Forensic DNA Mixture Interpretation

Statistical Approaches for DNA Mixtures

MAFS Workshop

Milwaukee, WI
September 25, 2012

Dr. Michael D. Coble
National Institute of Standards and Technology

michael.coble@nist.gov

Why Do Stats?

- A way to assess the weight of the statement that two profiles match (or cannot be excluded as originating from the same source)

- To be non-prejudicial

Stats Required for Inclusions

SWGDAM Interpretation Guideline 4.1:

> "**The laboratory must perform statistical analysis in support of any inclusion** that is determined to be relevant in the context of a case, irrespective of the number of alleles detected and the quantitative value of the statistical analysis."

Buckleton & Curran (2008): "There is a considerable aura to DNA evidence. Because of this aura **it is vital that weak evidence is correctly represented as weak or not presented at all**."

Buckleton, J. and Curran, J. (2008) A discussion of the merits of random man not excluded and likelihood ratios. *Forensic Sci. Int. Genet.* 2: 343-348.

Statistical Approaches with Mixtures

See Ladd *et al.* (2001) *Croat Med J.* 42:244-246

"Exclusionary" Approach

Random Man Not Excluded (RMNE)

Combined Prob. of Inclusion (CPI)

Combined Prob. of Exclusion (CPE)

"Inferred Genotype" Approach

Random Match Probability [modified] (mRMP)

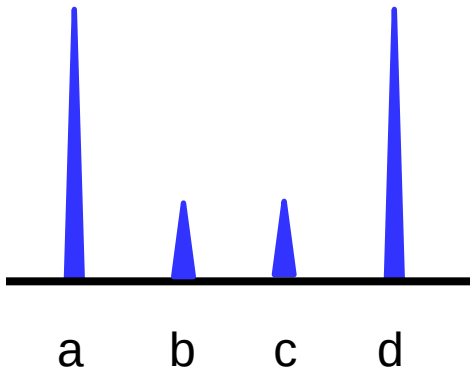
Likelihood Ratio (LR)

"Allele-centric"

"Genotype-centric"

Statistical Approaches with Mixtures

- **Random Man Not Excluded (CPE/CPI)** - The probability that a random person (unrelated individual) would be excluded as a contributor to the observed DNA mixture.



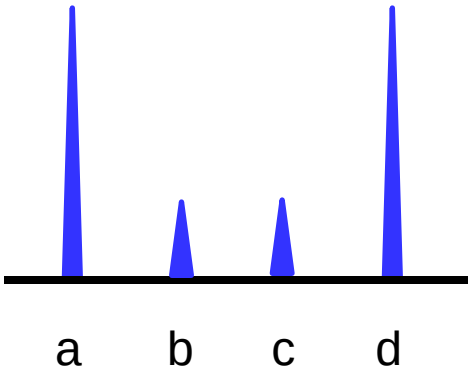
$$PI = (f(a) + f(b) + f(c) + f(d))^2$$

$$CPI = PI_{M1} \text{ X } PI_{M2} \cdots$$

$$CPE = 1 - CP1$$

Statistical Approaches with Mixtures

- **modified Random Match Probability (mRMP)**
  – The major and minor components can be successfully separated into individual profiles. A random match probability is calculated on the evidence as if the component was from a single source sample.



$$mRMP_{major} = 2pq$$

$$= 2f(a)f(d)$$

Statistical Approaches with Mixtures

- **Likelihood Ratio** - Comparing the probability of observing the mixture data under two (or more) alternative hypotheses

Probability

- Can be either…
- The frequency of observing an event in a large number of trials (Frequentist)
- The subjective degree of belief (Bayesian)

Probability

- Frequentist – What is the coincidental chance of the observed event.

- Bayesian – determine a *posterior probability* of observing the event based upon the data + the *prior probability* based upon knowledge of the source.

Laws of Probability

- Probabilities can range from 0 to 1.
- Events can be mutually exclusive (add)

$$P(G \text{ or } H|E) = P(G|E) + P(H|E)$$

- Events can be independent (multiply)

$$P(G \text{ and } H|E) = P(G|E) \times P(H|E)$$

Probabilities

- What is the probability of rolling a “5” using a six-sided die?
- P(rolling a 5) = 1/6
- What is the probability of rolling a “5” **or** “6”?
- P(rolling a 5) + P(rolling a 6) = 1/6+1/6 = 2/6 or 1/3.

Probabilities

- What is the probability of rolling a “5” on the first throw **and** rolling a “6” on the second roll?
- P(rolling a 5) * P(rolling a 6) = 1/6*1/6 = 1/36.

1,1	2,1	3,1	4,1	5,1	6,1
1,2	2,2	3,2	4,2	5,2	6,2
1,3	2,3	3,3	4,3	5,3	6,3
1,4	2,4	3,4	4,4	5,4	6,4
1,5	2,5	3,5	4,5	5,5	6,5
1,6	2,6	3,6	4,6	**5,6**	6,6

Conditioning

- **Probabilities are conditional**, which means that the probability of something is based on a hypothesis

- In math terms, conditioning is denoted by a vertical bar
  - Hence, **Pr(a|b)** means ‘the probability of *a* ***given*** that *b* is true”

- The probability of an event *a* is dependent upon various assumptions—and these assumptions or hypotheses can change…

Slide information from Peter Gill (ISFG 2007 workshop, Copenhagen, August 20-21, 2007)

Probability Example – Will It Rain? (1)

Defining the Event and Assumptions/Hypotheses

- Let's suppose that ***a*** is the probability of an event (e.g., **will it rain?**)
- What is the probability that it will rain in the afternoon – **Pr(*a*)**?

- This probability is dependent upon assumptions
  - We can look at the window in the morning and observe if it is sunny (s) or cloudy (c)
  - Pr(a) **if** it is sunny (s) is less than Pr(a) **if** it is cloudy (c)

- We can write this as **Pr(*a*|*s*)** and **Pr(*a*|*c*)**
  - Since sunny or cloudy are the only possibilities, Pr(s) + Pr(c) = 1
  - or **Pr(s) = 1 – Pr(c)**

Slide information from Peter Gill (ISFG 2007 workshop, Copenhagen, August 20-21, 2007)

Probability Example – Will It Rain? (2)

Examining Available Data

- Pr(a|s) and Pr(a|c) can be calculated from data

- How often does it rain in the afternoon when its sunny in the morning?
  - 10 out of 100 observations so **Pr(*a*|*s*) = 0.1**

- How often does it rain in the afternoon when it is cloudy in the morning?
  - 90 out of 100 observations so **Pr(*a*|*c*) = 0.9**

Slide information from Peter Gill (ISFG 2007 workshop, Copenhagen, August 20-21, 2007)

Probability Example – Will It Rain? (3)

Formation of the Likelihood Ratio (LR)

- The LR compares two probabilities to find out which of the two probabilities is the most likely

The probability that it will rain in the afternoon when it is cloudy in the morning or **Pr(*a*|*c*)** is divided by the probability that it will rain in the afternoon when it is sunny in the morning or **Pr(*a*|*s*)**

$$LR = \frac{\Pr(a \mid c)}{\Pr(a \mid s)} = \frac{0.9}{0.1} = 9$$

Slide information from Peter Gill (ISFG 2007 workshop, Copenhagen, August 20-21, 2007)

Probability Example – Will It Rain? (4)

Explanation of the Likelihood Ratio

$$LR = \frac{\Pr(a \mid c)}{\Pr(a \mid s)} = \frac{0.9}{0.1} = 9$$

- The probability that it will rain is 9 times more likely **if** it is cloudy in the morning than **if** it is sunny in the morning.

- The word **if** is very important here. It must always be used when explaining a likelihood ratio otherwise the explanation could be misleading.

Slide information from Peter Gill (ISFG 2007 workshop, Copenhagen, August 20-21, 2007)

Likelihood Ratios in Forensic DNA Work

- We evaluate the evidence (*E*) relative to alternative pairs of hypotheses

- Usually these hypotheses are formulated as follows:
  - The probability of the evidence if the crime stain originated with the suspect or Pr(*E*|*S*)
  - The probability of the evidence if the crime stain originated from an unknown, unrelated individual or Pr(*E*|*U*)

$$LR = \frac{\Pr(E \mid S)}{\Pr(E \mid U)}$$

The numerator

The denominator

Slide information from Peter Gill (ISFG 2007 workshop, Copenhagen, August 20-21, 2007)

The Likelihood Ratio Must Be Stated Carefully

- The probability of the evidence is *x* times more likely **if** the stain came from the suspect Mr. Smith than **if** it came from an unknown, unrelated individual.

- It is not appropriate to say: “The probability that the stain came from Mr. Smith.” because we must always include the conditioning statement – i.e., **always make the hypothesis clear in the statement.**

- Always use the word ‘**if**’ when using a likelihood ratio to avoid this trap

Slide information from Peter Gill (ISFG 2007 workshop, Copenhagen, August 20-21, 2007)

Likelihood Ratio (LR)

- Provides ability to express and evaluate both the prosecution hypothesis, H_p (the suspect is the perpetrator) and the defense hypothesis, H_d (an unknown individual with a matching profile is the perpetrator)

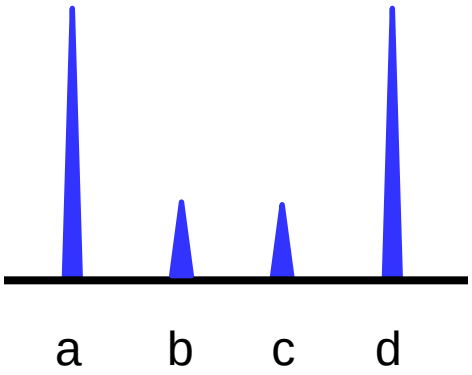
$$LR = \frac{H_p}{H_d}$$

- **The numerator, H_p, is usually 1** – since in theory the prosecution would only prosecute the suspect if they are 100% certain he/she is the perpetrator

- The denominator, **H_d**, is typically the profile frequency in a particular population (based on individual allele frequencies and assuming HWE) – i.e., **the random match probability**

Statistical Approaches with Mixtures

- **Likelihood Ratio** - Comparing the probability of observing the mixture data under two (or more) alternative hypotheses; in its simplest form LR = 1/RMP



$$\frac{P(E \mid H_1)}{P(E \mid H_2)}$$

E = Evidence
H_1 = Prosecutor's Hypothesis (the suspect did it) = 1
H_2 = Defense Hypothesis (the suspect is an unknown, random person)

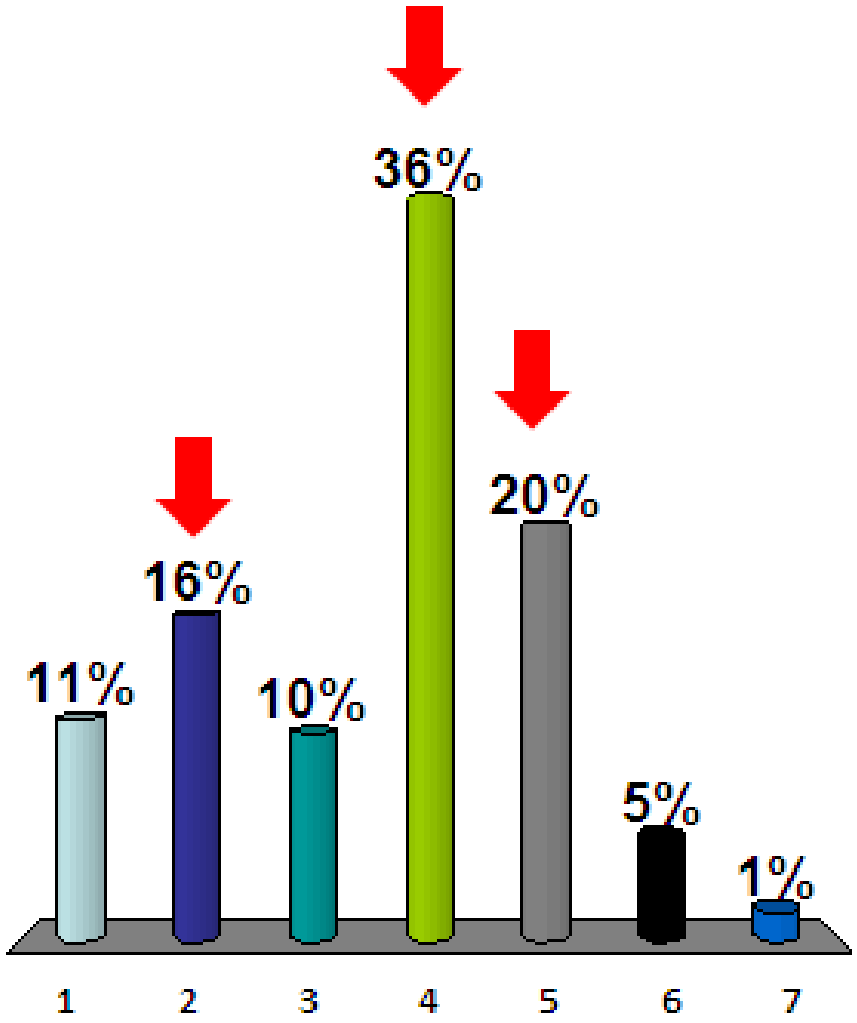
$$= \frac{1}{P(E \mid H_2)} = \frac{1}{2pq} = 1/RMP$$

Last Year's Response

What kind of mixture statistic does your lab use?

72% using CPI

1. LR
2. CPE (RMNE, CPI)
3. RMP
4. CPE or RMP
5. Other combinations
6. Probabilistic modeling (e.g., TrueAllele)
7. We don't use stats (contradicting the new guidelines – section 4.1)



Data from 138 responses
ISHI Mixture Workshop (Oct 2011)

FSI GENETICS

Forensic Science International: Genetics 2 (2008) 343–348

A discussion of the merits of random man not excluded and likelihood ratios

John Buckleton [a,*], James Curran [b]

[a] *ESR, PB 92021, Auckland, New Zealand*

[b] *Department of Statistics, University of Auckland, PB 92019, Auckland, New Zealand*

Received 15 January 2008; received in revised form 29 April 2008; accepted 1 May 2008

We conclude that the two matters that appear to have real force are:
(1) LRs are more difficult to present in court and
(2) the RMNE statistic wastes information that should be utilised.

CPE/CPI (RMNE) Limitations

- A CPE/CPI approach assumes that ***all alleles*** **are present** (i.e., cannot handle allele drop-out)
- Thus, statistical analysis of low-level DNA CANNOT be correctly performed with a CPE/CPI approach because some alleles may be missing
- Charles Brenner in his AAFS 2011 talk addressed this issue
- Research is on-going to develop allele drop-out models and software to enable appropriate calculations

Notes from Charles Brenner's AAFS 2011 talk

The Mythical "Exclusion" Method for Analyzing DNA Mixtures – Does it Make Any Sense at All?

1. The claim that it requires **no assumption about number of contributors** is mostly wrong.
2. The supposed **ease of understanding** by judge or jury is really an illusion.
3. **Ease of use** is claimed to be an advantage particularly for complicated mixture profiles, those with many peaks of varying heights. The truth is the exact opposite. **The exclusion method is completely invalid for complicated mixtures.**
4. The exclusion method is only **conservative** for guilty suspects.

Conclusion: "Certainly no one has laid out an explicit and rigorous chain of reasoning from first principles to support the exclusion method. It is at best guesswork."

Brenner, C.H. (2011). The mythical "exclusion" method for analyzing DNA mixtures – does it make any sense at all? *Proceedings of the American Academy of Forensic Sciences*, Feb 2011, Volume 17, p. 79

DAB Recommendations on Statistics

February 23, 2000

Forensic Sci. Comm. 2(3); available on-line at
http://www.fbi.gov/hq/lab/fsc/backissu/july2000/dnastat.htm

"The DAB finds either one or both PE or LR calculations acceptable and strongly recommends that one or both calculations be carried out whenever feasible and a mixture is indicated"

- Probability of exclusion (PE)
  - Devlin, B. (1993) Forensic inference from genetic markers. *Statistical Methods in Medical Research* 2: 241–262.
- Likelihood ratios (LR)
  - Evett, I. W. and Weir, B. S. (1998) *Interpreting DNA Evidence*. Sinauer, Sunderland, Massachusetts.

Statistical Methods in Medical Research 1993; **2**: 241–262

Forensic inference from genetic markers

B Devlin Department of Epidemiology and Public Health, Yale University School of Medicine

Section 5.1 Exclusion probability

- Discussion about exclusion probabilities in Paternity cases.

Two types:

(1) Conditional Exclusion Probability - excluding a random man as a possible father, given the mother-child genotypes for a particular case.

(2) Average Exclusion Probability – excluding a random man as a possible father, given a randomly chosen mother-child pair.

Statistical Methods in Medical Research 1993; **2**: 241–262

Forensic inference from genetic markers

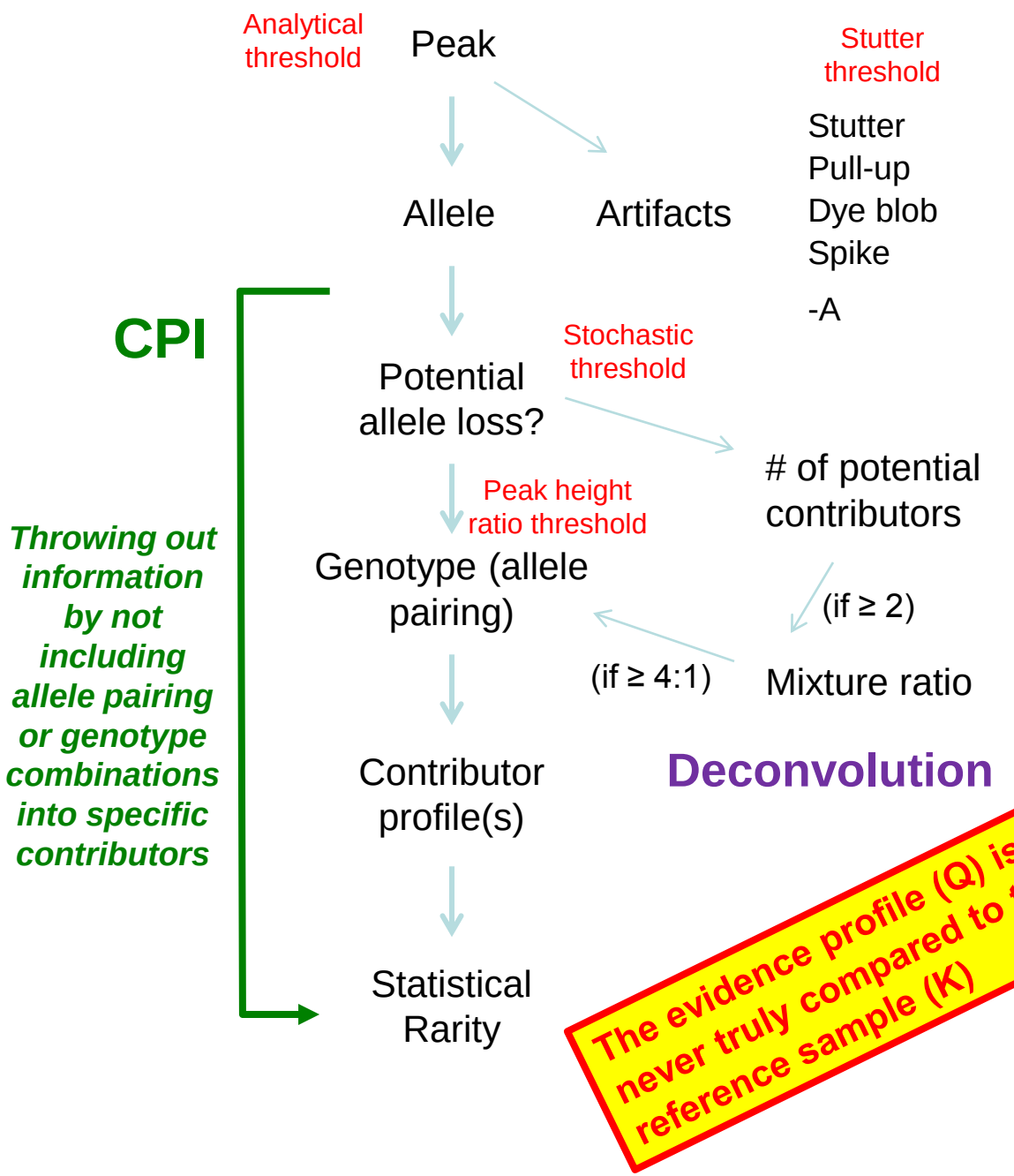
B Devlin Department of Epidemiology and Public Health, Yale University School of Medicine

Section 5.1 Exclusion probability

“The theoretical concept of exclusion probabilities, however, makes no sense within the framework of normal mixture models.”

“The interpretation of conditional exclusion probability is obvious, which accounts for its value in the legal arena. Unlike [LR], however, it is not fully efficient.”

Problem with CPI Approach



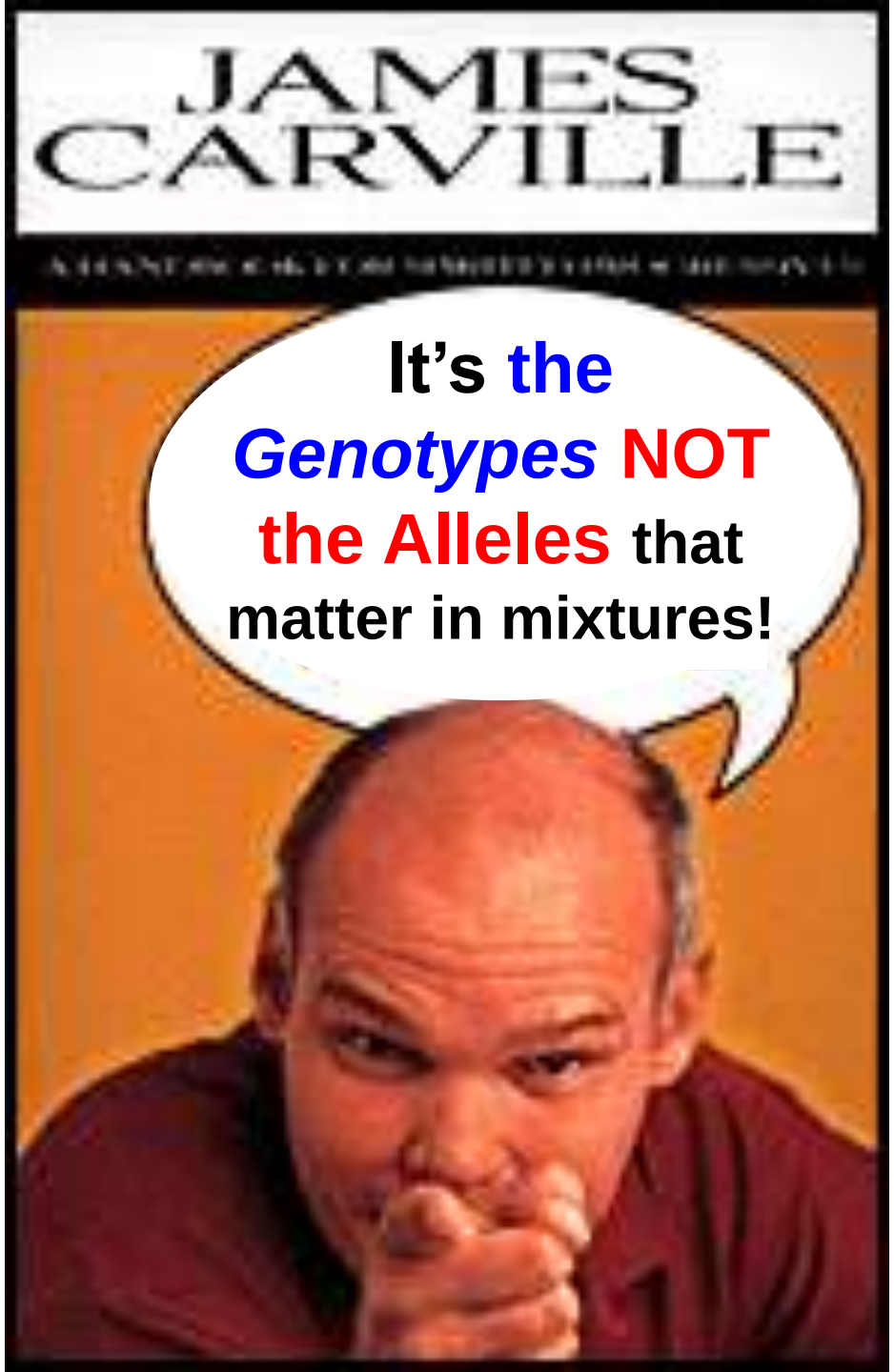
CPI deals with alleles NOT specific genotype combinations

Q → K Comparison

The evidence profile (Q) is never truly compared to the reference sample (K)

Report Issued with conclusions (inclusion, exclusion, inconclusive)

Just Remember



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

PAPER

CRIMINALISTICS; GENERAL

James M. Curran,[1] *M.Sc.(Hons.), Ph.D. and John Buckleton,*[2] *Ph.D.*

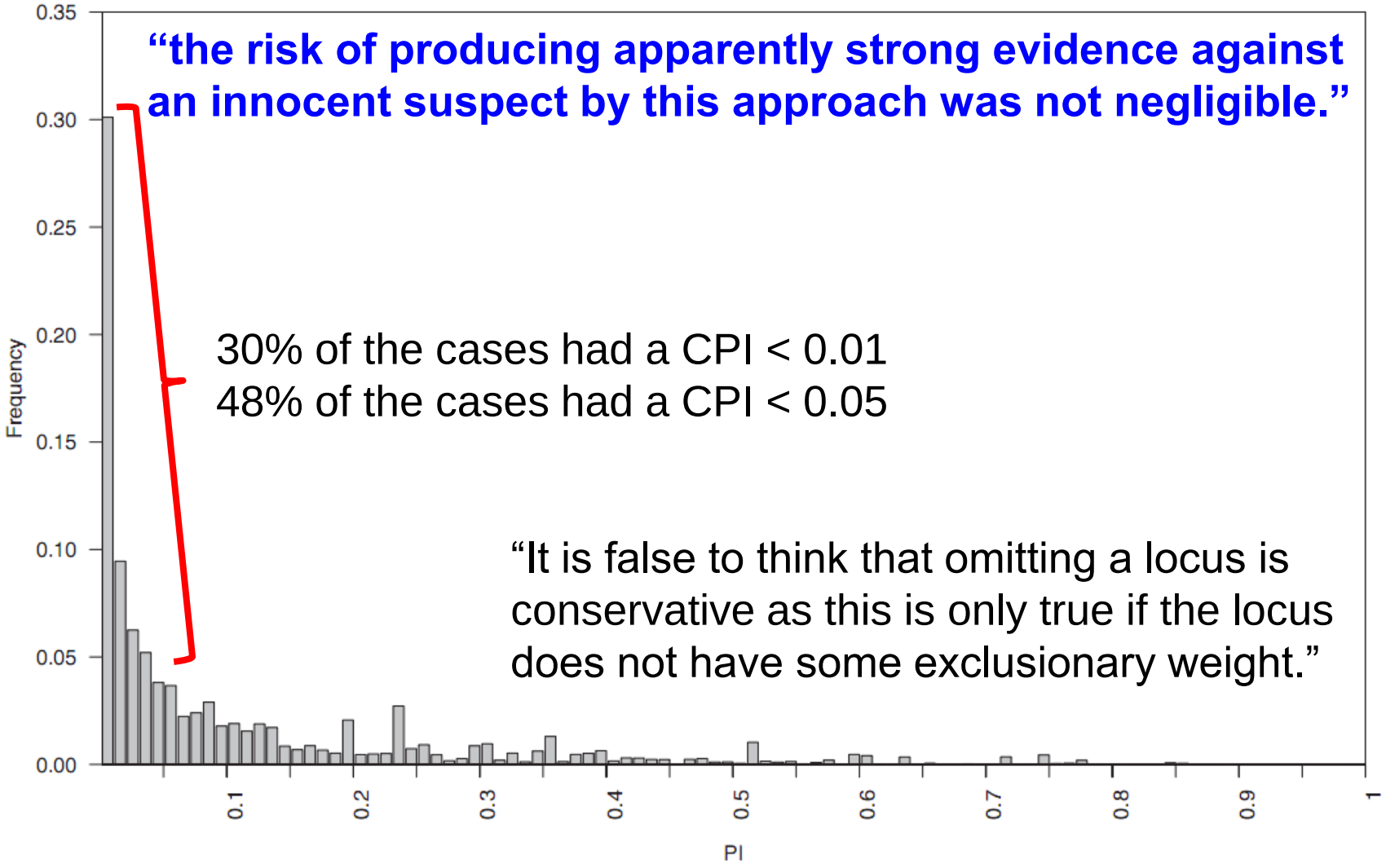
Inclusion Probabilities and Dropout

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)



Impact of Dropping Loci

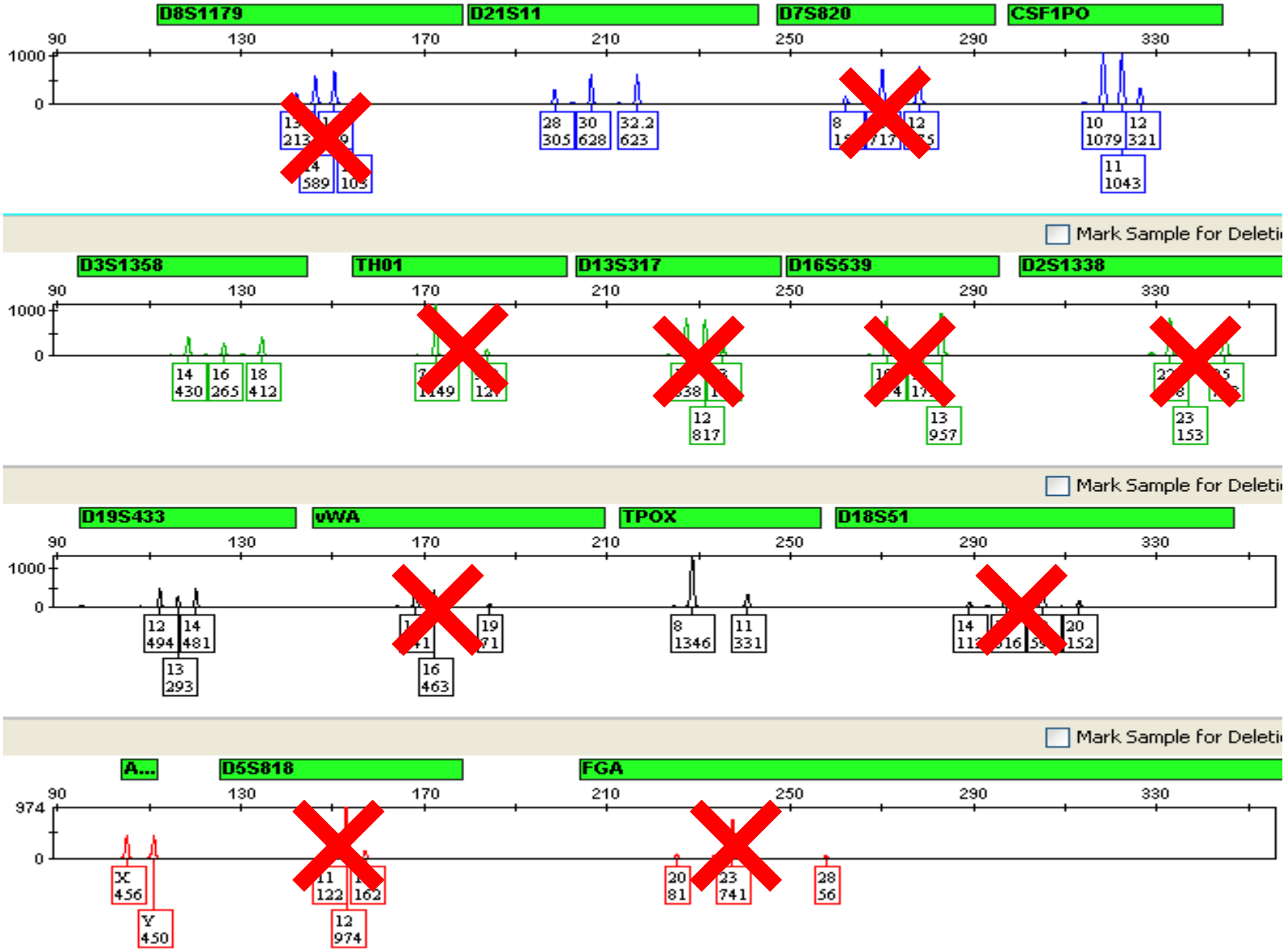
- The less data available for comparison purposes, the greater the chance of falsely including someone who is truly innocent

- Are you then being “conservative” (i.e., erring in favor of the defendant)?

If CPI/CPE Stats are Used

Since exclusionary statistics cannot adjust for the possibility of dropout, and does not take the number of contributors into account, any loci where alleles are below stochastic levels cannot be used in the CPI statistic.

If CPI/CPE Stats are Used



If CPI/CPE Stats are Used

Can use	Cannot use	
D21	D8	D2
CSF	D7	vWA
D3	TH01	D18
D19	D13	D5
TPOX	D16	FGA

If CPI/CPE Stats are Used

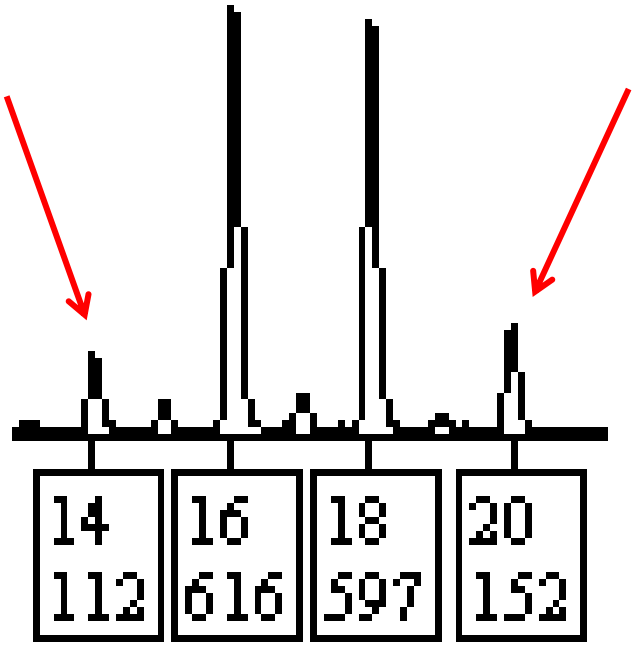
- CPI statistics using FBI Caucasian Frequencies

- 1 in 71 Caucasians included
- 98.59% Caucasians excluded

If RMP/LR Stats are Used

- Since there is an assumption to the number of contributors, it is possible to use data that falls below the ST.

RMP - D18S51



If Assume 2 Contributors….

Major	Minor
16,18	14,20

$RMP_{minor} = 2pq$

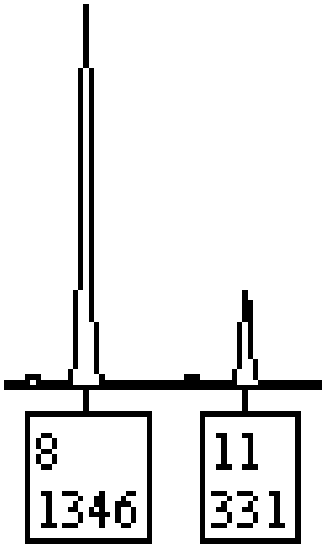
$= 2 \times f(14) \times f(20)$

$= 2 \times (0.1735) \times (0.0255)$

$= 0.00884$ or 1 in 113

(LR = 113)

RMP - TPOX



If Assume 2 Contributors….

Major	Minor
8,8	11,8 *OR* 11,11

RMP = 8,11 + 11,11

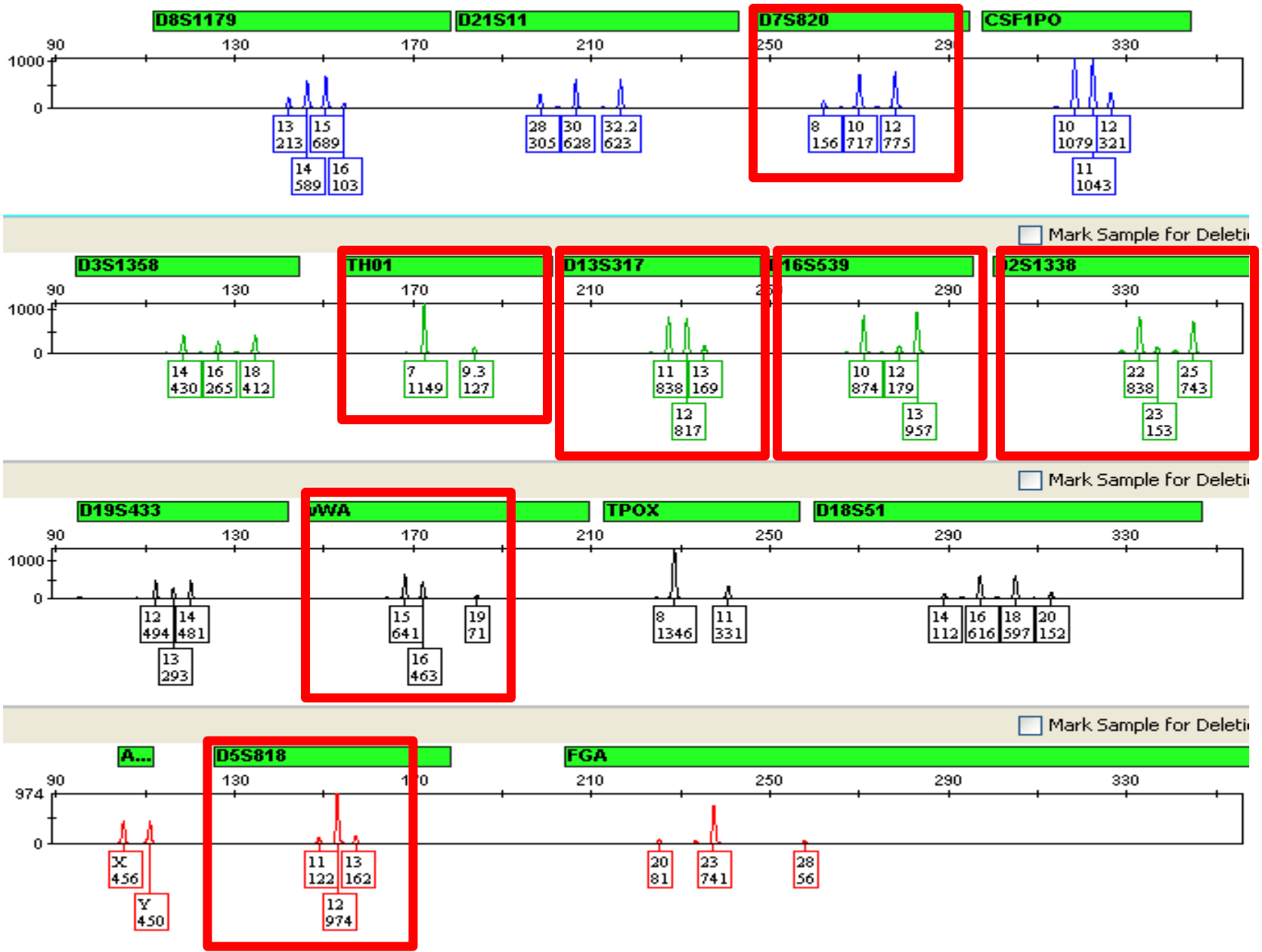
RMP = $2pq + (q^2 + q(1-q)\theta)$

RMP = $2(0.5443)(0.2537) + (0.2537)^2 + (0.2537)(0.7463)(0.01)$

= 0.3424 or 1 in 2.9

RMP/LR

Profile 1: ID_2_SCD_NG0.5_R4,1_A1_V1.2



If RMP/LR Stats are Used

Can use

D8

D21

D18

D3

D19

TPOX

FGA

CSF

Loci with potential D-out

D7

TH01

D13

D16

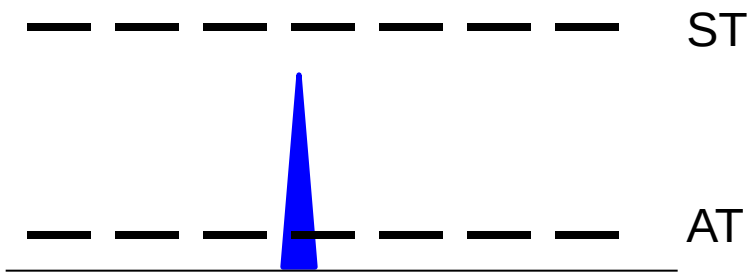
D2

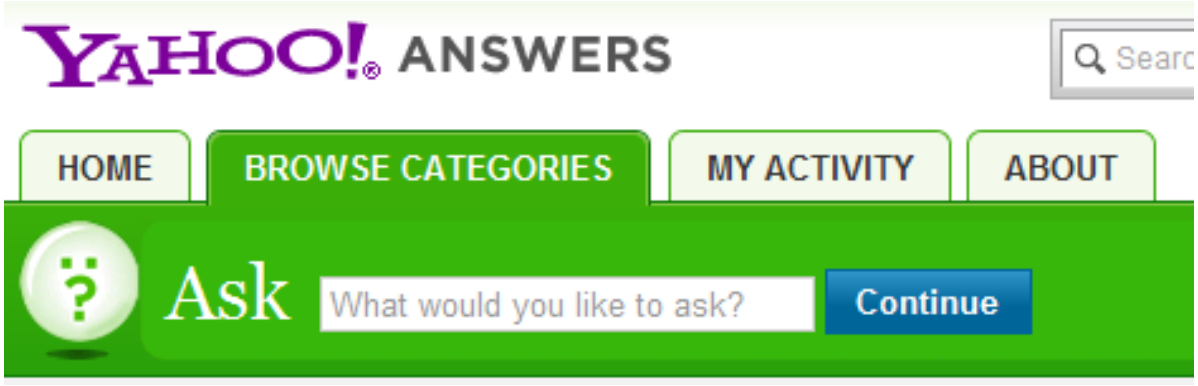
vWA

D5

The “2p” Rule

- The “2p” rule can be used to statistically account for zygosity ambiguity – i.e. is this single peak below the stochastic threshold the result of a homozygous genotype or the result of a heterozygous genotype with allele drop-out of the sister allele?





Resolved Question

Show me another »

To pee or not to pee? That is the question...?

"Drink sir, is a great provoker of three things…. nose painting, sleep and urine."

Macbeth: Act 2, Scene 3

http://answers.yahoo.com/question/index?qid=20100419211523AA8pQEJ

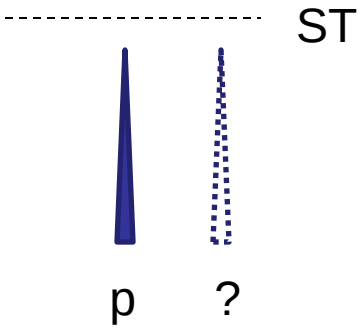
2p – SWGDAM Guidelines

- 5.2.1.3.1. The formula 2p, as described in recommendation 4.1 of NRCII, may be applied to this result.

- 5.2.1.3.2. Instead of using 2p, the algebraically identical formulae **$2p - p^2$** and **$p^2 + 2p(1-p)$** may be used to address this situation without double-counting the proportion of homozygotes in the population.

$2p - p^2$ and $p^2 + 2p(1-p)$

Suppose 5 allele system – P, Q, R, S & T



The possible genotype could be anything…

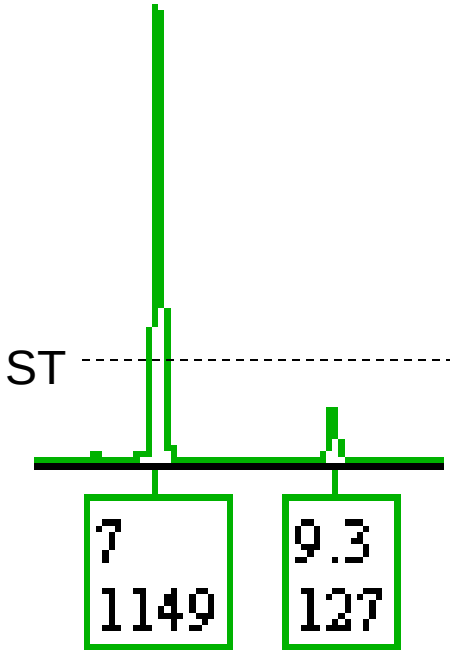
= PP + PQ + PR + PS + PT

= p^2 + 2pq + 2pr + 2ps + 2pt

= p^2 + 2p (q + r + s + t) ⟶ = (1-p)

= **p^2 + 2p (1-p)** = p^2 + 2p -$2p^2$ = **2p - p^2**

Profile 1 - TH01



Major – 7, 7

Possible Minor Contributors

7, 9.3 (2pq)

9.3, 9.3 p^2

9.3, ? 2p (or $p^2 + 2p(1 - p)$)

Profile 1 - TH01 (LR)

$$\frac{P(E \mid H_1)}{P(E \mid H_2)} = \frac{V \,\&\, S}{V \,\&\, U} = \frac{f_7^2 + f_7(1-f_7)\theta \,\&\, 1}{f_7^2 + f_7(1-f_7)\theta \,\&} \qquad p^2 + 2p(1-p)$$

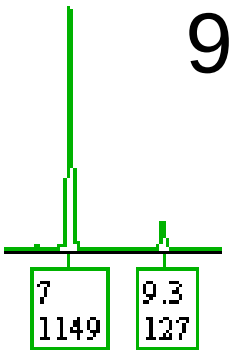
V = 7, 7

U = 7, 9.3

9.3, 9.3

9.3, ?

$$= \frac{1}{f_{9.3}^2 + 2f_{9.3}(1-f_{9.3})}$$



$f_{9.3} = 0.3054$

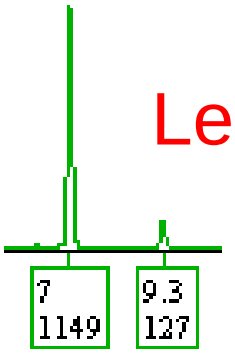
$= 1 / 0.5175 = \boxed{1.93}$

Profile 1 - TH01 (LR)

$$\frac{P(E \mid H_1)}{P(E \mid H_2)} = \frac{V \,\&\, S}{V \,\&\, U} = \frac{1}{p^2 + p(1-p)\theta + 2pq}$$

V = 7, 7
U = 7, 9.3
9.3, 9.3

$$= \frac{1}{f_{9.3}^2 + f_{9.3}(1-f_{9.3})\theta + 2f_{9.3}f_7}$$



Let ST = 125 RFU

$f_{9.3} = 0.3054$

$f_7 = 0.1724$

$= 1 / 0.2007 = \boxed{4.98}$

The “2p” Rule

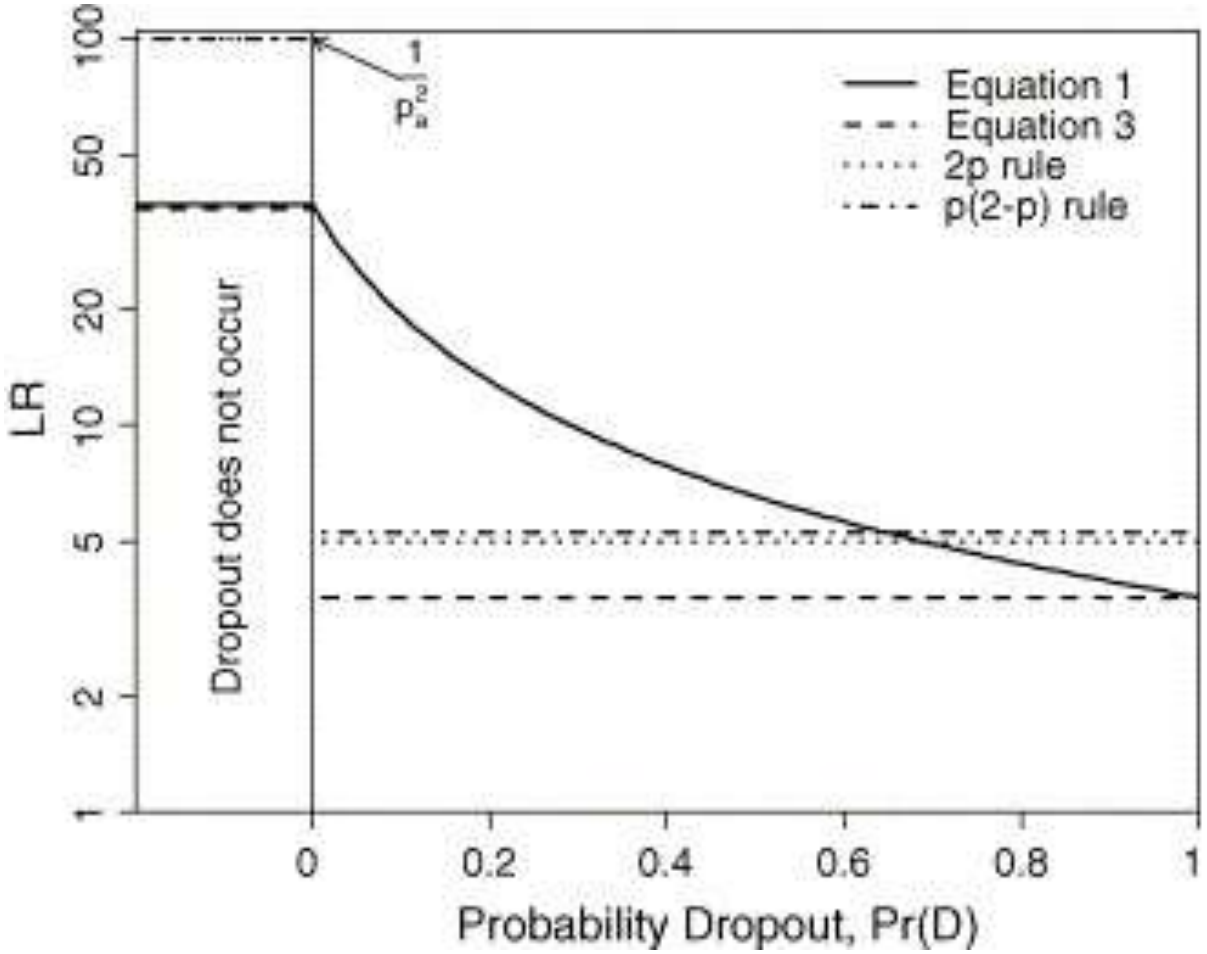
- “This rule arose during the VNTR era. At that time many smaller alleles “ran off the end of the gel” and were not visualised.”

- Buckleton and Triggs (2006)

“Is the 2*p* rule always conservative?”

The “2p” Rule

LR = 100



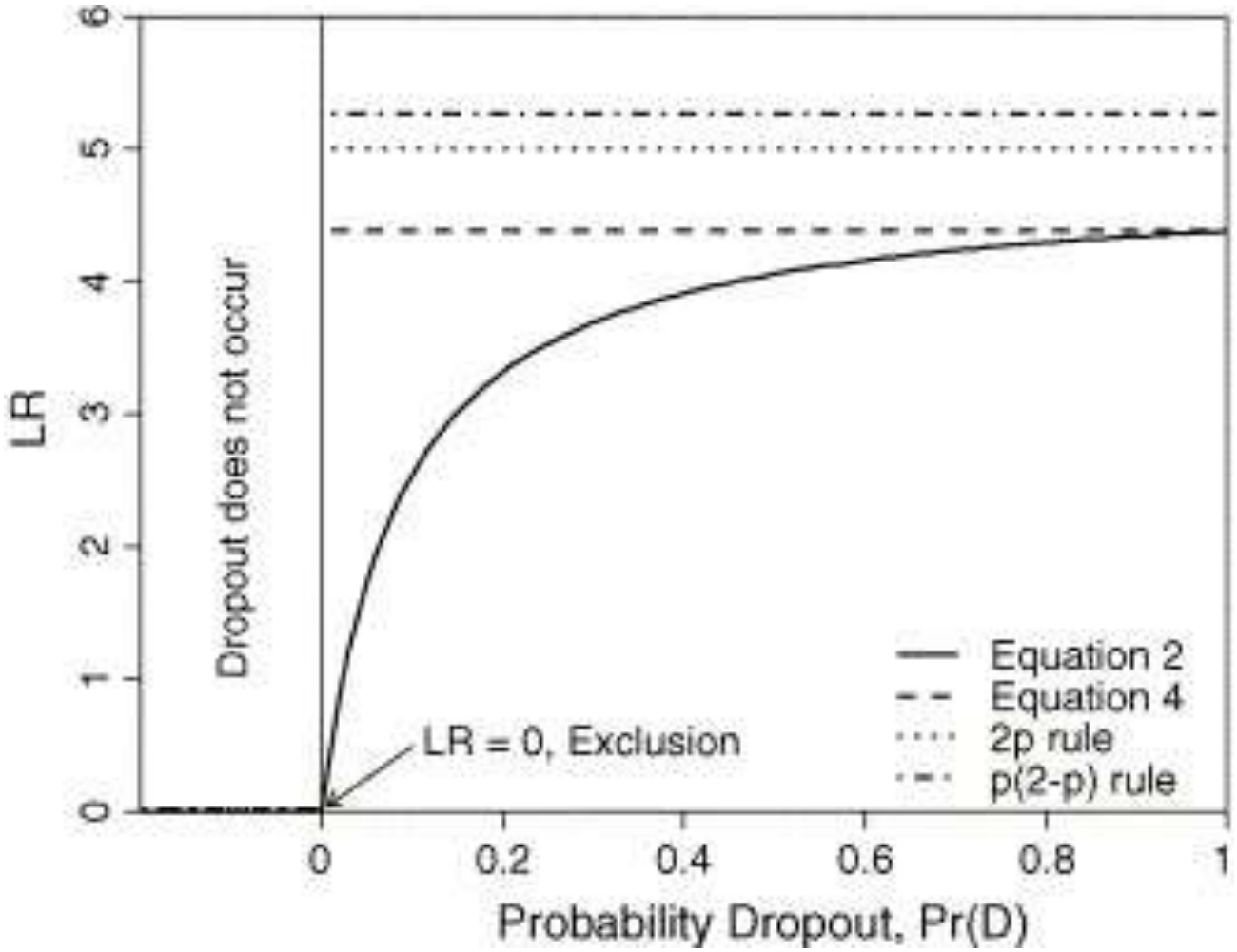
Stain = aa

Suspect = aa

LR = 4

f(a) = 0.10

The “2p” Rule

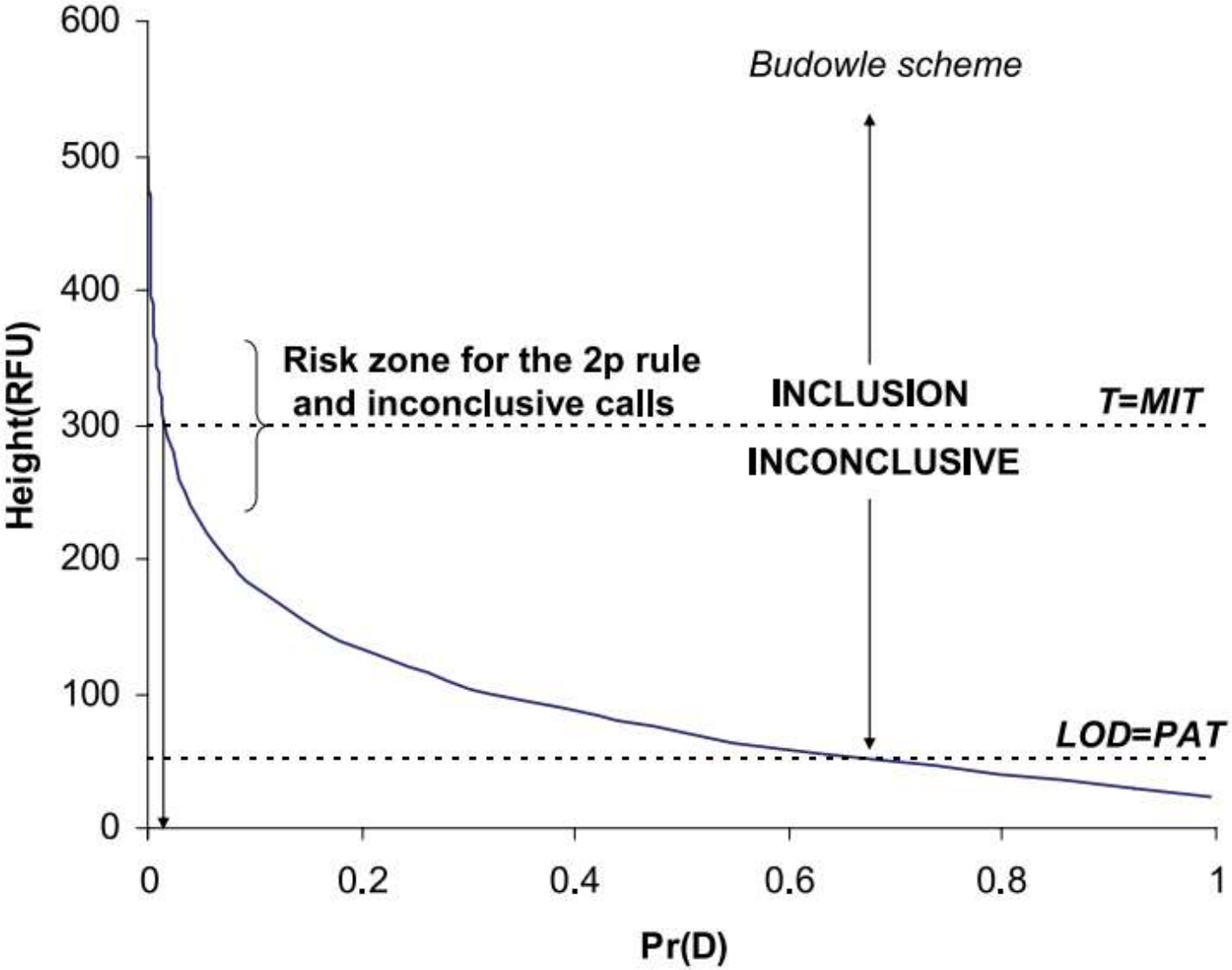


LR = 4

Stain = aa

Suspect = ab

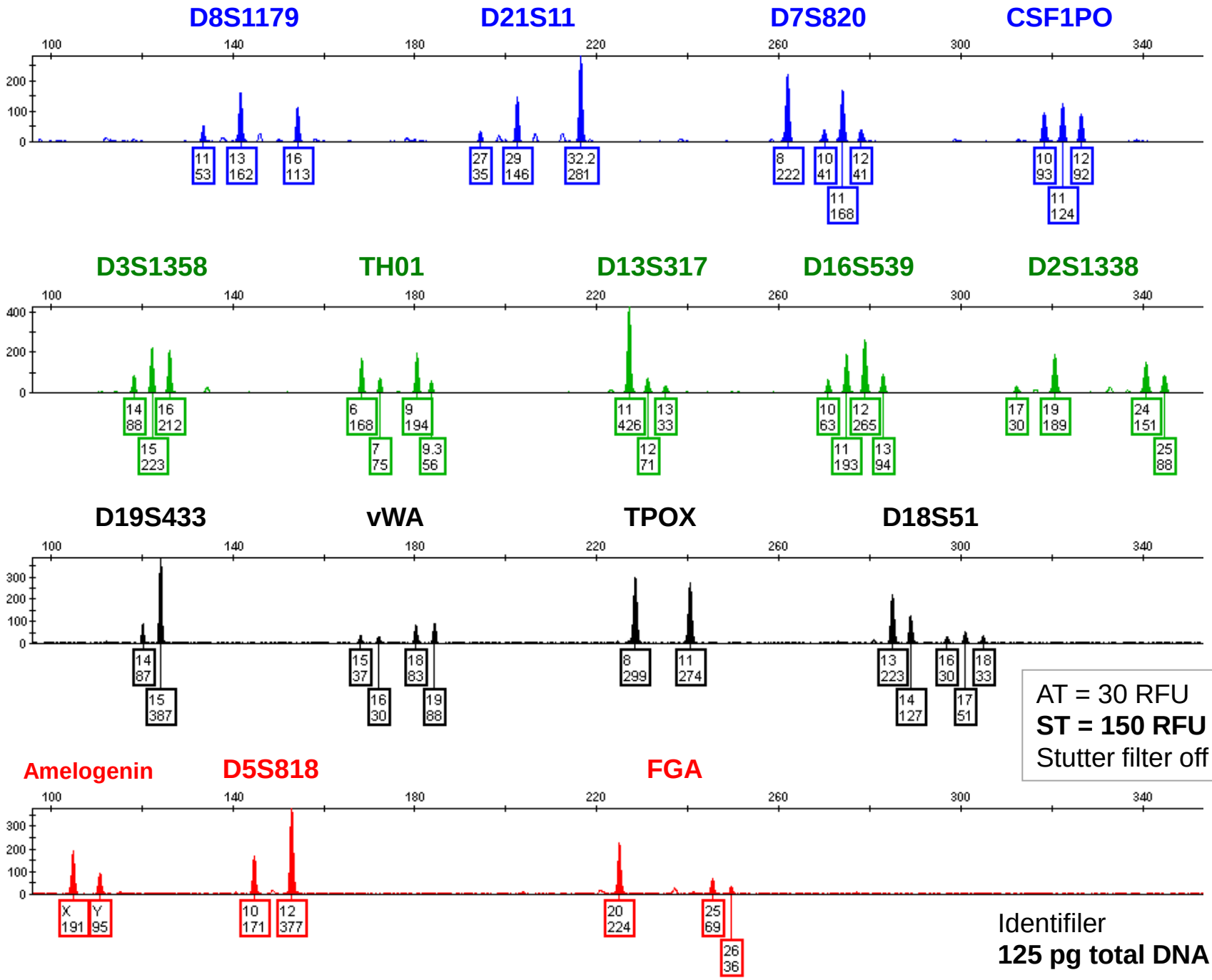
Gill and Buckleton (2010)



ST

AT

Challenges with low level, complex mixtures


D8S1179
D21S11
D7S820
CSF1PO
D3S1358
TH01
D13S317
D16S539
D2S1338
D19S433
vWA
TPOX
D18S51
Amelogenin
D5S818
FGA
AT = 30 RFU
ST = 150 RFU
Stutter filter off
Identifiler
125 pg total DNA

Clayton et al. (1998)
ISFG (2006) Rec. #4

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 Contributors

Step #5

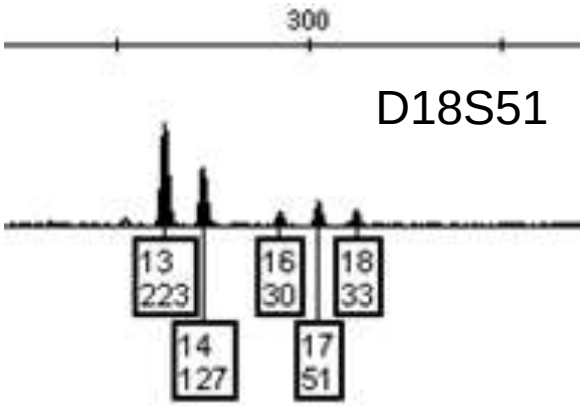
Consider All Possible Genotype Combinations

Step #6

Compare Reference Samples

Impact of Results with Low Level DNA

When amplifying low amounts of DNA (e.g., 125 pg), allele dropout is a likely possibility leading to **higher uncertainty** in the potential number of contributors and in the possible genotype combinations



Identifiler

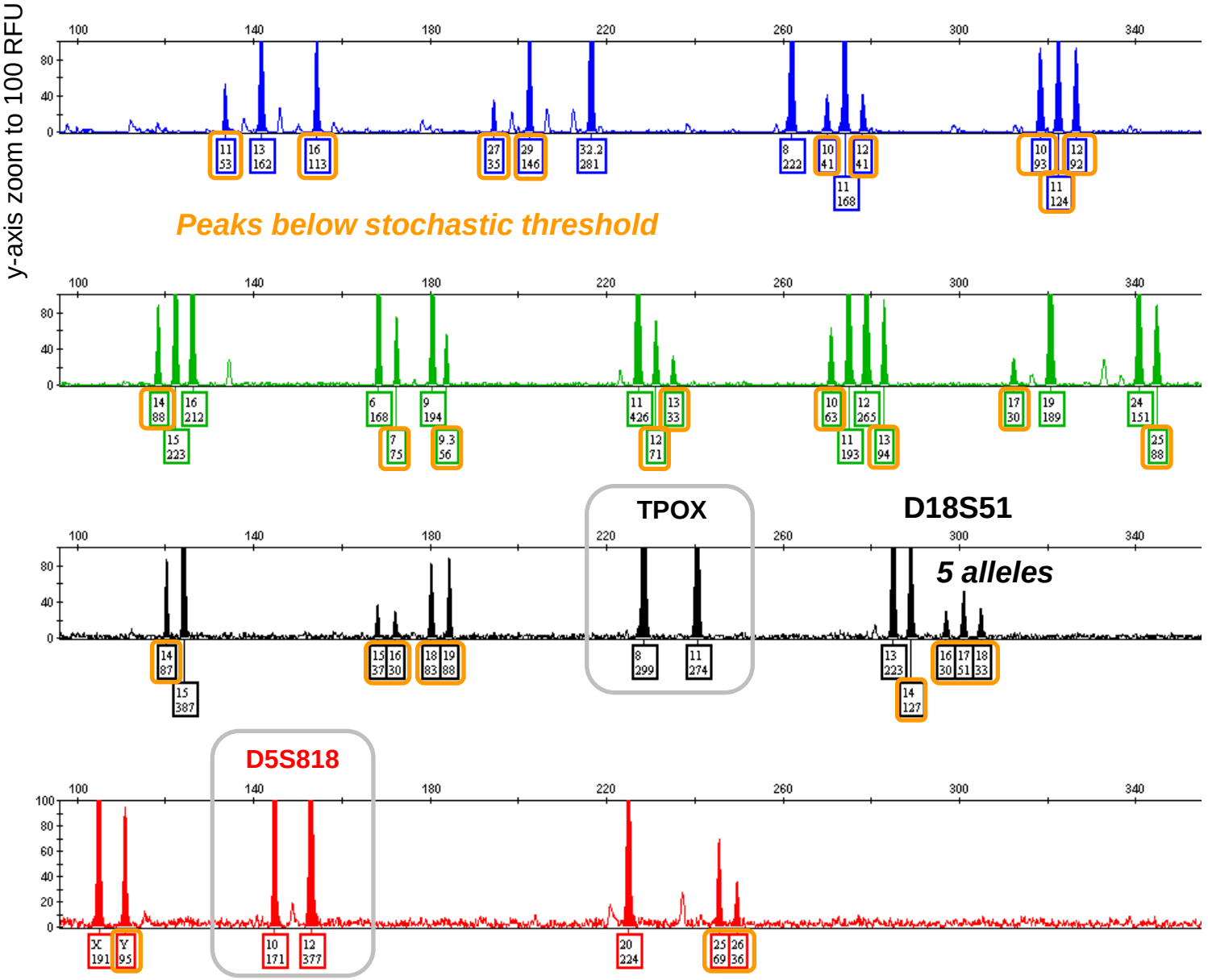
125 pg total DNA

Complex Mixture

AT = 30 RFU

ST = 150 RFU

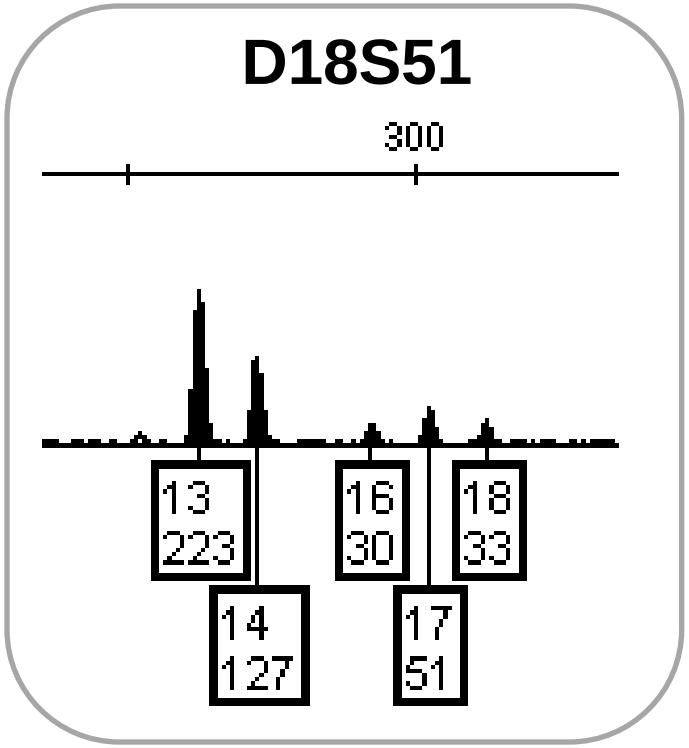
Stutter filter off



What Can We Say about this Result?

- Low level DNA (only amplified 125 pg total DNA)
  - likely to exhibit stochastic effects and have allele dropout
- Mixture of at least 3 contributors
  - Based on detection of 5 alleles at D18S51
  - If at equal amounts, ~40 pg of each contributor (if not equal, then less for the minor contributors); **we expect allele dropout**
- At least one of the contributors is male
  - Based on presence of Y allele at amelogenin
- Statistics if using CPI/CPE
  - Would appear that we can only use TPOX and D5S818 results with a stochastic threshold of 150 RFU *(will explore this further)*
- **Due to potential of excessive allele dropout, we are unable to perform any meaningful Q-K comparisons**

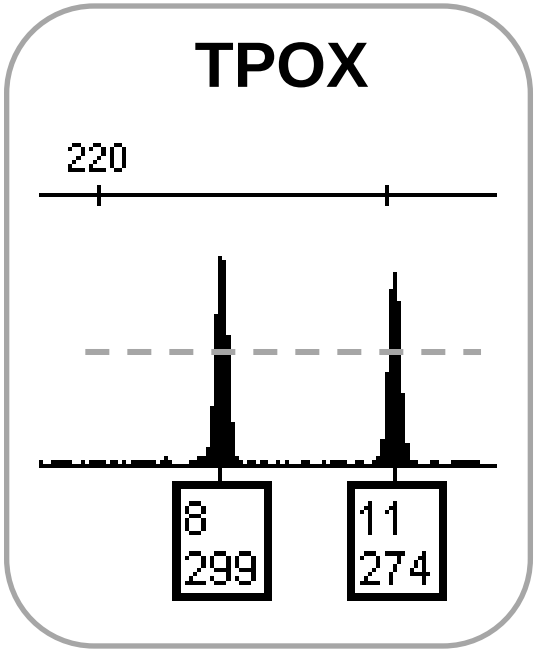
Uncertainty in the Potential Number of Contributors with this Result



5 alleles observed

- Several of the peaks are barely above the analytical threshold of 30 RFU
  - In fact, with an analytical threshold of 50 RFU or even 35 RFU, there would only be three detected alleles at D18S51
- Stochastic effects could result in a high degree of stutter off of the 17 allele making alleles 16 and 18 potential stutter products
- No other loci have >4 alleles detected

All Detected Alleles Are Above the Stochastic Threshold – **Or Are They?**



Stochastic threshold = 150 RFU

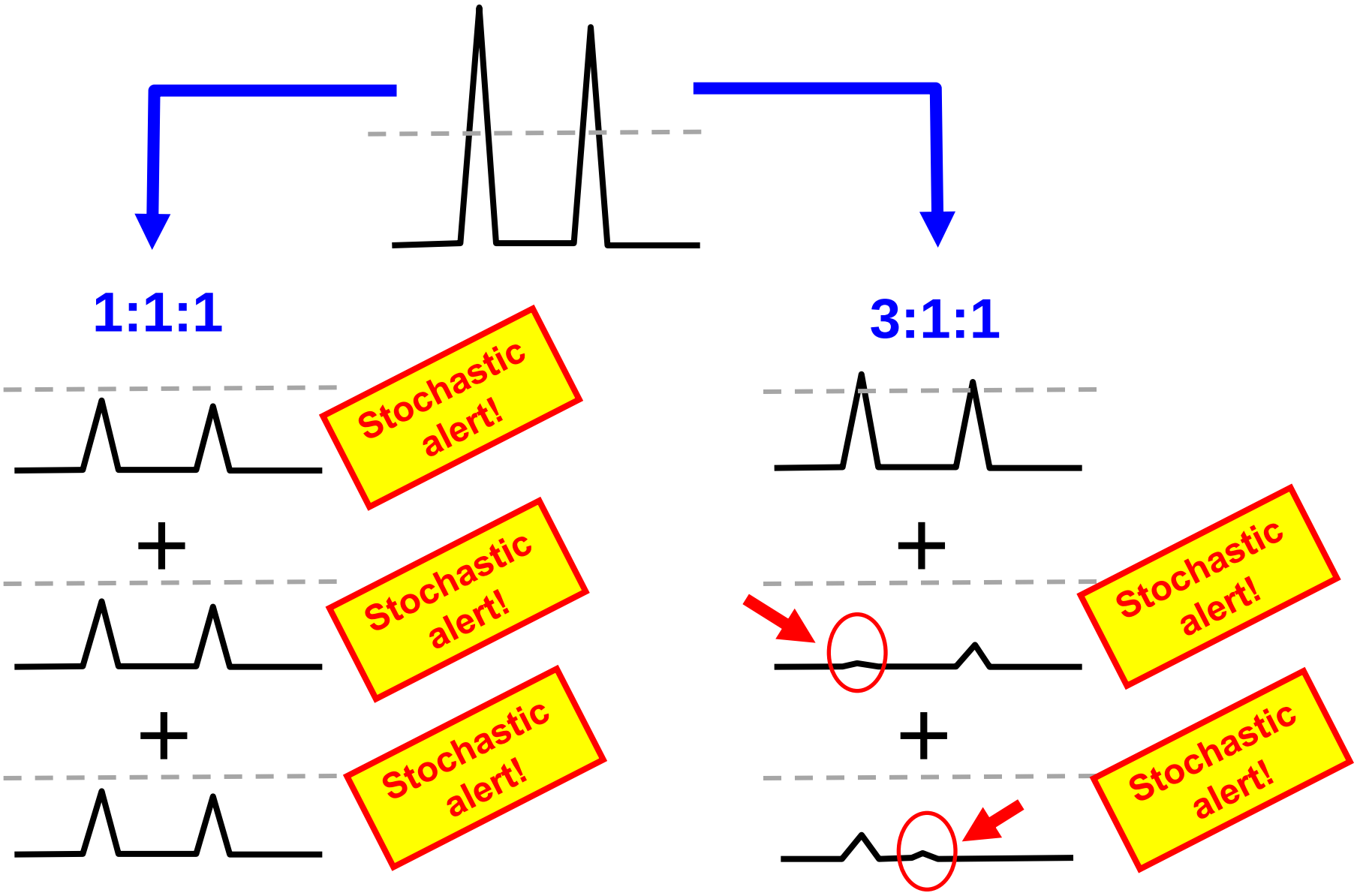
Does this result guarantee no allele drop-out?

We have assumed three contributors. If result is from an equal contribution of 3 individuals…

Then some alleles from individual contributors would be below the stochastic threshold and we could not assume that all alleles are being observed!

Assuming Three Contributors…

Some Possible Contributions to This Result

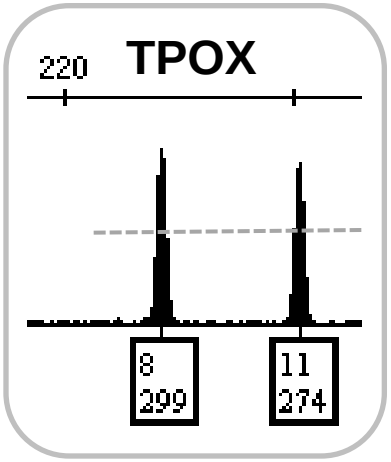


All Loci Are Not Created Equal
when it comes to mixture interpretation

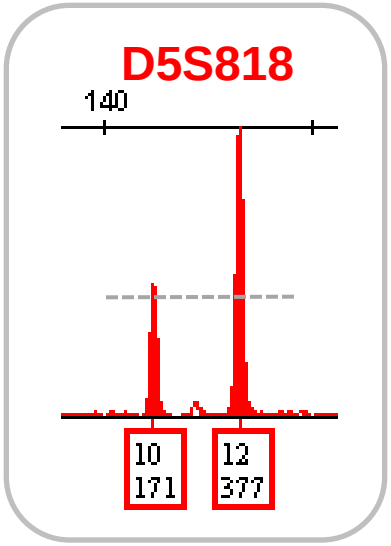
- In the case of less polymorphic loci, such as TPOX, there are fewer alleles and these occur at higher frequency. Thus, there is a greater chance of allele sharing (peak height stacking) in mixtures.

- **Higher locus heterozygosity is advantageous for mixture interpretation** – we would expect to see more alleles (within and between contributors) and thus have a better chance of estimating the true number of contributors to the mixture

Even if you did attempt to calculate a CPI/CPE statistic using loci with all observed alleles above the stochastic threshold on this result…



TPOX Allele Frequencies (NIST Caucasian, Butler et al. 2003)
8 = 0.53
11 = 0.24
CPI = $(0.53 + 0.24)^2$ = 0.59 or **59%**



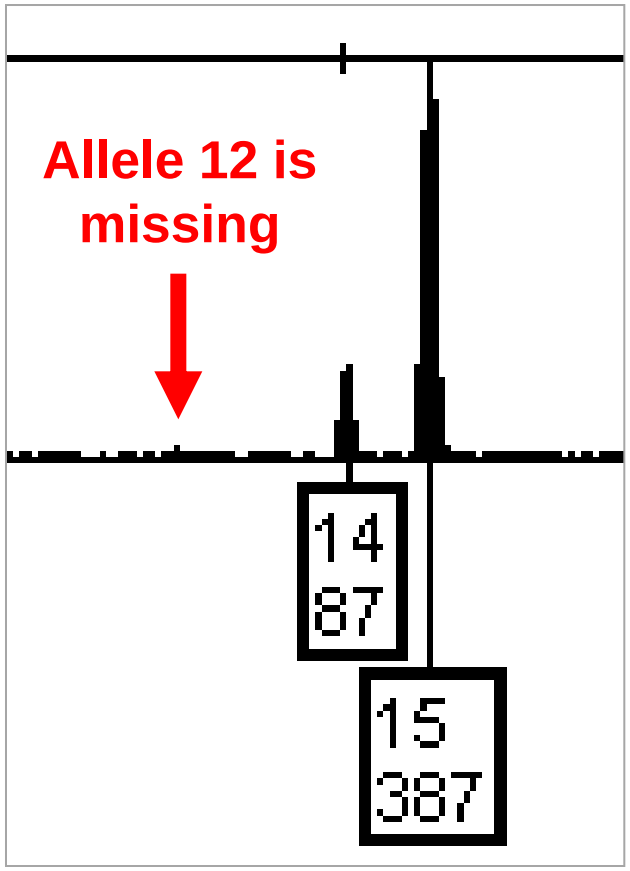
Combine loci = 0.59 x 0.18 = 0.11 or **11%**

Approximately 1 in every 9 Caucasians could be included in this mixture

D5S818 Allele Frequencies (NIST Caucasian, Butler et al. 2003)
10 = 0.05
12 = 0.38
CPI = $(0.05 + 0.38)^2$ = 0.18 or **18%**

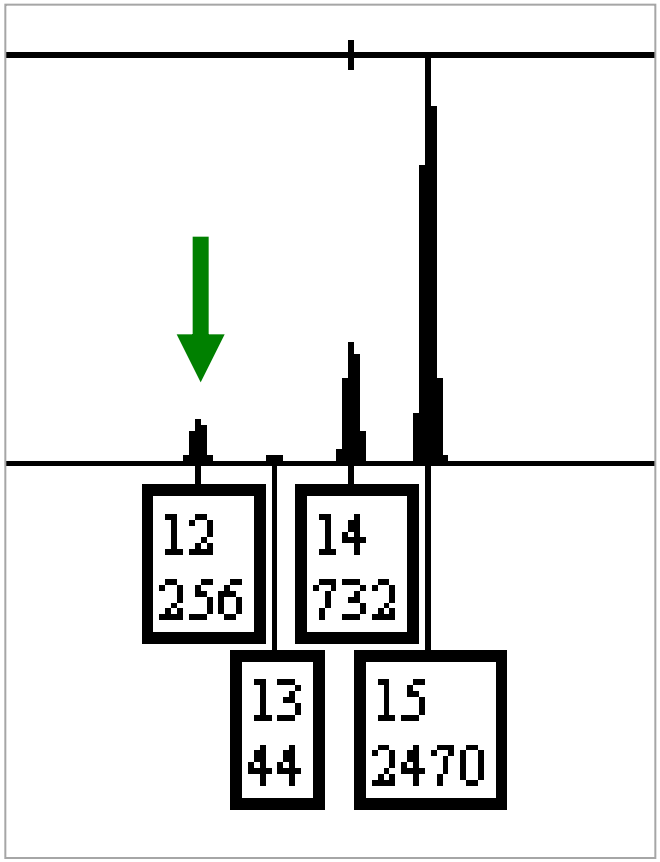
Impact of Amplifying More DNA

D19S433



125 pg total DNA amplified

D19S433



500 pg total DNA amplified

True Contributors

3 contributors
with a 2:1:1 mixture

15,15 (2x)
14,15 (1x)
12,14 (1x)

How should you handle the suspect comparison(s) with this case result?

- **No suspect comparisons should be made** as **the mixture result has too much uncertainty** with stochastic effects that may not account for all alleles being detected

- **Declare the result "inconclusive"**

How not to handle this result

- "To heck with the analytical and stochastic thresholds", **I am just going to see if the suspect profile(s) can fit into the mixture allele pattern observed** – and then if an allele is not present in the evidentiary sample try to explain it with possible allele dropout due to stochastic effects

- This is what Bill Thompson calls "painting the target around the arrow (matching profile)…"

Thompson, W.C. (2009) Painting the target around the matching profile: the Texas sharpshooter fallacy in forensic DNA interpretation. *Law, Probability and Risk* 8: 257-276

What to do with low level DNA mixtures?

- **German Stain Commission "Category C"**
  (Schneider et al. 2006, 2009)
  - Cannot perform stats because stochastic effects make it uncertain that all alleles are accounted for

- **ISFG Recommendations #8 & #9** (Gill et al. 2006)
  - Stochastic effects limit usefulness

- *Fundamentals of Forensic DNA Typing* (2010)
  Butler 3rd edition (volume 1), chapter 18
  - Don't go "outside the box" without supporting validation

ISFG Recommendations on Mixture Interpretation

http://www.isfg.org/Publication;Gill2006

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

A Complexity/Uncertainty Threshold

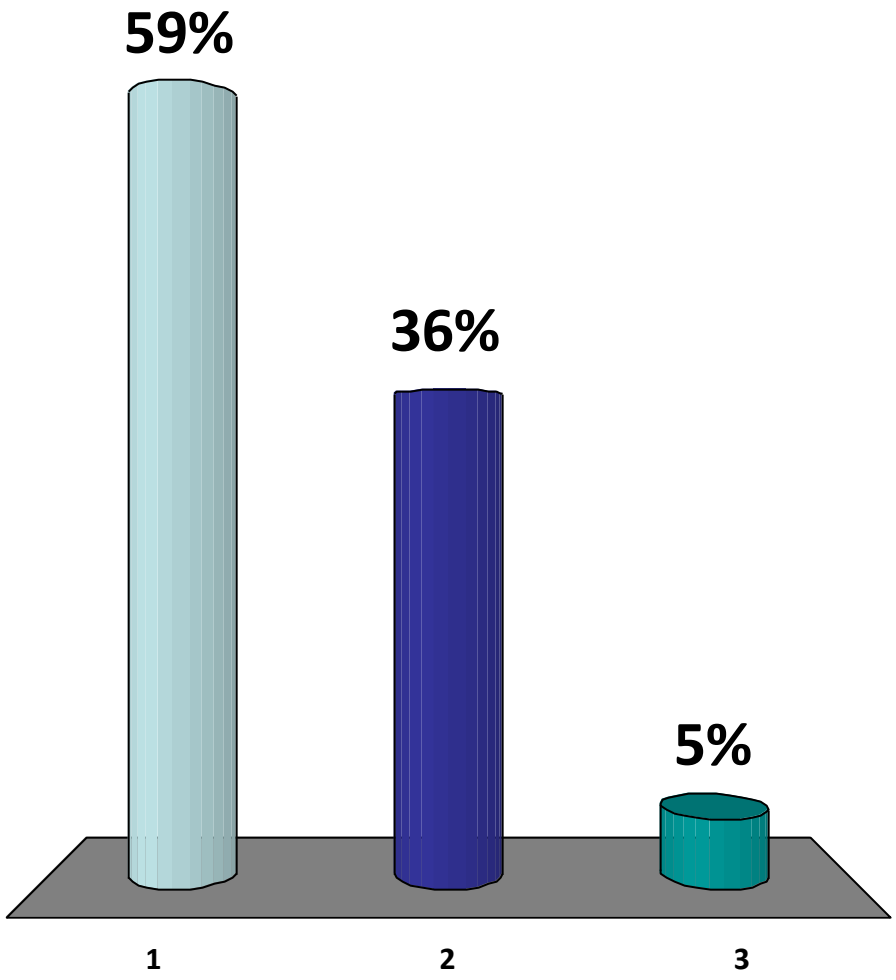
New Scientist article (August 2010)

- **How DNA evidence creates victims of chance**
  - 18 August 2010 by Linda Geddes
- From the last paragraph:
  - **In really complex cases, analysts need to be able to draw a line** and say "This is just too complex, I can't make the call on it," says Butler. "Part of the challenge now, is that every lab has that line set at a different place. But the honest thing to do as a scientist is to say: **I'm not going to try to get something that won't be reliable**."

http://www.newscientist.com/article/mg20727743.300-how-dna-evidence-creates-victims-of-chance.html

Has your laboratory implemented a “stop testing” approach with complex and/or low-level mixture?

1. Yes
2. No
3. I don’t work in a lab



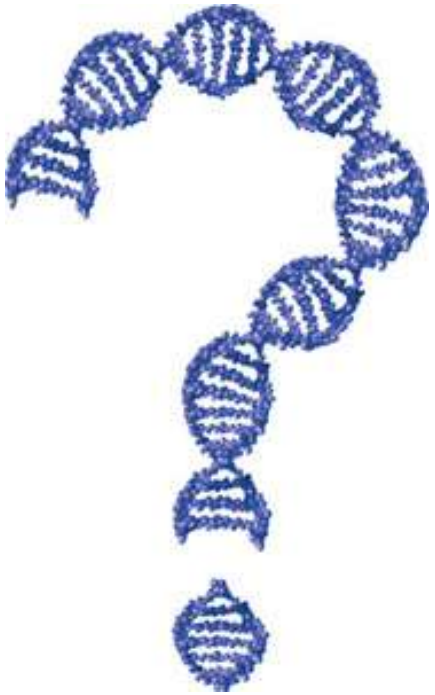
Is there a way forward?

Thank You!

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

Funding from the **National Institute of Justice (NIJ)** through NIST Office of Law Enforcement Standards

Questions?



john.butler@nist.gov
301-975-4049

michael.coble@nist.gov
301-975-4330