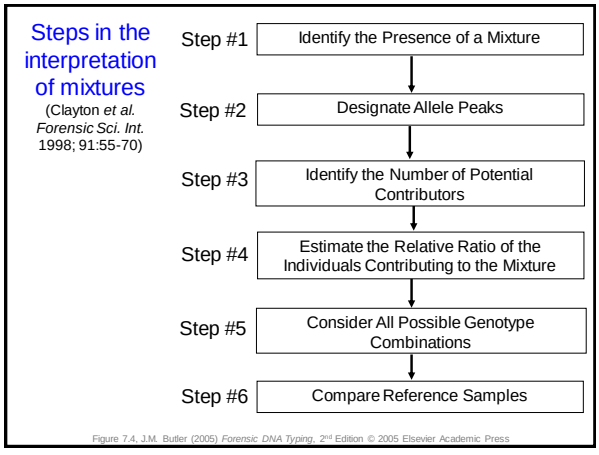


Figure 7.4, J.M. Butler (2005) *Forensic DNA Typing*, 2nd Edition © 2005 Elsevier Academic Press

Mixture Interpretation

- Criteria for mixture
- Criteria for determining number of contributors
- Criteria for classifying mixture
  - Distinguishable vs. indistinguishable
- Calculating mixture ratio and use
- Criteria for major/minor contribtors
- Determining genotypes

Minimum Number of Contributors

- Can be determined based on the locus that exhibits the greatest number of allelic peaks
- 2 loci have 4 alleles – maximum number alleles observed
- 2 = minimum number of contributors
- What is the ***true*** number of contributors?
  - Must make assumptions

Impact of Assumptions on Interpretation and Statistical Calculations

With assumptions for # of contributor:

- May be able to associate alleles into genotypes
- May be able to associate genotypes into single-source profiles
- Has an effect on the types of statistical calculations possible

Simulation Study Regarding Detecting the Number of Contributors to a Mixture

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

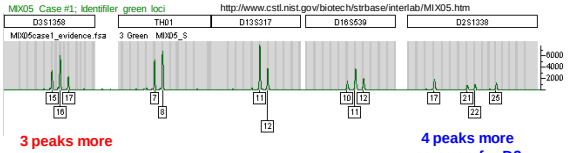
DNA evidence recovered from a scene or collected in relation to a case is generally declared as a mixture when more than two alleles are observed at several loci. However, in principle, all DNA profiles may be considered to be potentially mixtures, even those that show not more than two alleles at any 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 multiplexes. 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.

© 2006 Elsevier Ireland Ltd. All rights reserved.

Levels of Locus Heterozygosity Impact the Number of Alleles Observed in Mixtures

Loci	No. of alleles	Simulated 2-Person Mixture		
	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

Results from a 2-Person Mixture



3 peaks more common for D3

4 peaks more common for D2

Simulations with 2-person Mixtures

Table 1
The probability of observing a given number of alleles in a two-person mixtures for simulated profiles at the SGM+™ 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. (2007) Towards understanding the effect of uncertainty in the number of contributors to DNA stains. *FSI Genetics* 1:20-28

Simulations with 3-person Mixtures

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. (2007) Towards understanding the effect of uncertainty in the number of contributors to DNA stains. *FSI Genetics* 1:20-28

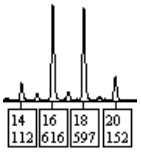
Simulations with 4-person Mixtures

Table 3
The probability of observing a given number of alleles in a four person mixtures for simulated profiles at the SGM^{+TM} loci

Loci	No. of alleles showing							
	1	2	3	4	5	6	7	8
D3	0.000	0.011	0.178	0.497	0.291	0.023	0.001	0.000
vWA	0.000	0.008	0.107	0.406	0.377	0.097	0.005	0.000
D16	0.000	0.027	0.240	0.458	0.238	0.036	0.001	0.000
D2	0.000	0.001	0.020	0.148	0.363	0.345	0.112	0.012
D8	0.000	0.009	0.103	0.340	0.377	0.151	0.019	0.001
D21	0.000	0.005	0.058	0.262	0.392	0.231	0.049	0.003
D18	0.000	0.000	0.023	0.166	0.382	0.321	0.101	0.008
D19	0.000	0.025	0.199	0.399	0.282	0.086	0.010	0.000
THO	0.000	0.020	0.222	0.501	0.241	0.016	0.000	0.000
FGA	0.000	0.001	0.034	0.215	0.398	0.281	0.068	0.004

Buckleton et al. (2007) Towards understanding the effect of uncertainty in the number of contributors to DNA stains. *FSI Genetics* 1:20-28

Determination of Genotypes (PHR)



D18S51

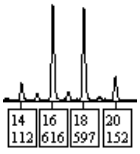
Possible Combinations

14, 16 and ~~18, 20~~
(18%) (25%)

14, 18 and ~~16, 20~~
(19%) (25%)

14, 20 and 16, 18
(74%) (97%)

Determination of Mixture Ratio



Major: 16,18
Minor: 14,20

D18S51

Total of all peak heights
= 112 + 616 + 597 + 152
= 1477 RFUs

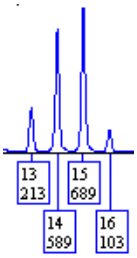
Minor component:
("14"+"20")/total = (112+152)/1477 = **0.179**

Major component:
("16"+"18")/ total = (616+597)/1477 = **0.821**

≈ 4.6 : 1

Four Peaks **(4 allele loci)**
heterozygote + heterozygote, no overlapping alleles (genotypes are unique)

Determination of Genotypes (PHR)



D8S1179

Possible Combinations

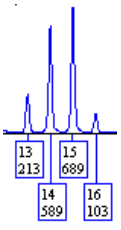
13, 14 (36%) and 15, 16 (15%)

13, 15 (31%) and 14, 16 (17%)

13, 16 (48%) and 14, 15 (85%)

Includes "stutter" from the 14 allele

Determination of Mixture Ratio



Major: 14,15
Minor: 13,16

D8S1179

Total of all peak heights
= 213 + 589 + 689 + 103
= 1594 RFUs

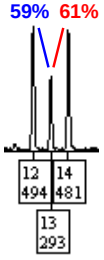
Minor component:
("13"+"16")/total = (213+103)/1594 = **0.198**

Major component:
("14"+"15")/ total = (589+689)/1594 = **0.802**

≈ 4 : 1

Four Peaks **(4 allele loci)**
heterozygote + heterozygote, no overlapping alleles (genotypes are unique)

Application of the Mixture Ratio



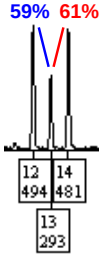
D19S433

Using peak height ratio, all genotypes possible:

12,12	12,13
13,13	12,14
14,14	13,14

Is there a major:minor here?

Application of the Mixture Ratio

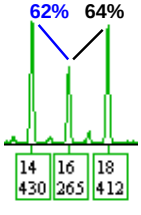


All possible genotype combinations:

12,12 + 13,14	1:1.6
13,13 + 12,14	1:3.3
14,14 + 12,13	1:1.6
12,13 + 12,14	1:1.4
~~12,13 + 13,14~~	~~1:1~~
12,14 + 13,14	1:1.4

Using MIXTURE RATIO calculations, can eliminate genotype pairs

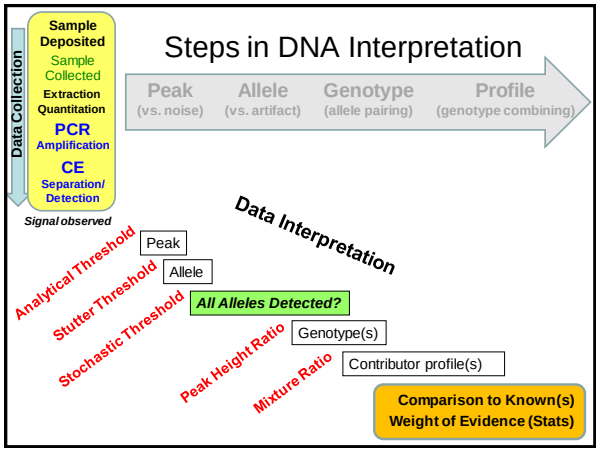
Application of the Mixture Ratio



All possible genotype combinations:

14,14 + 16,18	1:1.5
16,16 + 14,18	1:3
18,18 + 14,16	1:1.7
14,16 + 14,18	1:1.3
~~14,16 + 16,18~~	~~1:1~~
14,18 + 16,18	1.3:1

Using MIXTURE RATIO calculations, can eliminate genotype pairs



Thank you for your attention

Acknowledgments: NIJ & FBI Funding

Contact Information

Becky Hill
Research Biologist
becky.hill@nist.gov
301-975-4275

Mike Coble
Research Biologist
michael.coble@nist.gov
301-975-4330

Our team publications and presentations are available at:
http://www.cstl.nist.gov/biotech/strbase/NISTpub.htm