

## 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>

# Probabilistic Genotyping

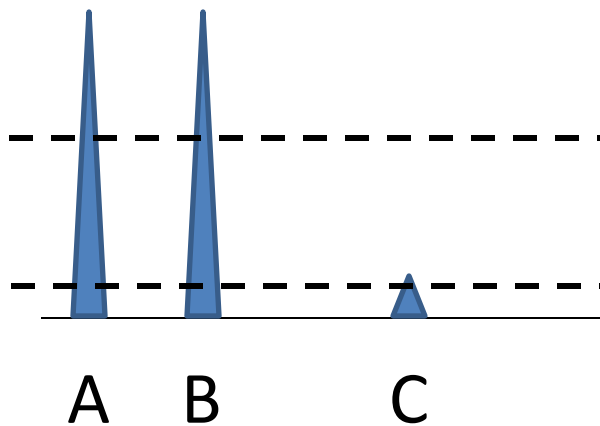
**Michael D. Coble**

National Institute of Standards and Technology



# What should we do with discordant data?

- Ignore/drop the locus – this is the “most conservative” option.



Complainant = AB  
 POI = CD

# Curran and Buckleton (2010)

JOURNAL OF **FORENSIC  
SCIENCES**



*J Forensic Sci*, September 2010, Vol. 55, No. 5  
doi: 10.1111/j.1556-4029.2010.01446.x  
Available online at: [interscience.wiley.com](http://interscience.wiley.com)

**PAPER**

**CRIMINALISTICS; GENERAL**

*James M. Curran,<sup>1</sup> M.Sc.(Hons.), Ph.D. and John Buckleton,<sup>2</sup> Ph.D.*

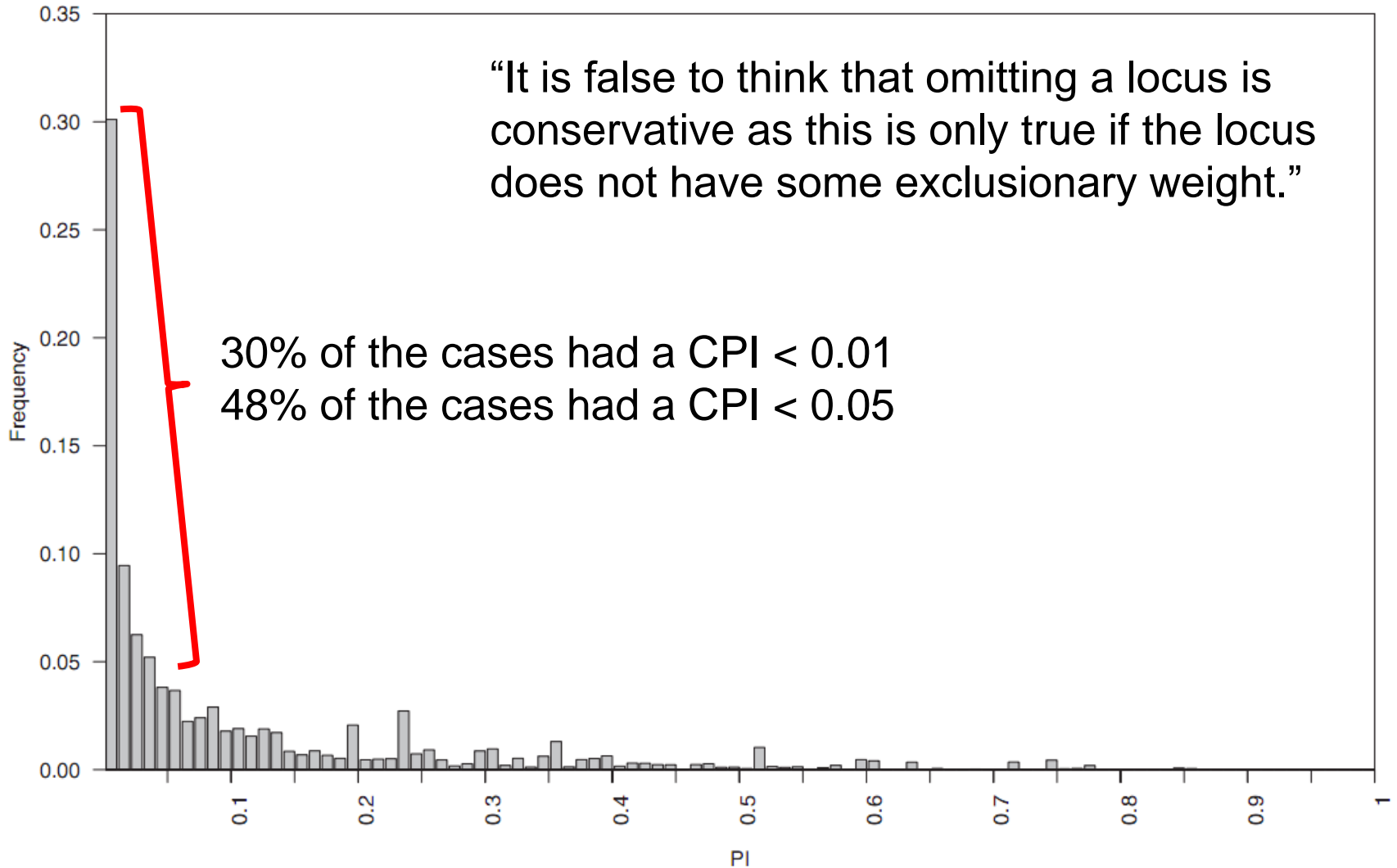
## Inclusion Probabilities and Dropout

Created 1000 Two-person Mixtures (Budowle *et al.* 1999 AfAm freq.).

Created 10,000 “third person” genotypes.

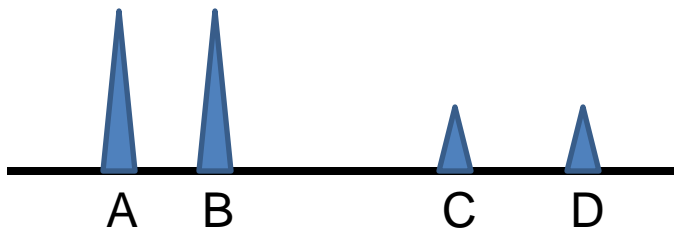
Compared “third person” to mixture data, calculated PI for included loci, ignored discordant alleles.

# Curran and Buckleton (2010)



# Curran and Buckleton (2010)

“It is false to think that omitting a locus is conservative as this is only true if the locus does not have some exclusionary weight.”



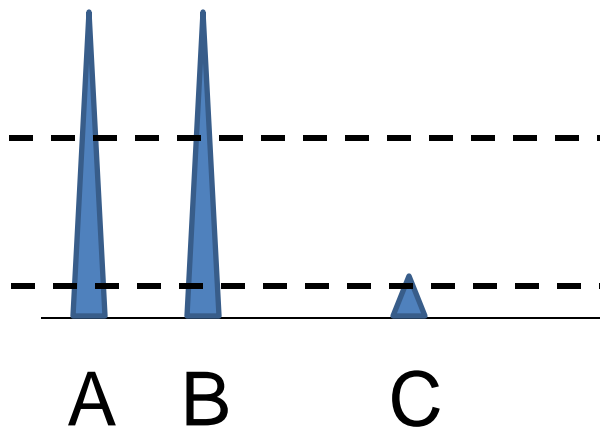
Dropping a locus is beneficial to the “guilty” and detrimental to the “innocent”.

POI = C,D

“Conservative”

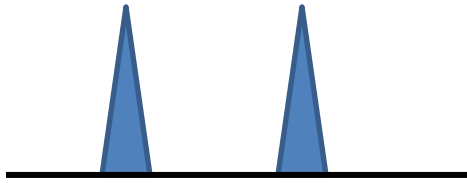
# What should we do with discordant data?

- Ignore/drop the locus – this is the “most conservative” option.

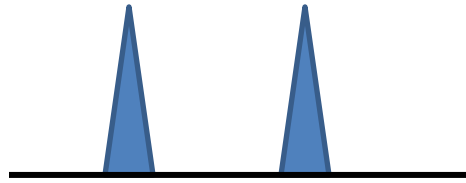


Complainant = AB  
 POI = CD

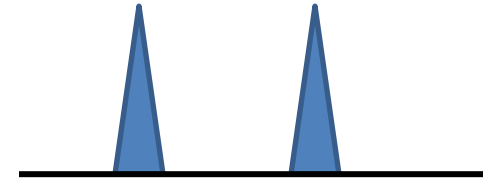
Suspect



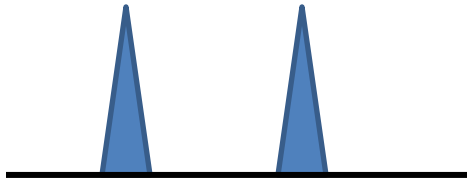
Suspect



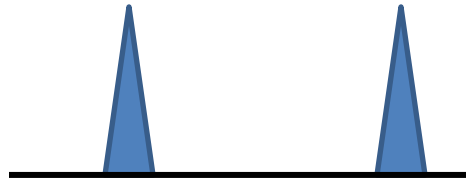
Suspect



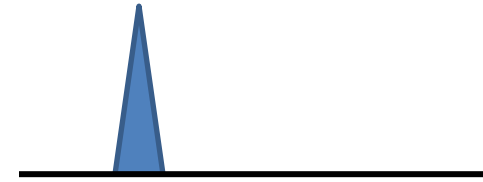
Evidence



Evidence



Evidence



$$LR = \frac{1}{2pq}$$

$$LR = \frac{0}{2pq}$$

$$LR = \frac{?}{2pq}$$

“2p”

Whatever way uncertainty is approached, probability is the *only* sound way to think about it.



-Dennis Lindley

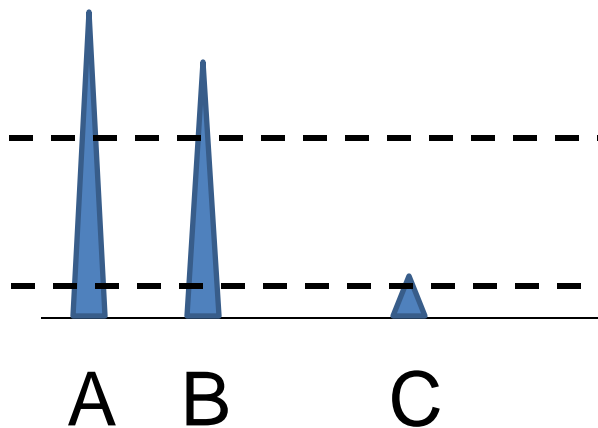


# What should we do with discordant data?

- Continue to use RMNE (CPI, CPE)
- Use the Binary LR with  $2p$
- Semi-continuous methods with a LR (Drop models)

# Drop Models

- Examine the alleles present and include a  $\text{Pr}(D)$  in the LR calculation



Alleles Present  
 ABC**F**

# December 2012 Issue of FSI-G

Forensic Science International: Genetics 6 (2012) 679–688



Contents lists available at [SciVerse ScienceDirect](#)

Forensic Science International: Genetics

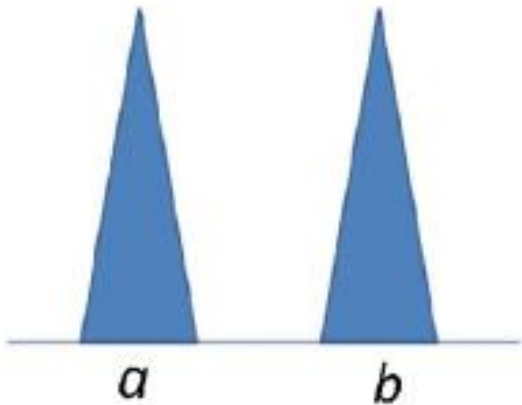
journal homepage: [www.elsevier.com/locate/fsig](http://www.elsevier.com/locate/fsig)



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

# ISFG Recommendations

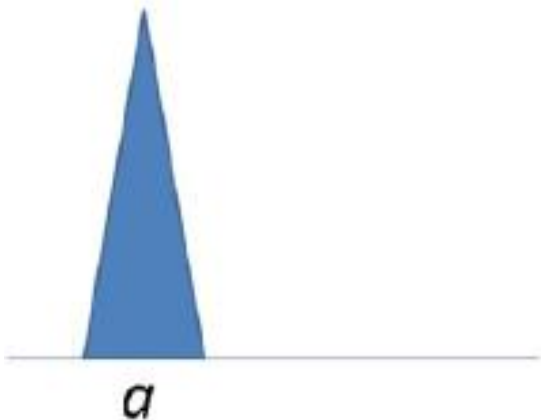


$\Pr(D) = \text{Prob. Drop-out (het)}$

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

$\Pr(D_2) = \text{Prob. Drop-out (hom)}$

$\Pr(\bar{D}_2) = \text{No Prob. Drop-out (hom)}$

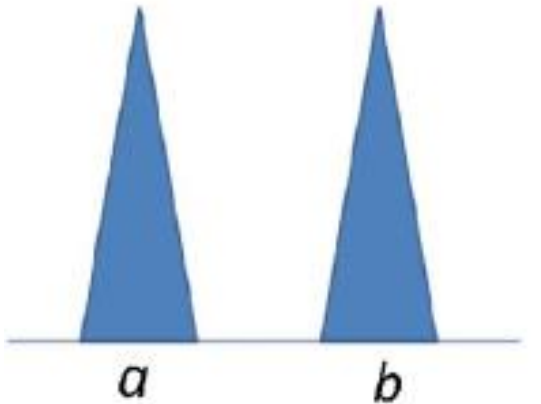


$\Pr(C) = \text{Prob. Drop-in}$

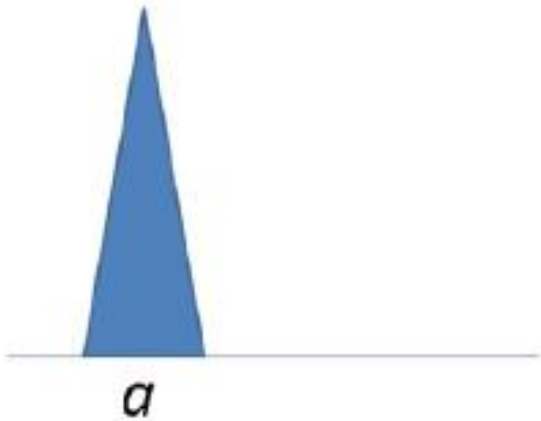
$\Pr(\bar{C}) = \text{No Prob. Drop-in}$

# Prosecutor's Explanation

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



Reference profile (S)



Crime stain profile (E)

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

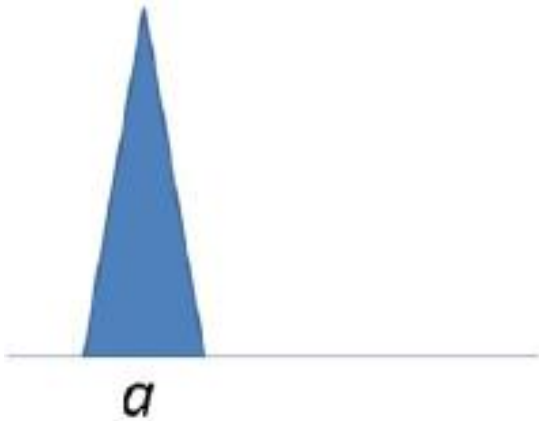
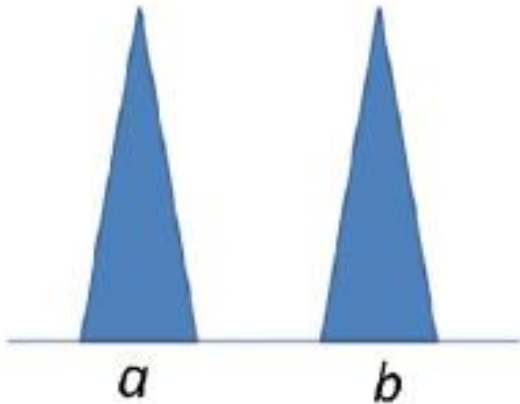
# The LR

$$LR = \frac{\Pr(\bar{D}) \Pr(D) \Pr(\bar{C})}{\text{—————}}$$

# Defense Explanation

## 4 possibilities

(1) The real culprit is a homozygote

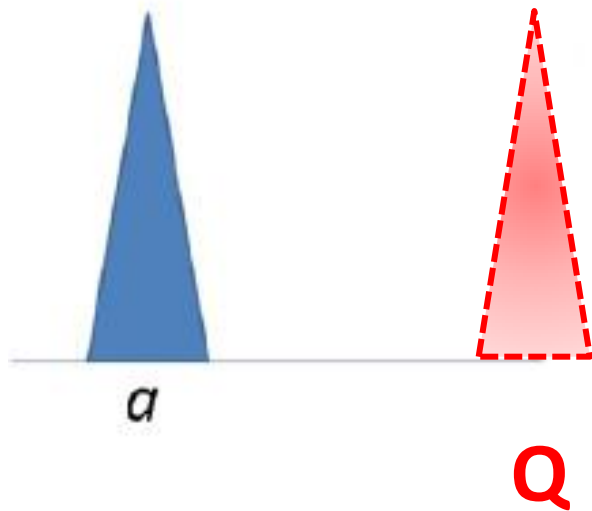
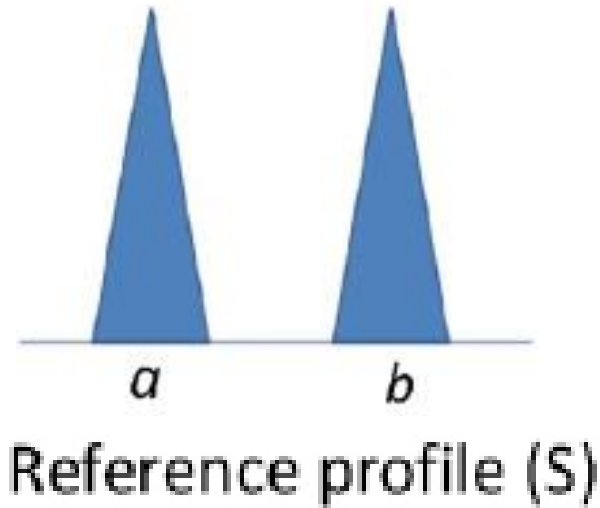


$$p_a^2 \Pr(\overline{D}_2) \Pr(\overline{C})$$

# Defense Explanation

4 possibilities

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



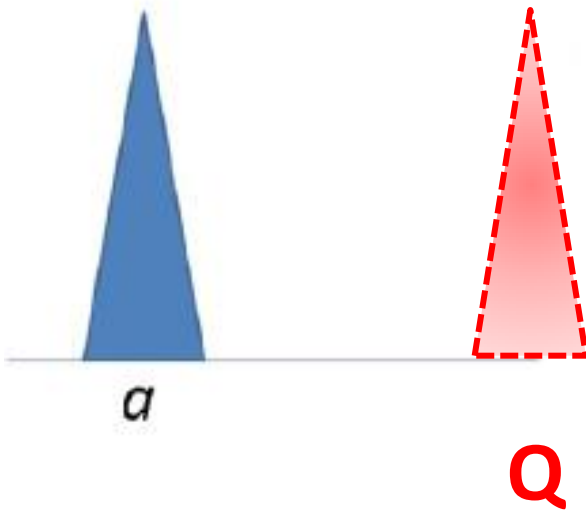
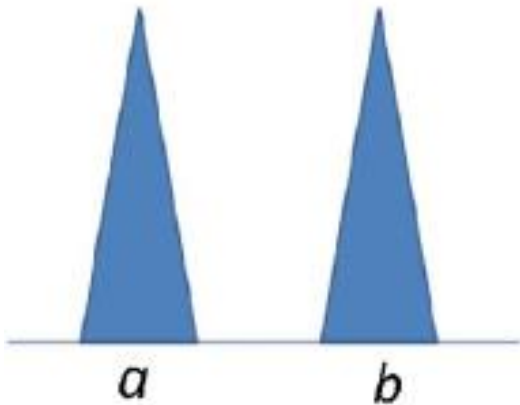
$$2p_a p_Q \Pr(\bar{D}) \Pr(D) \Pr(\bar{C})$$



# Defense Explanation

## 4 possibilities

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

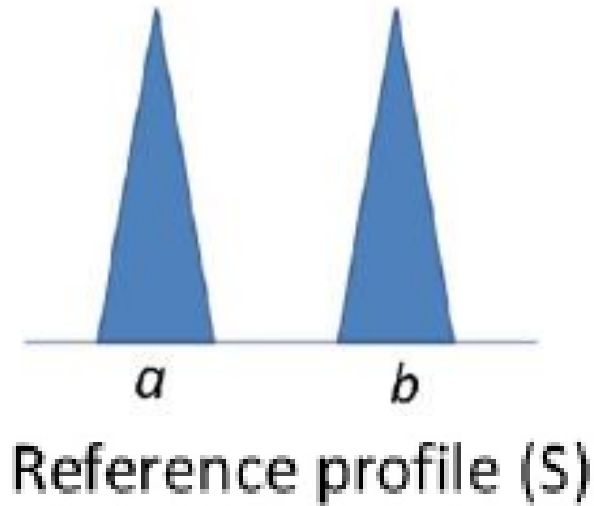


$$p_Q^2 \Pr(D_2) \Pr(C) p_a$$

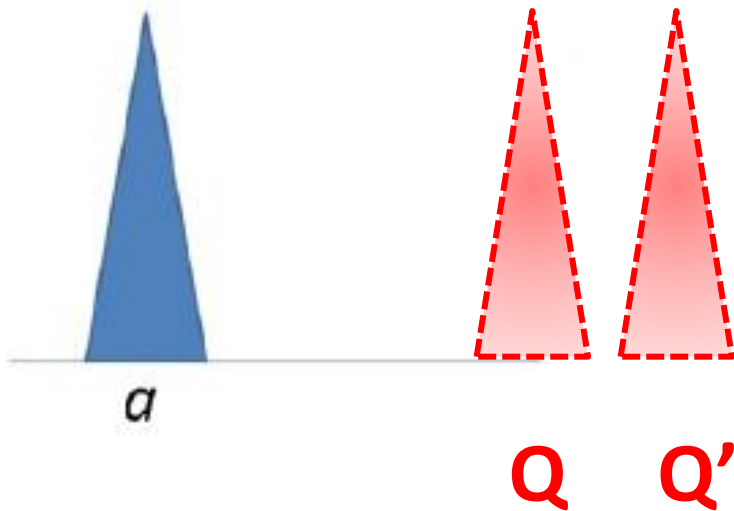
# Defense Explanation

## 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

$$\text{LR} = \frac{\text{Pr}(\bar{D}) \text{Pr}(D) \text{Pr}(\bar{C})}{
 \begin{aligned}
 & p_a^2 \text{Pr}(\bar{D}_2) \text{Pr}(\bar{C}) & + \\
 & 2p_a p_Q \text{Pr}(\bar{D}) \text{Pr}(D) \text{Pr}(\bar{C}) & + \\
 & p_Q^2 \text{Pr}(D_2) \text{Pr}(C) p_a & + \\
 & 2p_Q p_{Q'} \text{Pr}(D)^2 \text{Pr}(C) p_a
 \end{aligned}
 }$$

# Some Drop Model Examples

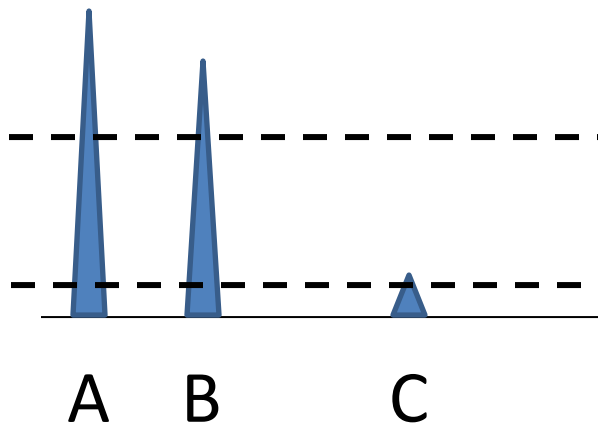
- LR mix (Haned and Gill)
- Balding and Buckleton (R program)
- FST (NYOCME, Mitchell *et al.*)
- Kelly *et al.* (University of Auckland, ESR)
- Lab Retriever (Lohmueller, Rudin and Inman)

# What should we do with discordant data?

- Continue to use RMNE (CPI, CPE)
- Use the Binary LR with 2p
- Semi-continuous methods with a LR (Drop models)
- Fully continuous methods with LR

# Continuous Models

- Mathematical modeling of “molecular biology” of the profile (mix ratio, PHR (Hb), stutter, etc...) to find optimal genotypes, giving WEIGHT to the results.



## Probable Genotypes

AC – 40%

BC – 25%

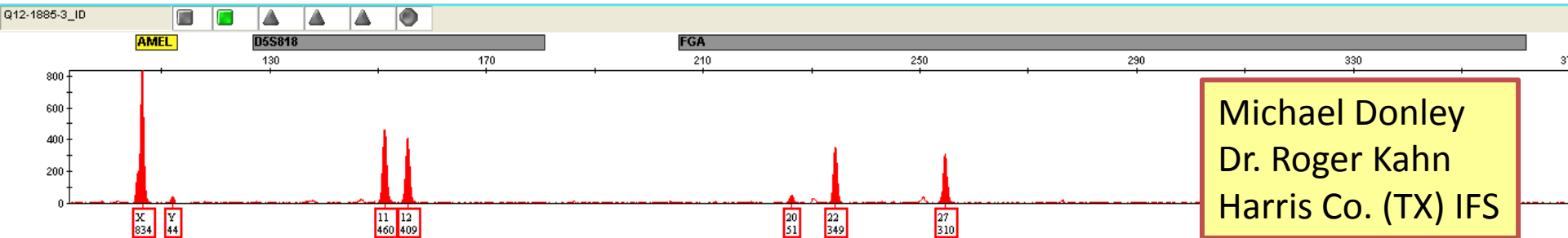
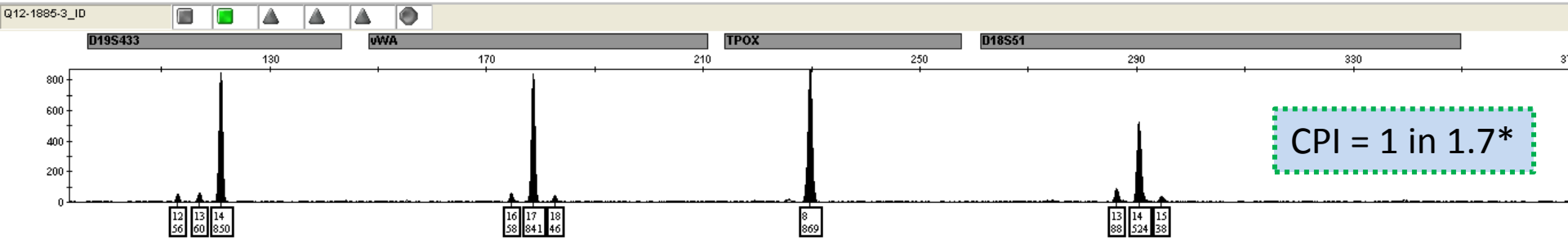
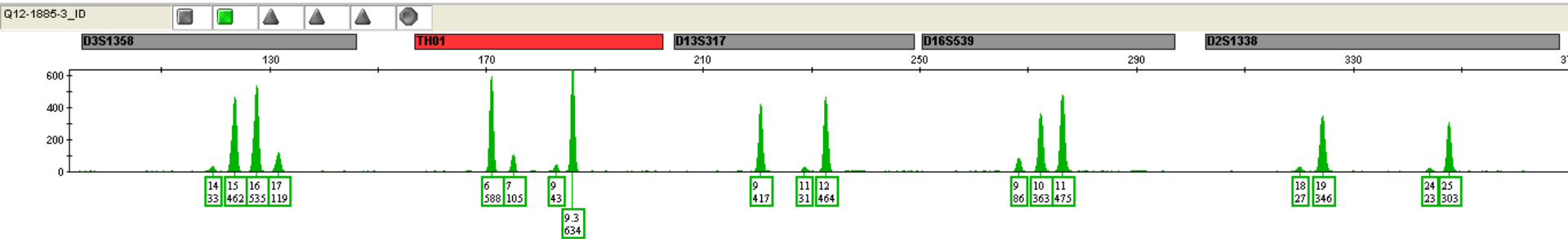
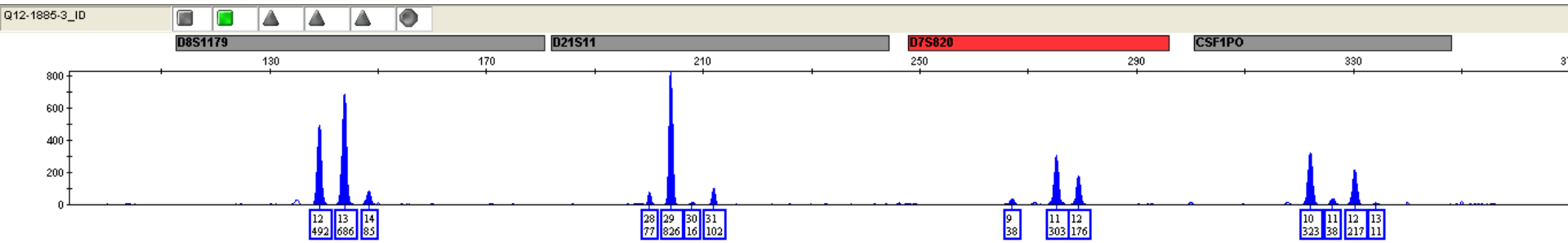
CC – 20%

CQ – 15%

# Some Continuous Model Examples

- TrueAllele (Cybergenetics)
- STRmix (ESR [NZ] and Australia)
- Cowell et al. (FSI-G (2011) 5:202-209)

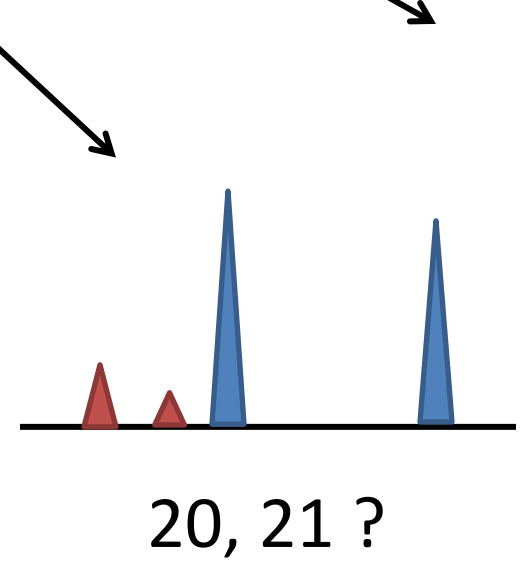
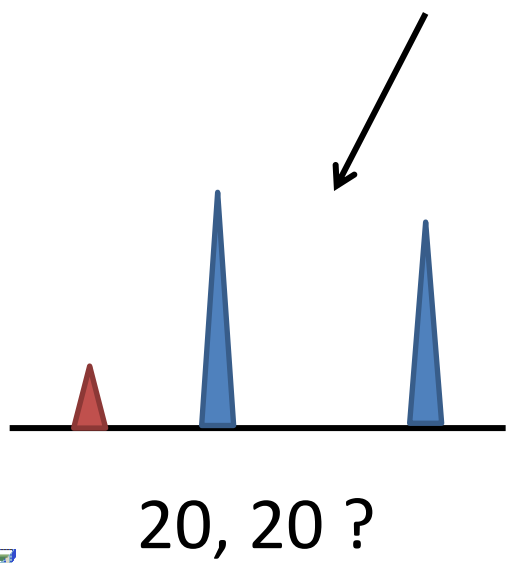
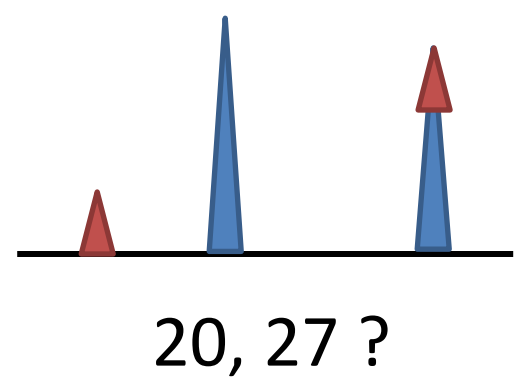
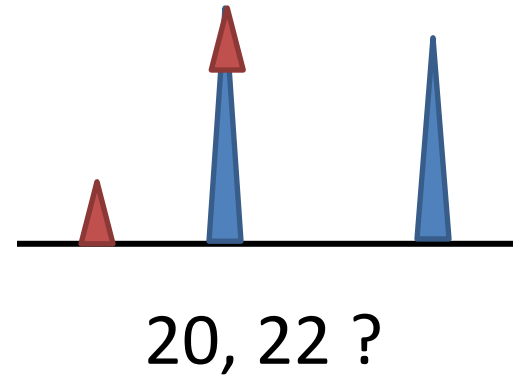
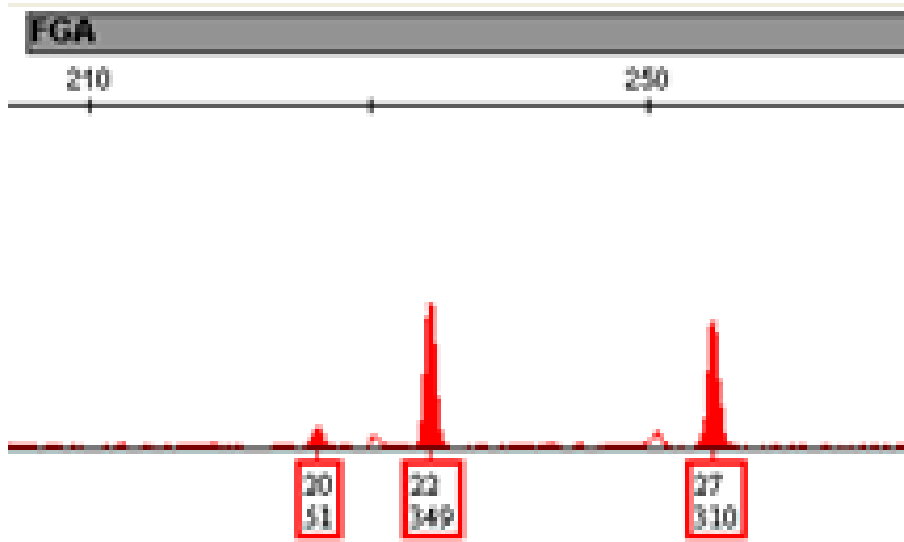
# Challenging Mixture



Michael Donley  
Dr. Roger Kahn  
Harris Co. (TX) IFS



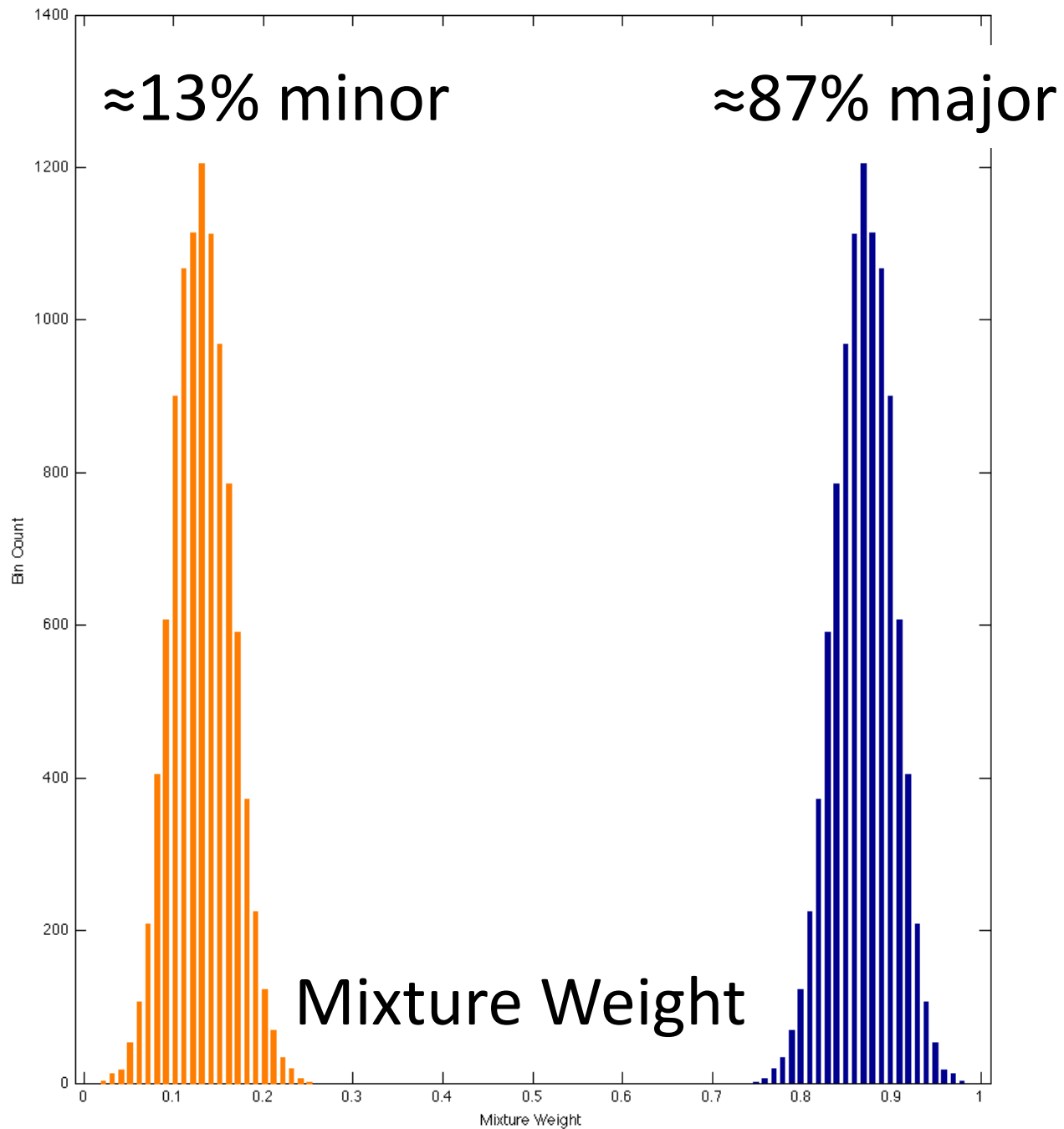
# Challenging Mixture

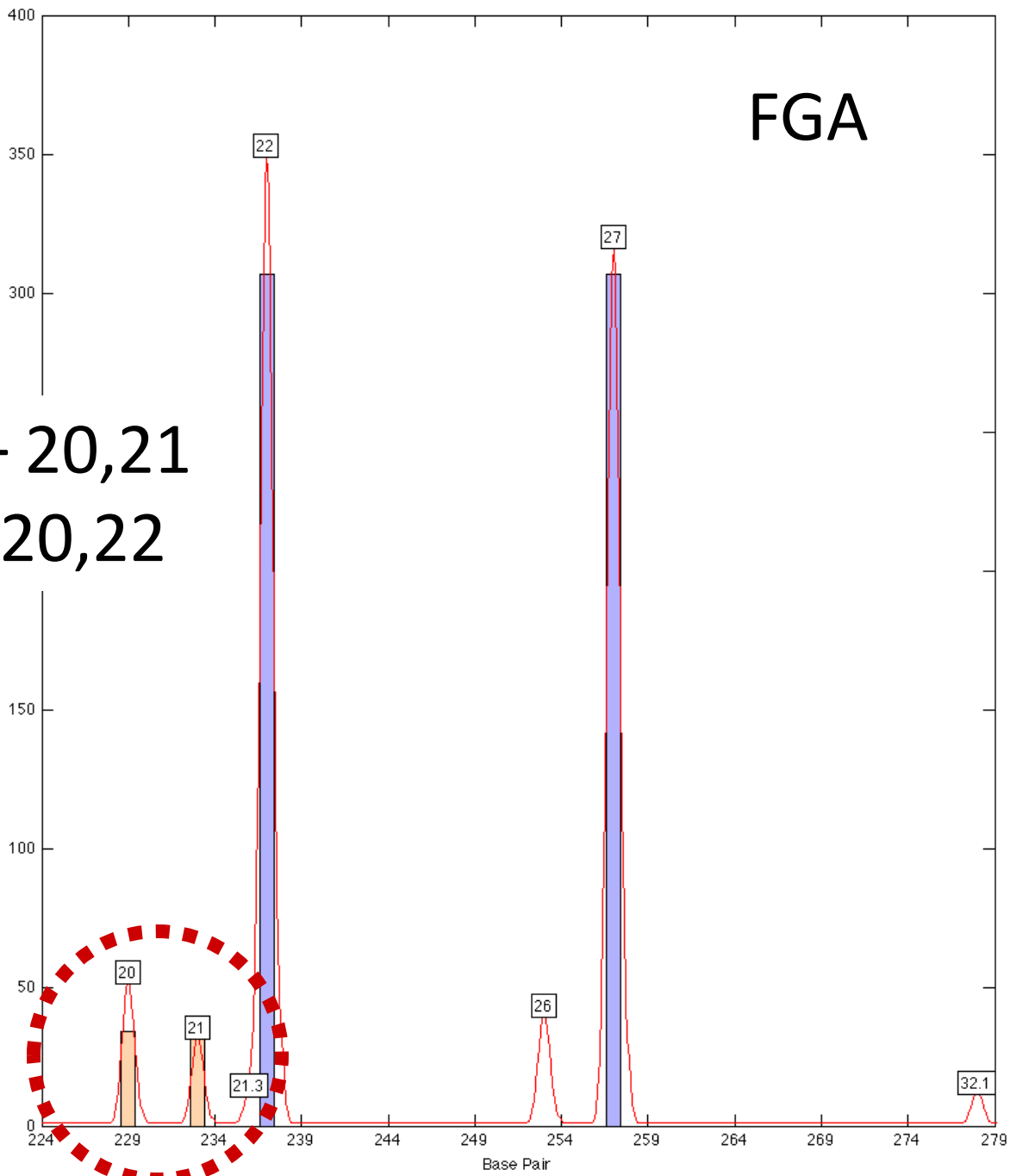


ETC...

# TrueAllele Results

# Bin Count





Inferred – 20,21  
Actual – 20,22

# Statistical Calculation

	Inferred	Prob.	HWE	Suspect
FGA	20, 22	0.1474	0.0543	1
	20, 21	0.0722	0.0461	0
	20, 26	0.1309	0.0058	0
	20, 20	0.0882	0.0156	0
	21, 22	0.0056	0.08	0
	21, 26	0.0176	0.0085	0
	22, 26	0.0077	0.01	0
	20, 27	0.0142	0.0008	0
	22, 22	0.001	0.0471	0

$H_p$

$$LR = \frac{0.1474}{\quad}$$

# Statistical Calculation

	Inferred	Prob.	HWE	Pr*HWE
FGA	20, 22	0.1474	0.0543	0.008
	20, 21	0.0722	0.0461	0.0033
	20, 26	0.1309	0.0058	0.0008
	20, 20	0.0882	0.0156	0.0014
	21, 22	0.0056	0.08	0.0004
	21, 26	0.0176	0.0085	0.0001
	22, 26	0.0077	0.01	0.0001
	20, 27	0.0142	0.0008	0
	22, 22	0.0001	0.0471	0

$\Sigma$  0.0143

$H_D$

$$LR = \frac{0.1474}{0.0143}$$

$$LR = 10.33$$

# STRmix

# STRmix

Start Mixture Analysis

LR from Previous Analysis

Search Database

Settings

Model Maker

Exit

About

STRmix V1.07 - User: Coble\_F...



Government of South Australia  
Forensic Science SA



## Mixture Proportions

Contributor 1 - 87%

Contributor 2 - 13%

---

GENOTYPE PROBABILITY DISTRIBUTION

FGA

[22, 27]	[20, 20]	0.3293750584474653
[22, 27]	[-1, 20]	0.17910853800572998
[22, 27]	[20, 22]	0.23609375398505444
[22, 27]	[20, 27]	0.25542264956175026

---

Locus 15 (FGA) :  $\Pr(E|H_p) = 0.05152$ ,  $\Pr(E|H_d) = 5.0E-4$ ,  $LR = 103.29885$

# Summary of the Issues

- New kits, new instruments will only increase the difficulties of interpreting low-level, challenging samples.
- If we are really serious about properly interpreting low level and complex mixtures, we must move away from the RMNE mentality. POPSTATS will not do!!
- Probabilistic methods are the way forward and a number of software programs are available ranging from “open source” to commercial packages.

# Thank you for your attention

## Contact Information

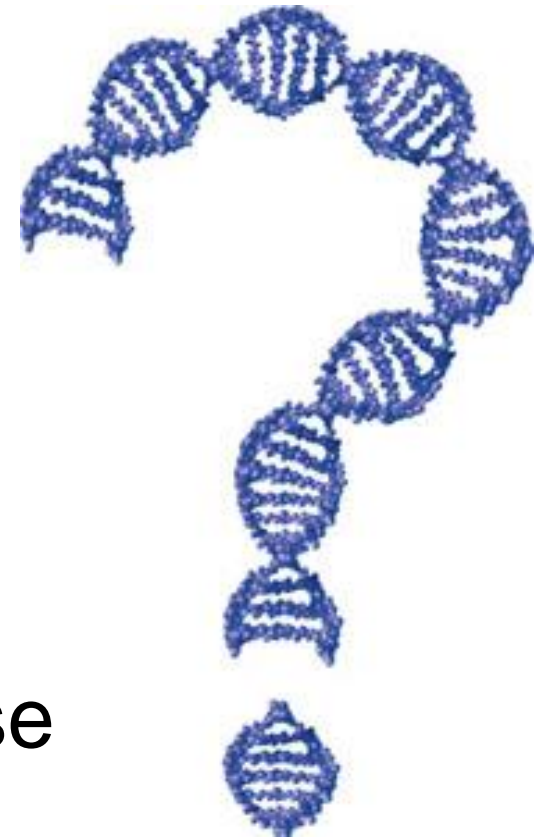
**Michael D. Coble**

Forensic Biologist

[michael.coble@nist.gov](mailto:michael.coble@nist.gov)

301-975-4330

<http://www.cstl.nist.gov/strbase>



**Additional DNA mixture information available at:**  
<http://www.cstl.nist.gov/strbase/mixture.htm>