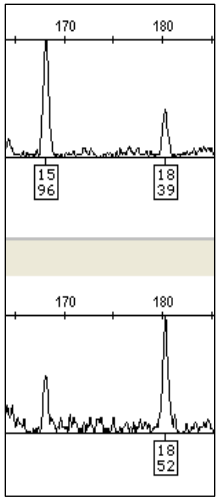
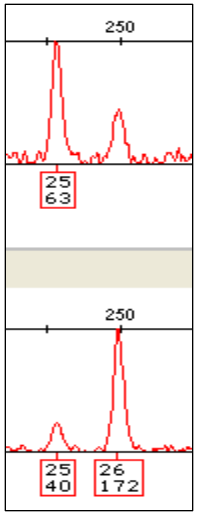
NIST DNA Analyst Webinar Series: Probabilistic Genotyping and Software Programs (Part 1)

May 28, 2014

Why Do We Need to Consider Probabilistic Modeling?

Charlotte J. Word, Ph.D.
Consultant





Why Now? What's Changed?

- 20+ years of DNA testing
- Generally accepted
  - Admitted in courts world-wide
- The "Gold Standard" of forensic sciences

- Changes in: Cases Accepted, Samples Tested, Test Kits and Instrumentation leading to changes in Profile Results and Interpretation

Cases Accepted in the Laboratory

BEFORE

- Homicides, rapes
- High Profile
- Strict laboratory acceptance policies

NOW

- Property crimes also
- Any crime with possible biological sample
- Few (or no) restrictions on samples accepted and tested

Samples Tested

BEFORE

- Blood, Semen, Saliva
- Single Source or Two Person Mixtures
- Large Visible Stains
  - Duplicate samples
  - Reproducibility of result
  - Do additional tests

NOW

- Handled items
- Complex Mixtures – many with Low Template (LT) DNA
- Small or unknown size
  - "Duplicate" swab may not be the same
  - Limited size often; unable to reproduce the results or do additional tests

Profiles Generated/Interpretation

BEFORE

- Artifacts (stutter, pull-up) easily recognized in most profiles
- Degradation not a significant issue usually

NOW

- Difficult to distinguish artifacts vs. true alleles in complex mixtures (especially with LT DNA)
- Degradation more common; more difficult to detect & deal with when have mixtures

Profiles Generated/Interpretation

BEFORE

- Analytical threshold (AT) can be high due to high peaks
- Generally one AT is sufficient
- Instrument differences generally not an issue

NOW

- Analytical threshold needs to be carefully determined
- May need 2 ATs – for Low Template (LT) DNA vs. high peaks
- Instrument differences may affect interpretation of the data

Mixture Interpretation

BEFORE

- Major and minor contributor profiles often can be readily determined in 2 person mixtures
- Can deduce using a known with indistinguishable mixtures

NOW

- Difficult to determine if there is a major contributor; multiple minor contributors a problem
- Generally cannot deduce any profile even if a contributor is known

Statistical Calculations

- Required in the US for all inclusions
  - Case law in many states
  - SWGDAM Guidelines

- If unable to provide a statistical frequency for a potential inclusion statement, may need to report "inconclusive" for the sample
  - Evidence is often deemed inadmissible in court when a "weight" determination is not provided

Statistical Calculations

BEFORE

- Random match probability (RMP) for single source & major or minor contributor
- Combined probability of inclusion (CPI) for many two person mixtures
- Likelihood ratio rarely used (in US)

NOW

- RMP generally not an option unless a clear major contributor is present
- CPI generally not suitable due to possible loss of alleles
- Likelihood ratios rarely used, but may be the next step

Improved Sensitivity of Amplification Kits

BEFORE

- PCR reaction buffer good; inhibition a possible problem
  - More DNA needed
  - 1-2 ng minimum
- Limited flexibility of use of kits
- Few Stochastic Effects

NOW

- Improved reaction buffer; inhibition less of a problem
  - Less DNA needed
  - 0.2-0.75 ng min/max
- Variable/Increased cycle number (LCN testing)
- Increased Stochastic Effects present

Quantity of DNA and Cell Count

Amount of DNA in PCR	Approximate Amount of DNA per Person in ng Approximate Number of Cells 1 Person
1 ng	1
	150
0.5 ng	0.5
	75
0.25 ng	0.25
	38
0.1 ng	0.1
	15
0.05 ng	0.05
	7

Quantity of DNA and Cell Count

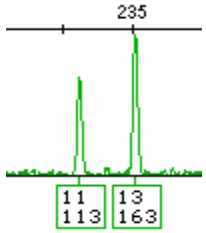
Amount of DNA in PCR	Approximate Amount of DNA per Person in ng		
	Approximate Number of Cells		
	1 Person	2 Person Mixture 1:1	2 Person Mixture 4:1
1 ng	1	0.5 + 0.5	0.8 + 0.2
	150	75 + 75	120 + 30
0.5 ng	0.5	0.25 + 0.25	0.4 + 0.1
	75	38 + 38	60 + 15
0.25 ng	0.25	0.125 + 0.125	0.2 + 0.05
	38	19 + 19	30 + 7
0.1 ng	0.1	0.05 + 0.05	0.075 + 0.025
	15	7 + 7	11 + 4
0.05 ng	0.05	0.025 + 0.025	0.04 + 0.01
	7	4 + 4	6 + 1

Quantity of DNA and Cell Count

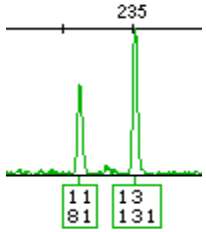
Amount of DNA in PCR	Approximate Amount of DNA per Person in ng / Approximate Number of Cells			
	3 Person 1:1:1	3 Person 5:2:1	4 Person 1:1:1:1	4 Person 5:2:2:1
1 ng	0.33 x 3	0.6+0.25+.125	0.25 x 4	0.5+0.2+0.2+0.1
	50 x 3	94 + 38 + 19	38 x 4	75 + 30 + 30 + 15
0.5 ng	0.16 x 3	0.31+0.12+0.06	0.125 x 4	0.25+0.1+0.1+0.05
	24 x 3	47 + 18 + 9	19 x 4	38 + 15 + 15 + 7
0.25 ng	0.08 x3	0.15+0.06+0.03	0.062 x 4	0.12+0.05+0.05+0.02
	12 x 3	23 + 9 + 4	9 x 4	18 + 7 + 7 + 3
0.1 ng	0.03 x3	0.062+0.02+0.01	0.025 x 4	0.05+0.02+0.02+0.01
	5 x 3	10 + 3 + 1	4 x 4	7 + 3 + 3 + 1
0.05 ng	0.016 x3	0.03+0.012+0.006	0.0125 x 4	0.025+0.01+0.01+0.005
	2 x 3	5 + <2 + <1	2 x 4	4 + 1 + 1 + <1

Low Template DNA Leads to Stochastic Effects

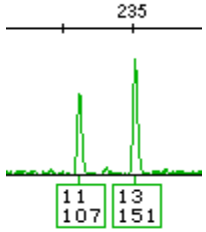
1



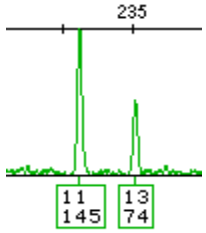
2



3



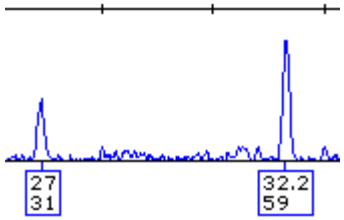
4



Peak Height Imbalance at Heterozygous Loci – Peak Height Ratio has less meaning

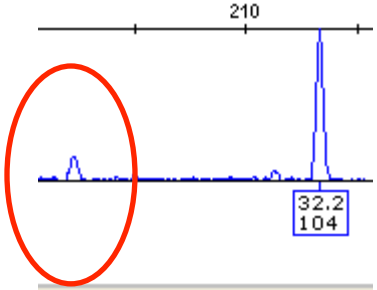
Allele Drop-Out

Allele Loss

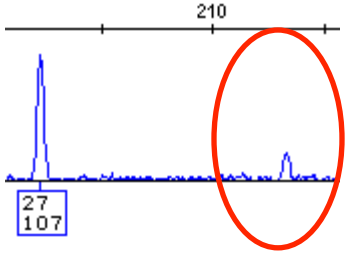


1

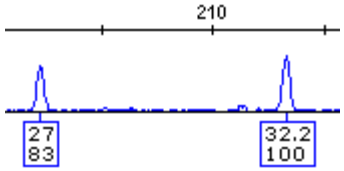
2



3

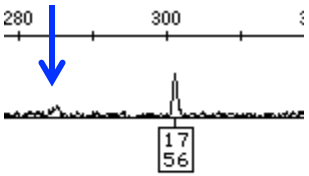


4

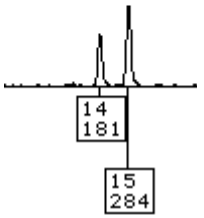


Low Template DNA Leads to Stochastic Effects

1

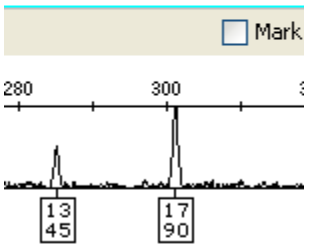


Allele Gain

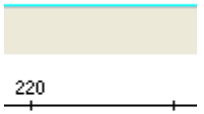


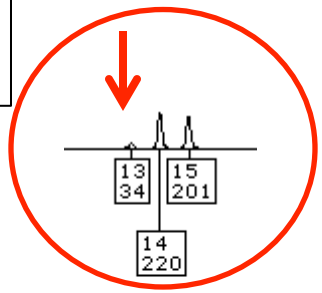
1

2



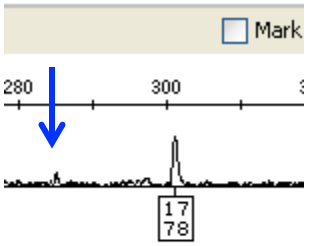
Elevated Stutter



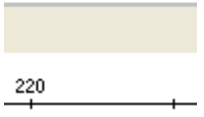


2

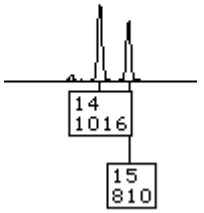
3



Allele Drop-In (& Drop-out)

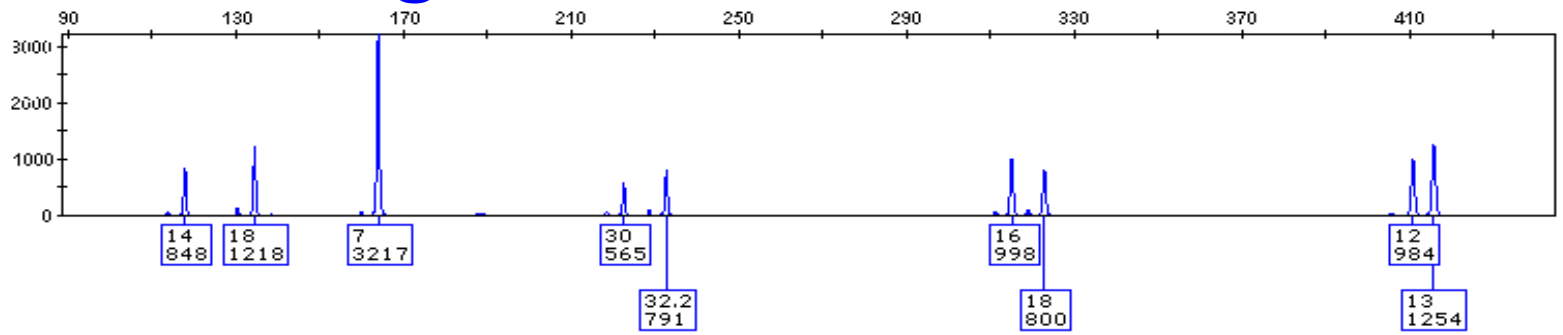


4



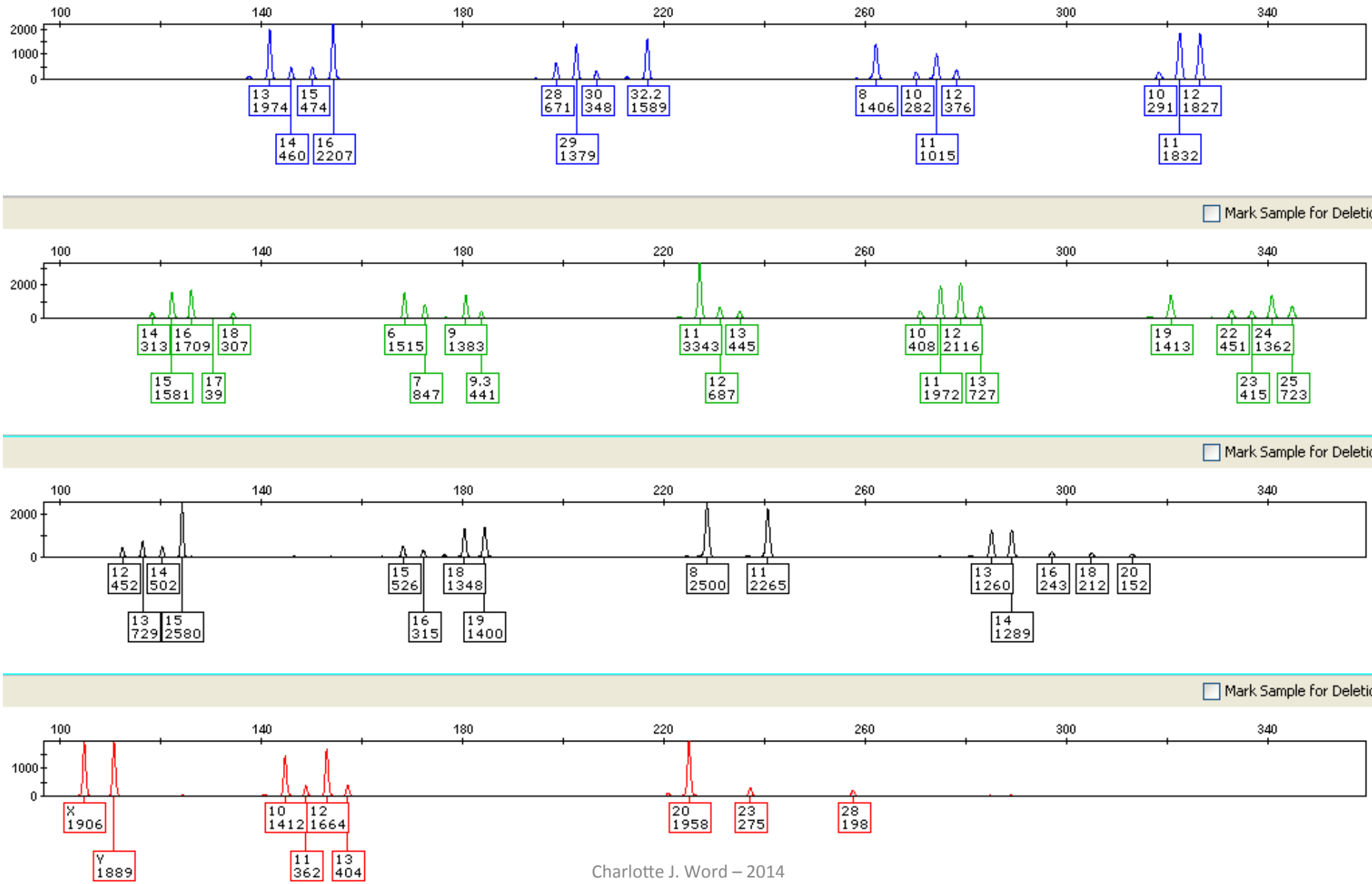
3

Single Source or a Mixture?

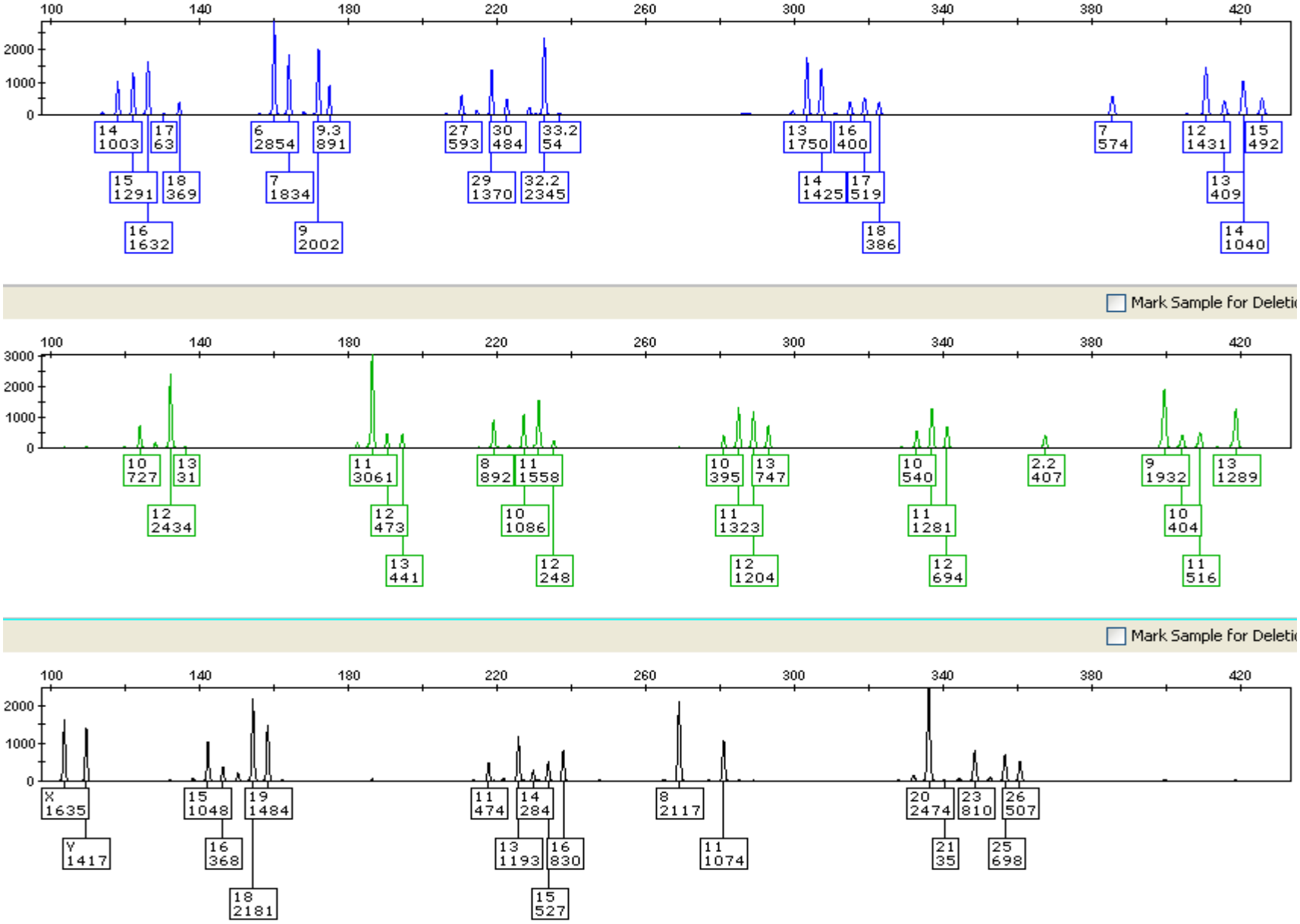


Single Source with Peak Height Imbalance and Drop-in

Number of Contributors? Major Contributor?

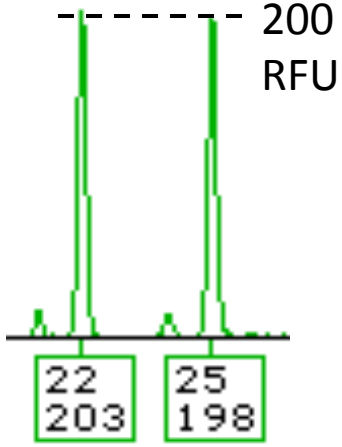


Are ALL Alleles Present? Artifacts?



How Can We Assess if ALL Alleles are Present?

- Stochastic threshold – not an ideal method
  - Gives a good estimate of where caution is needed
  - "Line drawn in the sand" – peaks of almost same height treated differently
  - Ignores the increased likelihood of stochastic effects as the amount of DNA amplified and peak heights decrease



- Estimate probability of drop-out

Confidence in Data

THEN		NOW
High Confidence	⟹	Decreased Confidence
Minimal Uncertainty	⟹	Higher Uncertainty

Real allele vs. stutter, pull-up artifact
Real allele vs. drop-in "allele"
Number of contributors
Major/minor contributors
Peak height ratios
Mixture ratios
Ability to exclude a non-contributor

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

-Dennis Lindley

Borrowed from Michael Coble, NIST

Now and Future for Complex Mixtures and LT DNA Profile Interpretation

Paradigm shift in the field requires a shift in the methods used

- Probabilistic Modeling of data with Likelihood Ratio calculations may be the answer for some of the profiles obtained today
  - Some profiles may still be uninterpretable
- Validations needed
- TRAINING needed (analysts, attorneys, judges, law enforcement)

THANK YOU!!

John Butler

Mike Coble

Robin Cotton

Catherine Grgicak

Bruce Heidebrecht

For many hours of discussions!

Catherine Grgicak

Robin Cotton

NIJ Grant to Boston University

For the profiles!

Crime lab analysts for excellent questions and comments from past presentations