DNA Mixture Interpretation

John M. Butler, Ph.D.

National Institute of Standards and Technology

Gaithersburg, Maryland

Acknowledgments and Disclaimers

Funding for research and training on forensic DNA performed by the NIST Applied Genetics Group has come from the National Institute of Justice and the NIST Law Enforcement Standards Office

Although I chaired the SWGDAM Mixture Committee that produced the 2010 STR Interpretation Guidelines**, I cannot speak for or on behalf of the Scientific Working Group on DNA Analysis Methods**

Points of view are mine and do not necessarily represent the official position or policies of the US Department of Justice or the National Institute of Standards and Technology.

Certain commercial equipment, instruments and materials are identified in order to specify experimental procedures as completely as possible. In no case does such identification imply a recommendation or endorsement by the National Institute of Standards and Technology nor does it imply that any of the materials, instruments or equipment identified are necessarily the best available for the purpose.

Steps in Forensic DNA Testing

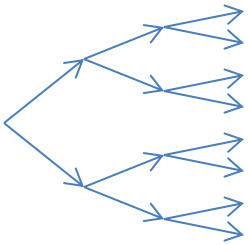
Collection/Storage/ Characterization → **Extraction/ Quantitation** → **Amplification/ Marker Sets** → **Separation/ Detection** → **Interpretation** → **Report**

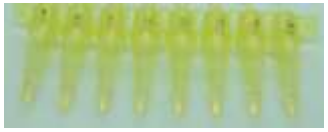
Mixture interpretation

Blood Stain Buccal swab

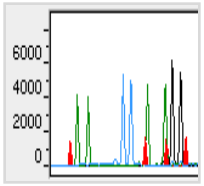
Sample Collection & Storage

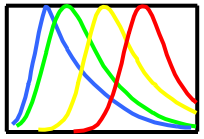
DNA Extraction & Quantitation



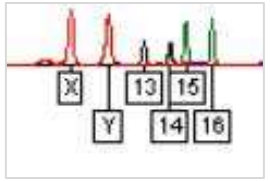


Multiplex PCR Amplification of STR Markers





CE with LIF Detection



Male: 13,14-15,16-…

Data Interpretation , Review & Reporting

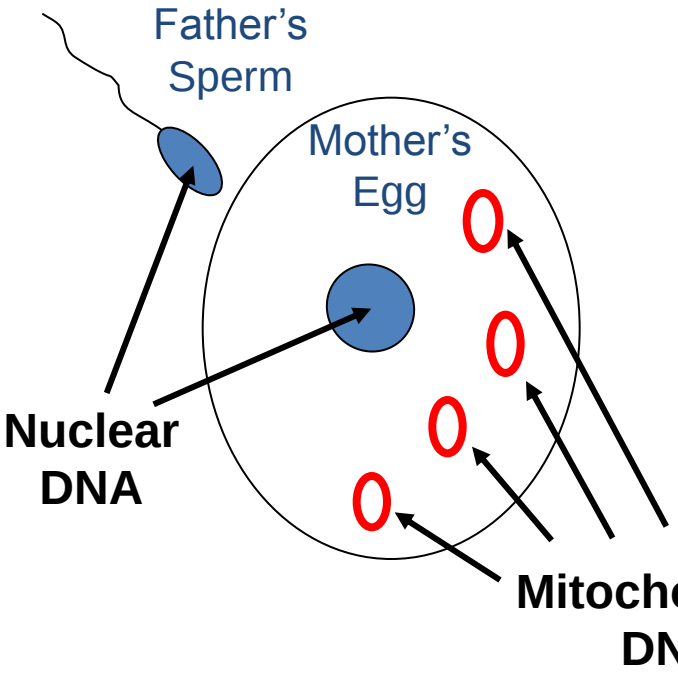
GeneAmp 9700 Thermal Cycler

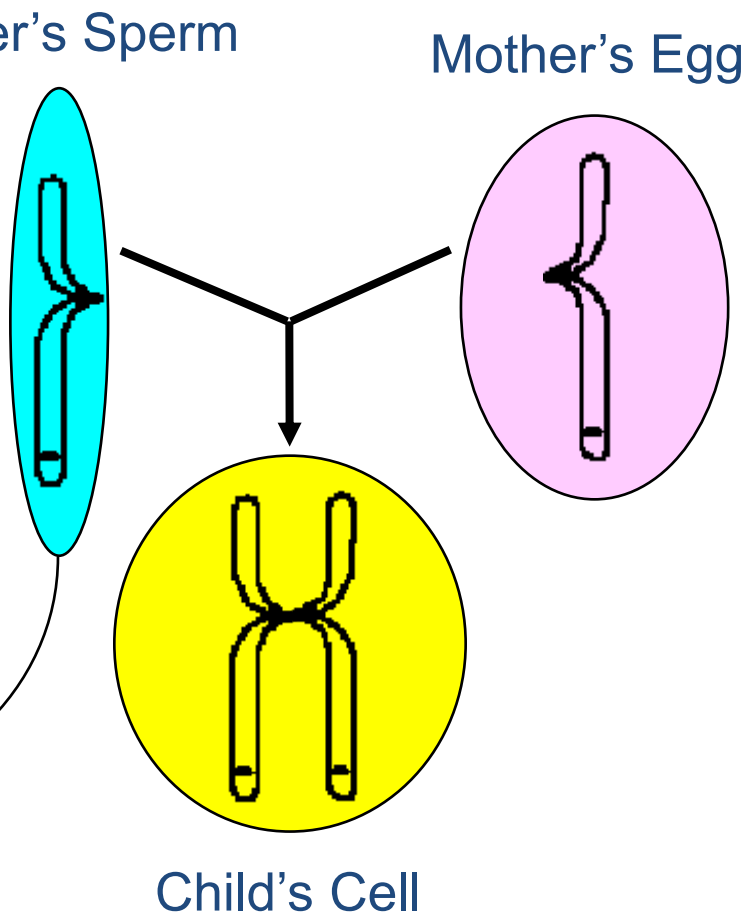
ABI 3500 Genetic Analyzer

GeneMapper*ID-X* software

capillary electrophoresis

Genetic Inheritance





Current scientific thinking:

- ~**99.9%** of 6 billion letters (2 x 3 billion bp) are **the same between people**
- This 0.1% is still **~6 million differences**

Father contributes: 22 autosomes (1 of each pair), X or **Y**

Mother contributes: 22 autosomes (1 of each pair), X and **mtDNA**

Punnett Square Showing Possible Genotype Combinations (from Genetic Inheritance)

Parental Alleles →

Child Genotypes

father \ mother	A	B
A	AA	AB
B	AB	BB

Observed Data

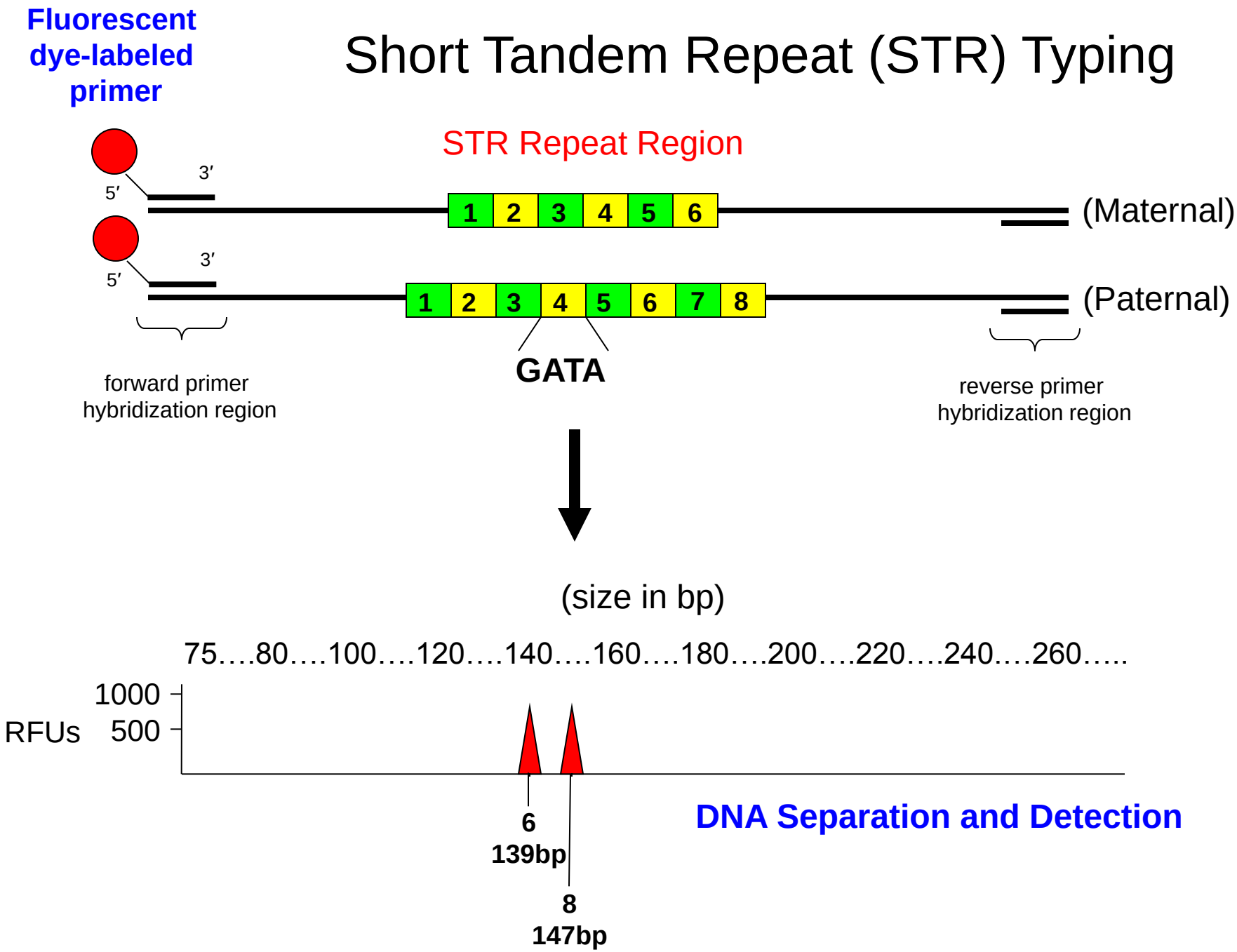
Allele Frequencies

father \ mother	p	q
p	p^2	pq
q	pq	q^2

pq + pq = **2pq**

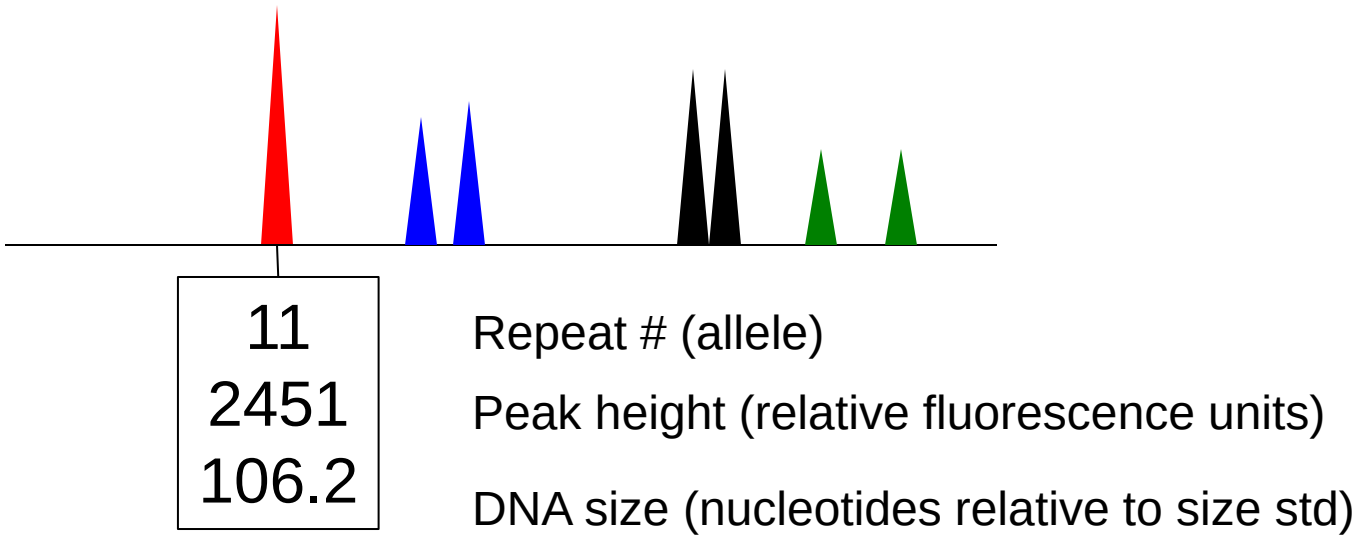
Calculated Statistics

Short Tandem Repeat (STR) Typing



Understanding an STR Electropherogram (E-gram; EPG)

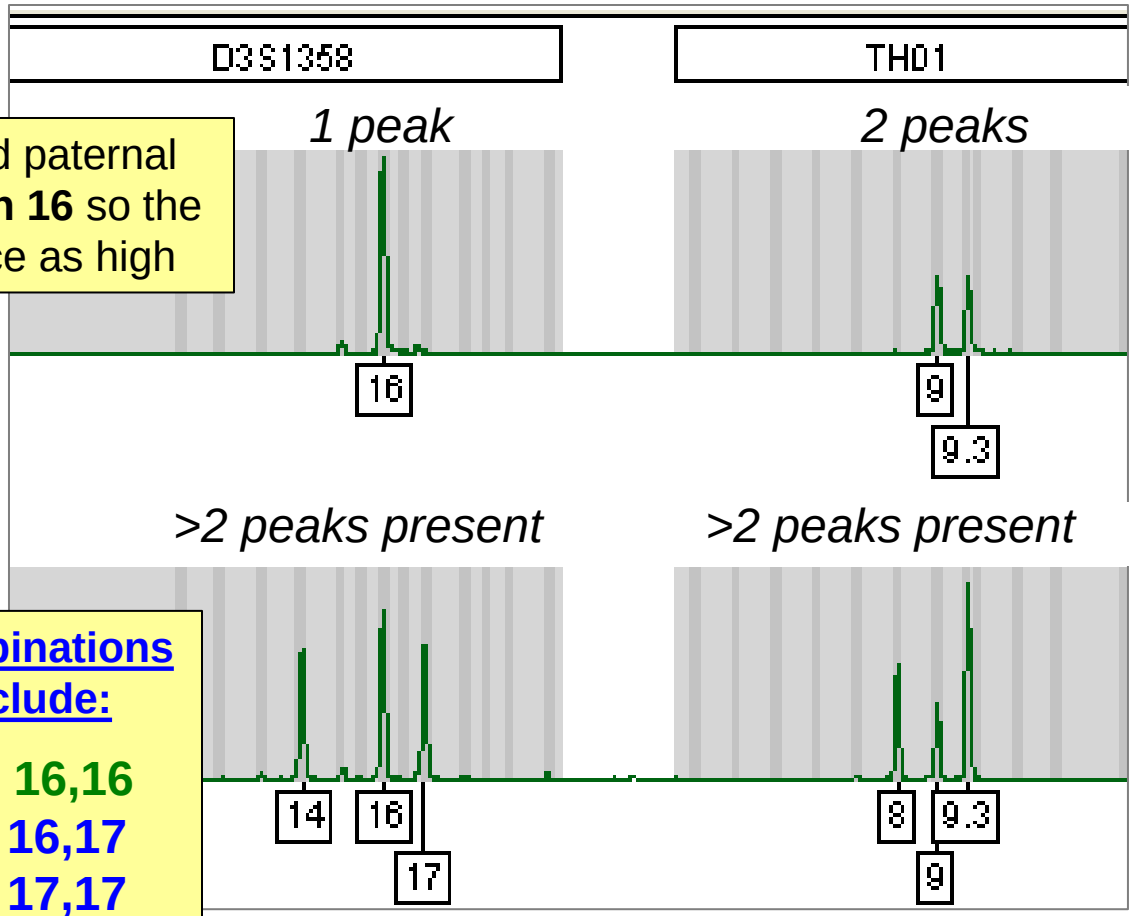
Peak height correlates to amount of DNA present (signal detected)
Peak position relates to the DNA size, which corresponds to STR allele repeat #
Peak color relates to the fluorescent dye label used to copy the specific DNA target



Alleles (peaks) are detected - but **Genotypes**, the specific combination of alleles, matter in terms of identifying individuals

Single-Source Sample vs Mixture Results

Maternal and paternal allele are **both 16** so the signal is twice as high



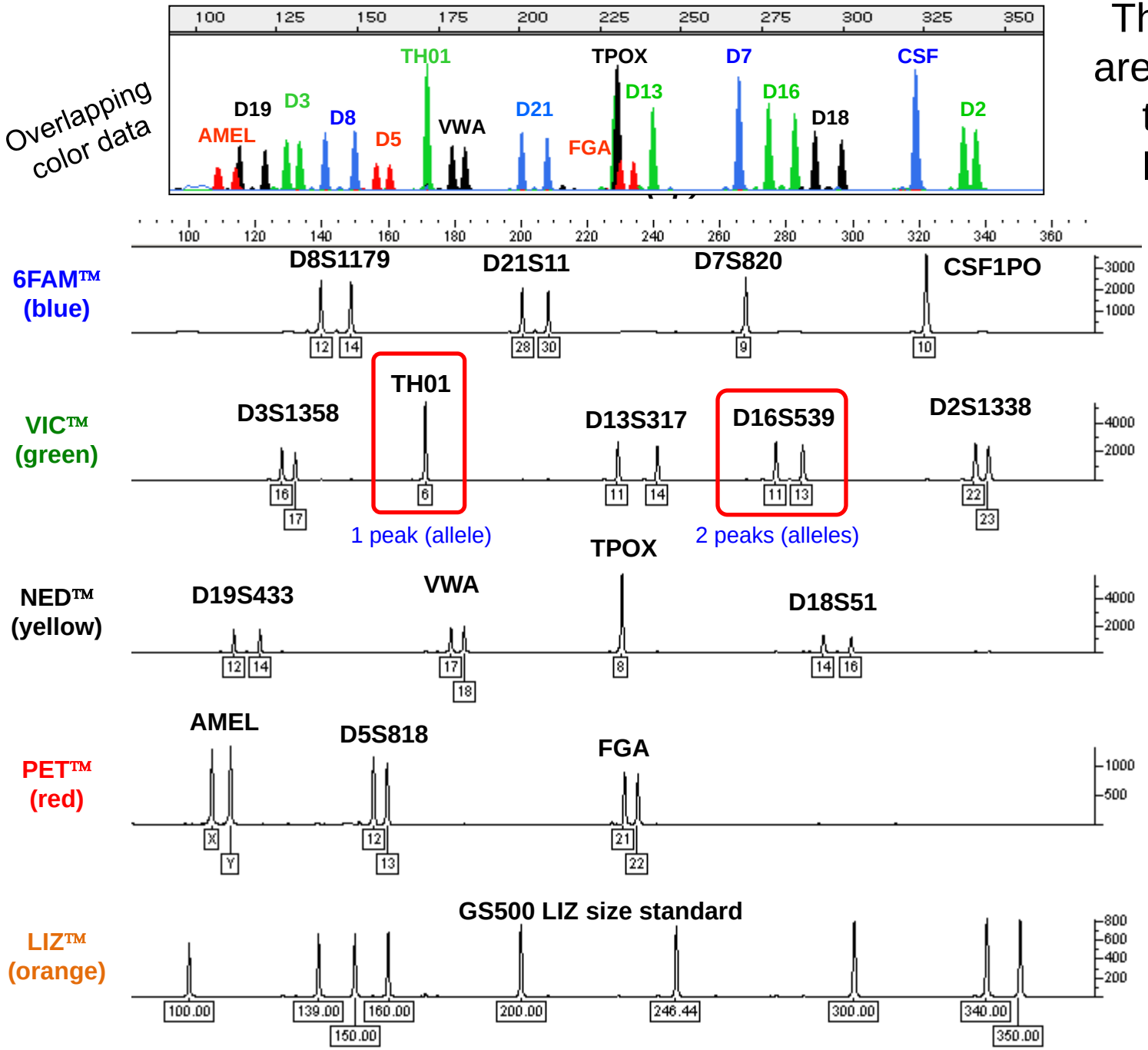
Single-Source

Mixture

Possible combinations at D3S1358 include:

14, 17 with 16,16
14,14 with 16,17
14,16 with 17,17

Multiple possible combinations could have given rise to the mixture observed here



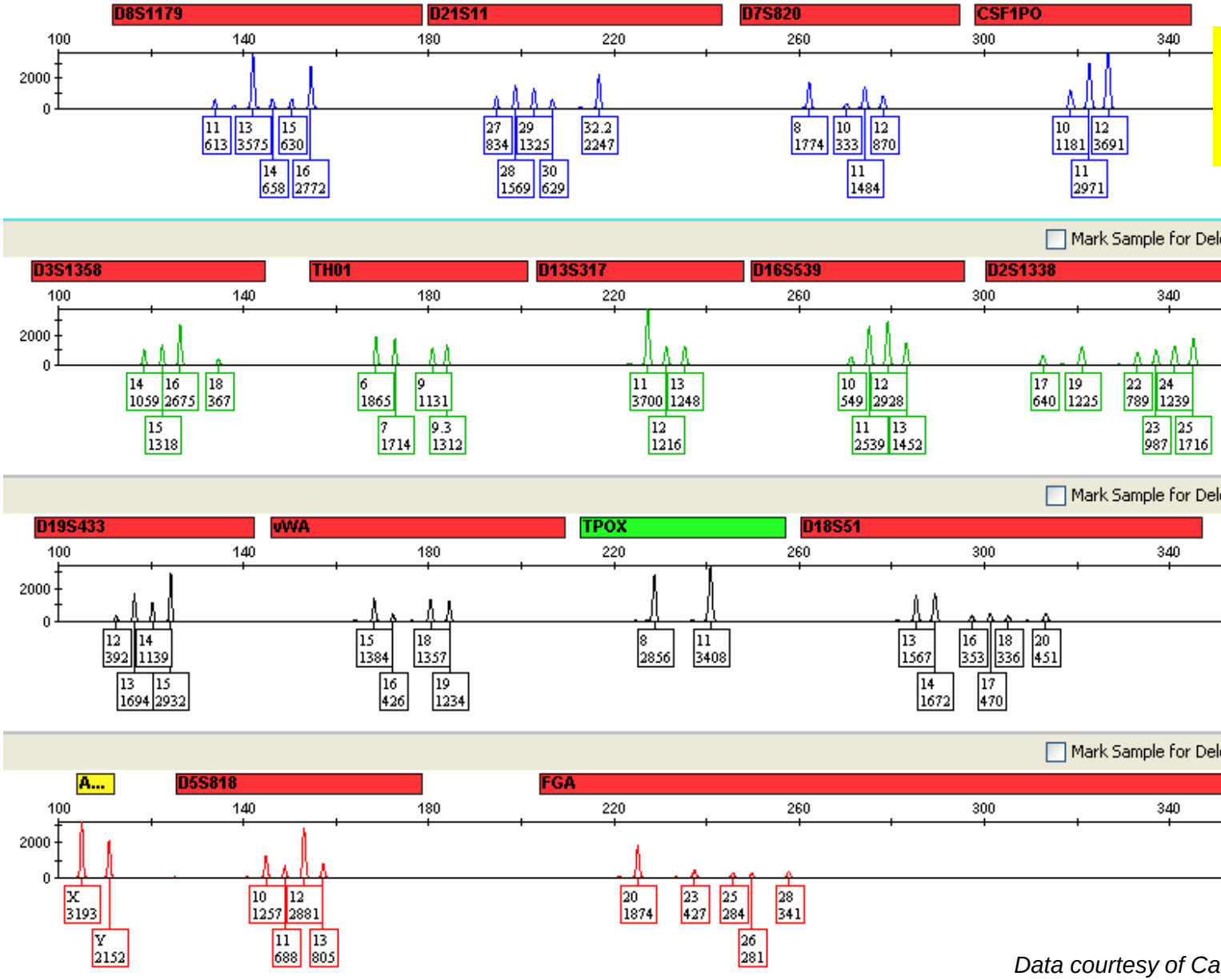
These results are from a DNA test called **Identifiler**

A **single-source** (reference) sample displays only 1 or 2 peaks per DNA site

Identifiler DNA test

DNA Mixture Result

Controlled mixture of **4 individuals**



More than two peaks per locus (DNA test site)

Data courtesy of Catherine Grgicak (Boston U.)

Different DNA Tests from Various STR Kits

Kit Name	# STR Loci Tested	Manufacturer	Why Used?
Identifiler, **Identifiler Plus***	**15 autosomal STRs** (aSTRs) & amelogenin	Life Technologies (Applied Biosystems)	Covers the 13 core CODIS loci plus 2 extra
PowerPlex 16 **PowerPlex 16 HS***	**15 aSTRs** & amelogenin	Promega Corporation	Covers the 13 core CODIS loci plus 2 extra
Profiler Plus & COfiler (2 different kits)	**13 aSTRs** [9 + 6 with 2 overlapping] & amelogenin	Life Technologies (Applied Biosystems)	Original kits used to provide 13 CODIS STRs
Yfiler	**17 Y-chromosome STRs**	Life Technologies (Applied Biosystems)	Male-specific DNA test
MiniFiler	**8 aSTRs** & amelogenin	Life Technologies (Applied Biosystems)	Smaller regions examined; helps with degraded DNA samples
GlobalFiler*	**21 aSTRs**, DYS391, Y indel, & amelogenin	Life Technologies (Applied Biosystems)	Addresses future US core loci
PowerPlex Fusion*	**22 aSTRs**, DYS391, & amelogenin	Promega Corporation	Addresses future US core loci

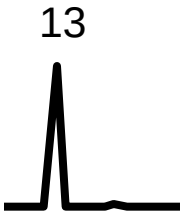
***Newer kits** that contain improved PCR buffers and DNA polymerases to yield more sensitive results and recover data from difficult samples

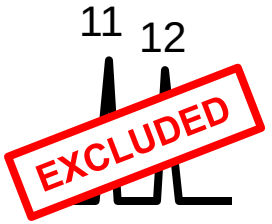
Three Possible Outcomes
of Evidence Examination

"Evidence"
Question (**Q**) Sample

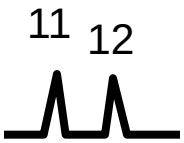
"Suspect"
Known (**K**) Sample

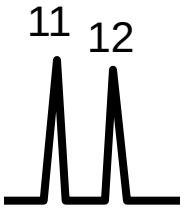
- **Exclusion** (no match)





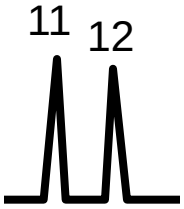
- **Inclusion** (match)





- **Inconclusive result**

No result
(or a complex mixture)



Unable to make Q → K comparison

DNA Mixture Basics

From J.M. Butler (2005) *Forensic DNA Typing, 2nd Edition*, p. 154

- Mixtures arise when two or more individuals contribute to the sample being tested.

- Mixtures can be challenging to detect and interpret without extensive experience and careful training.

Even more challenging with poor quality data when degraded DNA is present…

- Differential extraction can help distinguish male and female components of many sexual assault mixtures.

Y-chromosome markers can help here in some cases…

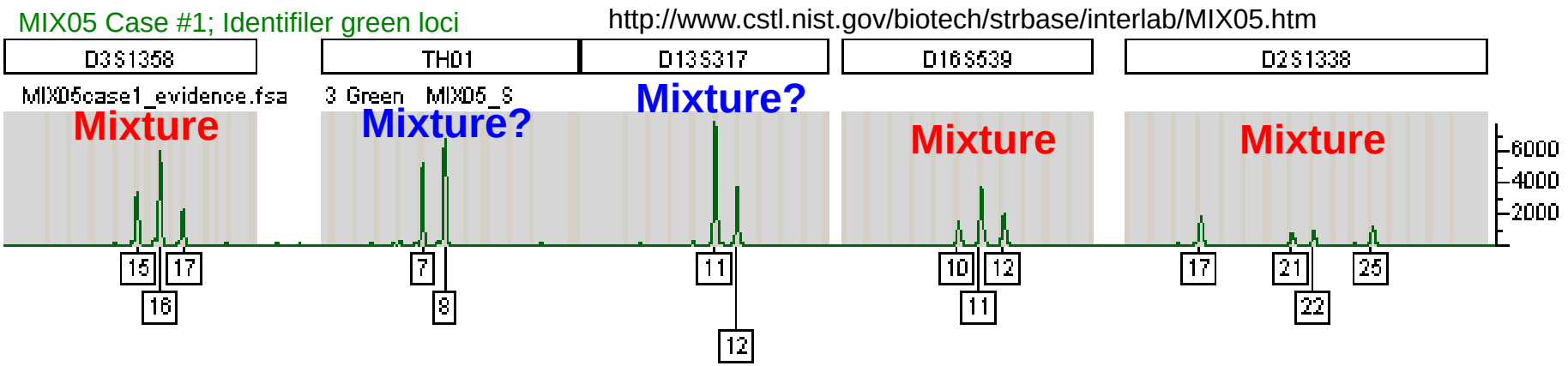
Mixtures: Issues and Challenges

From J.M. Butler (2005) *Forensic DNA Typing, 2nd Edition*, p. 155

- The probability that a mixture will be detected improves with the use of more loci and genetic markers that have a high incidence of heterozygotes.

- The detectability of multiple DNA sources in a single sample relates to the ratio of DNA present from each source, the specific combinations of genotypes, and the total amount of DNA amplified.

- Some mixtures will not be as easily detectable as other mixtures.



Sources of DNA Mixtures

- **Two (or more) individuals** contribute to the biological evidence examined in a forensic case (e.g., sexual assault with victim and perpetrator or victim, consensual sexual partner, and perp)

Victim Reference and Spouse or Boyfriend Reference

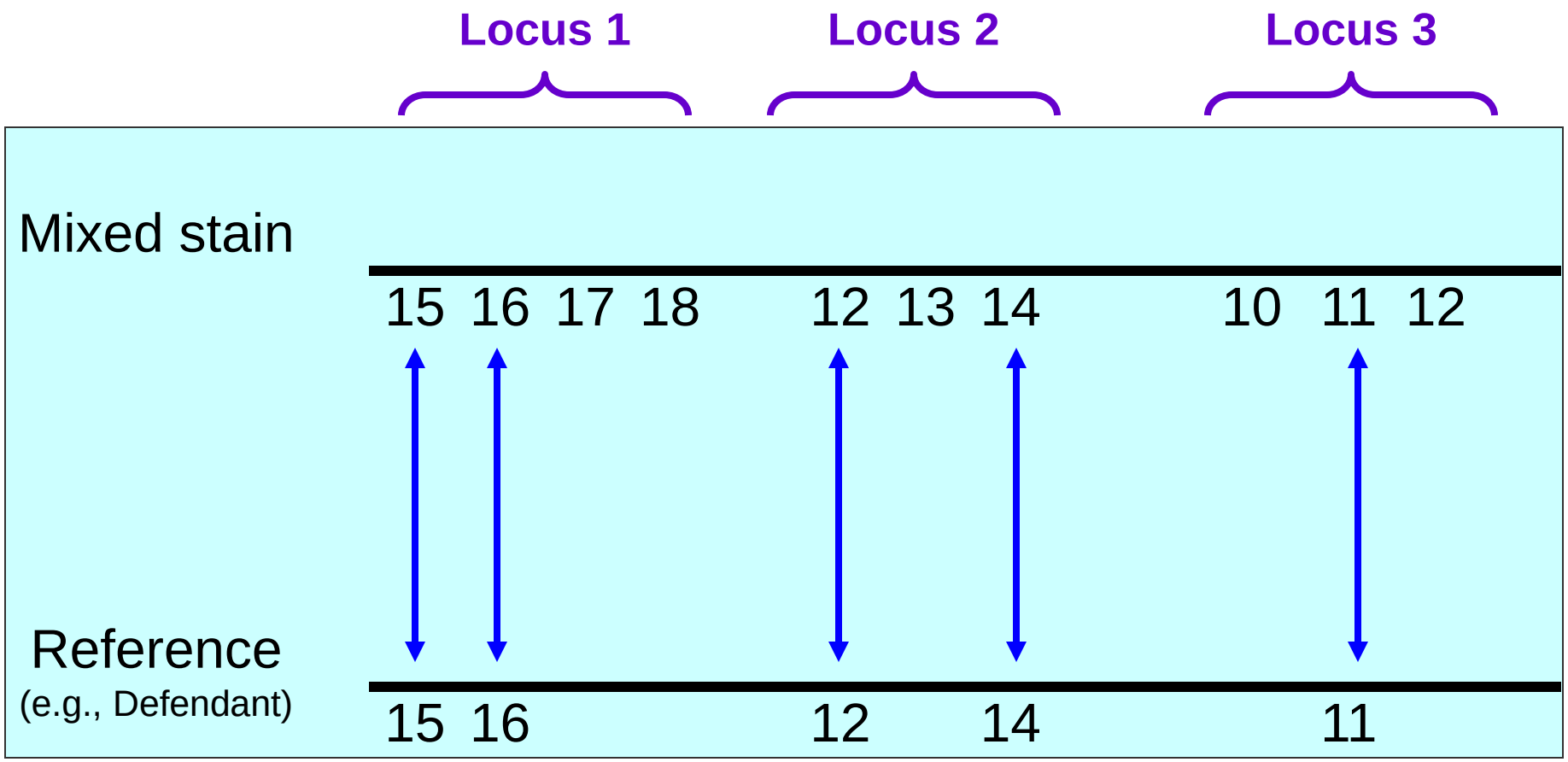
- **Contamination** of a single source sample from
  - evidence collection staff
  - laboratory staff handling the sample
  - Low-level DNA in reagents or PCR tubes or pipet tips

Examine Staff Profiles (Elimination Database), etc.

Reference elimination samples are useful in deciphering both situations due to possibility of intimate sample profile subtraction

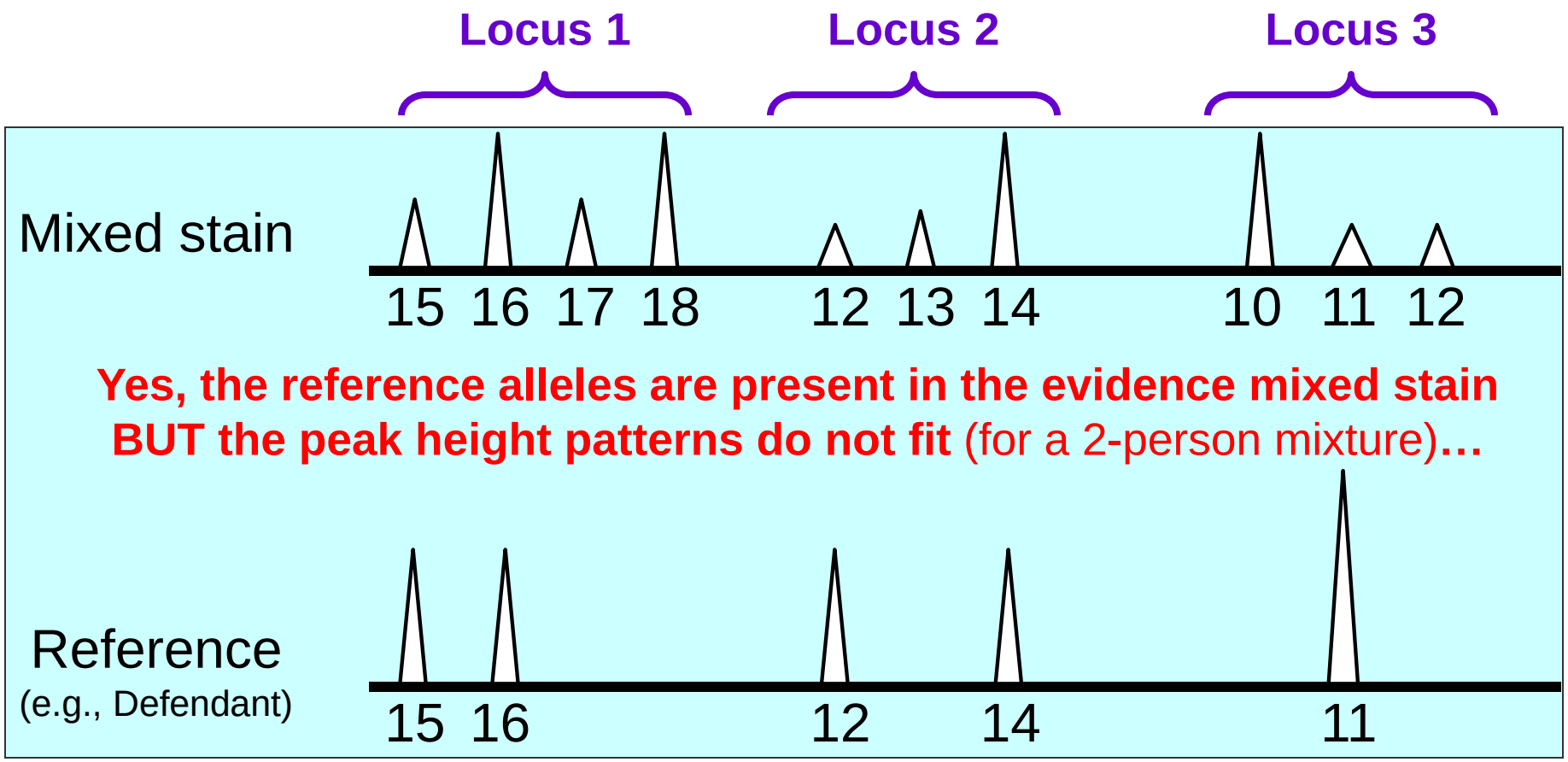
Mixture Example

Comparing Alleles Only



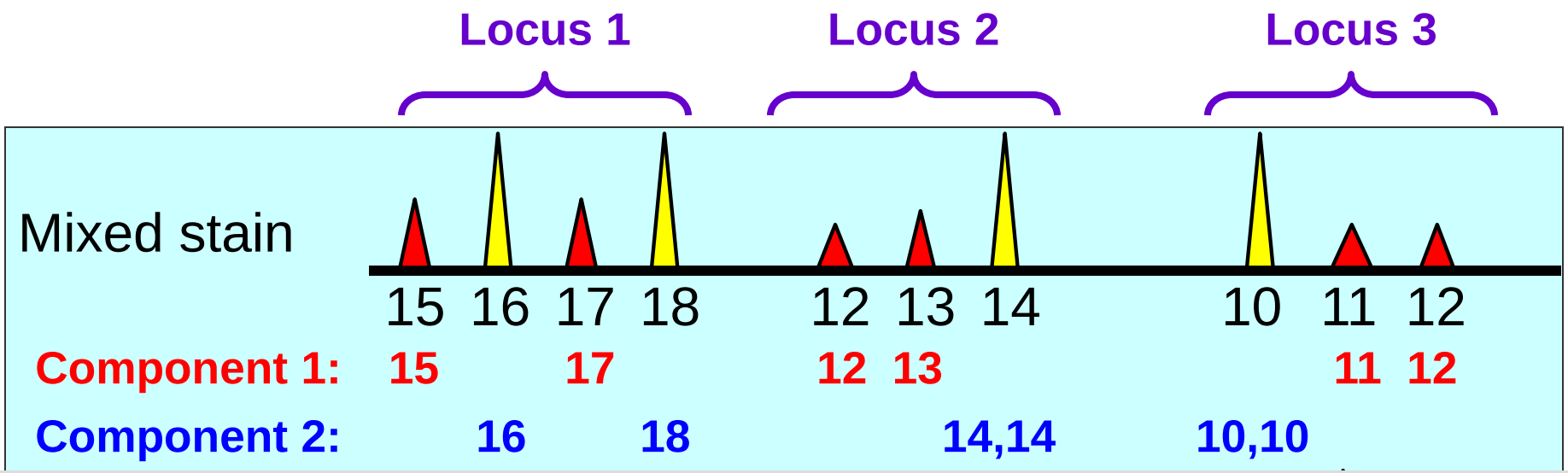
Mixture Example

Showing Importance of Using Peak Height Information



Mixture Example

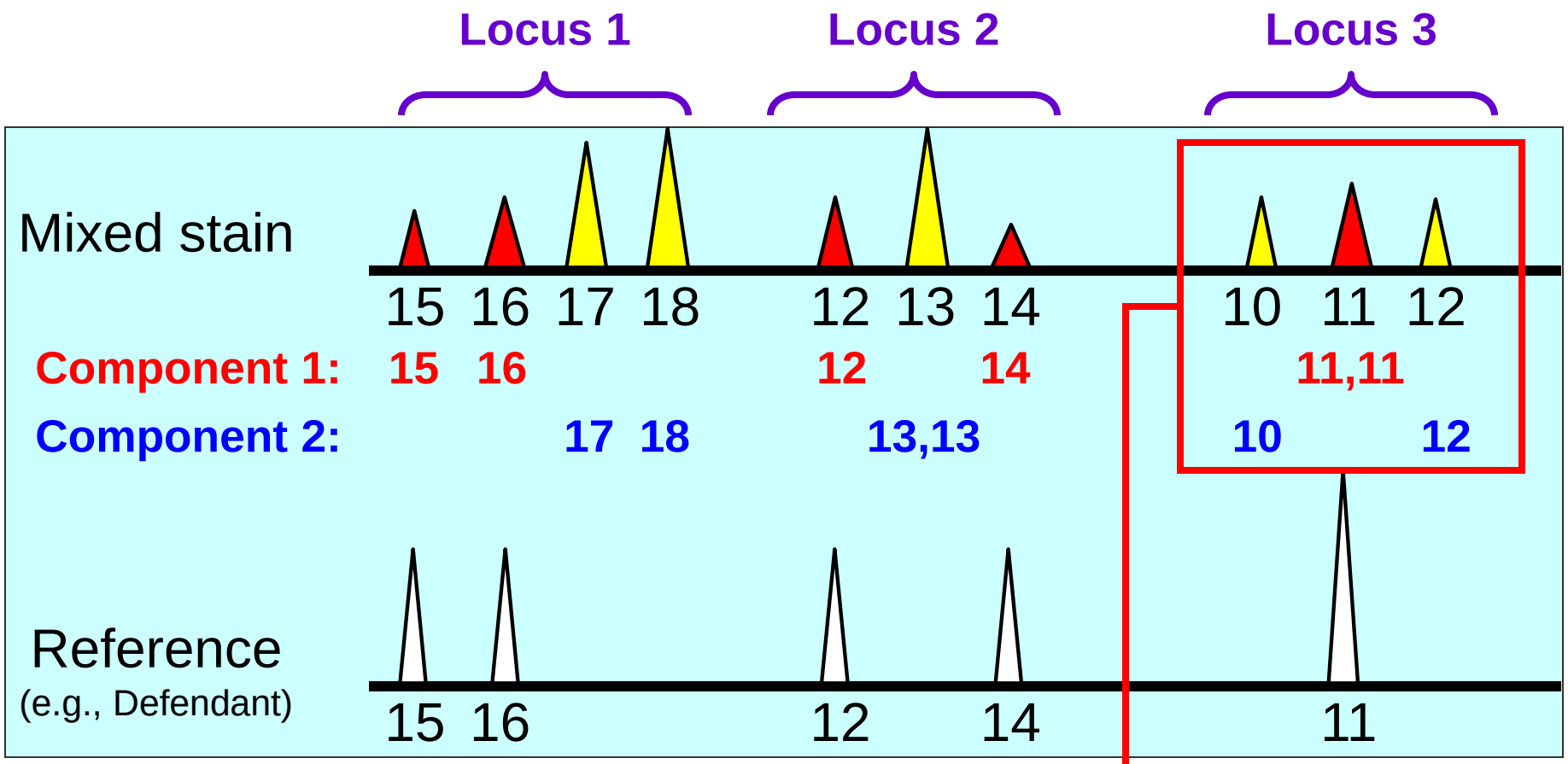
Solving Components Prior to Comparison to Suspect Reference



Reference (defendant) does not match either component of the mixed stain and therefore could not have contributed to the evidence sample (assuming 2-contributors)

Mixture Example

Different Evidence Sample…

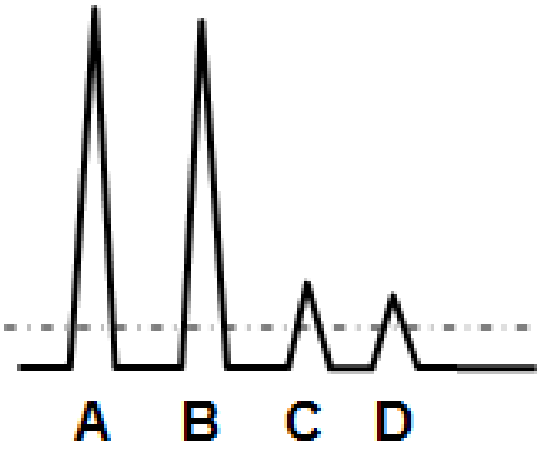


Here the Reference (defendant) does match solved Component 1 of the mixed stain and therefore could have contributed to the evidence sample

Possibilities include
10,10 with 11,12
11,11 with 10,12
12,12 with 10,11

Unrestricted vs. Restricted Genotype Combinations

Use of peak height information to select only certain combinations



Unrestricted

All combinations of alleles are deemed possible (relative peak height differences are not utilized)

AB + AC + AD + BC + BD + CD

Restricted

Based on relative peak heights, alleles are paired only where specific combinations of alleles are deemed possible

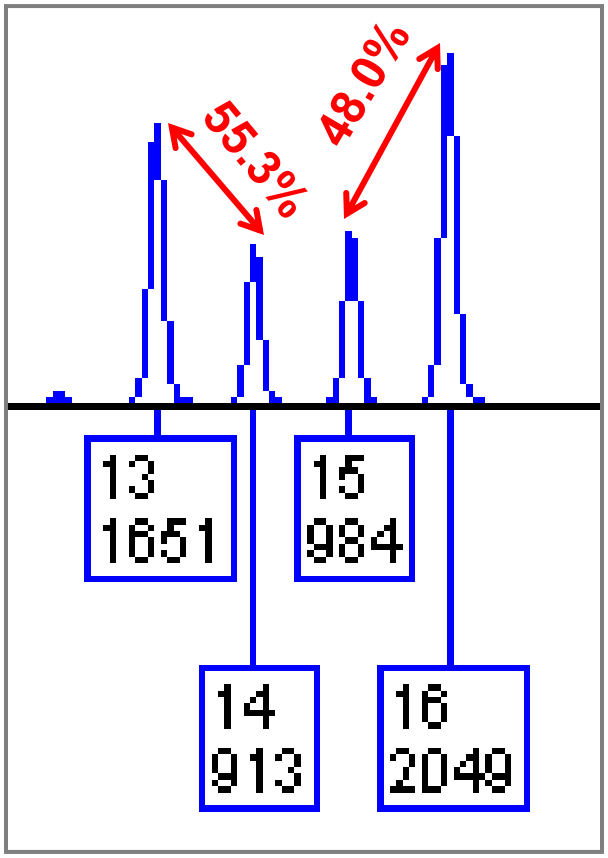
AB + ~~AC~~ + ~~AD~~ + ~~BC~~ + ~~BD~~ + CD

http://www.swgdam.org/Interpretation_Guidelines_January_2010.pdf

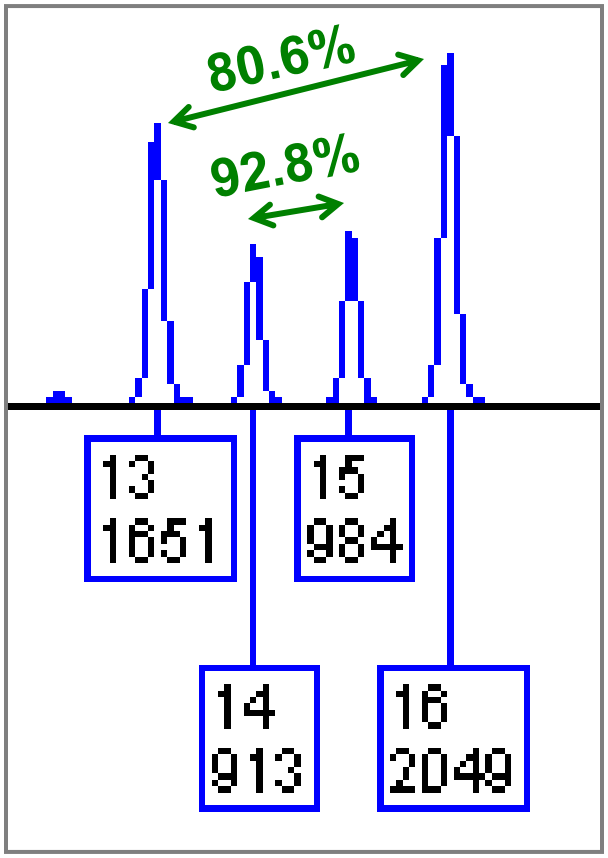
Peak Height Ratios Are Used in Mixture Component Deconvolution (Restricting Possible Genotypes)

Better Explanation of the Data

(assuming 2 contributors)

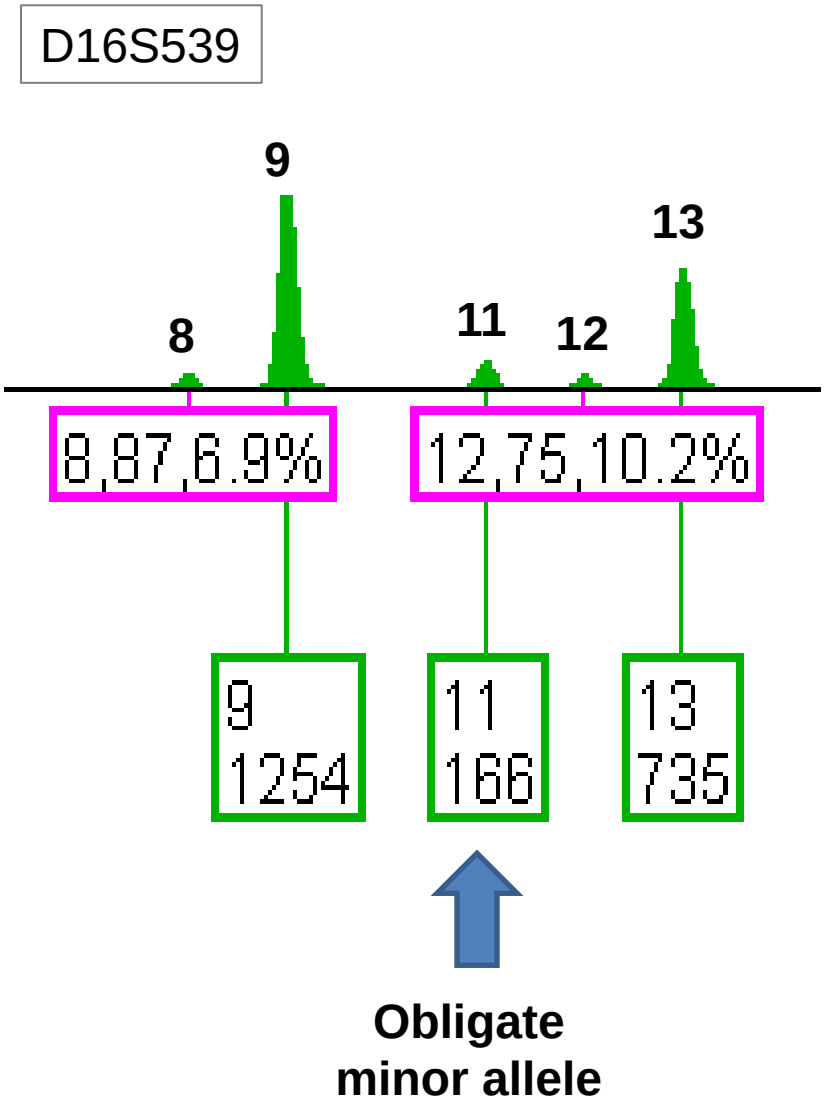


13,14 and 15,16



13,16 (major) and 14,15 (minor)

Uncertainty with Possible Genotypes



Genotype 9,13 is likely the major contributor (assuming a 2-person mixture)

The 11 allele is at 166 RFU (above a 150 ST)

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

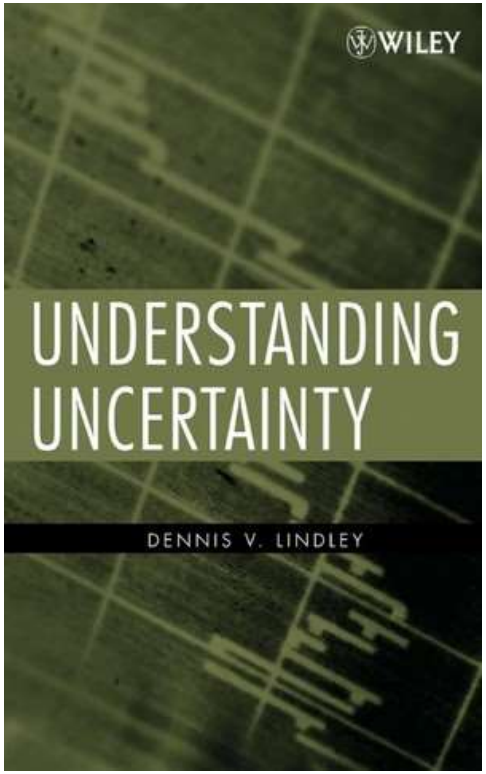
If we assume 8 and 12 are stutter peaks, then the possible genotypes of the minor contributor can be **9,11** or **11,11** or **11,13**

If we also include the 8 and 12 alleles in creating our genotype combinations, then the minor contributor possible genotypes expands to include **8,11** and **11,12**

Slide adapted from Mike Coble (NIST)

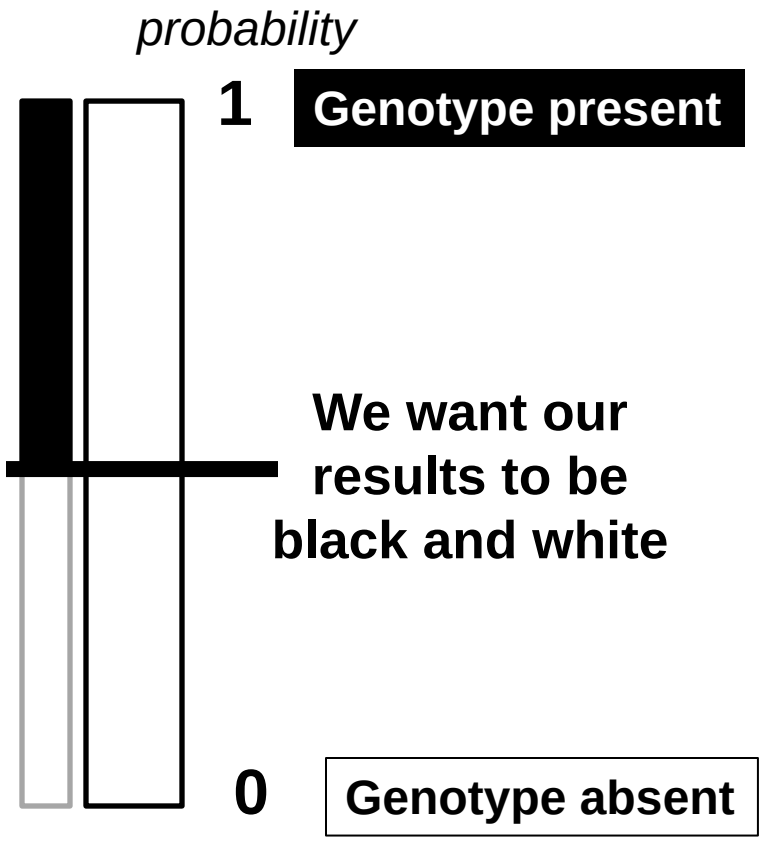
Whatever way uncertainty is approached, probability is the *only* sound way to think about it. *Understanding Uncertainty*, p. 71

- Dennis Lindley

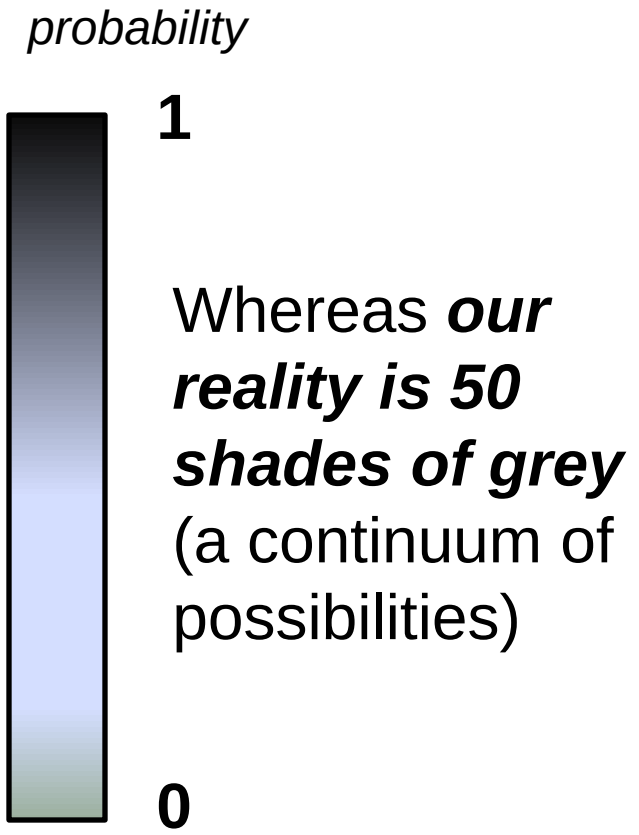


Wiley (2007)

Approaches to Data Interpretation:
Binary vs Probabilistic



Binary Approach

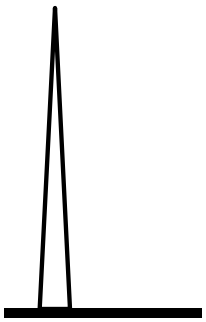


Whereas ***our reality is 50 shades of grey*** (a continuum of possibilities)

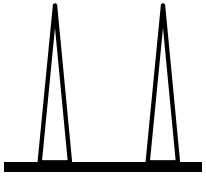
Probabilistic Approach

Allele Drop-out

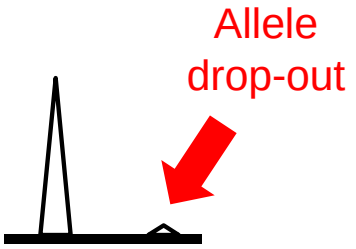
- If because of chemistry events sometimes associated with low levels of DNA (termed “stochastic effects”), one of the STR alleles “drop-out” and is not detected, then our sample at that locus looks like a homozygote instead of the heterozygote that it really is



True homozygote
(only a single peak)



True heterozygote
(both peaks detected)



False homozygote
(one peak has “dropped out” and fails to be detected)

Statistical treatment

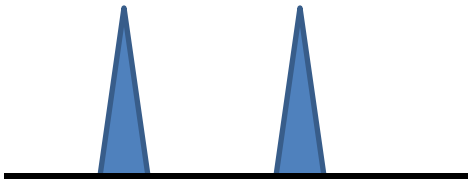
p^2

2pq

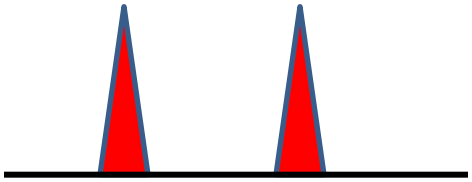
2p

Likelihood Ratios for Different Possibilities

Evidence



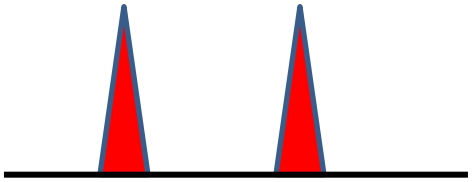
Suspect



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

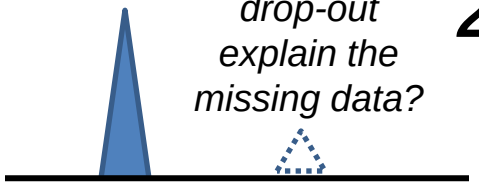
Evidence

Suspect

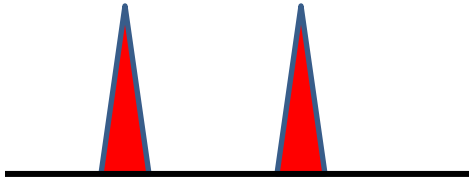


$$LR = \frac{0}{2pq}$$

Evidence



Suspect



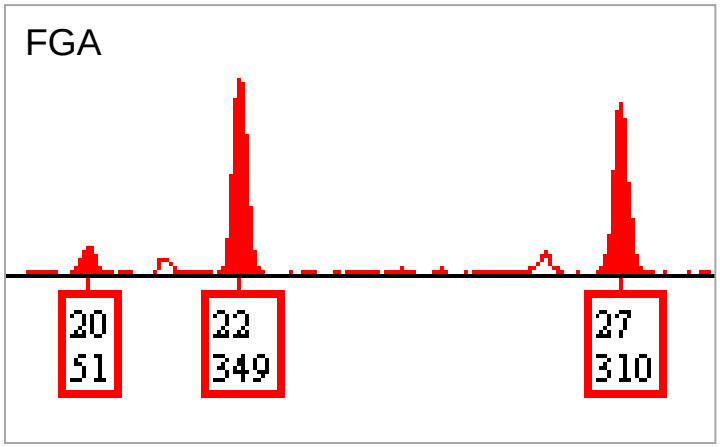
$$LR = \frac{?}{2pq}$$

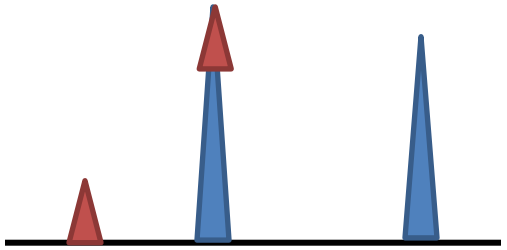
Binary LR approach (either 0 or 1)

Modified slide from Mike Coble (NIST)

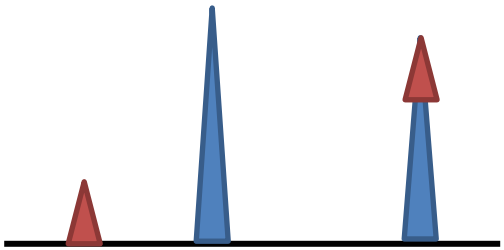
Probabilistic Genotyping Involves Exploring Multiple Possibilities to See Which One Best Fits the Data

Mixture Data (Evidence)

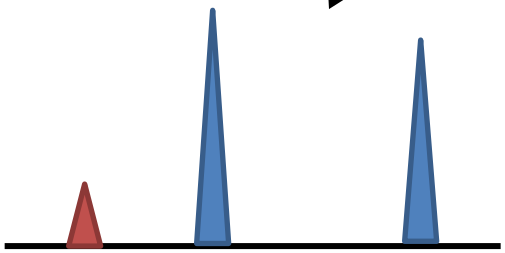




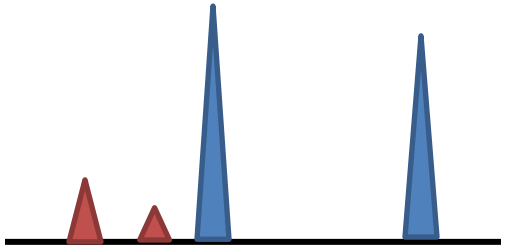
20, 22 ?



20, 27 ?



20, 20 ?



20, 21 ?

Thousands of **computer simulations** are performed to see which model is the best fit

Slide adapted from Mike Coble (NIST)

SWGDAM Interpretation Guidelines

SWGDAM Interpretation Guidelines for Autosomal STR Typing by Forensic DNA Testing Laboratories

SWGDAM = Scientific Working Group on DNA Analysis Methods (http://www.swgdam.org/)

- Approved January 14, 2010
- Available at: http://www.fbi.gov/about-us/lab/biometric-analysis/codis/swgdam.pdf or http://www.swgdam.org/Interpretation_Guidelines_January_2010.pdf

SWGDAM Mixture Interpretation Guidelines (2010)

- Provide guidance to labs for interpreting single-source and two-person mixtures
- **NOT** intended for Low Template DNA or >2 person mixtures
- Guidelines – NOT Standards
- Laboratories are not required to follow, but guidelines are STRONGLY RECOMMENDED

- Require statistics when DNA inclusions are made (SWGDAM 2010 section 4.1)

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.

FBI DNA Advisory Board (DAB) Recommendations on Statistics

February 23, 2000

"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.

Forensic Sci. Comm. 2(3); available on-line at
http://www.fbi.gov/about-us/lab/forensic-science-communications/fsc/july2000/dnastat.htm

Statistical Approaches with Mixtures

See Ladd *et al.* (2001) *Croat Med J.* 42:244-246; SWGDAM (2010) section 5

1. **Random Match Probability or RMP (after inferring genotypes of contributors)** – Separate major and minor components into individual profiles and compute the random match probability estimate as if a component was from a single source

2. **Combined Probability of Exclusion/Inclusion** – CPE/CPI (RMNE) – Calculation of the probability that a random (unrelated) person would be excluded/included as a contributor to the observed DNA mixture

RMNE = Random Man Not Excluded (same as CPI)
CPE = Combined Probability of Exclusion (CPE = 1 – CPI)
CPI = Combined Probability of Inclusion (CPI = 1 – CPE)

3. **Likelihood Ratio (LR)** – Compares the probability of observing the mixture data under two alternative hypotheses; in its simplest form LR = 1/RMP

$$LR = \frac{\Pr(E \mid H_1)}{\Pr(E \mid H_2)}$$

Assumptions for CPE/CPI Approach

- **There is no allele dropout** (i.e., all alleles are above stochastic threshold) – low-level mixtures can not reliably be treated with CPE
- All contributors are from the same racial group (i.e., you use the same allele frequencies for the calculations)
- **All contributors are unrelated**
- Peak height differences between various components are irrelevant (i.e., component deconvolution not needed) – this may not convey all information from the available sample data…

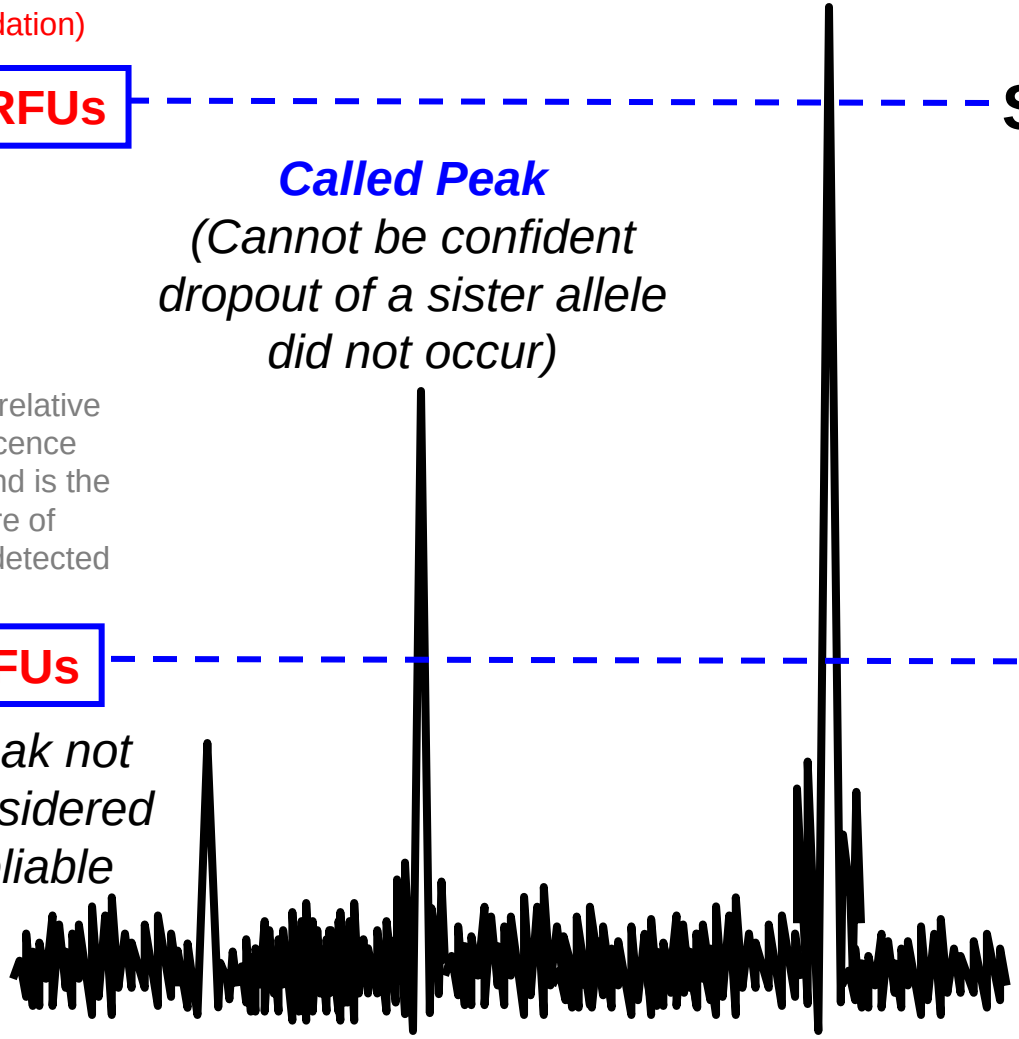
Coupling of Statistics and Interpretation

- **The CPE/CPI approach** for reporting an inclusionary statistic **requires that all alleles be observed** in the evidence sample

- If allele drop-out is suspected at a locus, then any allele is possible and the probability of inclusion goes to 100% -- in other words, the locus is effectively dropped from consideration for statistical purposes

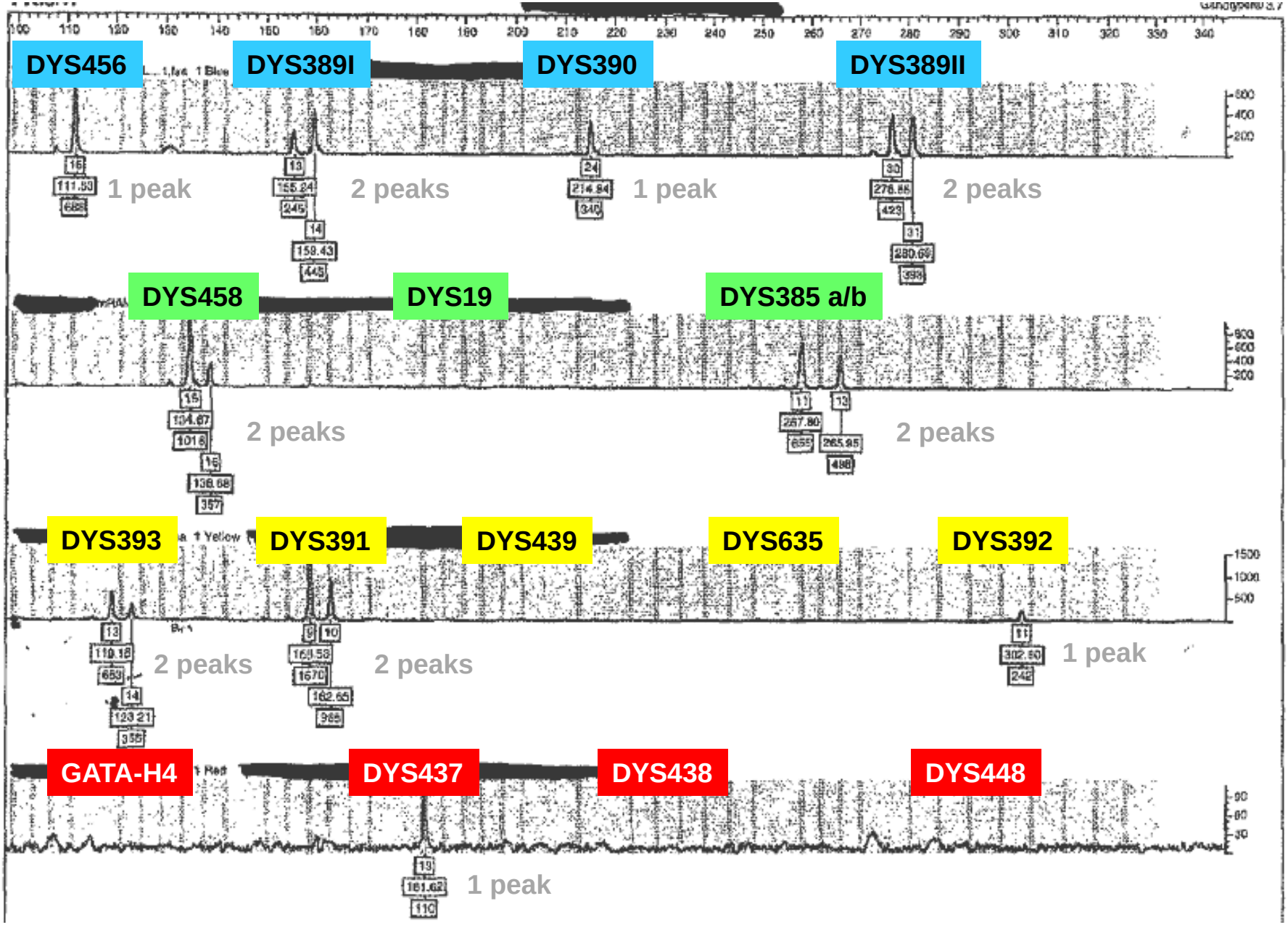
- If alleles are seen below the established stochastic threshold, then the locus is typically eliminated (“INC” – declared inconclusive) in many current lab SOPs

Overview of Two Thresholds



From Butler, J.M. (2010) *Fundamentals of Forensic DNA Typing*. Elsevier Academic Press: San Diego.

Yfiler (Y-STR) Data from a Knife Handle (1970s post-conviction case)



Comparison between Evidence (Q sample) and Defendant (K sample)

[redacted]

Report of Laboratory Examination

[redacted]

2/25/2009

[redacted]

[redacted]

AGENCY CASE NO: [redacted]

Table 1 Y-Filer

Sample Name	DYS456	DYS389I	DYS390	DYS389II	DYS458	DYS19	DYS385a/b	DYS393	DYS391	DYS439	DYS635	DYS392	Y GATA H4	DYS437	DYS438	DYS448
Swabbings: Black Handle of Folding Knife	15	13, 14	24	30, 31	15, 16	NR	11, 13	13, 14	9, 10	NR	NR	11	NR	13	NR	NR
[redacted] Defendant [redacted]	17	13	26	30	15	15	11, 15	13	10	10	22	11	12	14	10	20

[redacted]

NR = No Result

The results listed in the table do not depict intensity differences.

The lab <u>initially</u> issued a report stating that **the results excluded the defendant**.

Several months later, the lab changed its assessment and issued a <u>new report</u>:

"The partial Y-STR profile obtained …is a mixture consistent with originating from two males. **No determination can be made** as to whether or not [Defendant] is a contributor to this mixture."

Uncertainty in Evidence Result Leads to “Inconclusive” Report

- In my opinion, a high degree of uncertainty in the number of contributors (Y-STR loci with multiple alleles) and the true DNA types (due to extensive allele and locus dropout) makes comparison of this sample to ANY reference sample problematic

- **If evidence cannot be compared due to poor quality data, then the defendant cannot be excluded (and potentially exonerated) based on DNA results…**

- Poor quality DNA data (as well as potential, inadvertent contamination) may present challenges with reaching any conclusions on older Innocence Project cases

New Statistical Tools/Software for Mixtures

- **Lab Retriever** (David Balding → Norah Rudin et al.)
  - Uses likelihood ratios (LRs) and probability of dropout [Pr(D) or P(Do)]
- **FST** – Forensic Statistical Tool (NYC OCME)
  - Uses LRs and empirically determined Pr(D) based on DNA quantity
- **Armed Xpert** (USACIL → Niche Vision)
  - Originally developed by US Army Crime Lab (USACIL)
  - Performs calculations typically manually done by analysts
- **TrueAllele** (Mark Perlin/Cybergenetics)
  - Uses probabilistic genotyping approach with LRs

Scientific Collaboration, Innovation & Education Group

Lab Retriever Program

Beta-version is available for **free download from www.scieg.org**

Scientific Article - describes the math and statistical model

Balding, D.J., & Buckleton, J. (2009) Interpreting low template DNA profiles. *Forensic Science International: Genetics, 4,* 1-10.


Forensic Science International: Genetics

Interpreting low template DNA profiles

David J. Balding, John Buckleton

David Balding likeLTD – program written in R (computer language)

https://sites.google.com/site/baldingstatisticalgenetics/software/likeltd-r-forensic-dna-r-code

Norah Rudin and colleagues – prepare a GUI for likeLTD to make it more user-friendly

Lab Retriever

http://www.scieg.org/lab_retriever.html

Credits:

Based on the original work of:

- David Balding
- John Buckleton

Research and development:

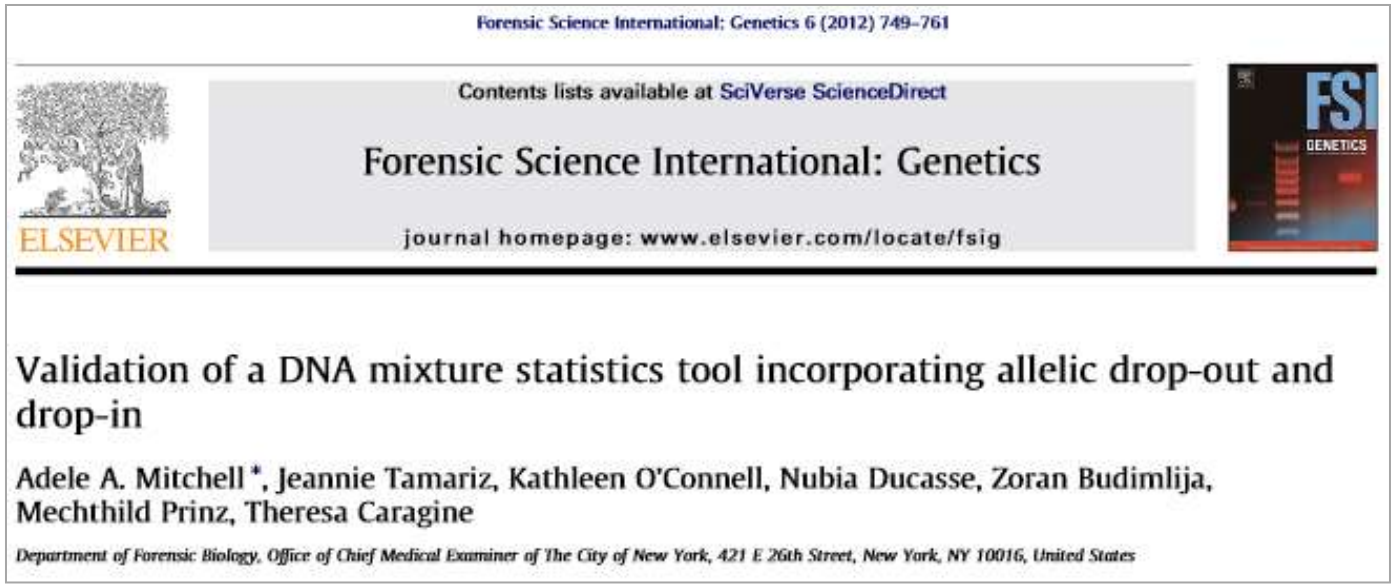
- Keith Inman
- Kirk Lohmueller
- Norah Rudin

Programmers:

- Ken Cheng
- Luke Inman-Semerau

FST (Forensic Statistical Tool)

Currently undergoing a Frye admissibility hearing in NYC



Forensic Science International: Genetics 6 (2012) 749–761

Contents lists available at SciVerse ScienceDirect

Forensic Science International: Genetics

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

ELSEVIER

FSI GENETICS

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

- "…FST does not deconvolute DNA mixtures, but simply **computes a LR for scenarios specified by the user**, allowing for mismatches between contributors' profiles and the DNA alleles labeled in the mixtures. The **mismatches are accounted for by incorporating drop-out and drop-in probabilities in the LR calculation**. While FST uses empirically determined drop-out and drop-in rates, [other programs] require the user to specify drop-out and drop-in probabilities…"

http://www.nyc.gov/html/ocme/html/hss/hssservices_provided.shtml

Armed Xpert

http://www.armedxpert.com/

http://www.nichevision.com/

- Developed by the US Army Crime Lab (USACIL) initially as a Virtual Basic program called “DNA_DataAnalysis”

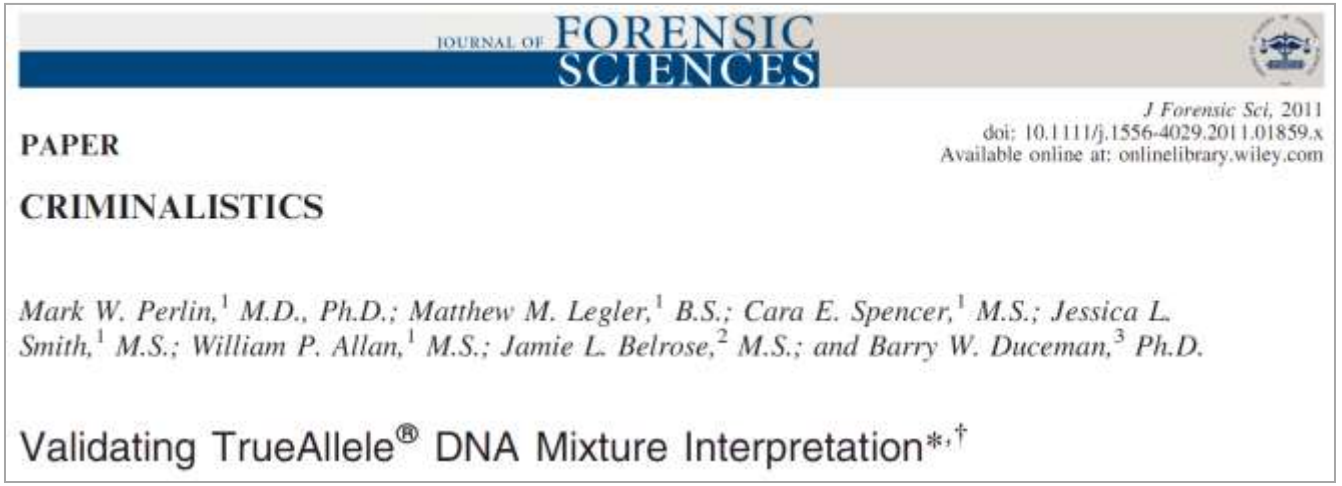
- Enables RMP, CPI, and LR calculations for 2-person and 3-person mixtures

- Plan to incorporate probability of drop-out models developed by John Buckleton (New Zealand)

True Allele Casework

http://www.cybgen.com/systems/casework.shtml

- Performs thousands of simulations to model mixture data
- Calculates a combined likelihood ratio
- A commercial product so not all of the mathematical details have been published
- Has been admitted in several states including PA and CA
- Validation work published with NYSP (JFS Nov 2011)



JOURNAL OF FORENSIC SCIENCES

J Forensic Sci, 2011
doi: 10.1111/j.1556-4029.2011.01859.x
Available online at: onlinelibrary.wiley.com

PAPER

CRIMINALISTICS

Mark W. Perlin,[1] M.D., Ph.D.; Matthew M. Legler,[1] B.S.; Cara E. Spencer,[1] M.S.; Jessica L. Smith,[1] M.S.; William P. Allan,[1] M.S.; Jamie L. Belrose,[2] M.S.; and Barry W. Duceman,[3] Ph.D.

Validating TrueAllele® DNA Mixture Interpretation*,†

Probabilistic Modeling of TrueAllele

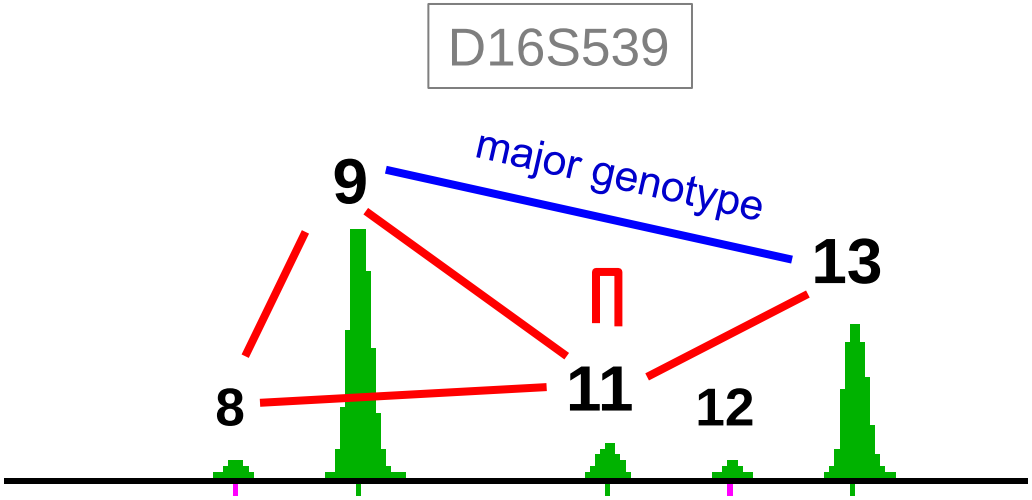
Mathematical Modeling of the Data

PHR, Mix Ratio, Stutter etc…

Typically 50,000 or 100,000 Simulations Performed

(MCMC)

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



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

- Quantitative computer interpretation using numerous Markov Chain Monte Carlo (MCMC) simulations
- Models peak uncertainty and infers possible genotypes
- Results are presented as the Combined LR

DNA Case Example

- Portions of **redacted results and lab report** were kindly provided by Olga Akselrod (Innocence Project)

- **Three pieces of evidence (mixtures)** plus victim and defendant DNA profiles to enable Q→K comparisons
  - Fingernail clippings (right hand & left hand) and jeans

- Testing was performed using **MiniFiler**
  - 8 STR loci + amelogenin (sex-typing marker)
  - MiniFiler is a miniSTR test that aids recovery of results from damaged DNA because it examines smaller portions of the DNA molecules than other STR typing kits

- Statistical analysis of mixtures were performed using **CPI** (combined probability of inclusion) and **FBI Popstats** computer program

Lab Report Wording

Focus of next slide

The Minifiler DNA profile obtained from the right hand fingernail clippings (Item #2.16) is a mixture of DNA from at least three individuals, including at least one male and one female individual.

The DNA profile of [redacted] (Item #5) cannot be excluded from the DNA in the mixture. For the loci D21S11, D7S820, CSF1PO, D13S317, D16S539, D2S1338, D18S51, and FGA, the probability of randomly selecting an unrelated individual as a possible contributor to the DNA profile of the mixture at the genetic loci above is at least 1 in 1,965 for U.S. individuals. Therefore, [redacted] cannot be excluded as a contributor to the Minifiler genetic material in this specimen.

The Minifiler DNA profile obtained from the left hand fingernail clippings (Item #2.17) is a mixture of DNA from at least three individuals.

The DNA profile of [redacted] (Item #5) cannot be excluded from the DNA in the mixture. For the loci D7S820, CSF1PO, D13S317, D16S539, D2S1338, D18S51, and FGA, the probability of randomly selecting an unrelated individual as a possible contributor to the DNA profile of the mixture at the genetic loci above is at least 1 in 358 for U.S. individuals. Therefore, [redacted] cannot be excluded as a contributor to the Minifiler genetic material in this specimen.

The Minifiler DNA profile obtained from the jeans (Item #2.7) is a mixture of DNA from at least four individuals.

The DNA profile of [redacted] (Item #5) cannot be excluded from the DNA in the mixture. For the loci D7S820, CSF1PO, D16S539, D18S51, and FGA, the probability of randomly selecting an unrelated individual as a possible contributor to the DNA profile of the mixture at the genetic loci above is at least 1 in 16 for U.S. individuals. Therefore, [redacted] cannot be excluded as a contributor to the Minifiler genetic material in this specimen.

Right hand fingernail clipping: **1 in 1965** (8 loci used)
Left hand fingernail clipping: **1 in 358** (7 loci used)
Jeans: **1 in 16** (5 loci used)

Breaking Down a Portion of the Report

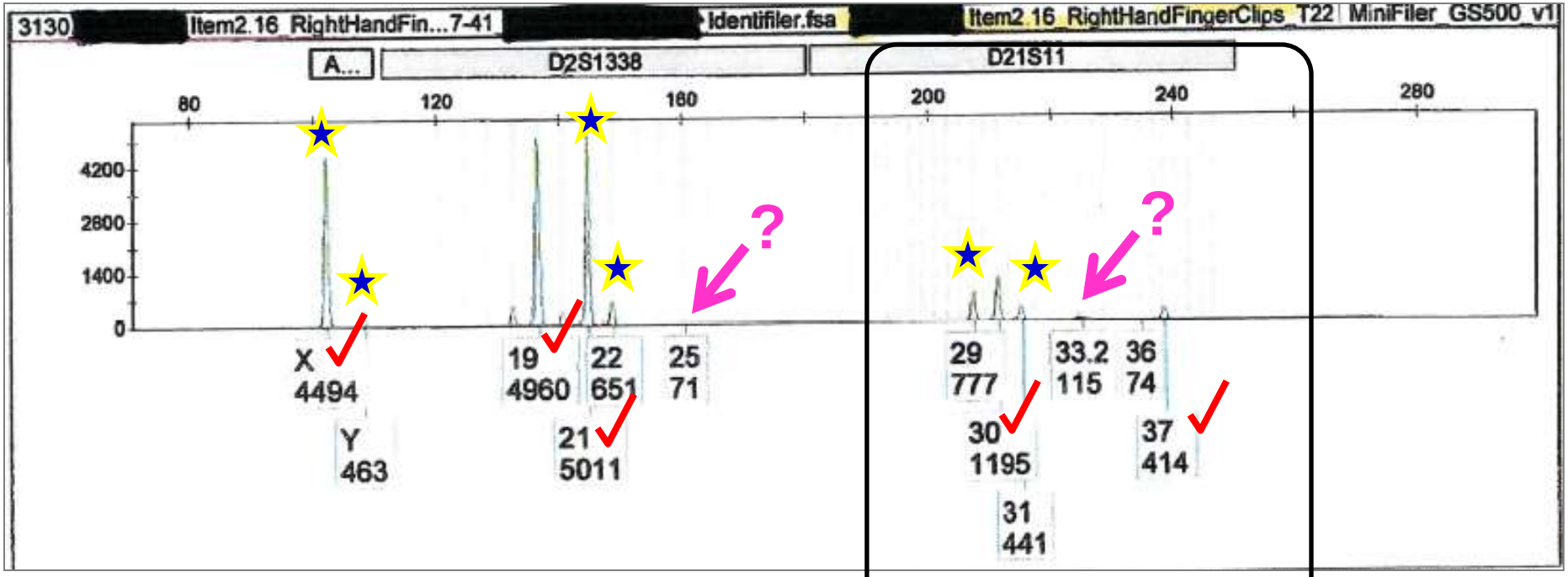
The Minifiler DNA profile obtained from the right hand fingernail clippings (Item #2.16) is **a mixture of DNA from at least three individuals**, including **at least one male and one female individual**.

The DNA profile of **DEFENDANT** (Item #5) **cannot be excluded** from the DNA in the mixture.

For the [MiniFiler] loci D21S11, …, **the probability of randomly selecting an unrelated individual as a possible contributor** to the DNA profile of the mixture at the [MiniFiler loci] is **at least 1 in 1,965** for U.S. individuals. …

MiniFiler Green Channel

Right Hand Fingernail Clippings (Item #2.16)



Amelogenin	D2S1338	D21S11
Victim: **X,X**	Victim: **19,21**	Victim: **30,37**
Defendant: **X,Y**	Defendant: **21,22**	Defendant: **29,31**

Lab Report Data

ITEM	Report #1 Item #2.16 Right Hand Fingernail Clippings **Minifiler**	Report #1 Item #2.17 Left Hand Fingernail Clippings **Minifiler**	Report #1 Item #2.7 Jeans **Minifiler**	Victim Report #1 Item #2.36 Dried Red Stain on Bra From [redacted] **Identifiler**	Defendant Item #5 [redacted] Known **Identifiler**
D8S1179	NT	NT	NT	14,15	14,14
D21S11	29,30,31,33.2,37*	30,31,37	27,29,30,31.2,37*	30,37	29,31
D7S820	8,9,10,11	8,10,11	8,10,11*	10,11	8,11
CSF1PO	7,8,10,11	7,10,11,12	7,8,10,11,12	10,11	7,10
D3S1358	NT	NT	NT	16,17	15,15
TH01	NT	NT	NT	7,9.3	8,9
D13S317	11,12,13	11,12,13	10,12,13*	12,13	11,11
D16S539	9,11,12,13*	9,11,12,13	9,11,12,13	11,12	9,13
D2S1338	19,21,22*	17,19,21,22	16,19,21,23,24*	19,21	21,22
D19S433	NT	NT	NT	14,15	12.2,13
vWA	NT	NT	NT	16,17	14,16
TPOX	NT	NT	NT	8,11	9,9
D18S51	10,15,16,20*	10,15,16,17,20	10,12,14,15,16,17,19	15,16	10,15
Amelo.	X,Y	X,Y	X,Y	X,X	X,Y
D5S818	NT	NT	NT	11,12	12,12
FGA	20,21,23,24,25	20,23,24	20,21,23,24	20,24	20,23

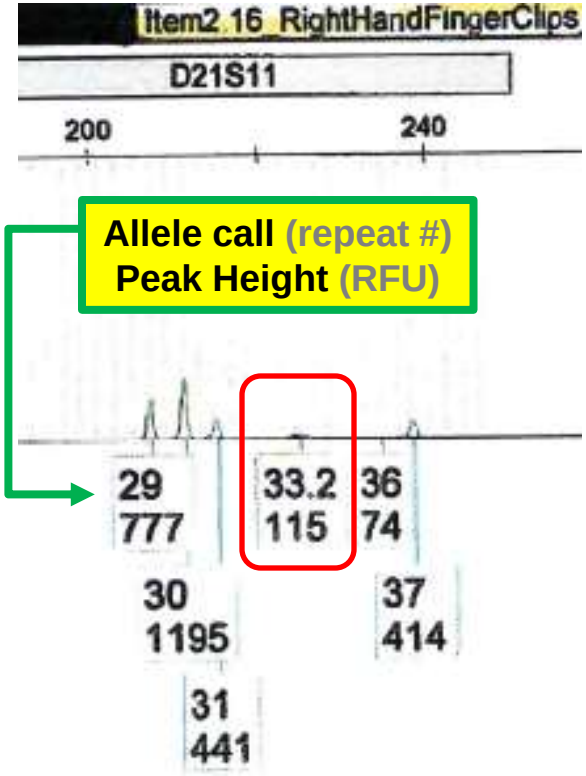
Item #2.16 (Right Hand Fingernail Clippings)

Information in Report Table

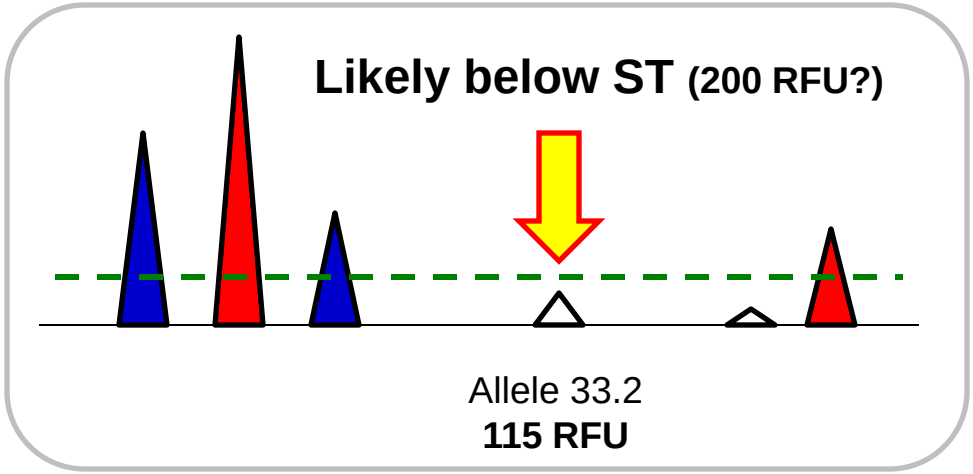
D21S11: 29, 30, 31, 33.2, 37*

Victim: 30, 37
Defendant: 29, 31

Electropherogram (mixture data observed)



Graphical representation of DNA data at D21S11



*Based on results at this single locus, we can assume **at least three individuals contributed to the DNA results** because there are more than 4 alleles*

Source of the Numbers for Right Hand Fingernail Clippings

Popstats 5.7.4 DNA Profile

Forensic Mixture Case: Probability of Inclusion
Database: F:\Popstats\Popdata\FBI\STR
Specimen: Reference
DNA Analyst: [redacted]
Lab ID: [redacted]
Date: 7/8/2011 2:22:19PM
Page 1 of 1

Allele Frequency

Locus	Allele	CAU	BLK	SWH
D21S11	29	0.1811	0.1899	0.2044
D21S11	30	0.2321	0.1788	0.3301
D21S11	31	0.0714	0.0922	0.069
D21S11	33.2	0.0306	0.0335	0.0419
D21S11	37	0.0128	0.014	0.0123
D7S820	8	0.1626	0.1738	0.0981
D7S820	9	0.1478	0.1571	0.0479
D7S820	10	0.2906	0.3238	0.3062
D7S820	11	0.202	0.2238	0.2895
CSF1PO	7	0.0123	0.0429	0.012
CSF1PO	8	0.0123	0.0857	0.012
CSF1PO	10	0.2537	0.2714	0.2536
CSF1PO	11	0.3005	0.2048	0.2656
D13S317	11	0.3189	0.2374	0.202
D13S317	12	0.3087	0.4832	0.2168
D13S317	13	0.1097	0.1257	0.1379
D16S539	9	0.104	0.1986	0.0793
D16S539	11	0.2723	0.2943	0.3149
D16S539	12	0.3391	0.1866	0.2861
D16S539	13	0.1634	0.1651	0.1034
D2S1338	19	0.1447	0.1377	0.2605
D2S1338	21	0.0197	0.1526	0.0176
D2S1338	22	0.0296	0.1377	0.0704
D18S51	10	0.0128	0.0139	0.0123
D18S51	15	0.1276	0.1667	0.1379
D18S51	16	0.1071	0.1889	0.1158
D18S51	20	0.0255	0.0556	0.0172
FGA	20	0.1454	0.0722	0.0714
FGA	21	0.1735	0.125	0.1305
FGA	23	0.1582	0.125	0.1404
FGA	24	0.1378	0.1861	0.1256
FGA	25	0.0689	0.1	0.1379
Amelogenin	X			
Amelogenin	Y			

CAU probability of inclusion 3.834E-05 = 1 in 2.608E+04
BLK probability of inclusion 5.089E-04 = 1 in 1.965E+03
SWH probability of inclusion 5.423E-05 = 1 in 1.844E+04

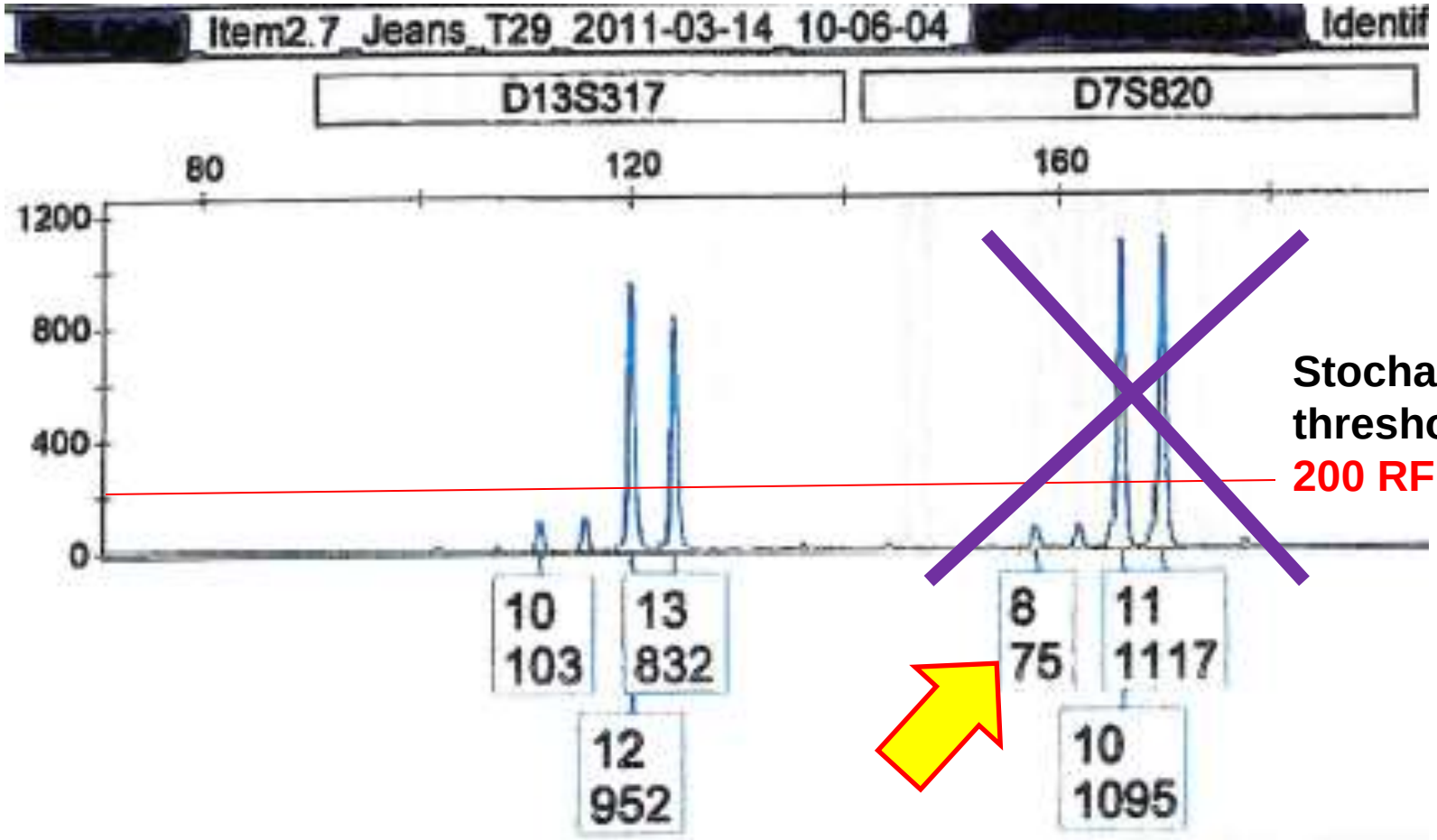
Allele Frequency

Locus	Allele	CAU	BLK	SWH
D21S11	29	0.1811	0.1899	0.2044
D21S11	30	0.2321	0.1788	0.3301
D21S11	31	0.0714	0.0922	0 069
D21S11	33.2	0.0306	0.0335	0.0419
D21S11	37	0.0128	0.014	0.0123

Using D21S11 alleles 29, 30, 31, 33.2, and 37 in statistical calculations

Population Group	CPI Stats Calculated
Caucasian (CAU)	1 in 26,080
Black (BLK)	**1 in 1,965**
Southwest Hispanic (SWH)	1 in 18,440

MiniFiler Blue Loci for Jeans (Item #2.7)



Allele below ST would invalidate D7S820 locus from use in CPI statistics

CPI Stats Calculations for the Jeans

Popstats 5.7.4 DNA Profile

Forensic Mixture Case: Probability of Inclusion
Database: F:\Popstats\Popdata\FBI\STR
Specimen: Reference
DNA Analyst: [redacted]
Lab ID:
Date: 7/8/2011
Page 1 of 1

Allele Fre

Locus	Allele	CAU	BLK	
D7S820	8	0 1626	0 1738	
D7S820	10	0 2906	0 3238	
D7S820	11	0 202	0 2238	
CSF1PO	7	0 0123	0 0429	
CSF1PO	8	0 0123	0 0857	
CSF1PO	10	0 2537	0 2714	
CSF1PO	11	0 3005	0 2048	
CSF1PO	12	0 3251	0 3	
D16S539	9	0 104	0 1986	
D16S539	11	0 2723	0 2943	
D16S539	12	0 3391	0 1866	
D16S539	13	0 1634	0 1651	
D18S51	10	0 0128	0 0139	
D18S51	12	0 1276	0 0583	
D18S51	14	0 1735	0 0639	
D18S51	15	0 1276	0 1667	
D18S51	16	0 1071	0 1889	
D18S51	17	0 1556	0 1639	
D18S51	19	0 0357	0 0778	0.037
FGA	20	0 1454	0 0722	0.0714
FGA	21	0 1735	0 125	0.1305
FGA	23	0 1582	0 125	0.1404
FGA	24	0 1378	0 1861	0.1256
Amelogenin	X			
Amelogenin	Y			

CAU probability of inclusion 5.913E-02 = 1 in 1.691E+01
BLK probability of inclusion 4.468E-02 = 1 in 2.238E+01
SWH probability of inclusion 3.084E-02 = 1 in 3.243E+01

Popstats 5.7.4

Forensic Mixture Case: Probability of Inclusion
Database: F:\Popstats\Popdata\FBI\ST
Specimen: Reference
DNA Analyst: [redacted]
Lab ID:
Date: 7/8/2011 2:27:37PM
Page 1 of 1

Allele Frequency

Locus	Allele	CAU	BLK	SWH
D7S820	8	0 1626	0 1738	0.0981
D7S820	10	0 2906	0 3238	0.3062
D7S820	11	0 202	0 2238	0.2895

D7S820 appears to have been used in CPI statistical calculation

Only 1 in 16.9 Caucasians

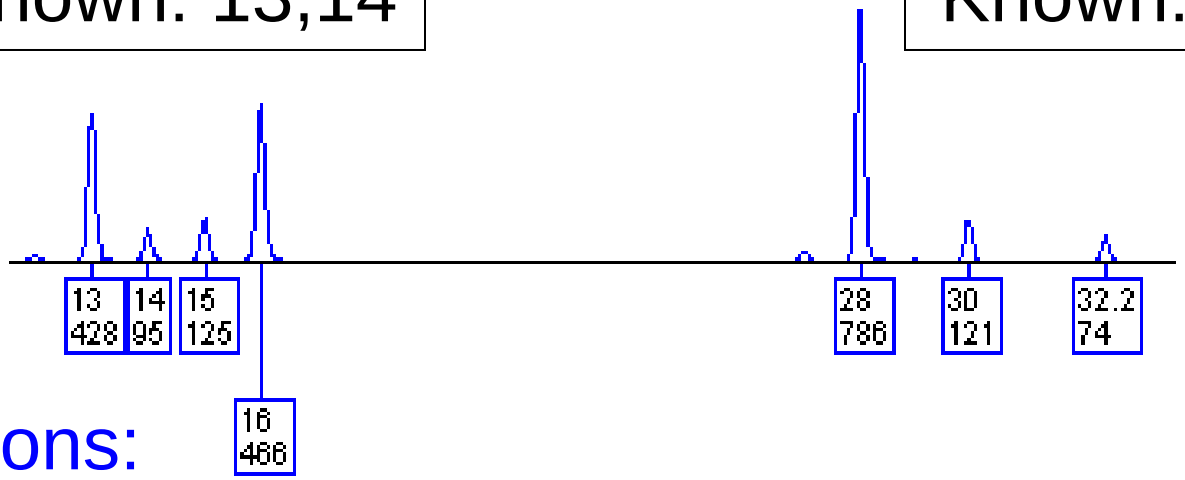
Where Can Potential Errors Occur in DNA Interpretation?

- Incorrect inclusion of an innocent person using allele drop-out as a reason for mismatch between evidence and suspect with a CPI approach
- Inclusion of loci in CPI calculations with alleles below stochastic threshold (CPI requires all alleles to be detected) could lead to an inflation of the match statistic
- Setting thresholds too high and thus losing relevant data that could be used to exclude
- Use of p^2 with single peaks (assuming genotype is a homozygote) instead of 2p (allowing for allele drop-out) will falsely inflate statistics
- Failure to exclude when alleles are present but genotypes do not fit

Is the Known Individual Included or Excluded?

Known: 13,14

Known: 28,30



Assumptions:

1) 2 contributors *and* all data are present →
2) 1 major and 1 minor contributor →
3) Major must have 13,16 and 28,28 genotypes and
4) Minor must have 14,15 and 30,32.2 genotypes

Based on these assumptions,
the individual is excluded

Genotypes are excluded even if alleles are included

Slide from Charlotte Word (consultant)

Different Experts→ Different Opinions

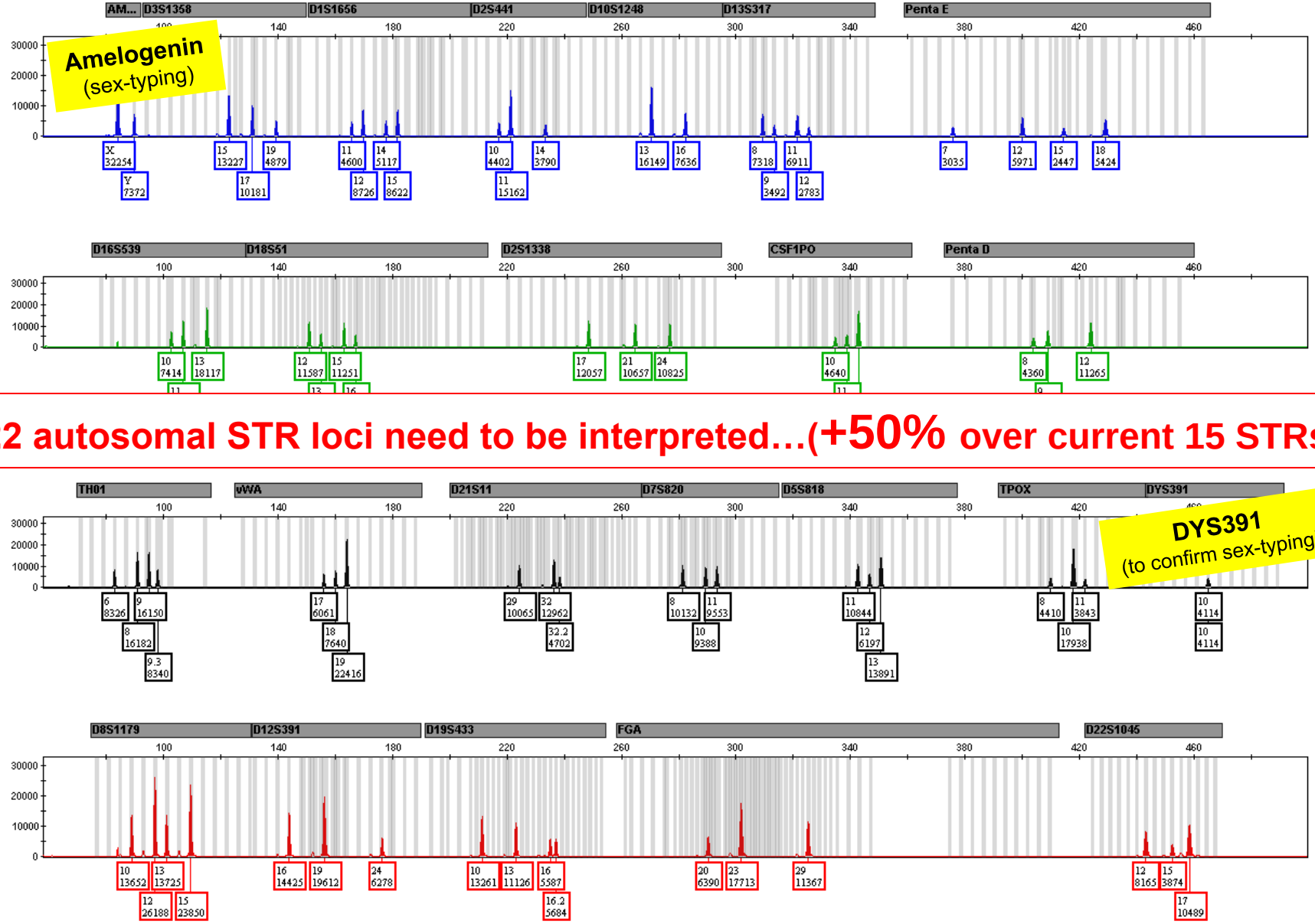
- Are the experts asking/answering the same question?
- Are they using the same information and data?
- Are they using the same interpretation methods?
- Are they using good scientific practices?
- Any possibility of bias?
- Are the differences meaningful or trivial?

Slide from Charlotte Word (consultant)

Some Thoughts on the Future…

- **PCR amplification**
  - Faster enzymes to enable rapid PCR
  - More robust enzymes and master mixes that work better
- **Instrumentation**
  - More dye colors to aid in analyzing more loci simultaneously
  - Rapid, integrated devices
  - Alternatives to capillary electrophoresis: next-generation sequencing
- **Marker systems**
  - Expanding sets of STR loci for growing DNA databases
  - Other marker systems: SNPs, InDels, X-STRs, RM Y-STRs
  - Body fluid identification using other molecules such as RNA
  - Phenotyping for external visible characteristics
  - Privacy challenges with additional genome information
- **Data interpretation**
  - Probabilistic genotyping for low-level DNA and mixture interpretation
  - Probability of dropout incorporated into DNA data interpretation

DNA Mixture Detected with PowerPlex Fusion (24plex STR kit)



22 autosomal STR loci need to be interpreted…(+50% over current 15 STRs)

Size standard not shown

Data courtesy of Becky Hill (NIST)

New Efforts to Improve DNA Interpretation
(especially low-level DNA and mixtures)

Forensic Science International: Genetics 6 (2012) 677–678

Contents lists available at SciVerse ScienceDirect

Forensic Science International: Genetics

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

Editorial

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

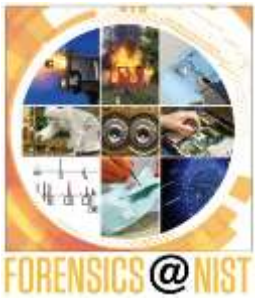
December 2012 – Forensic Science International: Genetics, volume 6, issue 6

Approaches to mixture data interpretation is in a state of change throughout the forensic DNA community

April 12, 2013 Webcast

http://www.nist.gov/oles/forensics/dna-analyst-training-on-mixture-interpretation.cfm

- **8-hours of DNA mixture interpretation training**
- **11 presentations** from **five different presenters**
  - John Butler, Mike Coble, Robin Cotton, Bruce Heidebrecht, Charlotte Word
- **20 poll questions** asked via SurveyMonkey (>600 participated)
  - Addressed additional questions sent via email or Twitter
- **>1000 participants** (almost entire U.S. represented and >10 countries)
- **Will be available for viewing or download** (by early May) for at least six months (storage costs may limit longer-term storage)



NIST FORENSIC SCIENCES

Acknowledgments

Case Examples and Input on This Presentation

- Olga Akselrod (Innocence Project)
- Jennifer Friedman (Los Angeles Public Defender's Office)

Slides and Discussions on DNA Mixtures

- Mike Coble (NIST Applied Genetics Group)
- Robin Cotton & Catherine Grgicak (Boston U.)
- Bruce Heidebrecht (Maryland State Police)
- Charlotte Word (consultant)

Contact info:

john.butler@nist.gov

301-975-4049