Topics and Techniques for Forensic DNA Analysis

Mixture Examples

Houston DNA
Training Workshop

Houston, TX
April 3-4, 2007

Dr. John M. Butler
National Institute of Standards and Technology

john.butler@nist.gov

Steps in the interpretation of mixtures

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

- Step #1: Identify the Presence of a Mixture
- Step #2: Designate Allele Peaks
- Step #3: Identify the Number of Potential Contributors
- Step #4: Estimate the Relative Ratio of the Individuals Contributing to the Mixture
- Step #5: Consider All Possible Genotype Combinations
- Step #6: Compare Reference Samples

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

Step #1: Is a Mixture Present in an Evidentiary Sample?

- Examine the number of peaks present in a locus
  - More than 2 peaks at a locus (except for tri-allelic patterns at perhaps one of the loci examined)
- Examine relative peak heights
  - Heterozygote peak imbalance <60%
  - Peak at stutter position >15%
- Consider all loci tested

Step #2: Designate Allele Peaks

- Use regular data interpretation rules to decipher between true alleles and artifacts
- Use stutter filters to eliminate stutter products from consideration (although stutter may hide some of minor component alleles at some loci)
- Consider heterozygote peak heights that are highly imbalanced (<60%) as possibly coming from two different contributors

Step #3: Identifying the Potential Number of Contributors

- **Important for some statistical calculations**
- Typically if 2, 3, or 4 alleles then 2 contributors
- If 5 or 6 alleles per locus then 3 contributors
- If >6 alleles in a single locus, then >4 contributors
- JFS Nov 2005 paper by Forensic Bioinformatics on number of possible contributors
  - Relies on maximum allele count alone
  - Does not take into account peak height information

Forensic Bioinformatics Article

http://www.bioforensics.com/articles/empirical_mixtures.pdf

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

TABLE 2—*Count and percent of three-person mixtures in which a particular number of unique alleles was the maximum observed across all loci, both for the original and randomized individuals*[a].

Unique Alleles	Count	Percent (%)
2	0	0.00%
3	78	0.00%
4	4,967,034	3.39%
5	93,037,010	63.49%
6	48,532,037	33.12%

Using 959 complete 13-locus STR profiles from FBI dataset

146,536,159 possible combinations **with 3-person mixtures**

3.39 % (4,967,034 combinations) would only show a maximum of four alleles (i.e., appear based on maximum allele count alone to be a 2-person mixture)

Recent Article by Buckleton *et al.*

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

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.

Two-Person Mixtures for Simulated Profiles:

Probability by Locus of A Particular Number of Alleles Being Observed

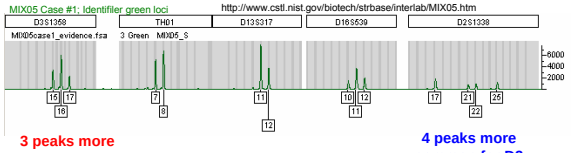
Table 1
The probability of observing a given number of alleles in a two-person mixtures for simulated profiles at the SGM^{+TM} loci

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

Levels of Locus Heterozygosity Impact Number of Alleles Observed in Mixtures

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



3 peaks more common for D3

4 peaks more common for D2

Three-Person Mixtures for Simulated Profiles:

Probability by Locus of A Particular Number of Alleles Being Observed

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

Number of Alleles Observed with Simulated **Four-Person Mixtures**

- The simulation of four person mixtures suggests that 0.014% of four person mixtures would show four or fewer alleles and that 66% would show six or fewer alleles for the SGM Plus loci.
- The results for the Profiler Plus loci were 0.6% and 75%.
- The equivalent values for the CODIS set from Paoletti et al. were 0.02% showing four or fewer and 76.35% showing six or fewer.

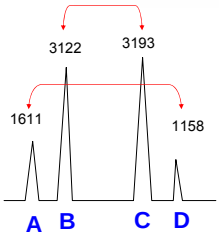
Buckleton et al. (2007) Towards understanding the effect of uncertainty in the number of contributors to DNA stains. *FSI Genetics* 1:20-28

Step #4: Estimation of Relative Ratios for Major and Minor Components to a Mixture

- Mixture studies with known samples have shown that the mixture ratio between loci is fairly well preserved during PCR amplification
- Thus it is generally thought that the peak heights (areas) of alleles present in an electropherogram can be related back to the initial component concentrations
- Start with loci possessing 4 alleles…

Estimating Mixture Proportion (M_x) or Mixture Ratio (M_r)

$$M_x = \frac{\phi_a + \phi_b}{\phi_a + \phi_b + \phi_c + \phi_d}$$



$$\frac{A + D}{A + D + B + C}$$

$$\frac{1611 + 1158}{1611 + 1158 + 3122 + 3193}$$

= 2769/9084 = 0.305

Step #5: Consider All Possible Genotype Combinations

Table 3
Pairwise combinations of two, three and four alleles

Four alleles (a,b,c,d)		Three alleles (a,b,c)		Two alleles (a,b)	
a,b	c,d	a,a	b,c	a,a	a,b
a,c	b,d	b,b	a,c	a,b	a,b
a,d	b,c	c,c	a,b	a,a	b,b
c,d	**a,b**	a,b	a,c	a,b	b,b
b,d	**a,c**	b,c	a,c	**a,b**	**a,a**
b,c	**a,d**	a,b	b,c	**b,b**	**a,a**
		b,c	**a,a**	**b,b**	**a,b**
		a,c	**b,b**		
		a,b	**c,c**		
		a,c	**a,b**		
		a,c	**b,c**		
		b,c	**a,b**		

Key: bold entries represent reciprocal combinations.

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

Step #5: Consider All Possible Genotype Combinations

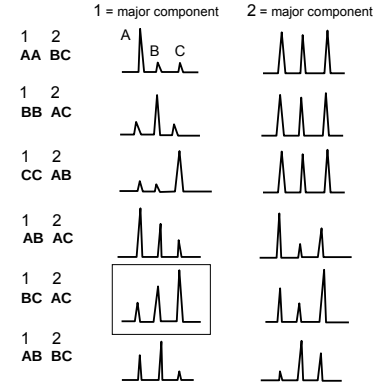


Figure 7.7, J.M. Butler (2005) *Forensic DNA Typing*, 2nd Edition © 2005 Elsevier Academic Press

Possible Genotype Combinations for a two person mixture

Four Peaks
- heterozygote + heterozygote, no overlapping alleles (genotypes are unique)

Three Peaks
- heterozygote + heterozygote, one overlapping allele
- heterozygote + homozygote, no overlapping alleles (genotypes are unique)

Two Peaks
- heterozygote + heterozygote, two overlapping alleles (genotypes are identical)
- heterozygote + homozygote, one overlapping allele
- homozygote + homozygote, no overlapping alleles (genotypes are unique)

Single Peak
- homozygote + homozygote, overlapping allele (genotypes are identical)

Genotype Combinations and Allele Overlap Observed in One Forensic Lab

Table 4
Resolution of genotypes in mixed samples depending on knowledge of the contributors' profile (no. of observations = 210)

	Total (%)
Four peaks	
Heterozygote + heterozygote (non-overlapping alleles)	28 (13)
Three peaks	
Heterozygote + heterozygote (one overlapping allele)	72 (34)
Heterozygote + homozygote (non-overlapping alleles)	33 (16)
Two peaks	
Heterozygote + homozygote (one overlapping allele)	39 (18)
Heterozygote + heterozygote (two overlapping alleles)	15 (7)
Homozygote + homozygote (non-overlapping allele)	7 (3)
One peak	
Homozygote + homozygote (overlapping allele)	5 (2)
Allele dropouts	11 (5)

Torres *et al.* (2003) DNA mixtures in forensic casework: a 4-year retrospective study. *Forensic Sci. Int.* 134:180-186

ISFG (2006) Table 2

Table 2
Assessment of major (*ab*)/minor (*cd*) genotypes of a mixture of two contributors relative to $\hat{M}_x$ and H_b calculated using ϕ_a = 1200 rfu, ϕ_b = 100 rfu, ϕ_c = 400 rfu, ϕ_d = 380 rfu, where rfu is relative fluorescence units (allele peak height)

Genotypes		M_x|major, minor genotypes	Heterozygous balance		Comment
Major	Minor		$H_{b\ major}$	$H_{b\ minor}$	
ab	*cd*	0.70	0.9	0.9	Passes H_b, $\hat{M}_x$
ac	*bd*	0.53	0.3	0.3	Fails H_b
ad	*bc*	0.51	0.3	0.3	Fails H_b
cd	*ab*	0.30	0.9	0.9	Fails $\hat{M}_x$
bd	*ac*	0.48	0.3	0.3	Fails H_b
bc	*ad*	0.49	0.3	0.3	Fails H_b

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

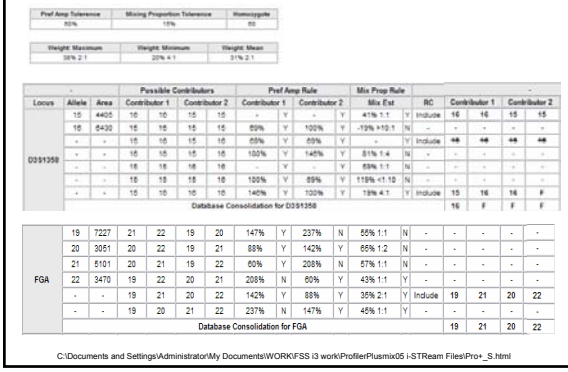
The Defense Hypothesis will include all possible combinations

Table 1
Evaluation of $\Pr(E|H_d)$; two person mixture with four discrete alleles present

Individual 1	Individual 2	Genotype probability
ab	*cd*	$4p_ap_bp_cp_d$
ac	*bd*	$4p_ap_bp_cp_d$
ad	*bc*	$4p_ap_bp_cp_d$
cd	*ab*	$4p_ap_bp_cp_d$
bd	*ac*	$4p_ap_bp_cp_d$
bc	*ad*	$4p_ap_bp_cp_d$
Sum		$24p_ap_bp_cp_d$

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

FSS-i3 iSTReam Mixture Interpretation Tool



C:\Documents and Settings\Administrator\My Documents\WORK\FSS i3 work\ProfilerPlusmix05 i-STReam Files\Pro+_S.html

Step #6: Compare Reference Samples

- If there is a suspect, a laboratory must ultimately decide to include or exclude him…
- If no suspect is available for comparison, does your laboratory still work the case? (Isn't this a primary purpose of the national DNA database?)
- Victim samples can be helpful to eliminate their allele contributions to intimate evidentiary samples and thus help deduce the perpetrator

- Tom Overson's USACIL program permits reference sample comparisons

Kit Comparability on the Same Mixture Samples

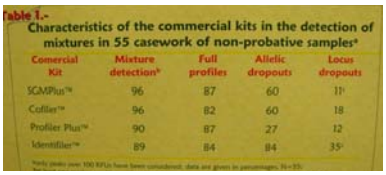


Table 1.- Characteristics of the commercial kits in the detection of mixtures in 55 casework of non-probative samples[a]

Comercial Kit	Mixture detection[b]	Full profiles	Allelic dropouts	Locus dropouts
SGMPlus™	96	87	60	11[c]
Cofiler™	96	82	60	18
Profiler Plus™	90	87	27	12
Identifiler™	89	84	84	35[c]

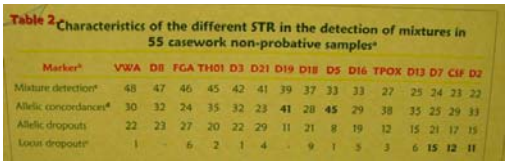


Table 2. Characteristics of the different STR in the detection of mixtures in 55 casework non-probative samples[a]

Marker[b]	VWA	D8	FGA	TH01	D3	D21	D19	D18	D5	D16	TPOX	D13	D7	CSF	D2
Mixture detection[a]	48	47	46	45	42	41	39	37	33	33	27	25	24	23	22
Allelic concordances[d]	30	32	24	35	32	23	**41**	28	**45**	29	38	35	25	29	33
Allelic dropouts	22	23	27	20	22	29	11	21	8	19	12	15	21	17	15
Locus dropouts[c]	1	-	6	2	1	4	-	9	1	5	3	6	**15**	**12**	**11**

Y. Torres and P. Sanz, Poster P-290 at ISFG Sept 2005 meeting (Ponta Delgada, Azores)

ISFG DNA Commission on Mixture Interpretation

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
ELSEVIER
Forensic Science International 160 (2006) 90–101
Forensic Science International

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]

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

ISFG (2006) Recommendations

- Recommendation 1: **The likelihood ratio is the preferred approach to mixture interpretation.** The RMNE (probability of exclusion) approach is restricted to DNA profiles where the profiles are unambiguous. If the DNA crime stain profile is low level and some minor alleles are the same size as stutters of major alleles, and/or if drop-out is possible, then the RMNE method may not be conservative.

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

ISFG (2006) Recommendations

- Recommendation 2: Even if the legal system does not implicitly appear to support the use of the likelihood ratio, it is recommended that the scientist is trained in the methodology and routinely uses it in case notes, advising the court in the preferred method before reporting the evidence in line with the court requirements. The scientific community has a responsibility to support improvement of standards of scientific reasoning in the court-room.

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

ISFG (2006) Recommendations

- Recommendation 3: The methods to calculate likelihood ratios of mixtures (not considering peak area) described by Evett *et al.* (*J. Forensic Sci. Soc.* 1991;31:41-47) and Weir *et al.* (*J. Forensic Sci.* 1997;42:213-222) are recommended.

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

ISFG (2006) Recommendations

- Recommendation 4: If peak height or area information is used to eliminate various genotypes from the unrestricted combinatorial method, this can be carried out by following a sequence of guidelines based on Clayton *et al.* (*Forensic Sci. Int.* 1998;91:55-70).

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

ISFG (2006) Recommendations

- Recommendation 5: The probability of the evidence under H_p is the province of the prosecution and the probability of the evidence under H_d is the province of the defense. The prosecution and defense both seek to maximize their respective probabilities of the evidence profile. To do this both H_p and H_d require propositions. There is no reason why multiple pairs of propositions may not be evaluated (Appendix C).

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

ISFG (2006) Recommendations

- Recommendation 6: If the crime profile is a major/minor mixture, where minor alleles are the same size (height or area) as stutters of major alleles, then stutters and minor alleles are indistinguishable. Under these circumstances alleles in stutter positions that do not support H_p should be included in the assessment.
- In general, stutter percentage is <15%

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

ISFG (2006) Recommendations

- Recommendation 7: If drop-out of an allele is required to explain the evidence under Hp: (S = ab; E = a), then the allele should be small enough (height/area) to justify this. Conversely, if a full crime stain profile is obtained where alleles are well above the background level, and the probability of drop-out approaches Pr(D) ≈ 0, then Hp is not supported.

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

ISFG (2006) Recommendations

- Recommendation 8: If the alleles of certain loci in the DNA profile are at a level that is dominated by background noise, then a biostatistical interpretation for these alleles should not be attempted.

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

ISFG (2006) Recommendations

- Recommendation 9: In relation to low copy number, stochastic effects limit the usefulness of heterozygous balance and mixture proportion estimates. In addition, allelic drop-out and allelic drop-in (contamination) should be taken into consideration of any assessment.

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