2012 Mixture Interpretation Workshop:

Mixtures **U**sing *SOUND* **S**tatistics, **I**nterpretation, & **C**onclusions

Probabilistic Genotyping

Michael D. Coble

October 15, 2012
Nashville, TN

BOSTON UNIVERSITY

NIST

Next Issue of FSI-Genetics

Forensic Science International: Genetics xxx (2012) xxx–xxx

Contents lists available at SciVerse ScienceDirect

Forensic Science International: Genetics

journal homepage: www.elsevier.com/locate/fsig

ELSEVIER

Editorial

Focus issue—Analysis and biostatistical interpretation of complex and low template DNA samples

Article in press…

ELSEVIER

Contents lists available at SciVerse ScienceDirect

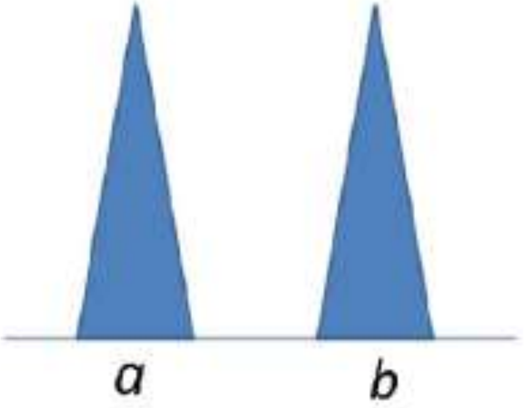
Forensic Science International: Genetics

journal homepage: www.elsevier.com/locate/fsig

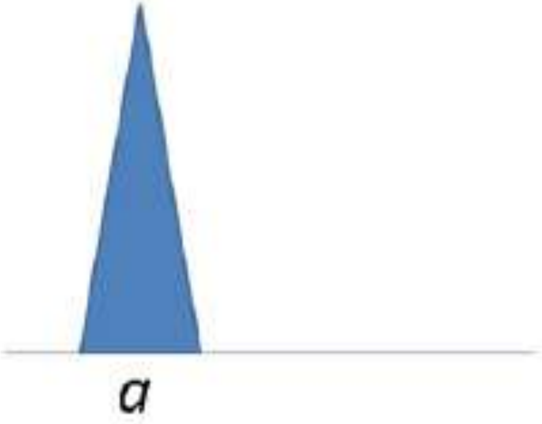
DNA commission of the International Society of Forensic Genetics: Recommendations on the evaluation of STR typing results that may include drop-out and/or drop-in using probabilistic methods

P. Gill[a,b,*], L. Gusmão[c], H. Haned[d], W.R. Mayr[e], N. Morling[f], W. Parson[g], L. Prieto[h], M. Prinz[i], H. Schneider[j], P.M. Schneider[k], B.S. Weir[l]

ISFG Recommendations



Reference profile (S)



Crime stain profile (E)

Pr(D) = Prob. Drop-out (het)

Pr($\overline{D}$) = No Prob. Drop-out (het)

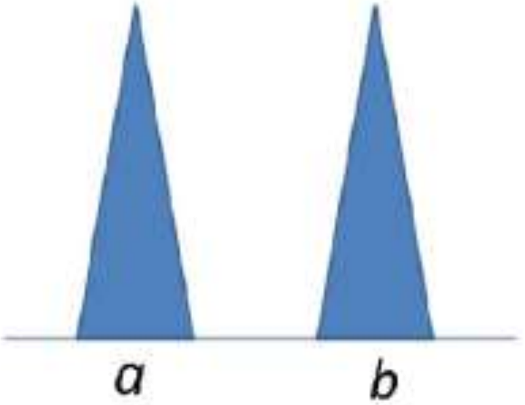
Pr(D_2) = Prob. Drop-out (hom)

Pr($\overline{D_2}$) = No Prob. Drop-out (hom)

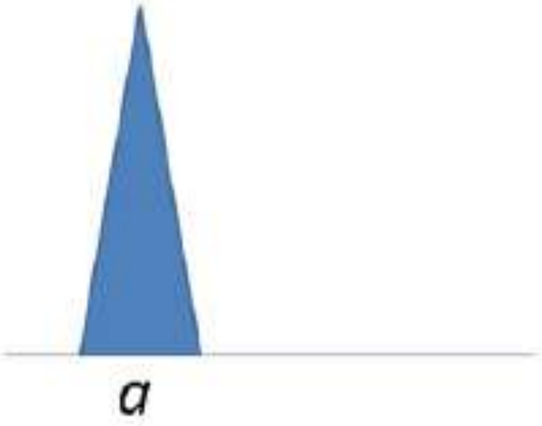
Pr(C) = Prob. Drop-in

Pr($\overline{C}$) = No Prob. Drop-in

Prosecutor’s Explanation



Reference profile (S)



Crime stain profile (E)

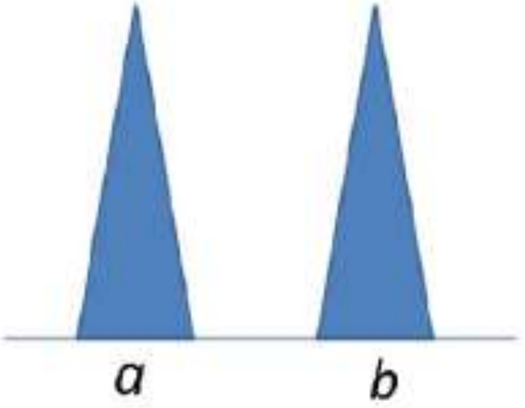
No Drop-out of the “A” allele
The “B” allele dropped out
No other Drop-in

$$\Pr(\overline{D})\ \Pr(D)\ \Pr(\overline{C})$$

The LR

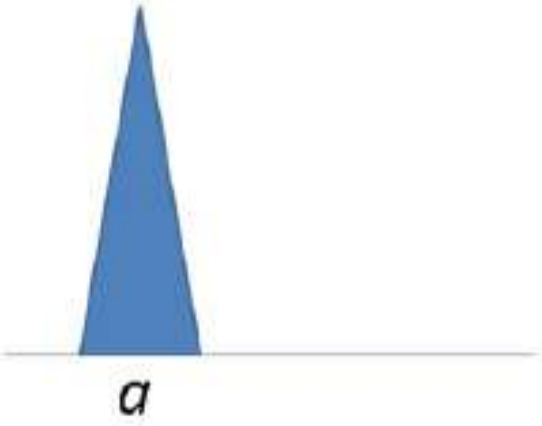
$$\mathrm{LR} = \frac{\Pr(\overline{D})\ \Pr(D)\ \Pr(\overline{C})}{\quad}$$

Defense Explanation



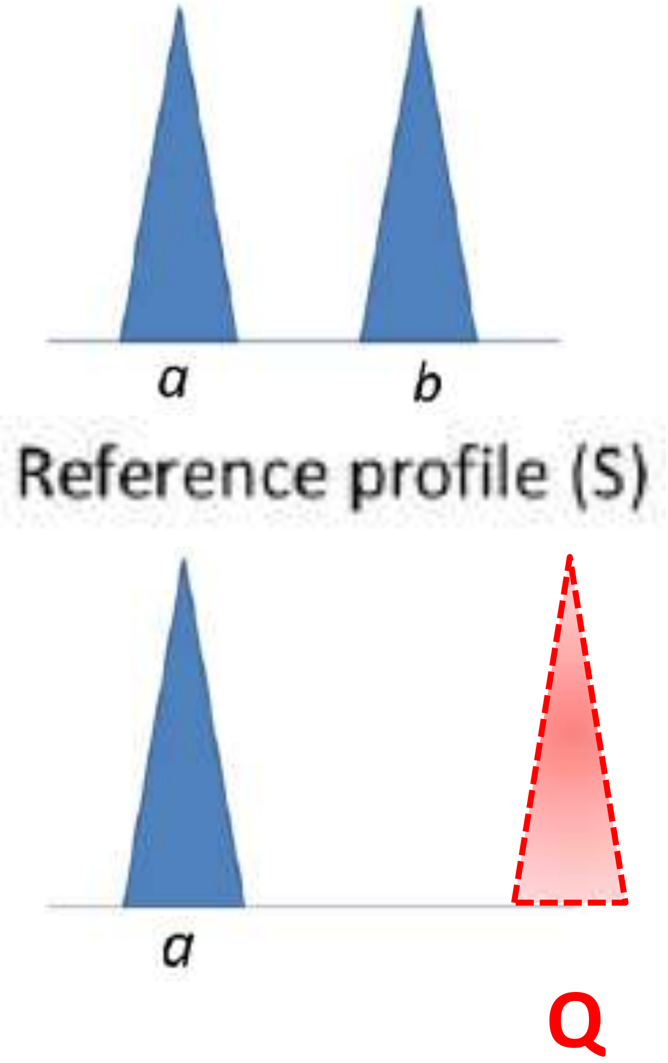
4 possibilities

(1) The real culprit is a homozygote



$$p_a^{\,2}\mathrm{Pr}(\overline{D}_2)\ \mathrm{Pr}(\overline{C})$$

Defense Explanation

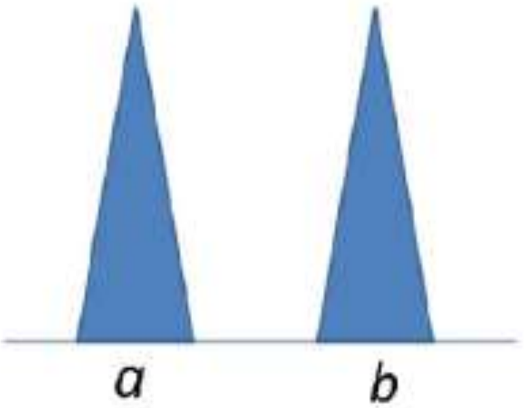


4 possibilities

(2) Drop out of a heterozygote (not B)
No drop-in of “A”

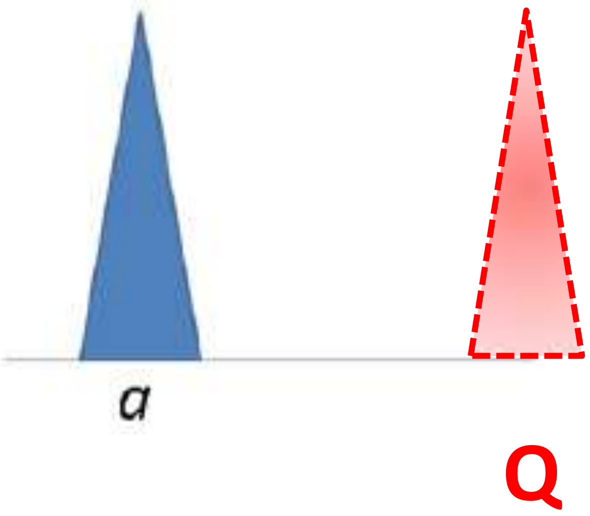
$$2p_a p_Q \Pr(\overline{D})\Pr(D)\Pr(\overline{C})$$

Defense Explanation



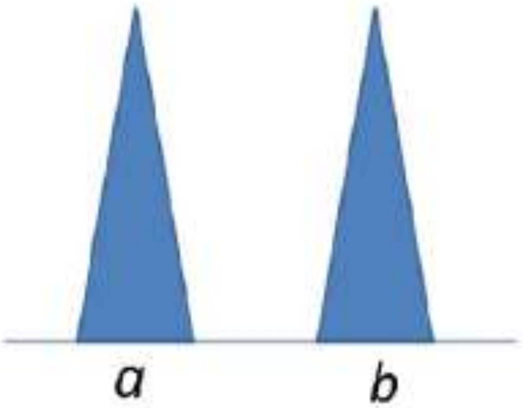
4 possibilities

(3) Drop out of a homozygote (not B)
Drop in of “A”

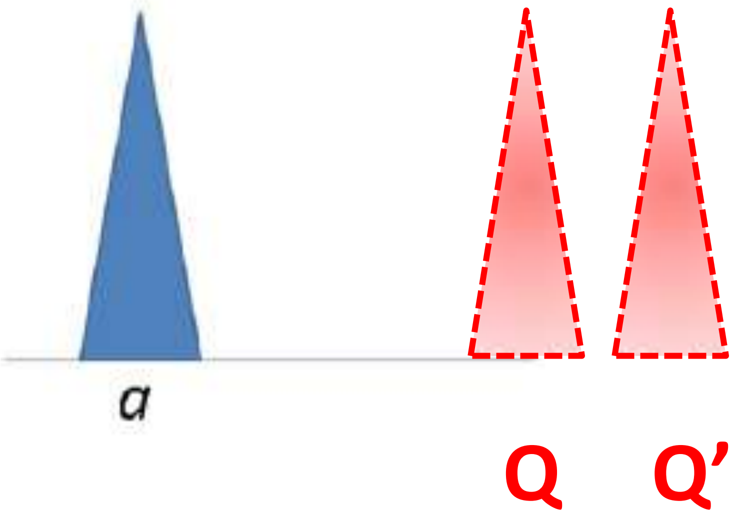


$$p_Q{}^2 Pr(D_2)\ Pr(C) p_a$$

Defense Explanation



Reference profile (S)



4 possibilities

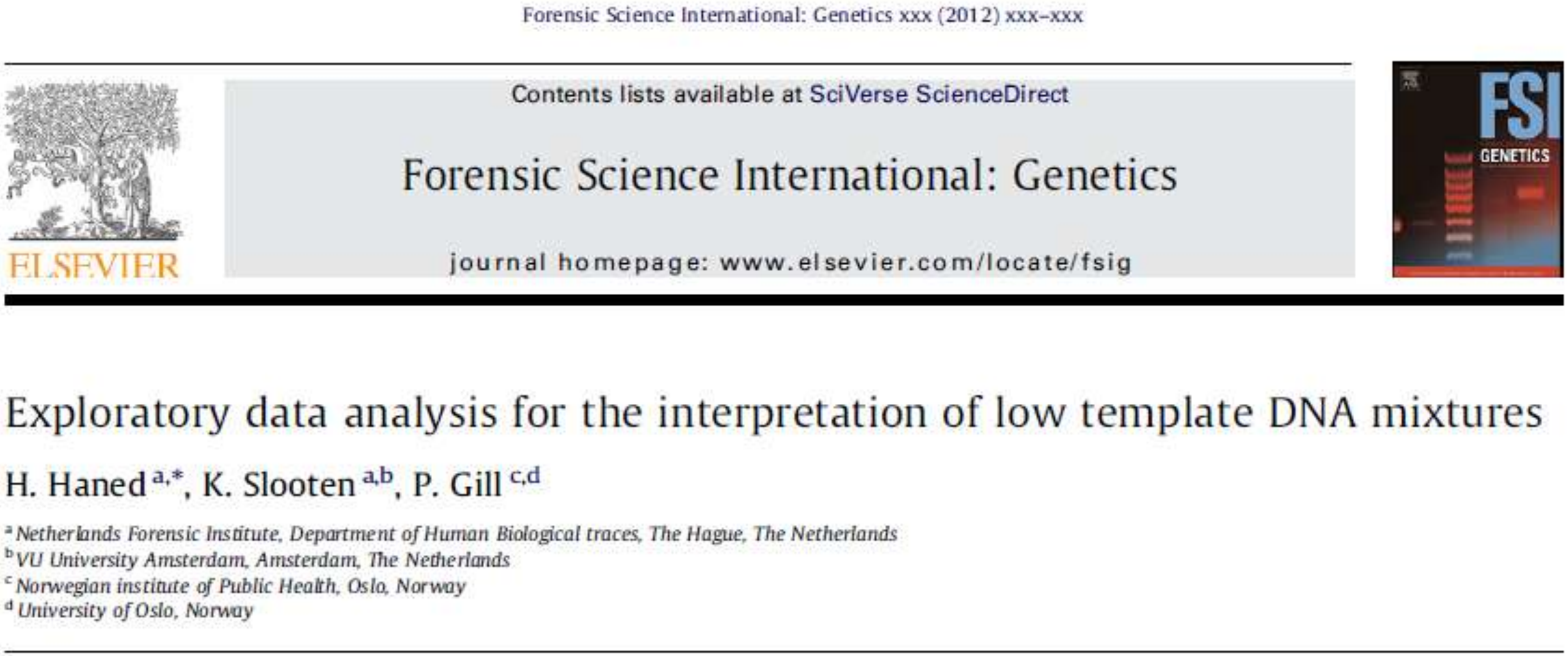
(4) Drop out of a homozygote (not AB)
Drop in of “A”

$$2p_Q p_{Q'} \Pr(D)^2 \Pr(C) p_a$$

The LR

$$LR = \frac{\Pr(\overline{D})\ \Pr(D)\ \Pr(\overline{C})}{\begin{array}{l} p_a{}^2\Pr(\overline{D}_2)\ \Pr(\overline{C}) \quad + \\ 2p_a p_Q \Pr(\overline{D})\Pr(D)\Pr(\overline{C}) \quad + \\ p_Q{}^2\Pr(D_2)\ \Pr(C)p_a \quad + \\ 2p_Q p_{Q'}\Pr(D)^2\ \Pr(C)p_a \end{array}}$$

Haned *et al.*



Forensic Science International: Genetics xxx (2012) xxx–xxx

Contents lists available at SciVerse ScienceDirect

Forensic Science International: Genetics

ELSEVIER

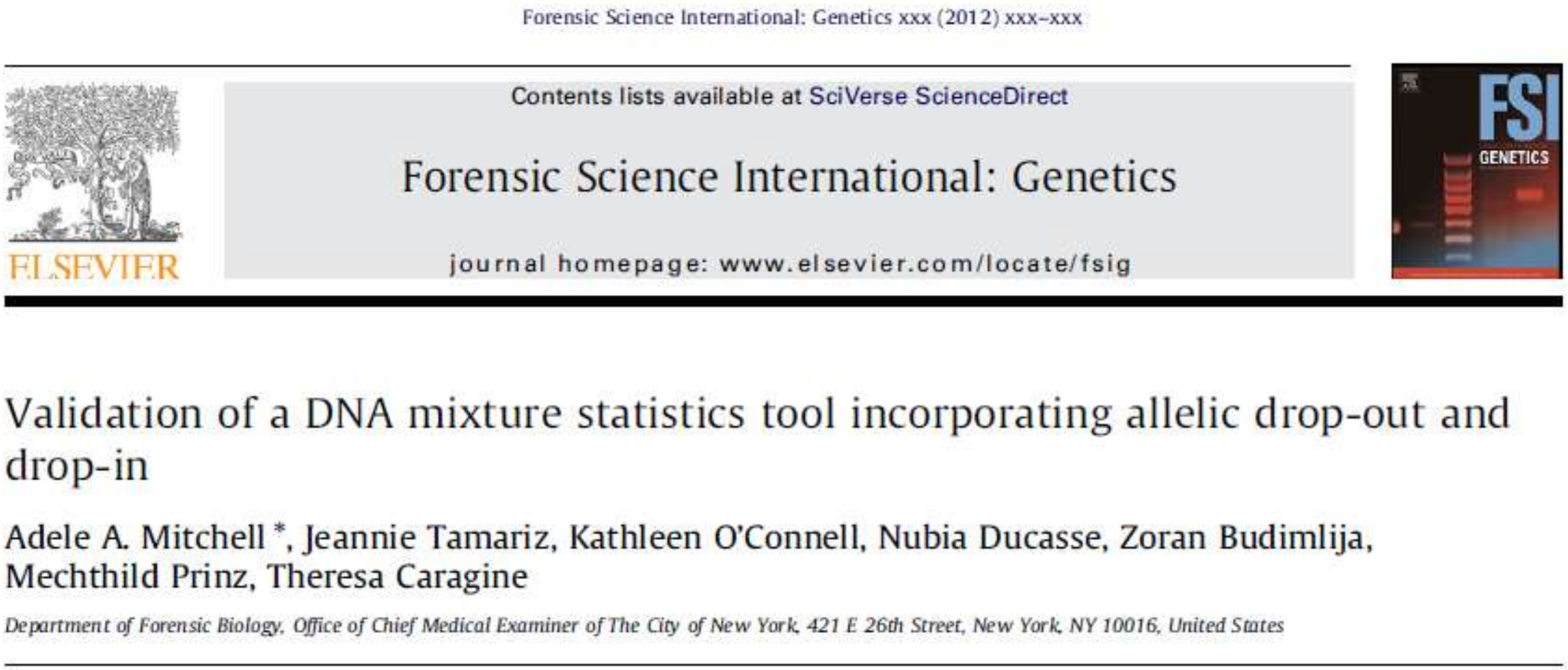
journal homepage: www.elsevier.com/locate/fsig

Exploratory data analysis for the interpretation of low template DNA mixtures

H. Haned [a,*], K. Slooten [a,b], P. Gill [c,d]

[a] *Netherlands Forensic Institute, Department of Human Biological traces, The Hague, The Netherlands*
[b] *VU University Amsterdam, Amsterdam, The Netherlands*
[c] *Norwegian institute of Public Health, Oslo, Norway*
[d] *University of Oslo, Norway*

Mitchell *et al.*



Forensic Science International: Genetics xxx (2012) xxx–xxx

Contents lists available at SciVerse ScienceDirect

Forensic Science International: Genetics

journal homepage: www.elsevier.com/locate/fsig

Validation of a DNA mixture statistics tool incorporating allelic drop-out and drop-in

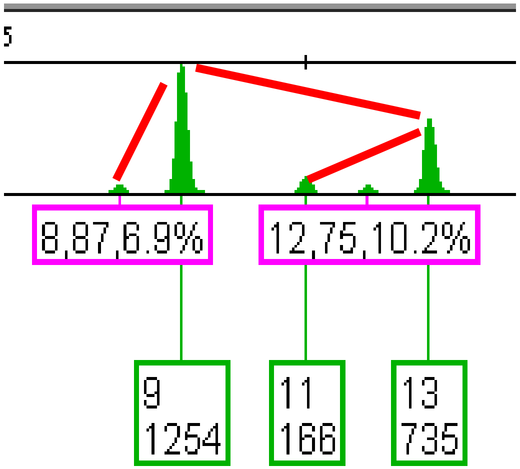
Adele A. Mitchell *, Jeannie Tamariz, Kathleen O'Connell, Nubia Ducasse, Zoran Budimlija, Mechthild Prinz, Theresa Caragine

Department of Forensic Biology, Office of Chief Medical Examiner of The City of New York, 421 E 26th Street, New York, NY 10016, United States

Probabilistic Modeling of TA

Mathematical Modeling of the Data

PHR, Mix Ratio, Stutter etc…



50-100,000 Simulations

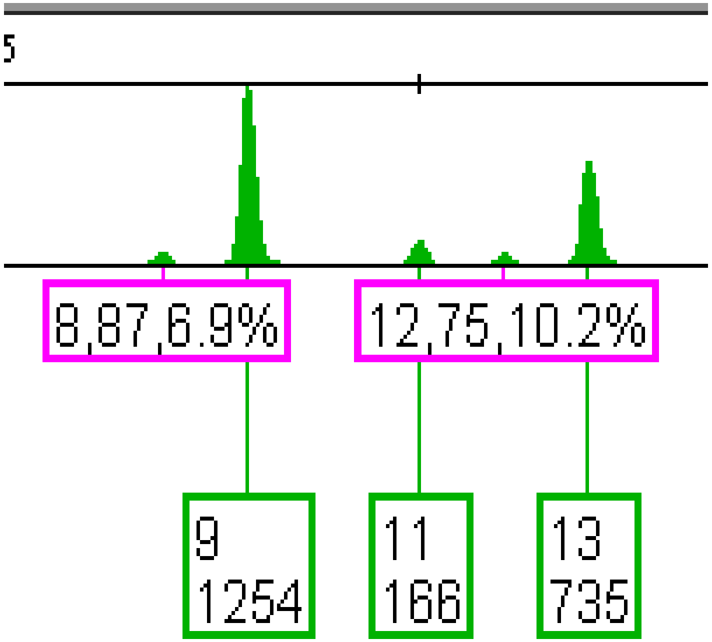
→

(MCMC)

Probable ***Genotypes*** to explain the mixture

Genotypes	Probability
9,11	76%
11,11	15%
11,13	2%
8,11	2%
11,12	2%
9,9	1%
9,12	<1%
10,11	<1%
8,12	<1%
8,9	<1%

Uncertainty with D16S539



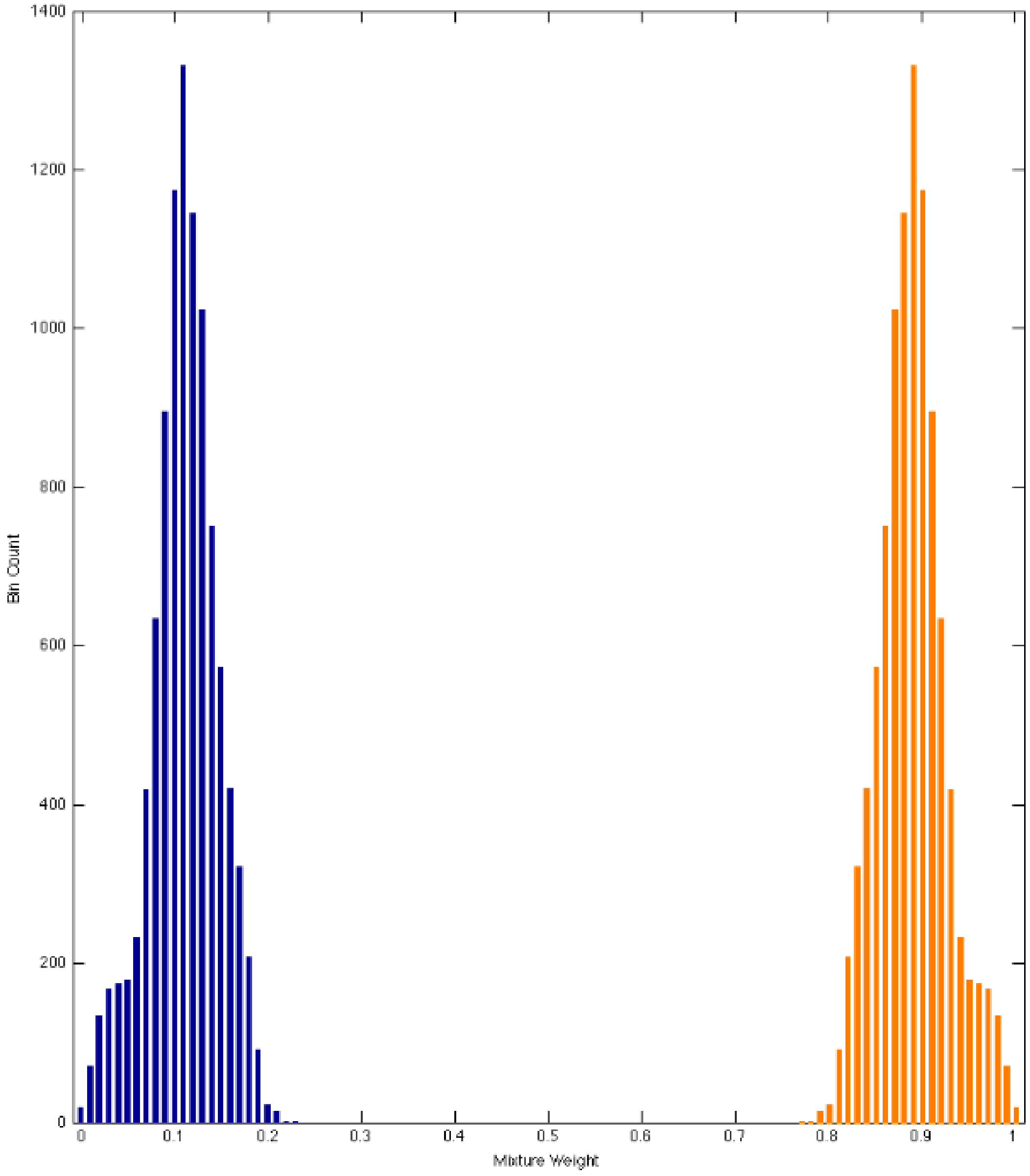
The 11 allele is at 169 RFU (above 150 ST)

The “12” peak in the stutter position is only slightly below our stutter threshold of 10.4%

If we assume 8 and 12 are stutter peaks, then the possible genotypes of the minor contributor are - **9,11 11,11 11,13**

Should we also include the 8 and 12 alleles in Creating our genotype combinations?

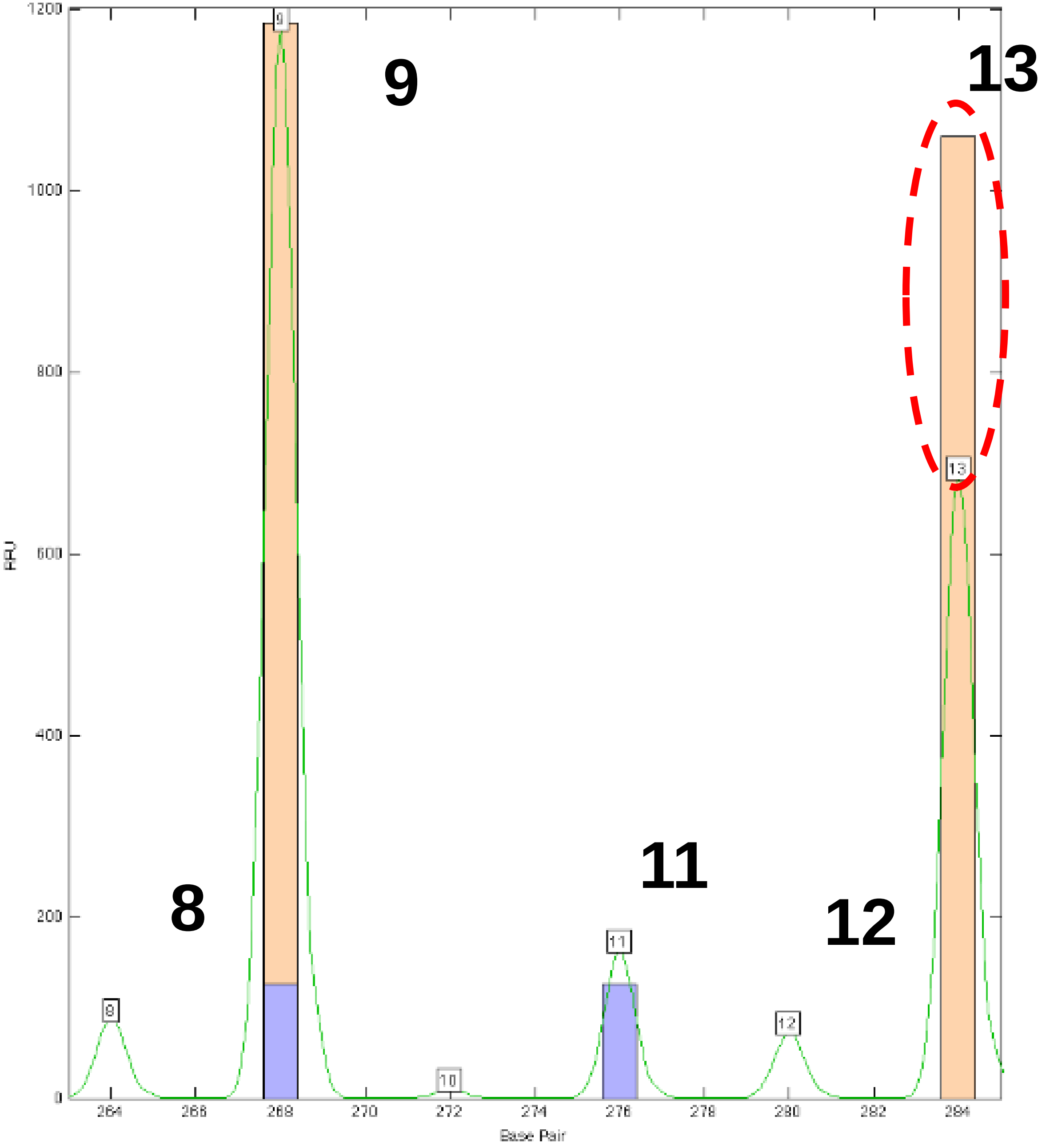
Summary – Mixture Weight



100,000 MCMC examinations of the data.

2 unknowns (no conditioning)

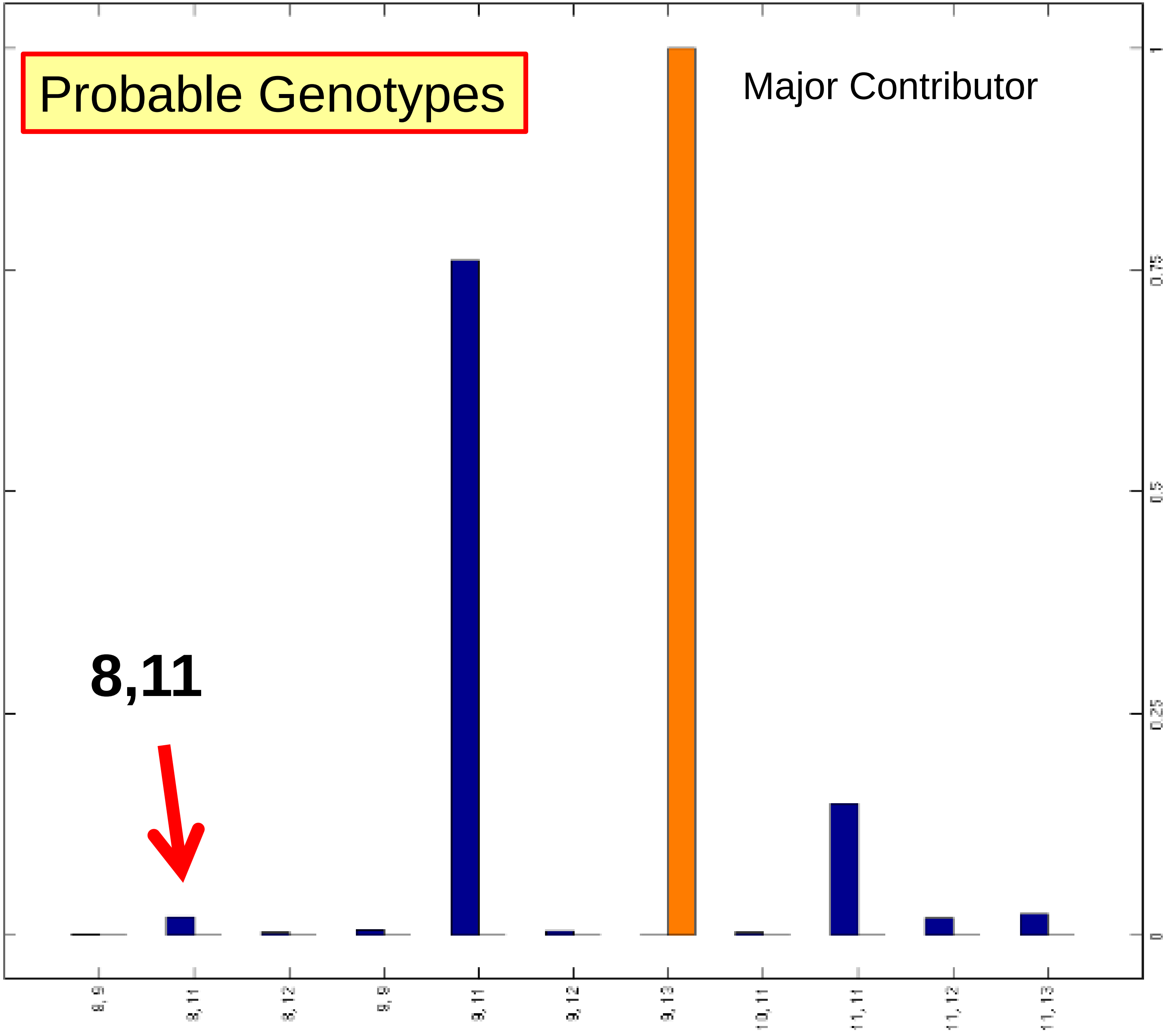
Clear separation of the two contributors.



Model doesn’t exactly fit the data

Most of the time (76%), 9,11 is predicted to be the genotype of the minor contributor

Probable Genotypes



Major Contributor

Genotypes	Probability
9,11	76%
11,11	15%
11,13	2%
8,11	2%
11,12	2%
9,9	1%
9,12	<1%
10,11	<1%
8,12	<1%
8,9	<1%

Determining the LR for D16S539 (H_P)

Genotypes	Probability (Before Conditioning)	Genotype Freq (HWE)	(Prob) x (HWE)
9,11	0.431	0.0719	0.031
11,11	0.098	0.1025	0.01
11,13	0.014	0.093	0.0013
8,11	0.092	0.0106	0.001
11,12	0.008	0.2093	0.0016
9,9	0.013	0.0126	0.0002
9,12	0.003	0.0734	0.0002
10,11	0.003	0.036	0.0001
8,12	0.014	0.0108	0.0002
8,9	0.015	0.0037	0.0001
		(sum)	0.046

$$LR = \frac{0.092}{}$$

Suspect = 8,11

Determining the LR for D16S539 (H_D)

Genotypes	Probability (Before Conditioning)	Genotype Freq (HWE)	(Prob) x (HWE)
9,11	0.431	0.0719	0.031
11,11	0.098	0.1025	0.01
11,13	0.014	0.093	0.0013
8,11	0.092	0.0106	0.001
11,12	0.008	0.2093	0.0016
9,9	0.013	0.0126	0.0002
9,12	0.003	0.0734	0.0002
10,11	0.003	0.036	0.0001
8,12	0.014	0.0108	0.0002
8,9	0.015	0.0037	0.0001
		(sum)	0.046

$$LR = \frac{0.092}{0.046}$$

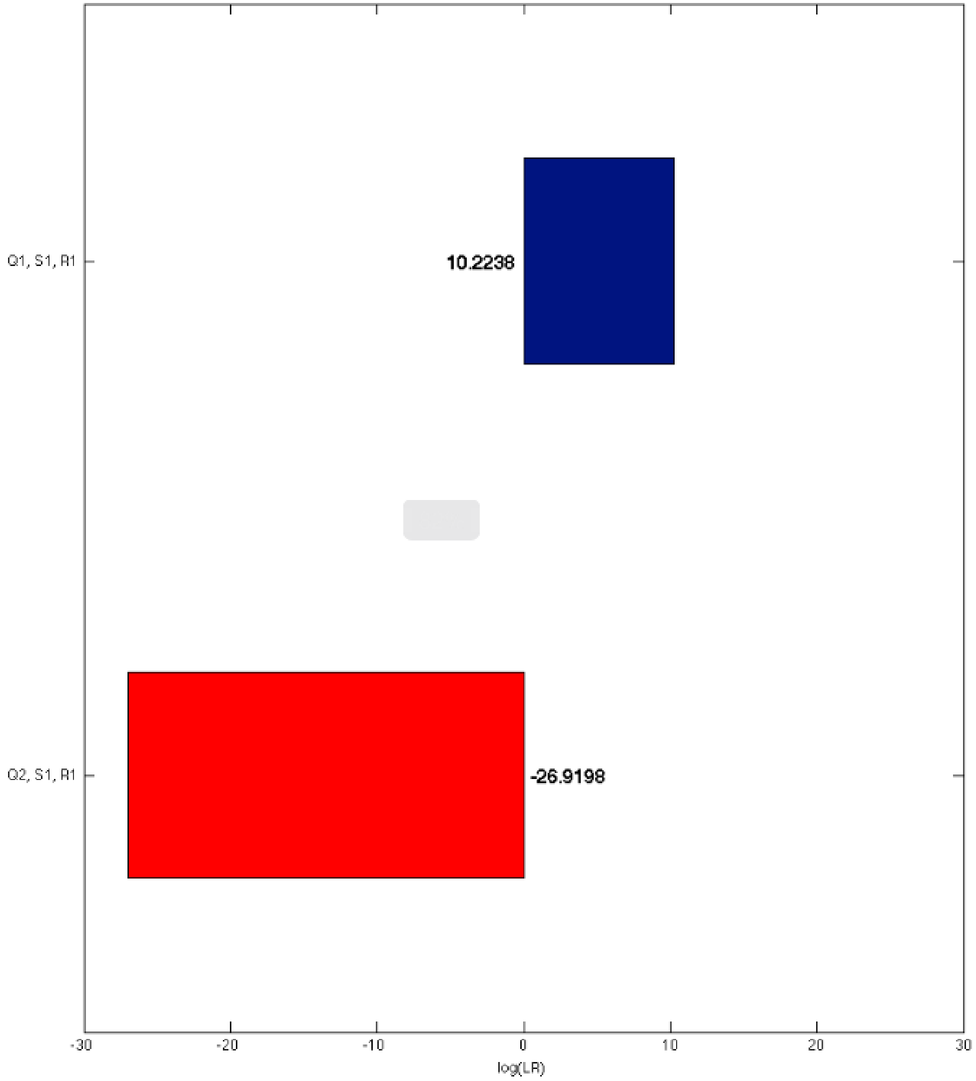
$$= 2.0$$

Suspect = 8,11

D16S539 Results

	LR	
Assume Stutter @8,12	3.6	(fails to capture 8,11)
Include 8,12	2.3	
True Allele	2.0	

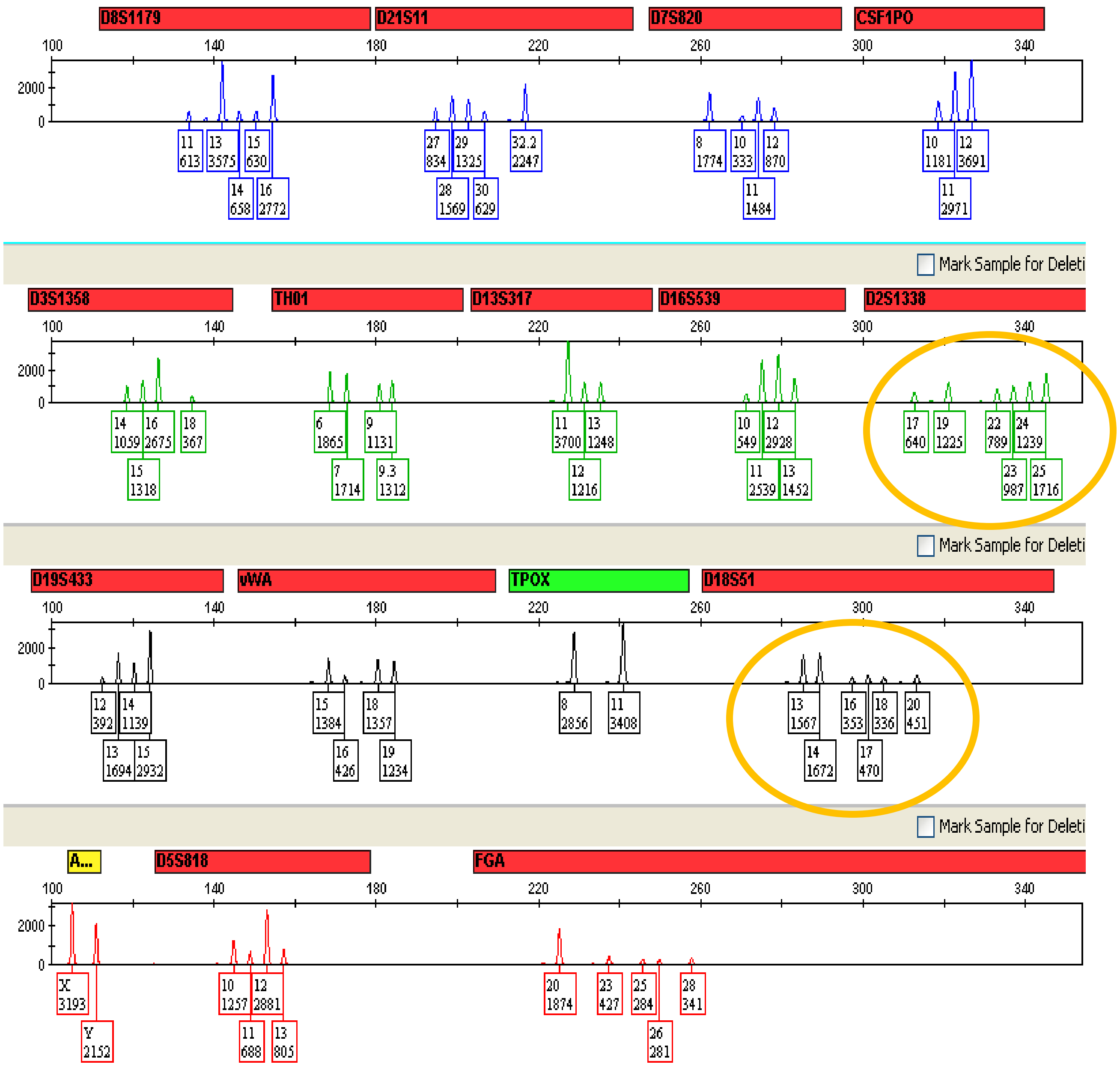
D16S539 Results



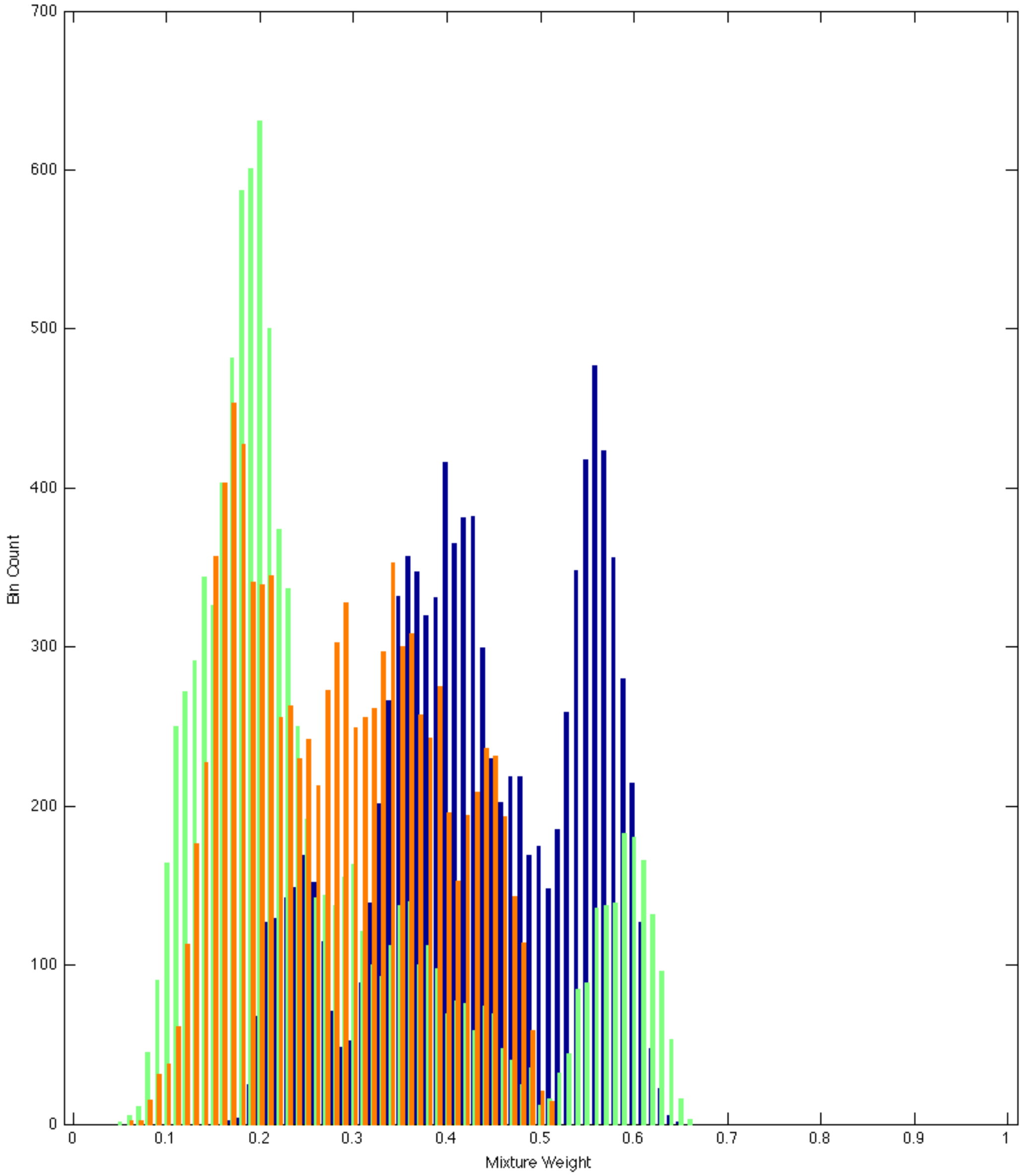
Joint LR = 16.7 Billion
(using True Allele, 2unk)

Using “2P” = 26.5 Trillion

Complex Mixture



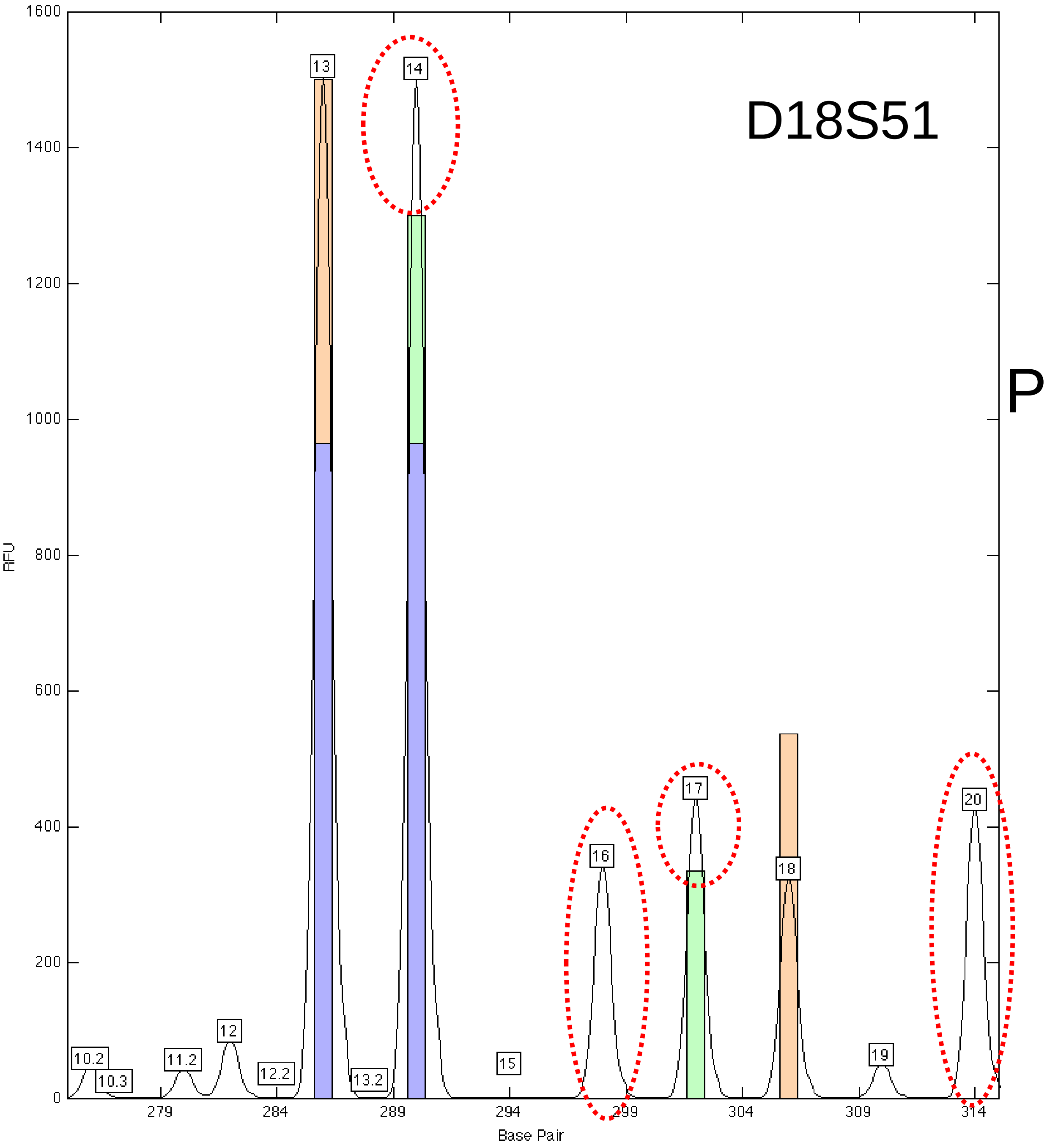
True Allele Results – 3 person mixture



100K examinations
3 unknowns
(no conditioning)

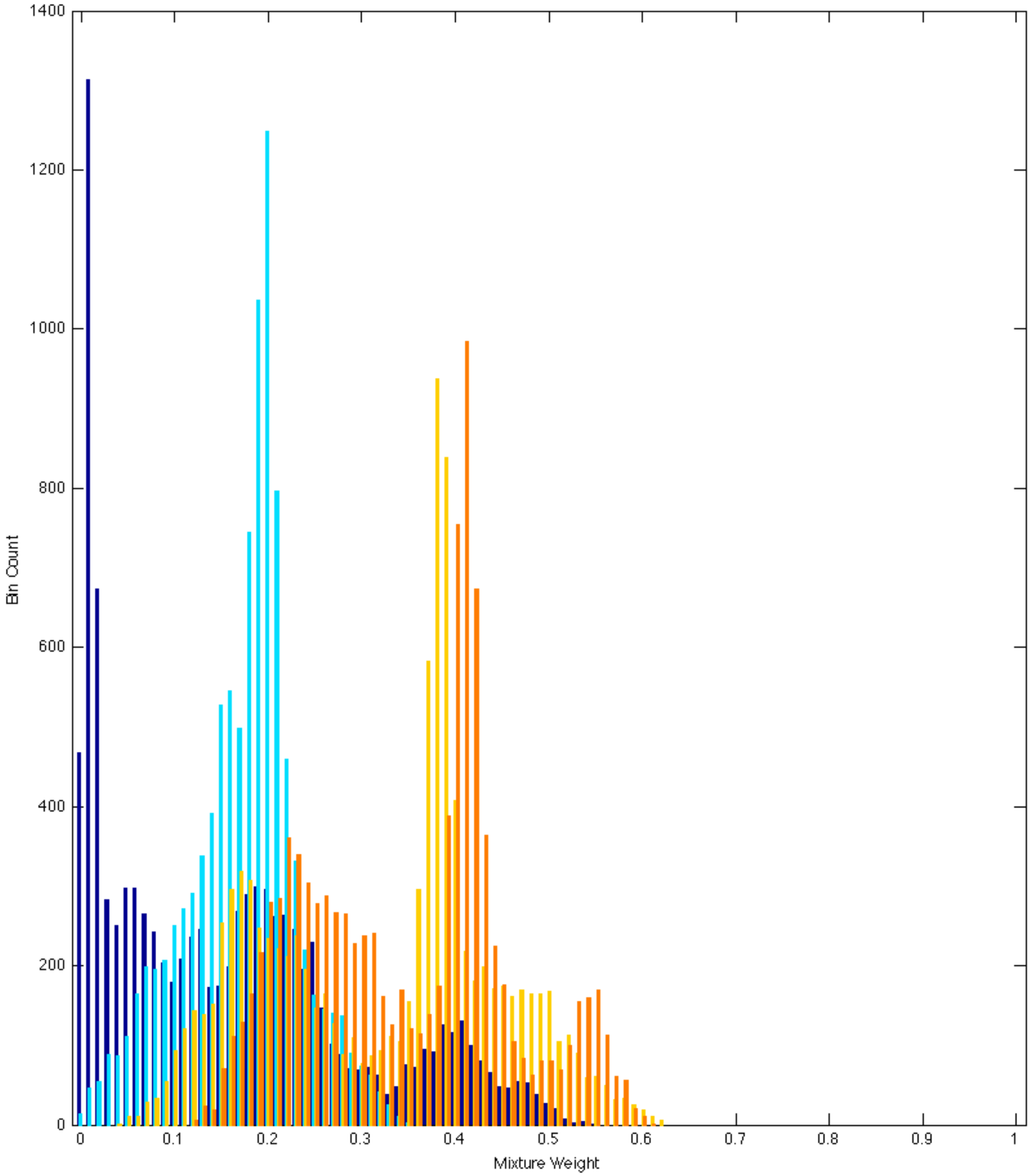
No clear separation

Mix ratio (green)
10-60%



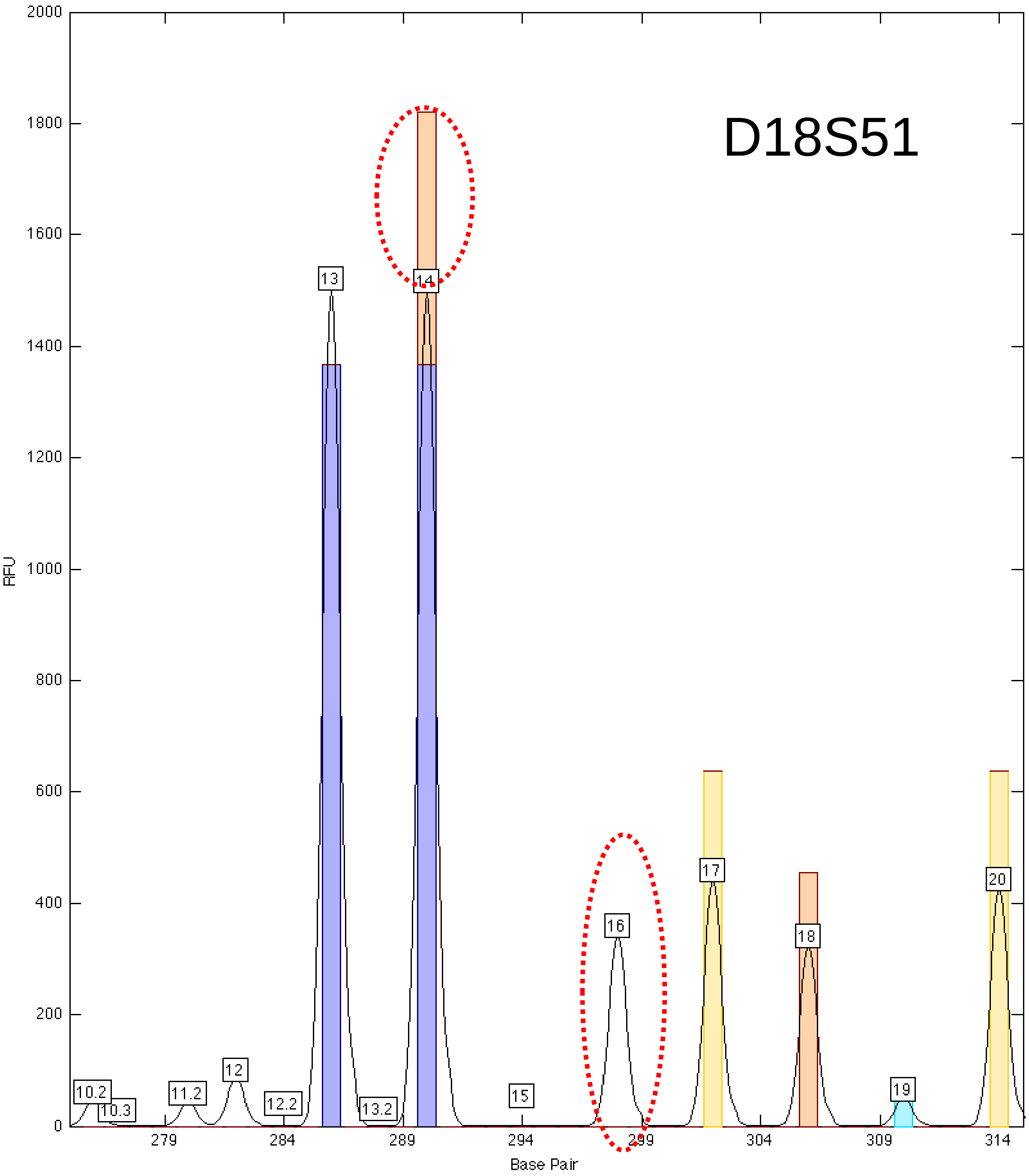
VERY
Poor fit of the data
to the model

True Allele Results – 4 person mixture

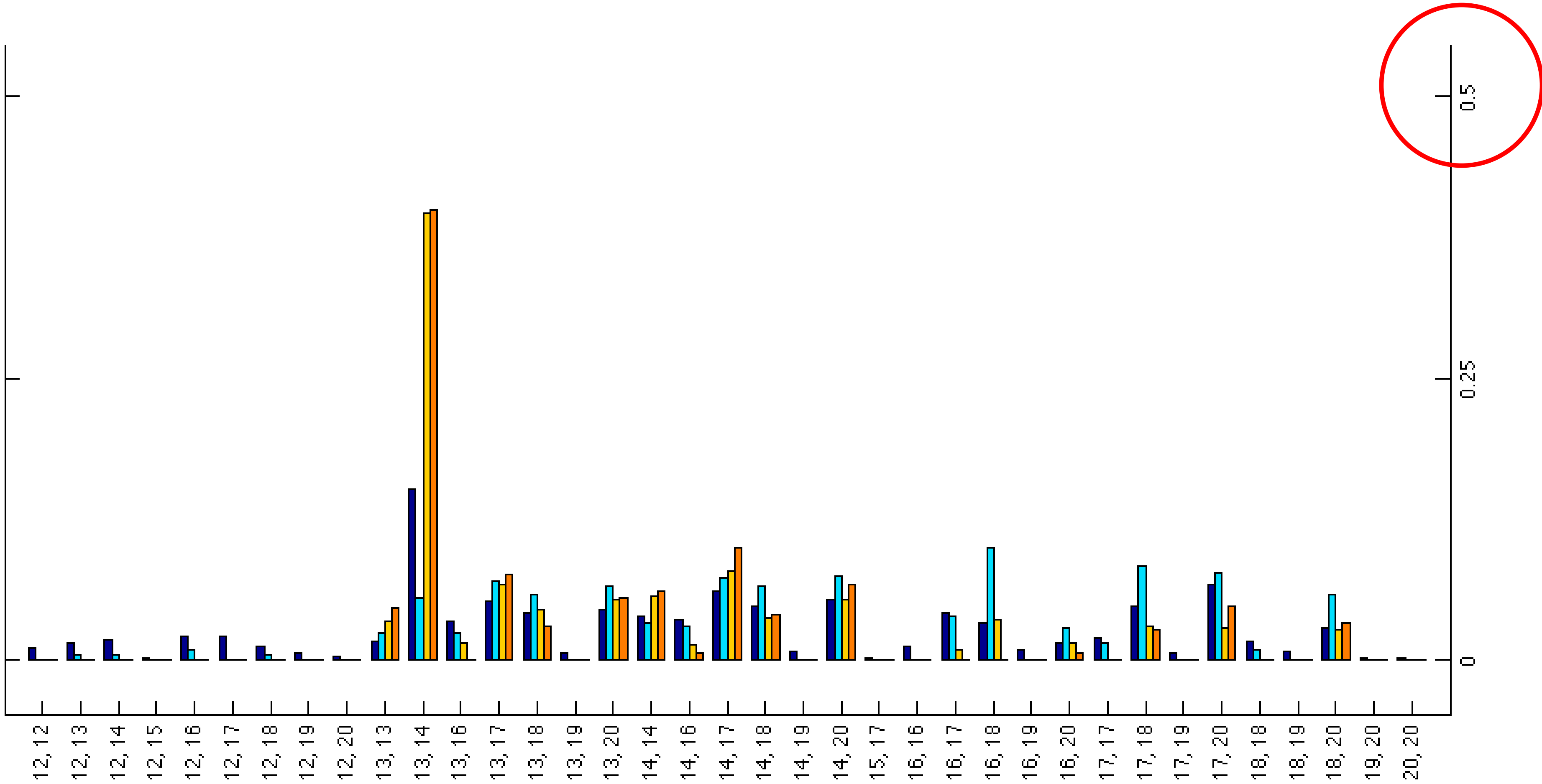


100K examinations
4 unknowns
(no conditioning)

Better separation,
Still uncertainty.

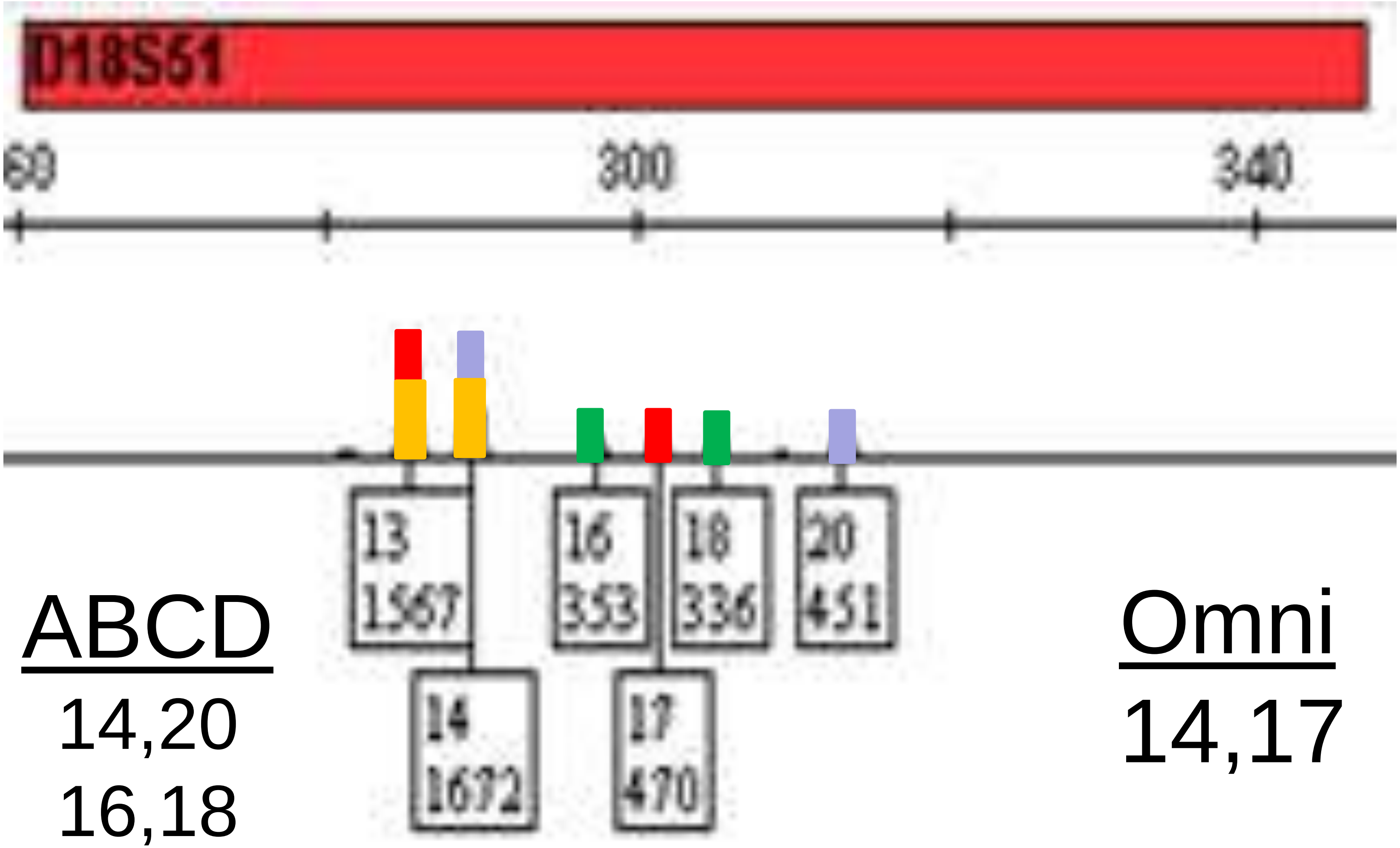


Still a poor fit of the data to the model


0.5
0.25
0
12, 12
12, 13
12, 14
12, 15
12, 16
12, 17
12, 18
12, 19
12, 20
13, 13
13, 14
13, 16
13, 17
13, 18
13, 19
13, 20
14, 14
14, 16
14, 17
14, 18
14, 19
14, 20
15, 17
16, 16
16, 17
16, 18
16, 19
16, 20
17, 17
17, 18
17, 19
17, 20
18, 18
18, 19
18, 20
19, 20
20, 20

Potential Suspects

- A, B, C and D are the four individuals in the mixture.
- John Butler is also a suspect (The Butler did it).
- “Omni man” is also a possible suspect.



ABCD

14,20

16,18

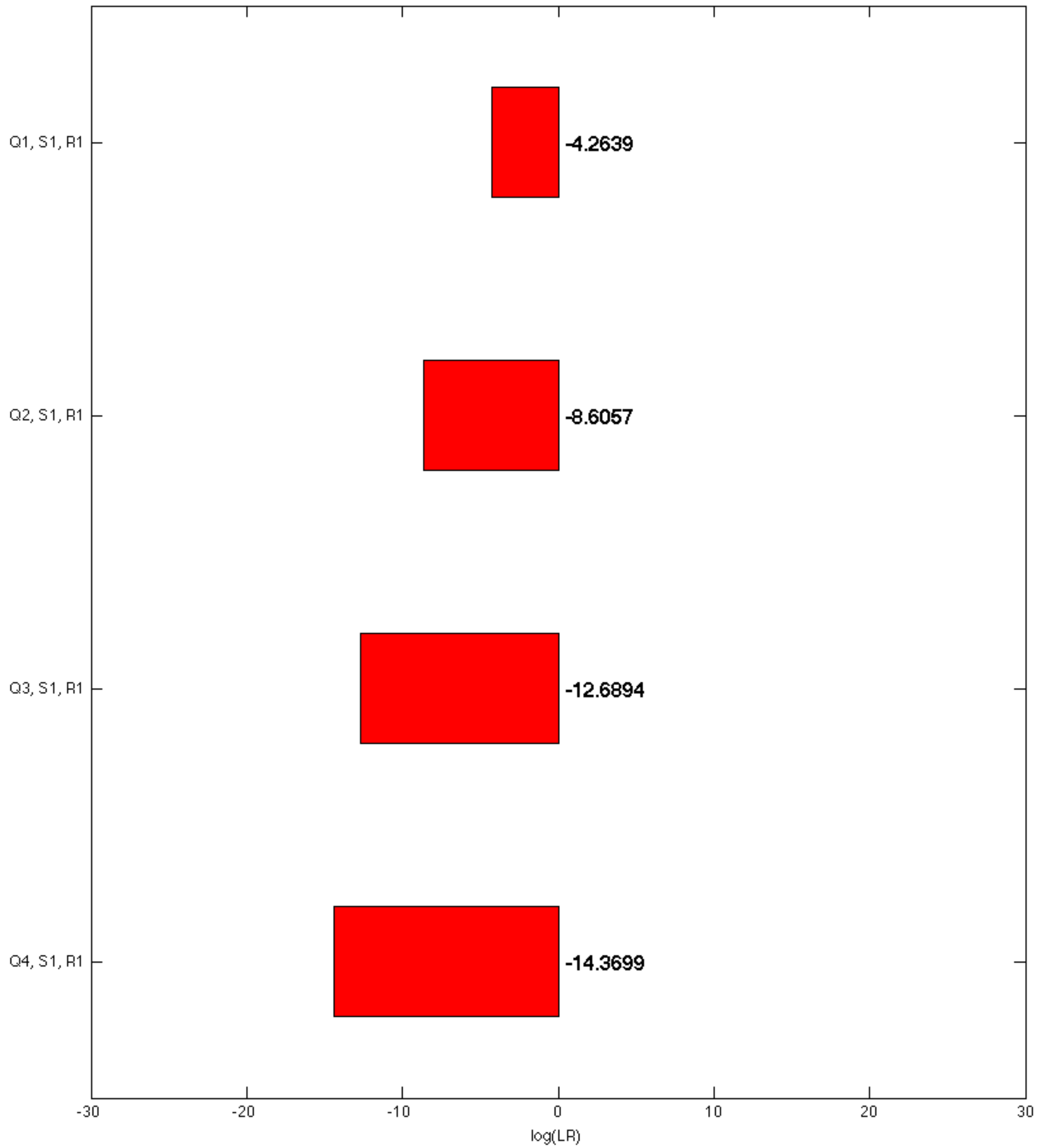
13,17

13,14

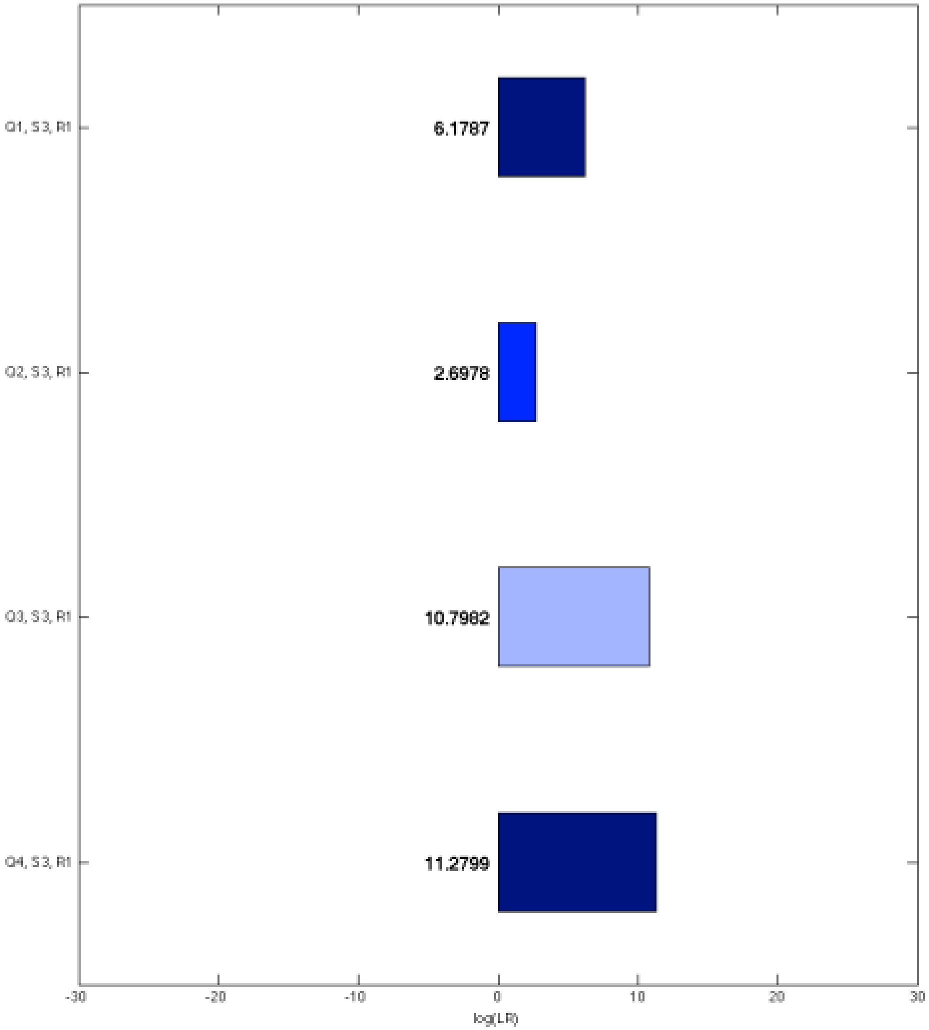
Omni

14,17

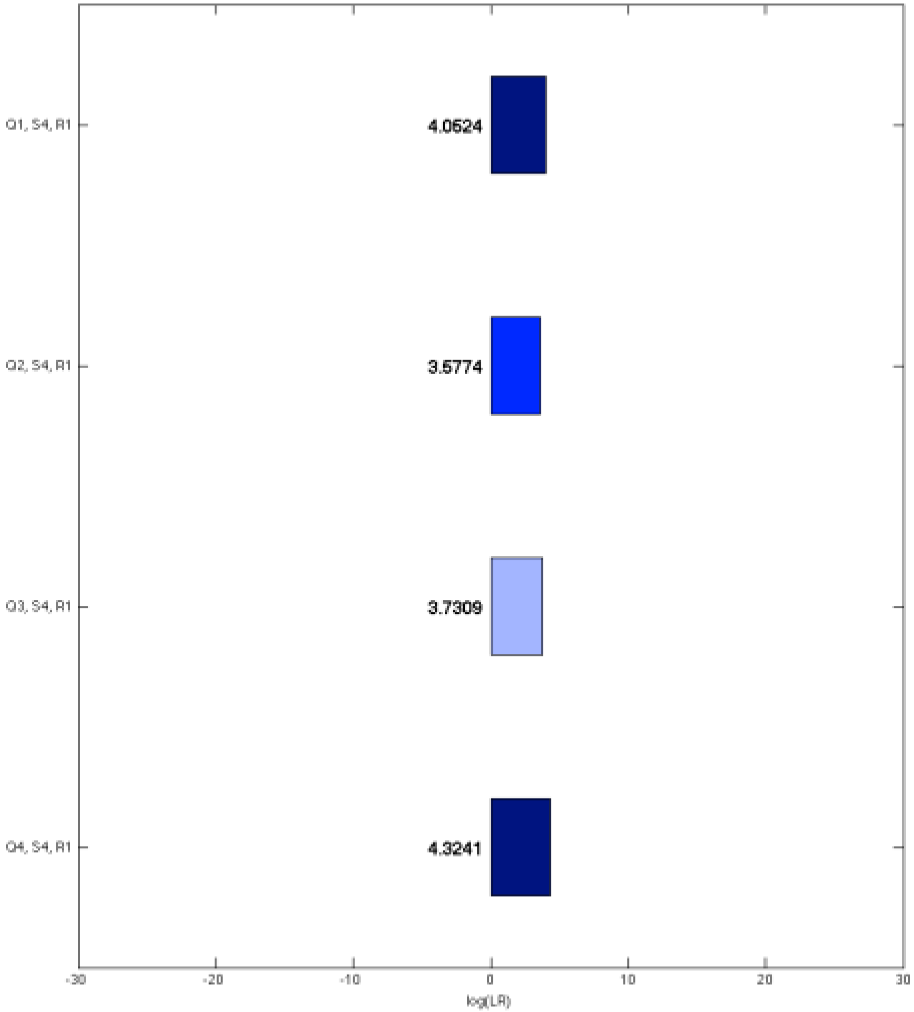
“The Butler”


Q1, S1, R1
-4.2639
Q2, S1, R1
-8.6057
Q3, S1, R1
-12.6894
Q4, S1, R1
-14.3699
-30
-20
-10
0
10
20
30
log(LR)

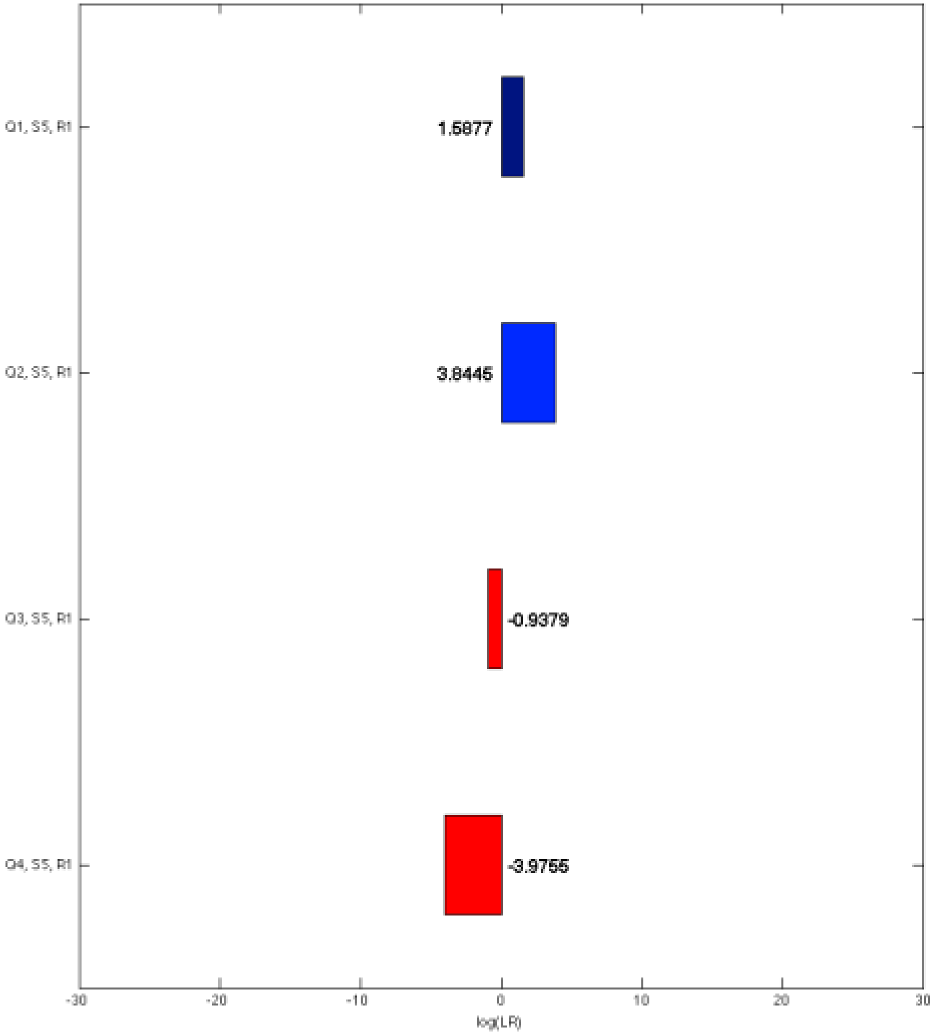
Suspect A



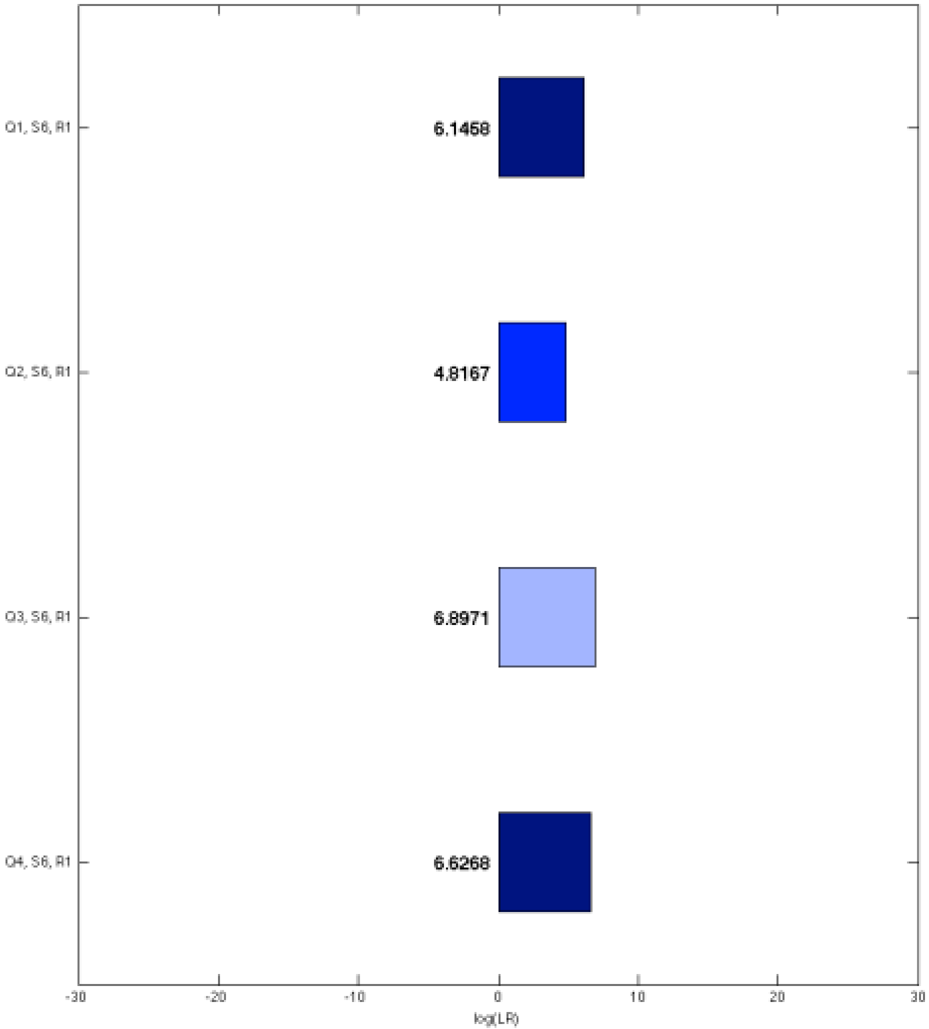
Suspect B



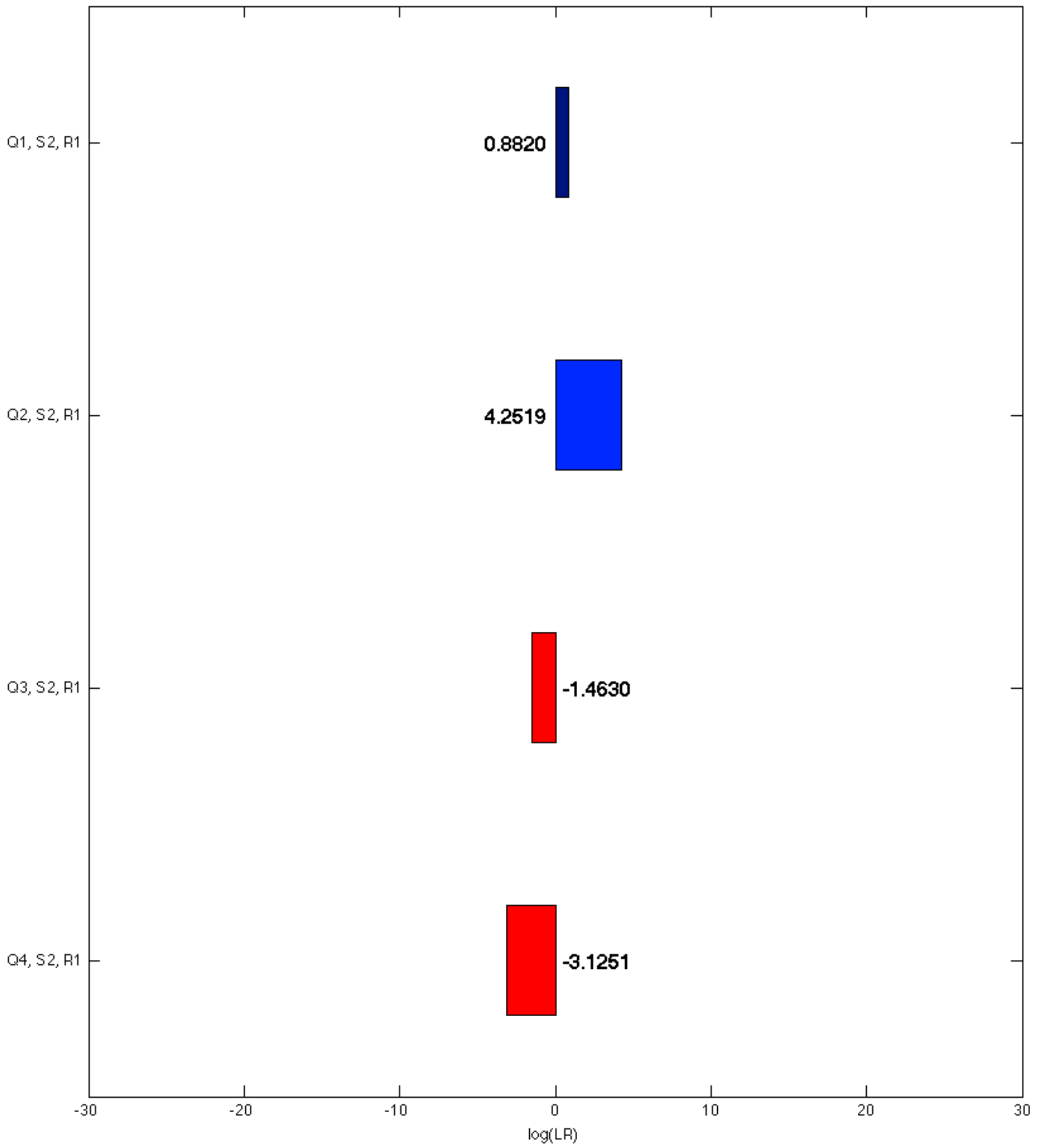
Suspect C



Suspect D



Omni Man


Q1, S2, R1
Q2, S2, R1
Q3, S2, R1
Q4, S2, R1
0.8820
4.2519
-1.4630
-3.1251
-30
-20
-10
0
10
20
30
log(LR)

Strategies

- Conditioning will help…
- This may not be possible.
- Multiple replicates will be necessary.
- There is a need to determine an appropriate method for an inclusion log(LR).

Summary of the Issues

- New kits, new instruments can only increase the difficulties of interpreting low-level, challenging samples.

- Probabilistic methods will be necessary to interpret low level samples with drop-out potential (or contaminating alleles) since classical approaches to interpretation such as RMNE or mRMP (even the classic LR) will not suffice.

Thanks to NIJ for Support of BU and NIST

NIJ
National Institute of Justice

BOSTON UNIVERSITY

NIST
Applied Genetics

- NIJ Forensic Science Training Development and Delivery Program Grant # 2008-DN-BX-K158, awarded to Biomedical Forensic Science Program at Boston University School of Medicine
- NIJ has an Interagency Agreement (IAA) with the NIST Office of Law Enforcement Standards (OLES)