National Institute of Standards and Technology • U.S. Department of Commerce

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

FORENSICS @ NIST

Different Assumptions & Different Interpretations
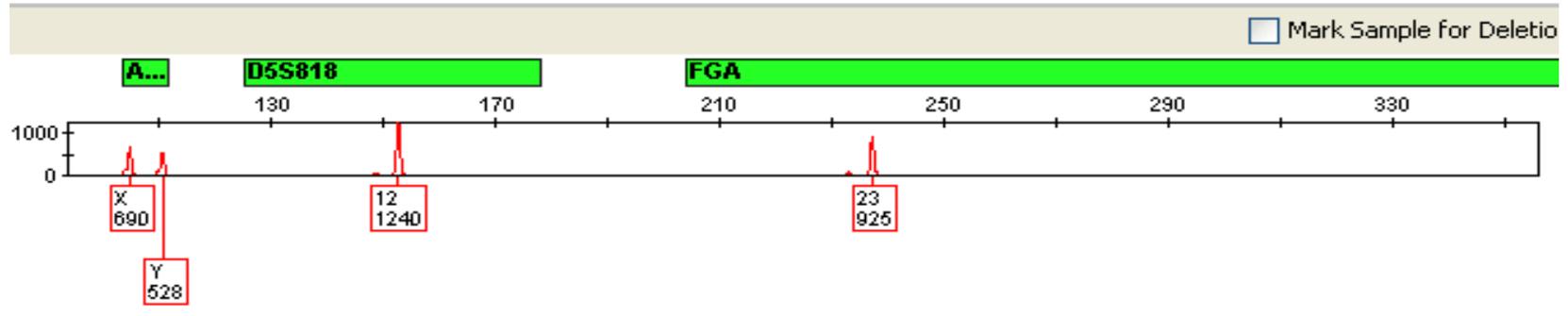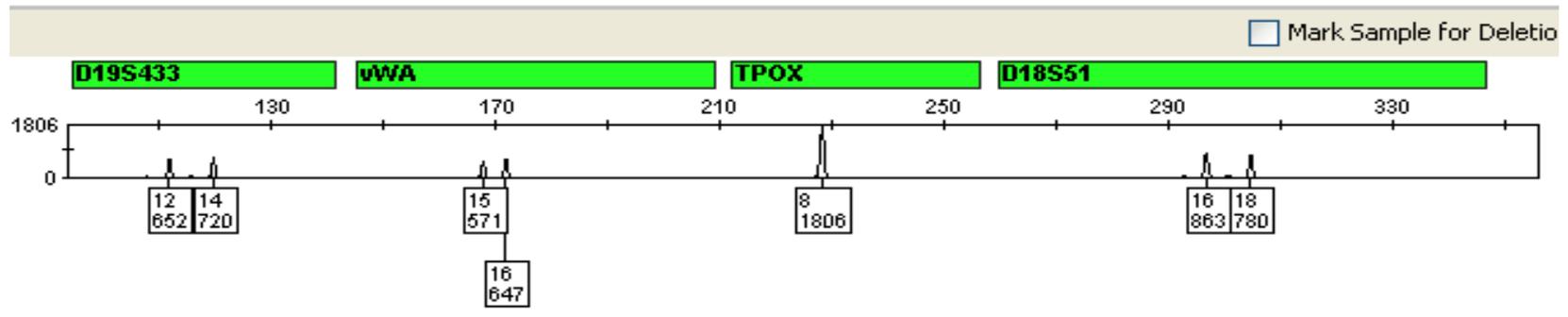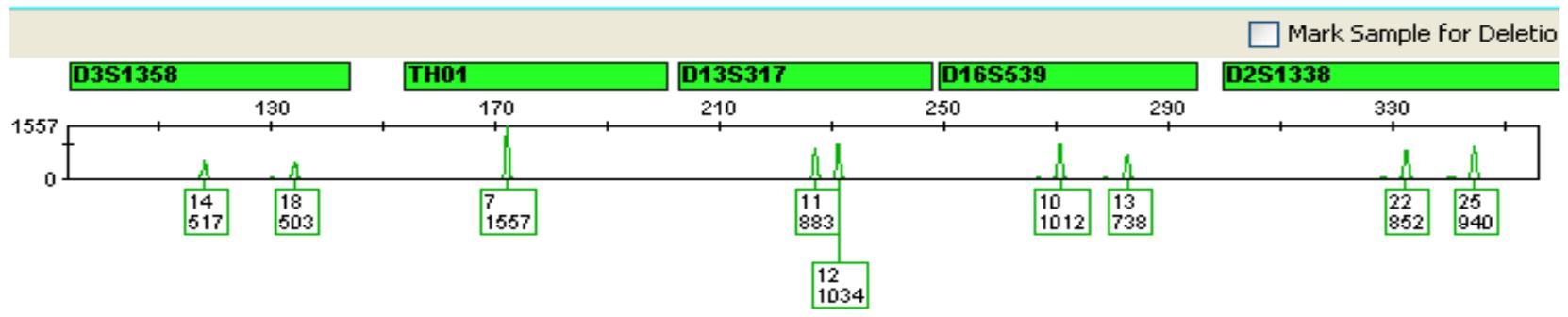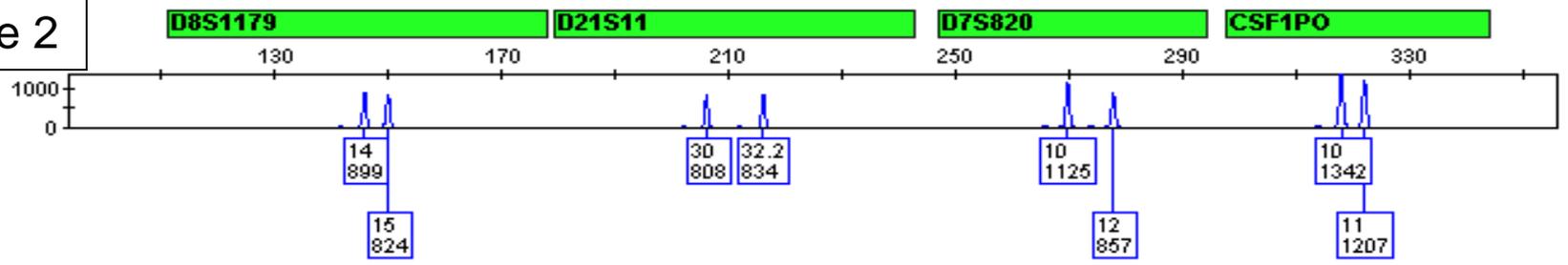
Charlotte J. Word

Consultant

MYTH

No assumptions are needed for interpreting DNA profiles from good quality single source samples.
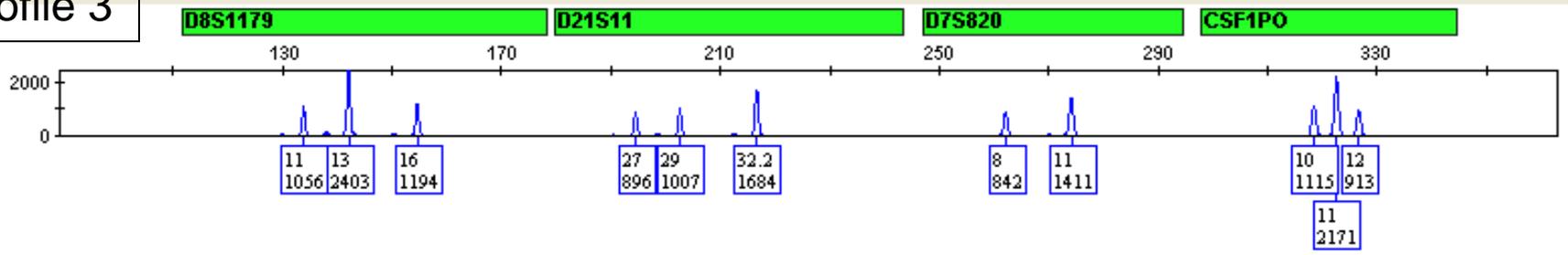
Profile 2

Single Source Sample

Assumptions Made
Single Source

- Peaks above the analytical threshold are alleles from the contributor
  - Stutter peaks, other peaks are assumed to be artifacts and can be ignored
- All alleles from the contributor are present since all peaks are above the stochastic threshold
- There is a single DNA contributor
  - No more than two alleles at any locus
  - Genotypes are easy to assume
    - Balanced peak heights where heterozygous
    - Double peak height where homozygous

Profile 3

Two-Person Mixture

Assumptions Made
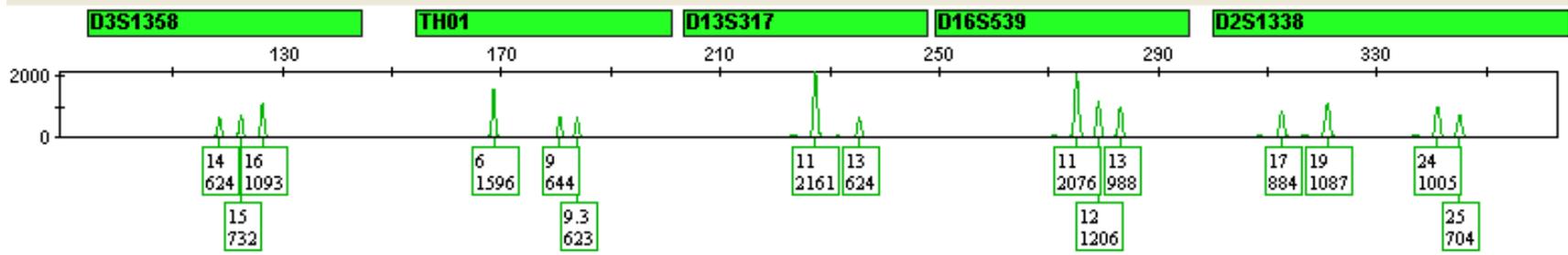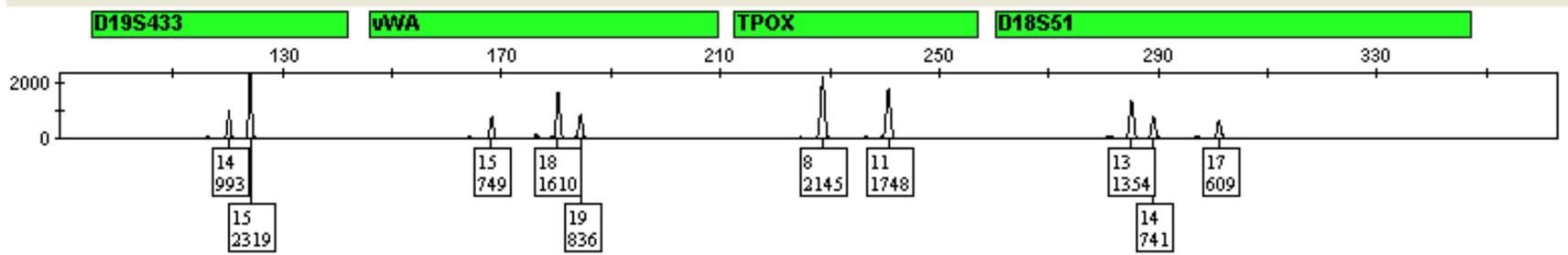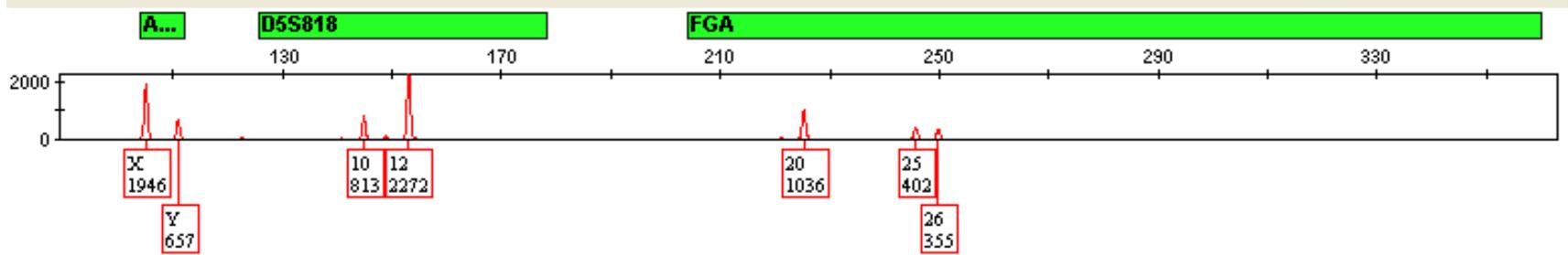## Two Person Mixture

- Peaks above the analytical threshold are alleles from the contributors
  - Stutter peaks, other peaks are assumed to be artifacts and can be ignored
- All alleles from the contributors are present since all peaks are above the stochastic threshold
- There are (only) two DNA contributors
  - No more than four alleles at any locus
  - Data consistent with mixture validation studies and experience with two person mixtures

Assumptions Made
Two Person Mixture

- Genotypes may be easily assumed
  - If have major:minor scenario, can use mixture ratio and peak height ratios to associate alleles into genotypes and associate genotypes into complete profiles
  - Can assume one known is a contributor and deduce the second contributor
  - If have indistinguishable mixture, can assume a limited number of possible genotypes and genotype combinations at each locus: (e.g., alleles 13,14,15,16 = genotypes of 13,14 + 15,16 or 13,15 + 14,16 or 13,16 + 14,15)

Assumptions

- Assumptions are made with all data analyses and with all interpretations of data

- We may not always clearly state those assumptions or even be aware that we are making those assumptions

- We may not always report those assumptions

But we MUST be aware of what assumptions we are making

MYTH

No assumptions are needed for interpreting DNA profiles from good quality single source samples.

Assumptions

- We have a lot of familiarity and experience making reasonable assumptions for high quality single source and two person mixtures

- High quality profile leads to high confidence in data and high certainty regarding interpretations and conclusions

But what about REAL Casework Profiles?!

REAL Casework

Situations with increased uncertainty, and therefore decreased confidence:

➢ Alleles vs. artifacts? (LT or high level DNA)

➢ Stochastic effects possible? (Low peak heights; all or some below stochastic threshold)

  ➢ Sure all alleles are present (drop-out)?

  ➢ Elevated stutter & drop-in present?

➢ Number of contributors? 1, 2, 3 or more?

➢ Inability to associate all alleles into reasonable genotypes with high confidence

➢ Degradation?

MYTH

It may be useful to consider some DNA profiles under different assumptions.

Profile 4
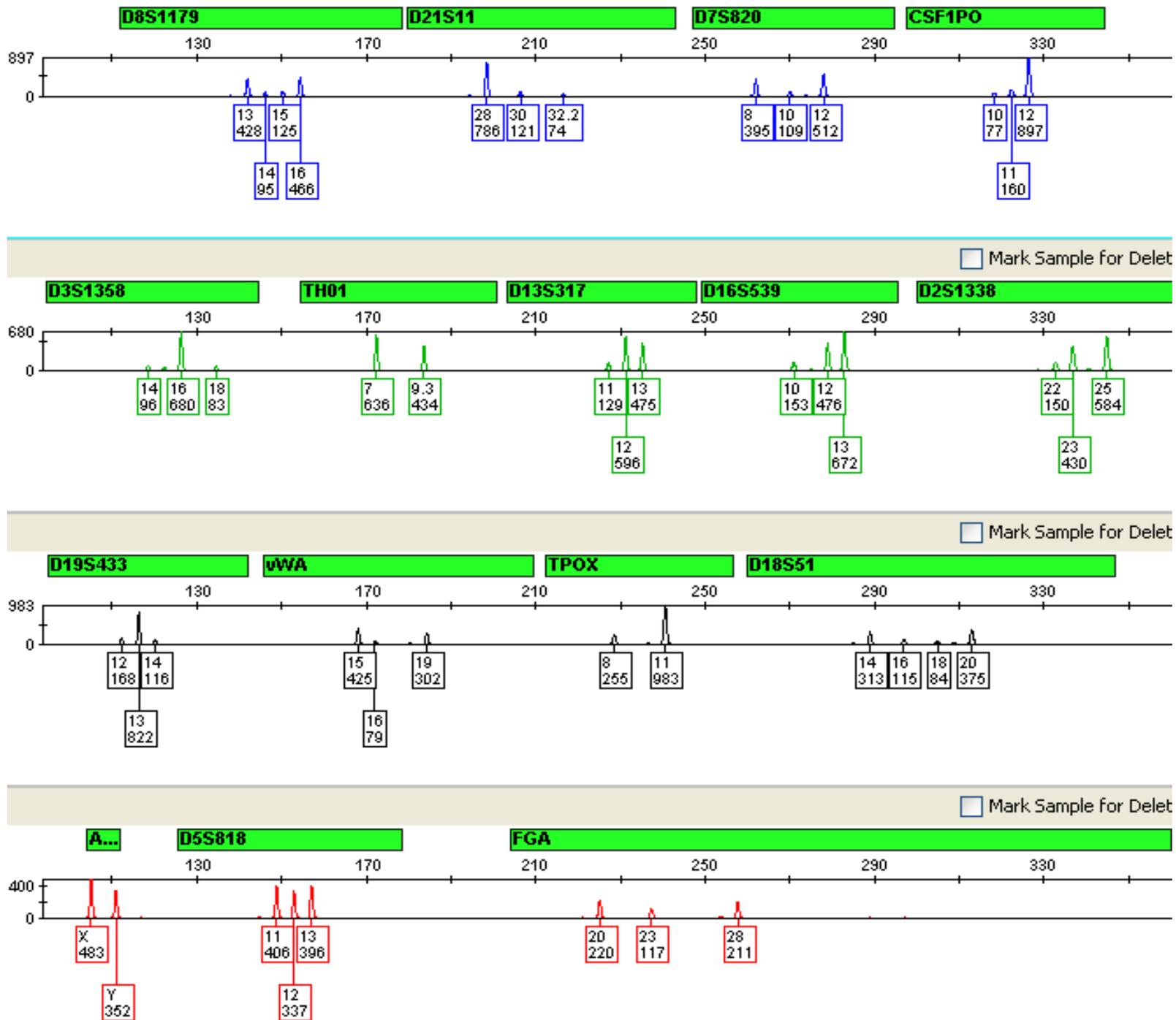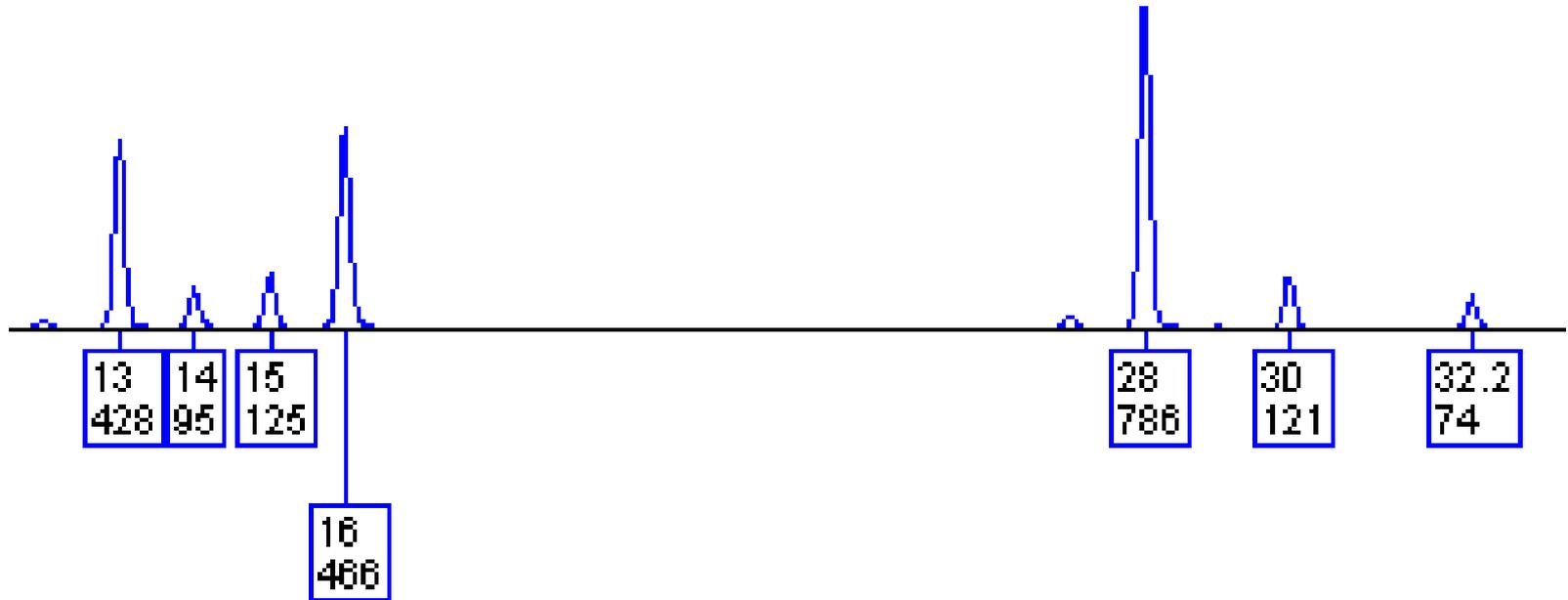
Single Source vs. Mixture?
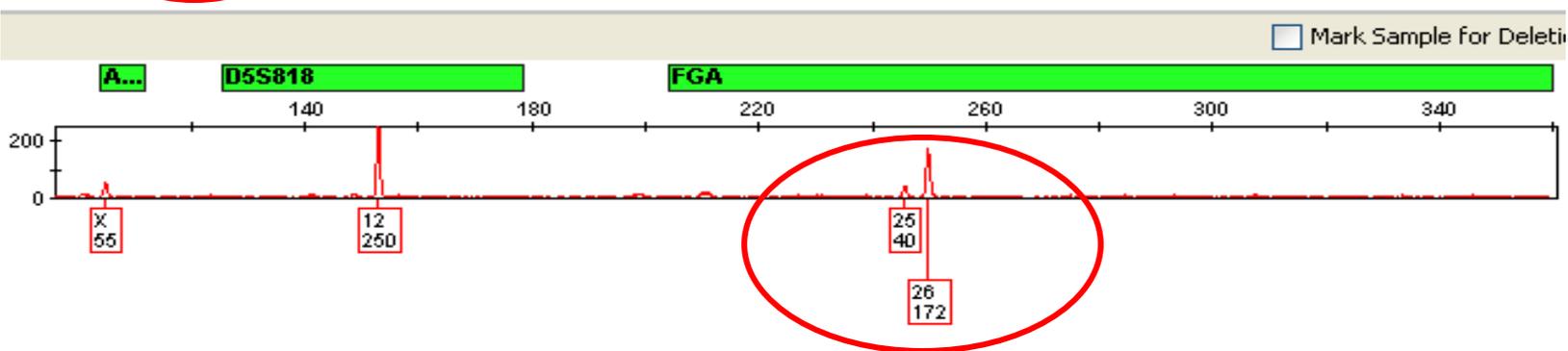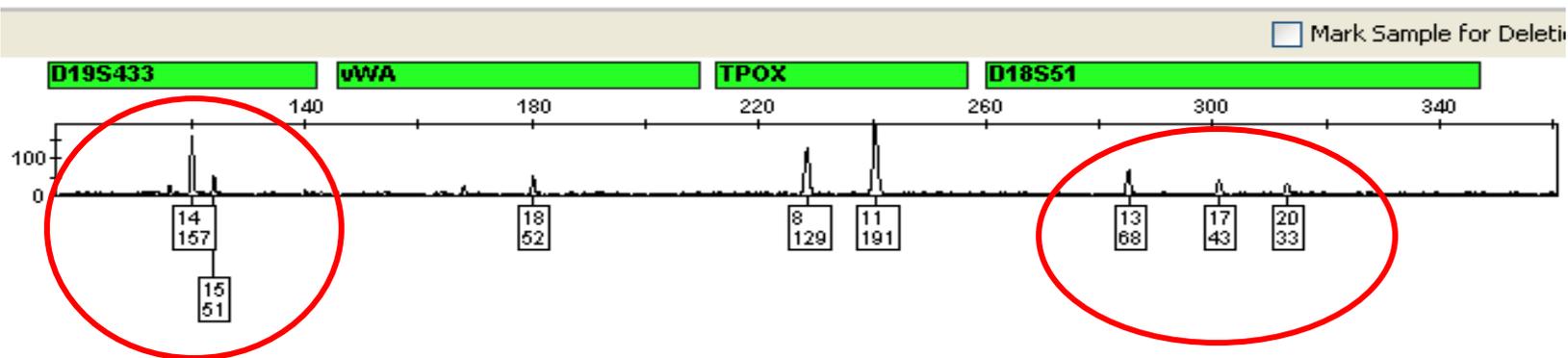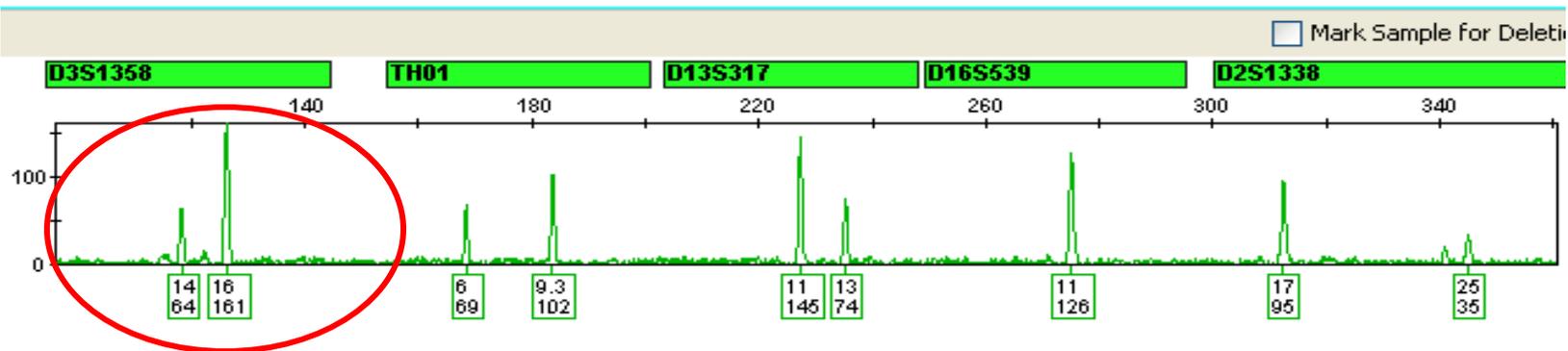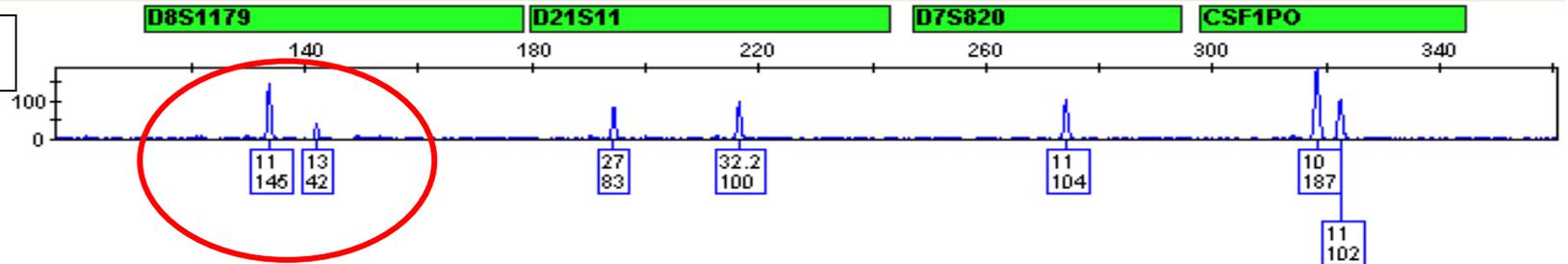
Profile 5

All Alleles vs. Allelic Drop-out?
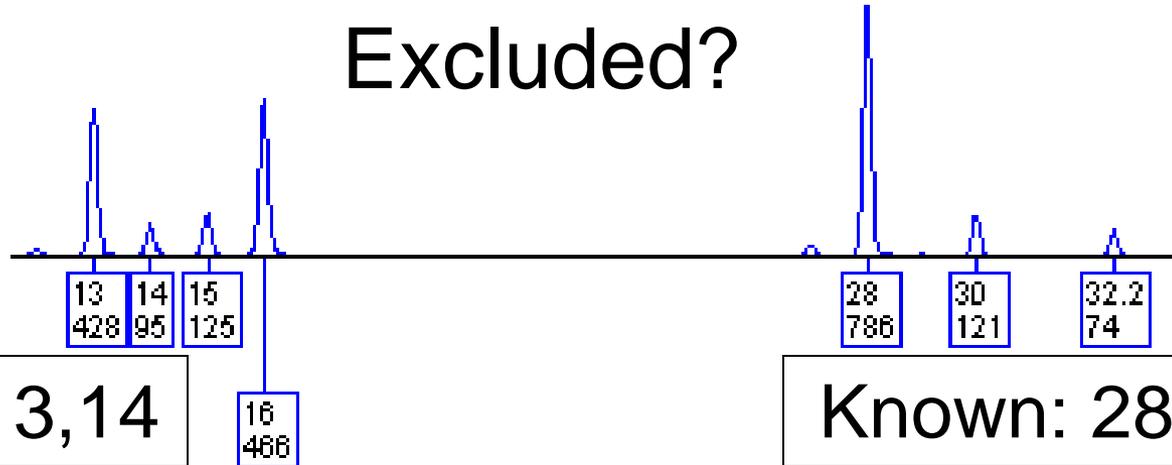Two or More Contributors?

Profile 5

Is Known Individual Included or Excluded?

13 428 | 14 95 | 15 125

16 466

28 786 | 30 121 | 32.2 74

Known: 13,14

Known: 28,30

Is Known Individual Included or Excluded?

Profile 5


```
13    14  15
428   95  125
```

```
28    30    32.2
786   121   74
```

```
16
466
```

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

Genotype is excluded even if alleles are included

# Is Known Individual Included or Excluded?

Profile 5

Known: 13,14

Known: 28,30

Peak labels: 13/428, 14/95, 15/125, 16/466, 28/786, 30/121, 32.2/74
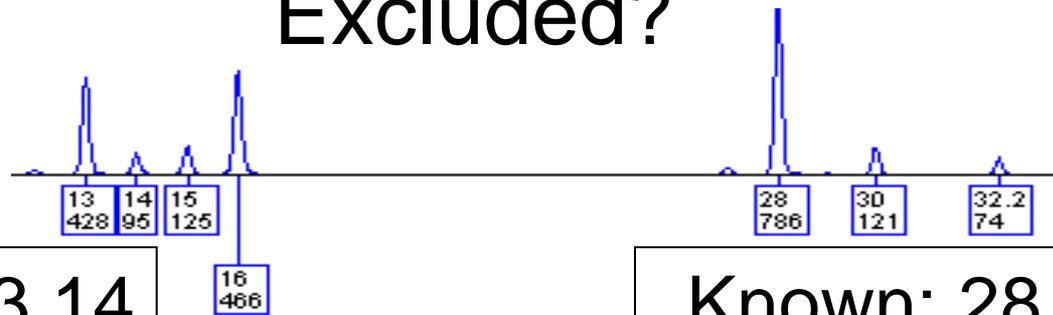
New Assumption:
1) 3 contributors *and* possible LT DNA →
2) 1 major and 2 minor contributors →
3) Major must have genotype of 13,16 and 28,28
4) One or other or both minor contributors have 14 and/or 15 and 30 and/or 32.2, but cannot associate alleles to genotypes
5) Possible genotype list is long due to stochastic effects

# Is Known Individual Included or Excluded?

Profile 5

Known: 13,14

Known: 28,30

13 428
14 95
15 125
16 466
28 786
30 121
32.2 74
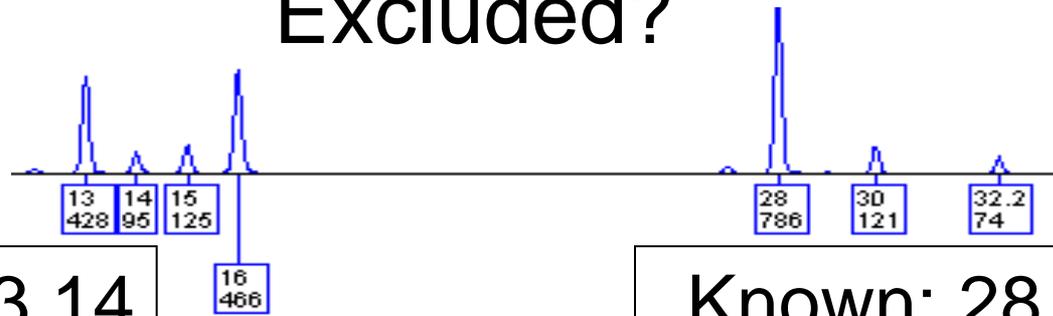
Other Possible Assumptions:
1) 3 or more contributors *and* possible LT DNA→
2) 1 major and 3 or more minor contributors → OR
3) 2 major contributors and 1 minor contributor → OR
4) 2 majors and 2 or more minor contributors →
5) Decreased ability to associate alleles to genotypes
6) Possible genotype list is long

Is Known Individual Included or Excluded?

Profile 5

13	14	15
428	95	125

16
466

| 28	30	32.2
786	121	74

Known: 13,14

Known: 28,30
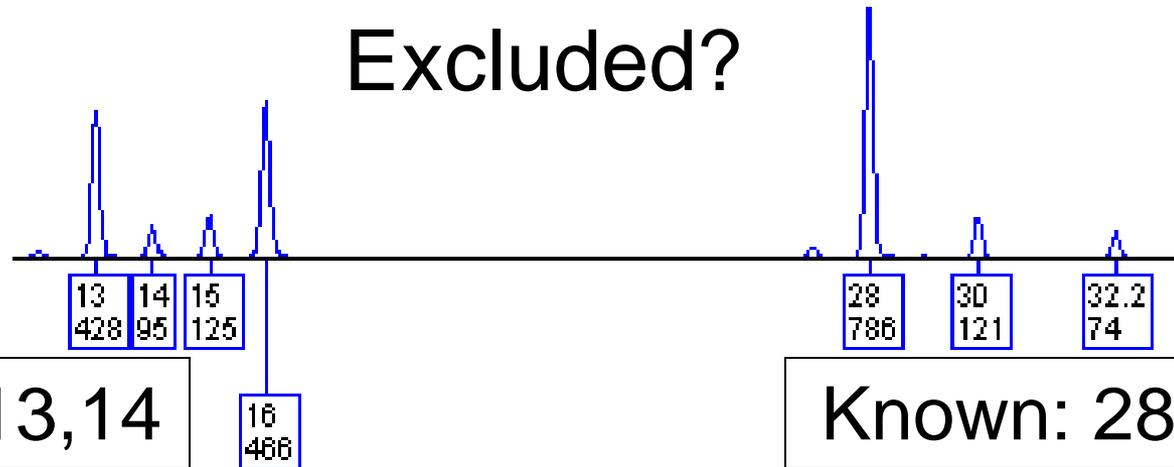
Based on the assumption of 3 or more contributors, there is insufficient information to exclude known genotypes. What do you report?

Inclusion – but statistics MUST take into account possible stochastic effects (may not be meaningful)
Inconclusive – but throwing away possibly exculpatory or inculpatory data

Is Known Individual Included or Excluded?

Profile 5

Known: 13,14

Known: 28,30
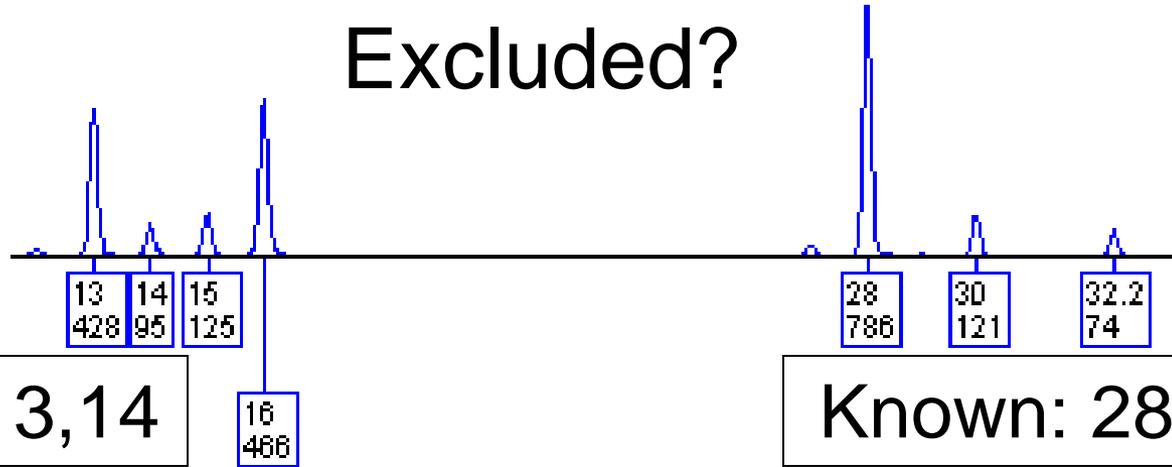
Which set of assumptions is "correct"?

May need to report using more than one assumption set!

Reporting Multiple Conclusions

Different conclusions may result from using different assumptions.

If 2 contributors: **EXCLUDED**

BUT

If ≥3 contributors: **INCLUDED**
INCONCLUSIVE

REPORT ALL CONCLUSIONS!

MYTH

It may be useful to consider some DNA profiles under different assumptions.

Indistinguishable Mixture Profile

13 197   14 304   15 280   16 247

28 471   30 223   32.2 287

Known: 13,14

Known: 28,30

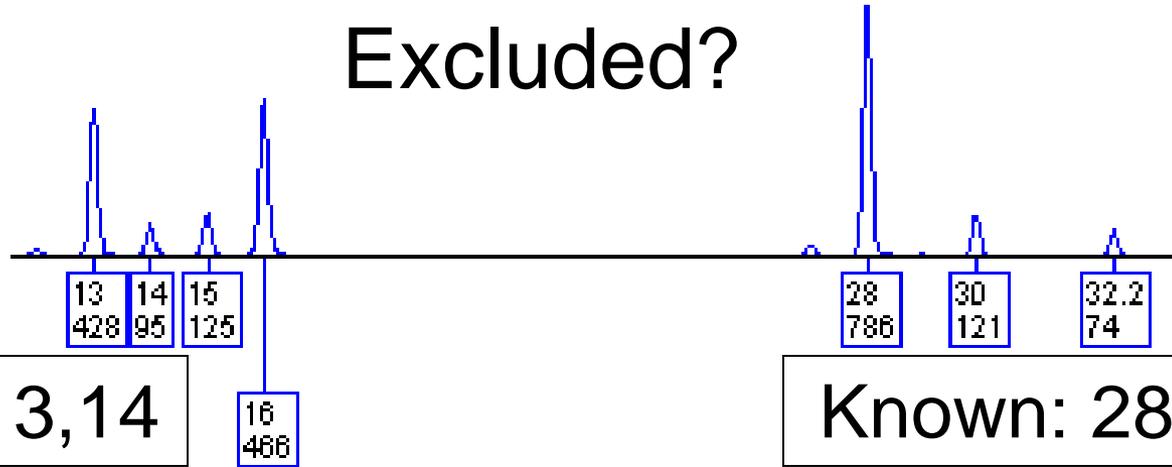## What if the genotypes CANNOT be distinguished?

Alleles are included, BUT are genotypes?

We know from previous data this person is excluded!
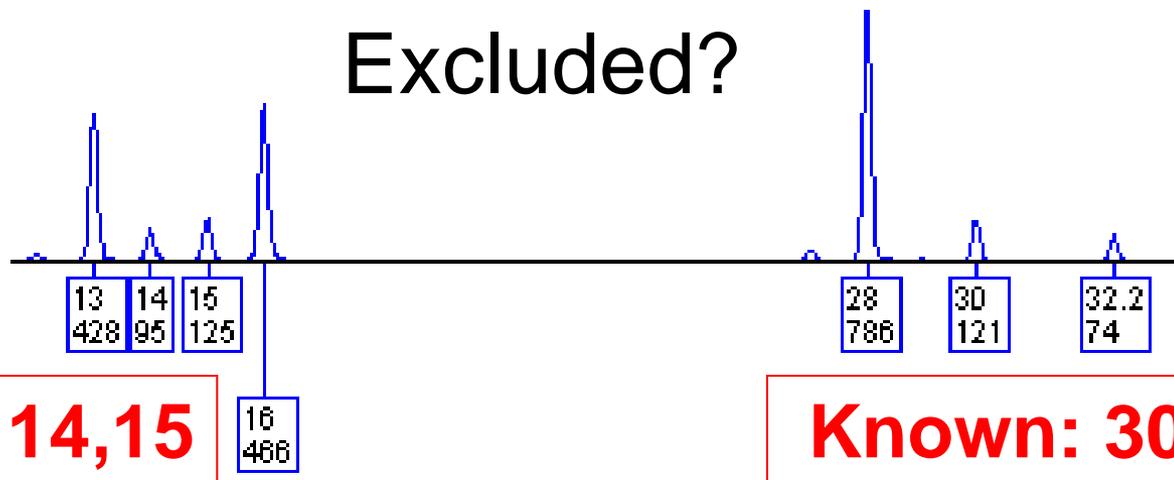(assuming 2 contributors)

Is Known Individual Included or Excluded?

Profile 5

13 428 14 95 15 125

16 466

28 786 30 121 32.2 74

Known: 13,14

Known: 28,30

Which set of assumptions is "correct "?

Is Known Individual Included or Excluded?

Profile 5

Known: 14,15

Known: 30,32.2

13 428   14 95   15 125

16 466

28 786   30 121   32.2 74
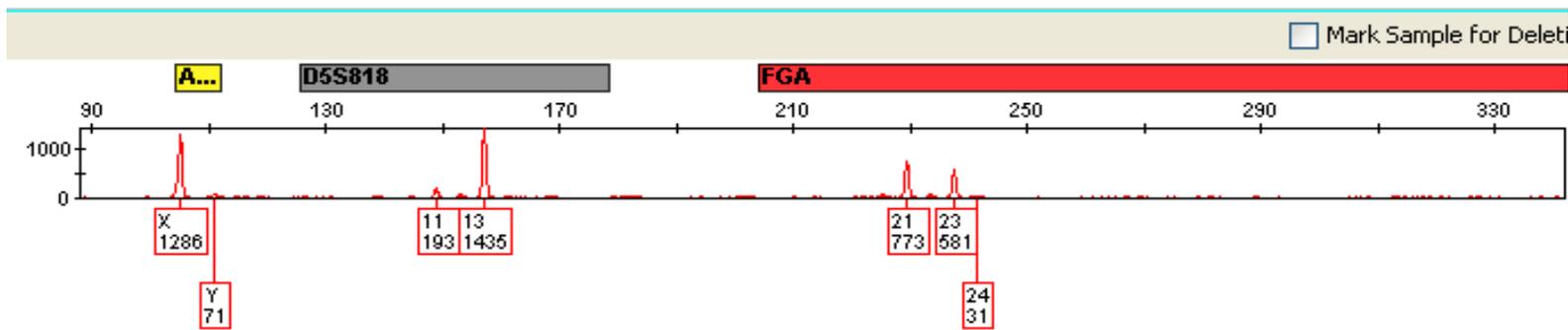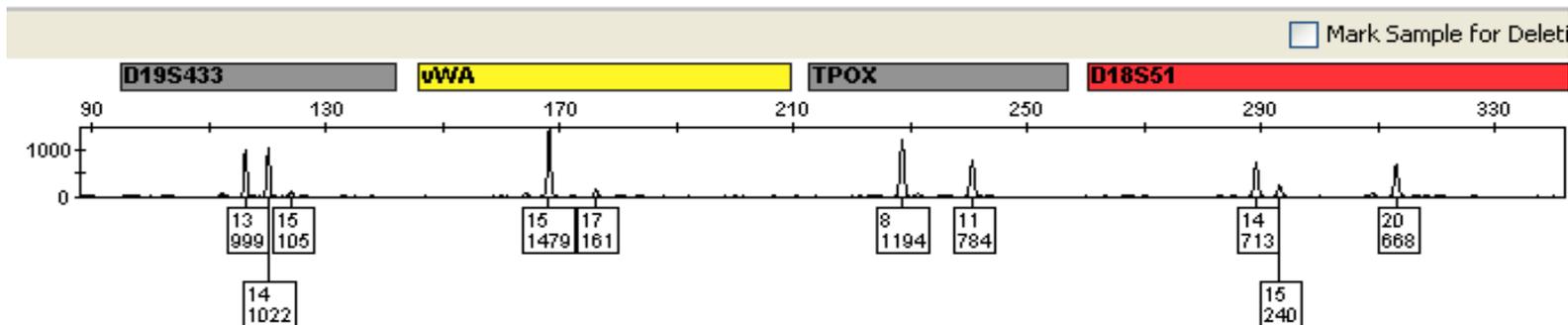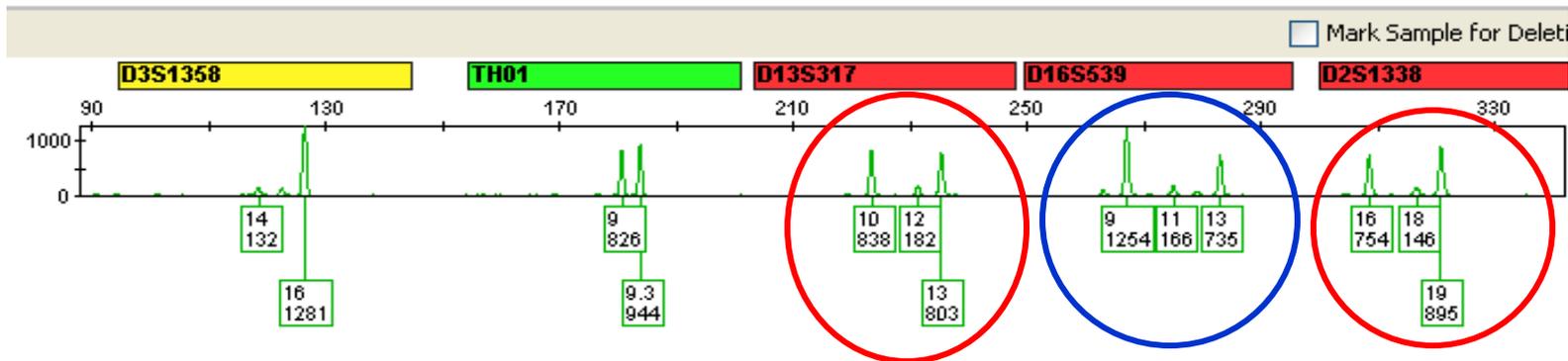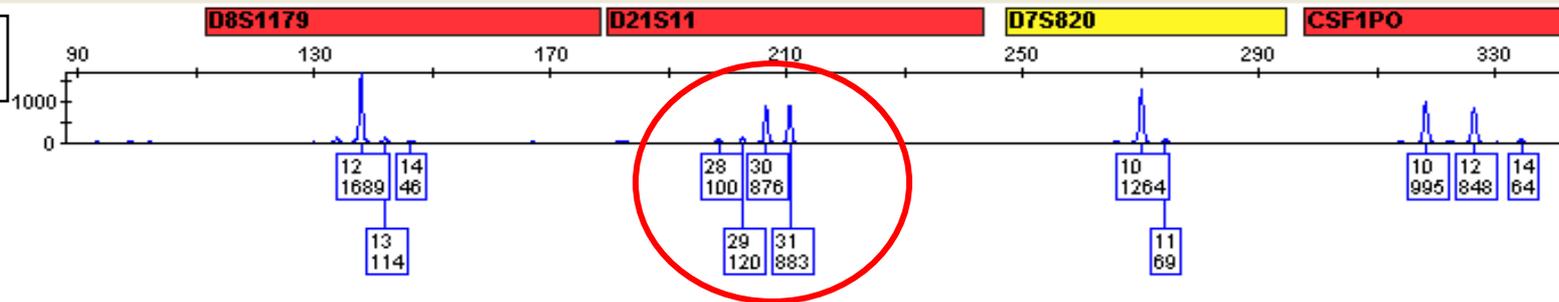
Which set of assumptions is "correct"?

What if known genotypes are different and included as the single minor contributor under the assumption of only two contributors?
Include with appropriate statistics
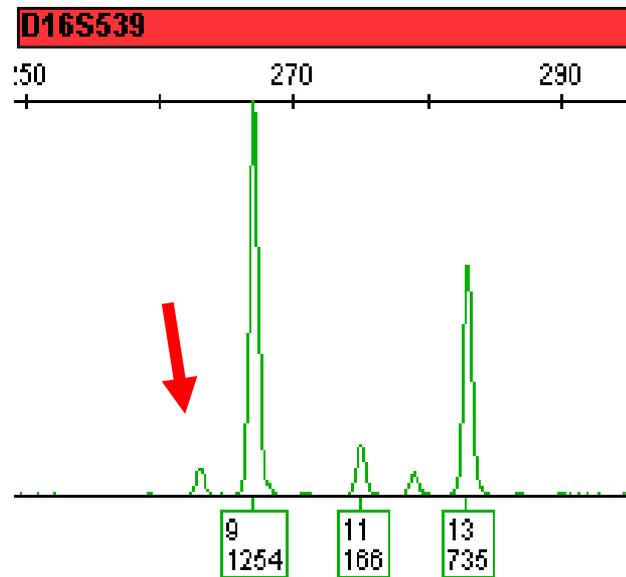What if ≥3 contributors? Include? Exclude?
Inconclusive?

Profile 1

Stutter or true allele?
All alleles present?

Profile 1

Stutter or true allele? All alleles present?

8,11 = true minor contributor

8 allele filtered out by software

If assume 8 is a stutter peak and assume all peaks are present, would exclude the true contributor!

Uncertainty in evaluating the presence or absence of alleles leads to false inclusions and exclusions

Inclusion/Exclusion Criteria

- Must have a good interpretation procedure for excluding individuals who are non-contributors to the DNA sample

- If fail to exclude an individual as a possible contributor, you **MUST** have a statistical approach that embraces all of the possible included alleles and genotypes

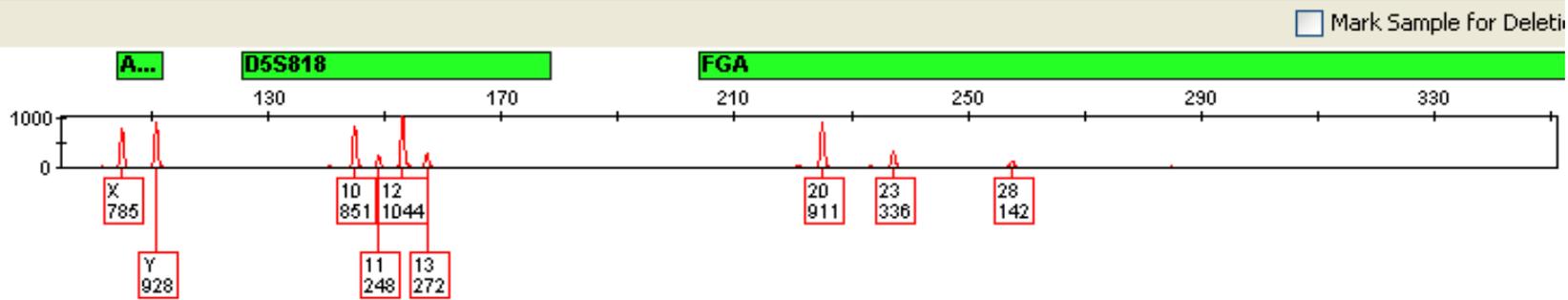- Must consider possible reasonable alternatives

Profile 7

Major vs. Indistinguishable?
2 contributors?

Profile 8
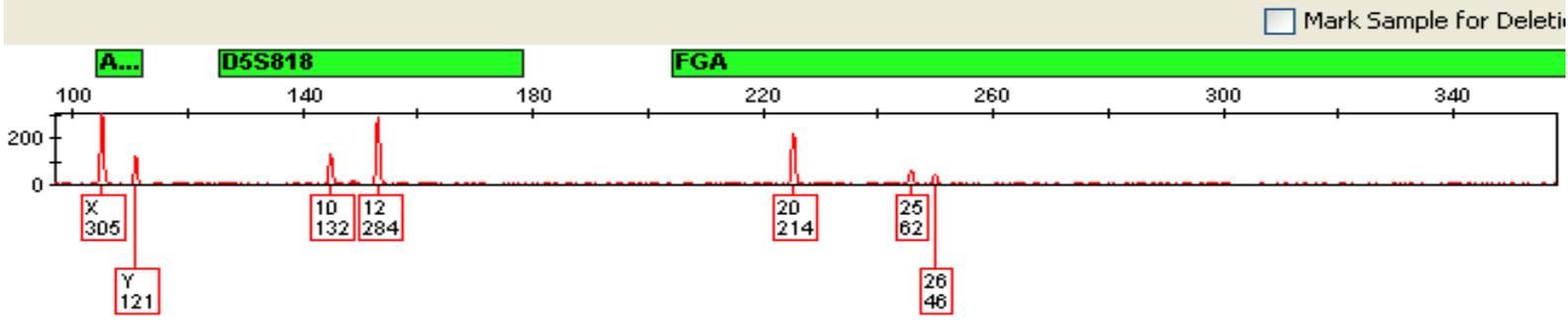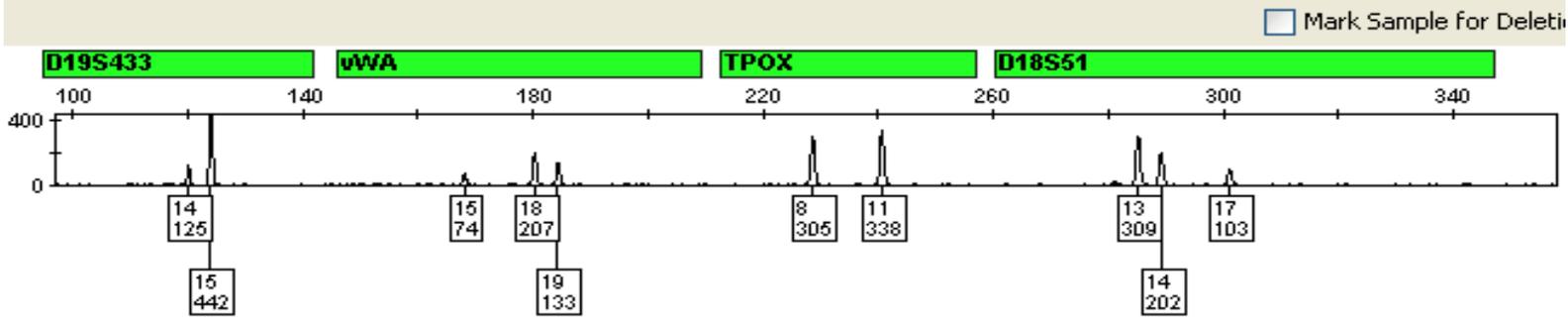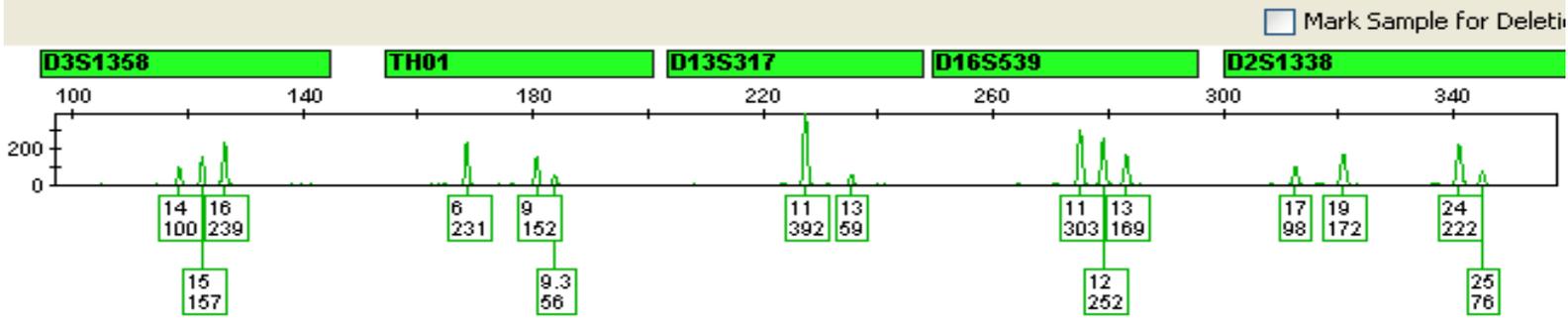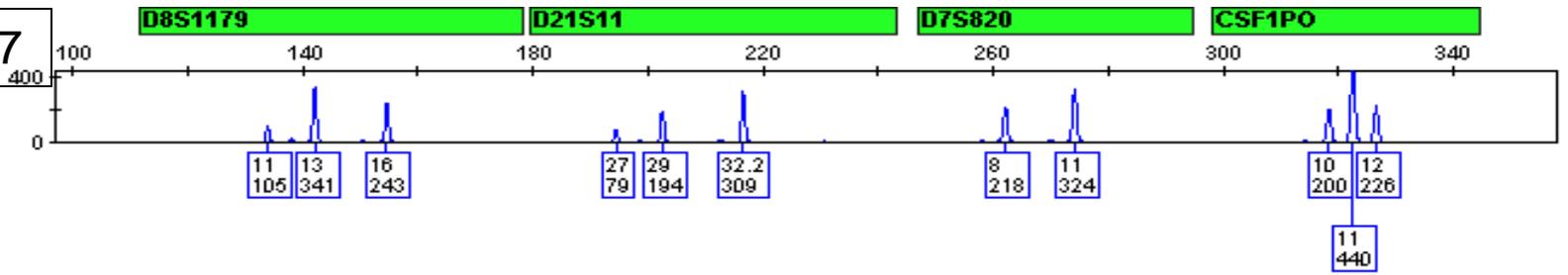
Major vs. Indistinguishable? Number of Contributors?
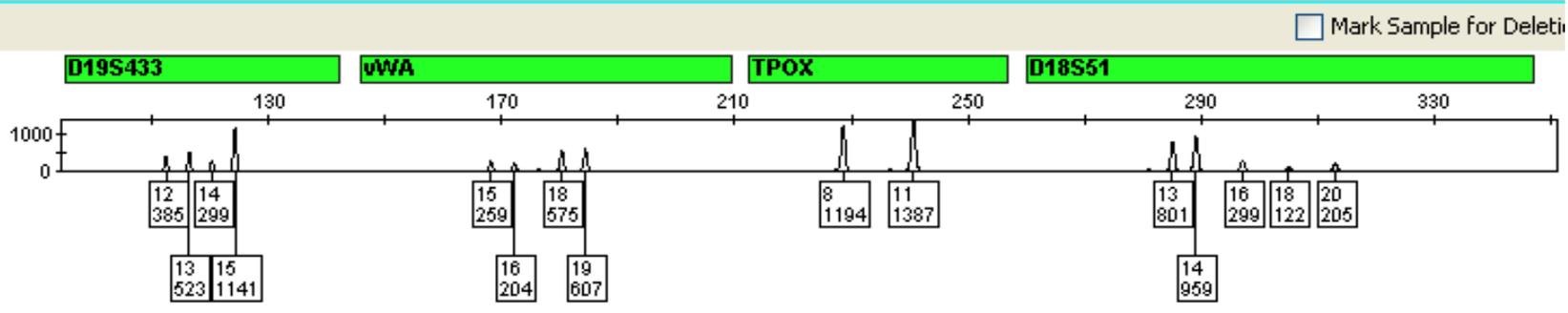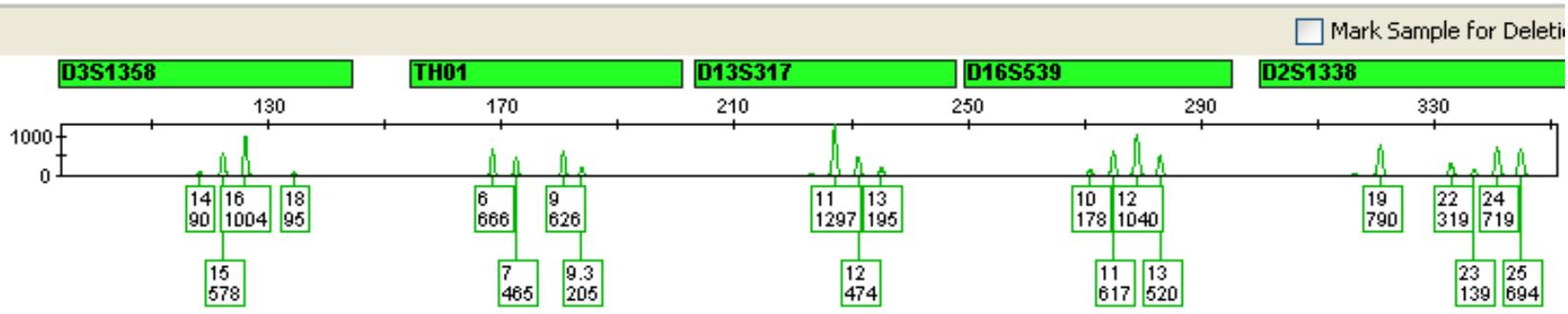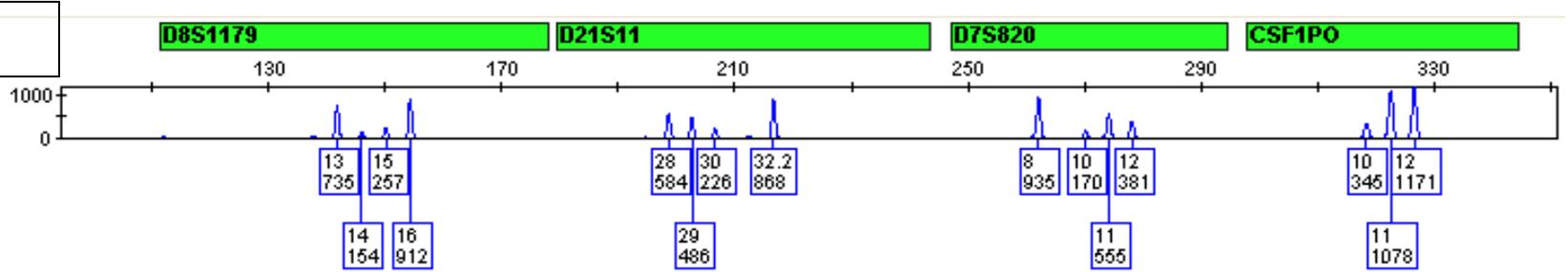
When to Consider Different Assumptions

May need to consider multiple assumptions for data interpretation when:

➢Possible LT DNA profile

➢Stochastic effects (allelic drop-in, allelic drop-out, elevated stutter)

➢Possible artifact vs. true allele

➢Possible minor contributor in mixed DNA profile

➢Possible known contributor(s) and deducing

➢More than 2 contributors (later today)

What do you do when…

You have increased uncertainty, and therefore decreased confidence?

Options for interpreting and reporting:

1. Do not interpret the data → report inconclusive
   - When uncertainty is too high

2. Pick one interpretation to report
   - When have minimal uncertainty

3. Interpret and report the data under two or more different assumptions
   - When certainty is medium-to-high but possible scientifically sound alternatives exist

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?

Reporting

- Consider the data from several scientific perspectives – for conclusions *and* statistical calculations

- Report all appropriate scientific conclusions and opinions in the laboratory report

- ESPECIALLY if the conclusions differ under different reasonable assumptions

Why Report?

- Opinions may be important to different individuals reading the report (e.g., law enforcement, prosecutor, defense attorney, client, judge, jury)

- Reports should be **neutral** to the case yet address the question(s) asked by the client

Why Report?

- Not all cases (<10%) make it to court

- Critical decisions often based on report and (mis)understandings alone

- If not provided in advance to all parties, opinions may not be admissible in court

Summary

- EVERY interpretation requires assumptions
- Assumptions MUST be made from the data alone and prior to knowing the profiles of the known contributors
  - Artifact, stutter vs. true alleles
  - Number of contributors
  - Major:minor contributors
- All assumptions must be documented and should be reported
- Just because the known profile "fits" the data under one assumption set does not mean those are the correct assumptions and the correct conclusion

THANK YOU!!

John Butler

Mike Coble

Robin Cotton

Catherine Grgicak

Bruce Heidebrecht

& Workshop attendees

For many hours of discussions!

Catherine Grgicak

Robin Cotton

NIJ Grant to Boston University

For all of the profiles!