DNA Mixture Analysis:

Principles and Practice of Mixture Interpretation and Statistical Analysis Using the SWGDAM STR Interpretation Guidelines

---

Complex Mixtures – Strategies and Challenges

Gary Shutler

Washington State Police Crime Lab

AAFS 2011 Workshop #17

Chicago, IL

February 22, 2011

- In general deconvolution strategies and statistical weight approaches are for 2 person mixtures since that is currently the most common condition encountered in casework.

So what do you do when you get difficult/complex sample results?

- Your first thoughts:
  - Is there any help in the new SWGDAM Guidelines?

Section 3.5 Interpretation of DNA Typing Results from Mixed Samples

- Alternatively, if the amounts of biological material from multiple donors are similar, it may not be possible to further refine the mixture profile. When major or minor contributors cannot be distinguished because of similarity in signal intensities, the sample is considered to be an indistinguishable mixture. **The classification as indistinguishable may be limited to some, not all, of the loci for which DNA typing results are obtained and does not imply that the profile is uninterpretable. Individuals may still be included or excluded as possible contributors to an indistinguishable mixture.**

also

- 3.5.5 Mixture with Multiple Major Contributors and one or more minor contributors.

  - The laboratory should establish guidelines based on peak height ratio assessments and/or mixture ratios for determining whether multiple major contributors are present in a mixed sample.

So what do you do when you get difficult/complex sample results?

- Your second thoughts:
  - Annual leave, followed by extended sick leave,
  - maternity/paternity leave,
  - go back to school and get an easier better paying job,
  - retire and live off the land,
  - become a crazy street person and yell at traffic (side effect of mixture interpretation).

Well maybe there is still some hope

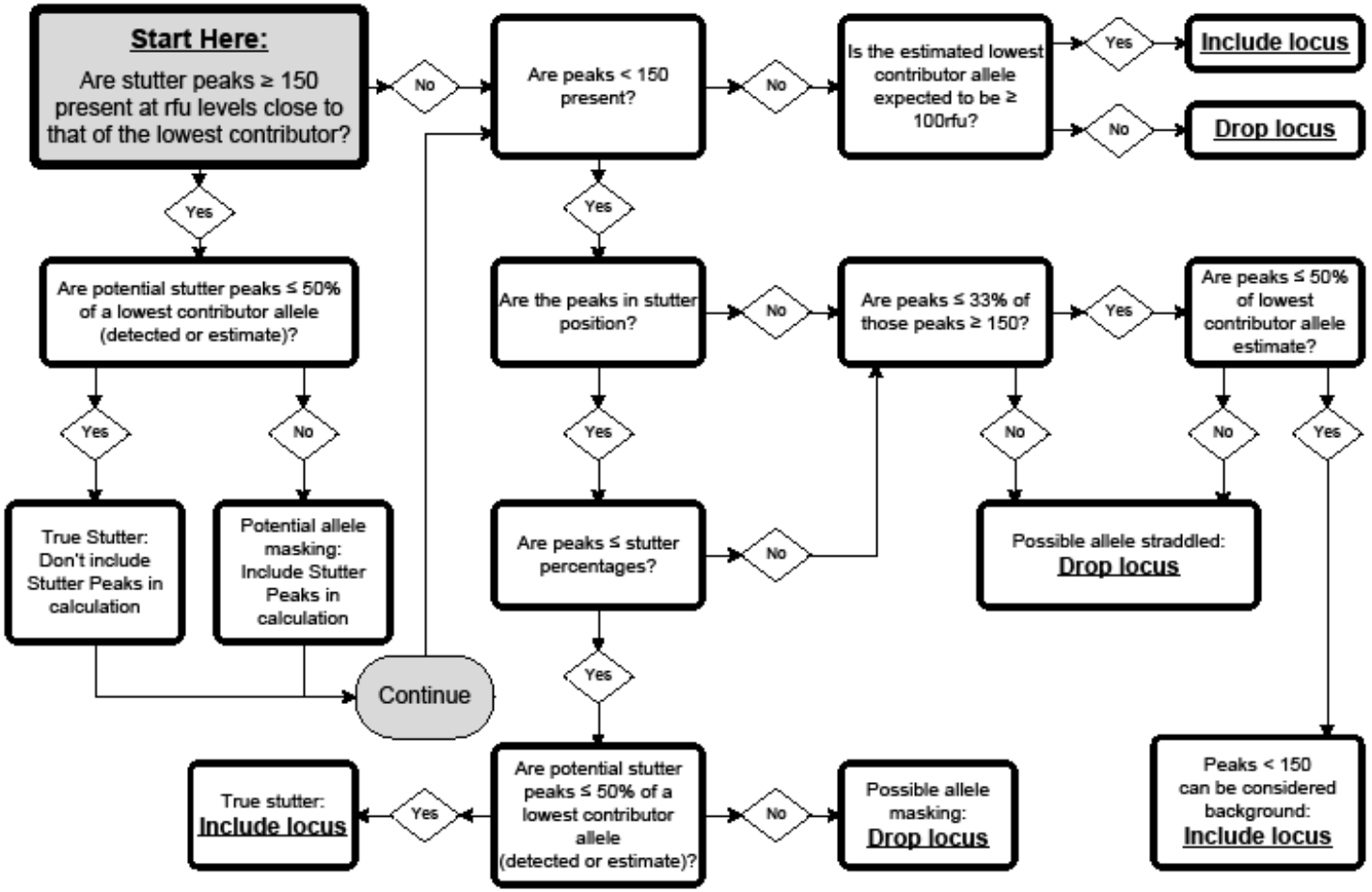
- Two approaches:
  1. Weed out the usable loci from the not (i.e. check for drop out and stutter masking) and use regular unrestricted stats approaches.

     or

  2. Use some sort of statistical compensations (i.e. assume the number of contributors and compensate for possible dropout – restricted LR or modified RMP)

- First an example of the weeding and unrestricted stats approach

The analysis of stutter is not required if one can account for all alleles of the lowest contributor
(i.e. 4 peaks in a two person mix)



Is there dropout?

As a prep for Combined Probability of Inclusion

- Determine what can be used for an inclusionary profile.
- Do this before comparison to Knowns
- The inclusionary profile contains the loci and alleles that must be present (mandatory) before a known sample is called included.
- Then a comparison can be done to the known.
    - If the mandatory alleles and loci include those found in the Known, then the known is included as a possible contributor.
    - If the known is missing mandatory alleles from the inclusionary profile, the Known is excluded.
    - If there is a scientific explanation for the missing alleles (straddle/dropout) then the comparison is inconclusive (not inclusionary).

Factoring for dropout

As a prep for Combined Probability of Inclusion

- Inclusionary profile determination is especially important for CPI since future comparisons depend on it.
- Method must be able to exclude loci where alleles from one component/contributor may be missing
- We use the concept of "do you expect to see the contributor's allele" to accomplish this.
    - Use a calculation of contributor proportion to determine if we expect to see alleles.

Factoring for dropout

As a prep for Combined Probability of Inclusion

- RFUs of all alleles are added at a locus to determine the total RFU contribution of all sources
- Loci are picked out where peak heights suggest that the lowest allele is present above the interpretation threshold.
- Lowest allele % is calculated and averaged for all loci where this is possible.
- This % is then applied to loci where it is not obvious that all alleles must be present and an 'expected lowest allele height' is calculated.

Factoring for dropout

As a prep for Combined Probability of Inclusion

- This expected lowest peak height calculation is used to decide if the alleles from a lesser contributor will be observed at the locus or not.
  - For example, a calculation of 60 RFU suggests that one would not necessarily expect to see the allele.
  - A calculation of 300 suggests that the alleles from that source will be observed.
- If one does not 'expect' to see the alleles, this locus is dropped from the Inclusion Profile.
- Having applied this to all loci, one is left with the Inclusion Profile and the LR or CPI is calculated based on this set of loci.

Stutter and Stochastic Considerations

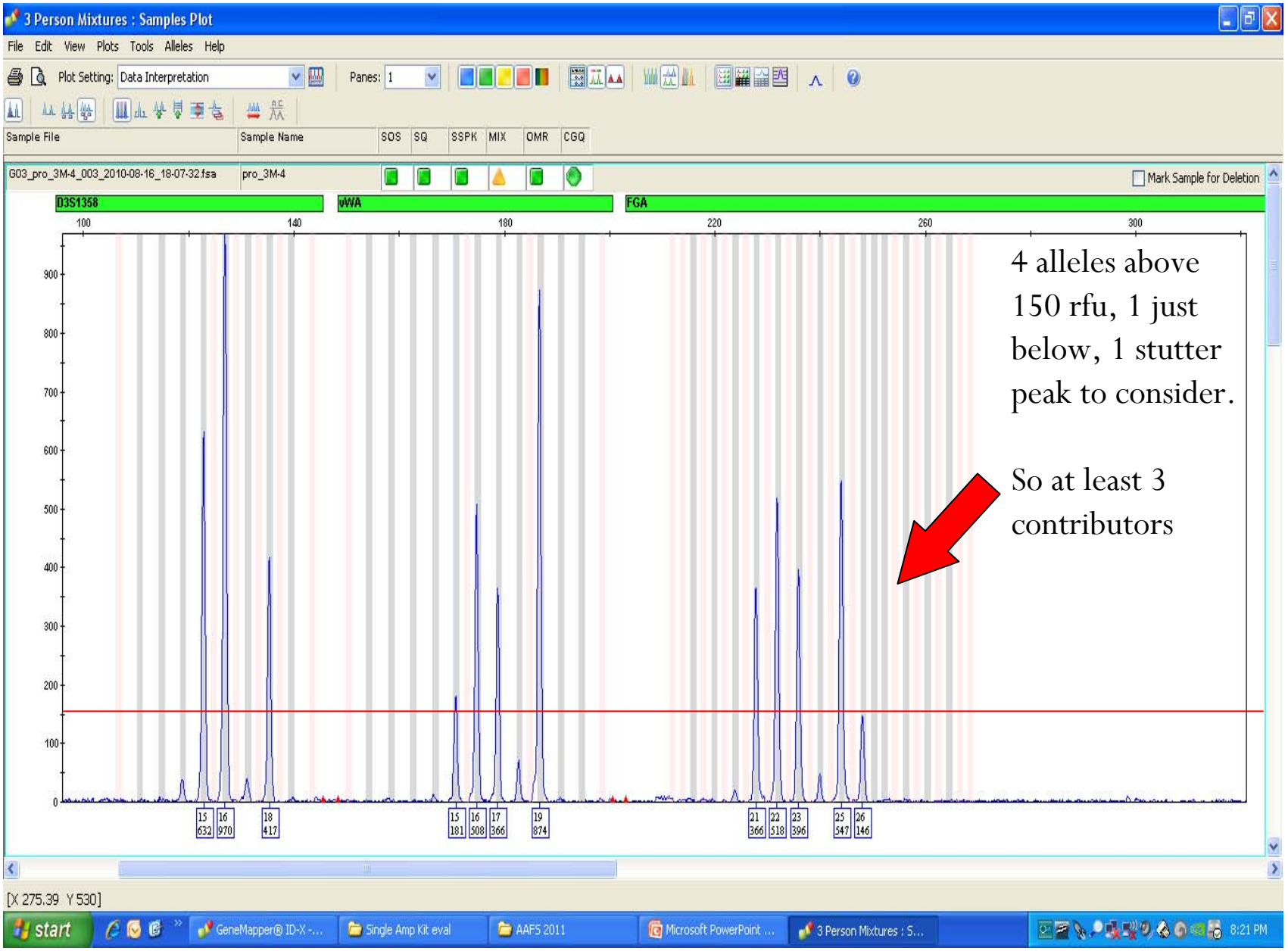
As a prep for Combined Probability of Inclusion

- Stutter is an important consideration since true alleles may be buried or masked by it.
- Stutter above interpretation threshold and below must be considered.
  - Stutter above interpretation threshold must be included in the alleles for calculation of LR or CPI unless their peak heights are such that they are not from a possible source.
  - Stutter below interpretation threshold will require that the locus is removed from the Inclusion Profile unless their peak heights are such that they are not from a possible source.

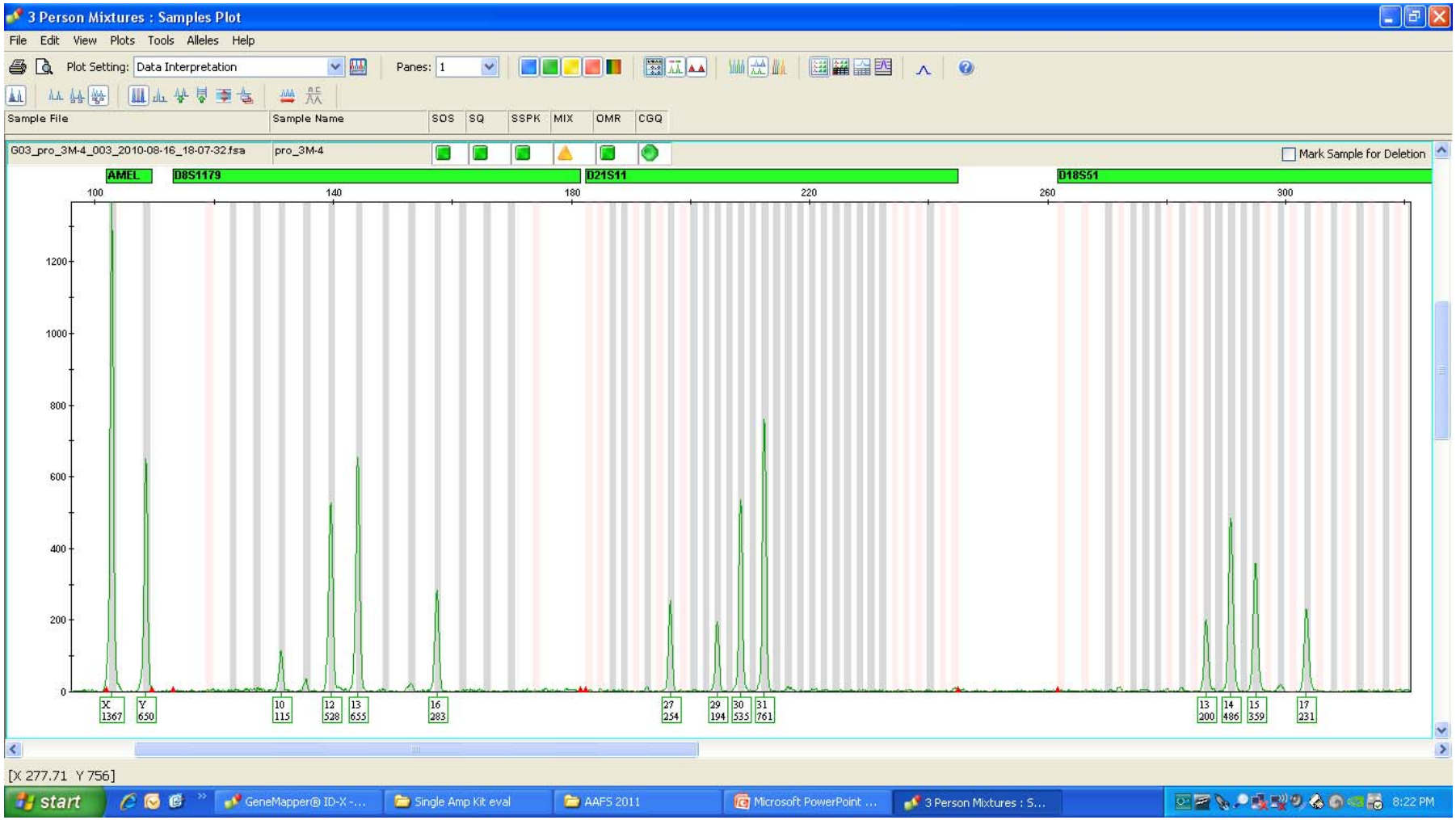
Putting it all together

- 33% calculations (as in Major/Minor calculations) are used to determine if allele peaks below threshold can be discounted.
  - Determined empirically from validation work
- 50% calculations are used in determining if stutter and or Estimated lowest peak heights can be discounted.
  - 50% seems to work.
- Putting it all together, WSP uses a flow chart to help visualize the decision process used to determine if a locus will be included in the Inclusion Profile (and CPI calculation) or not.

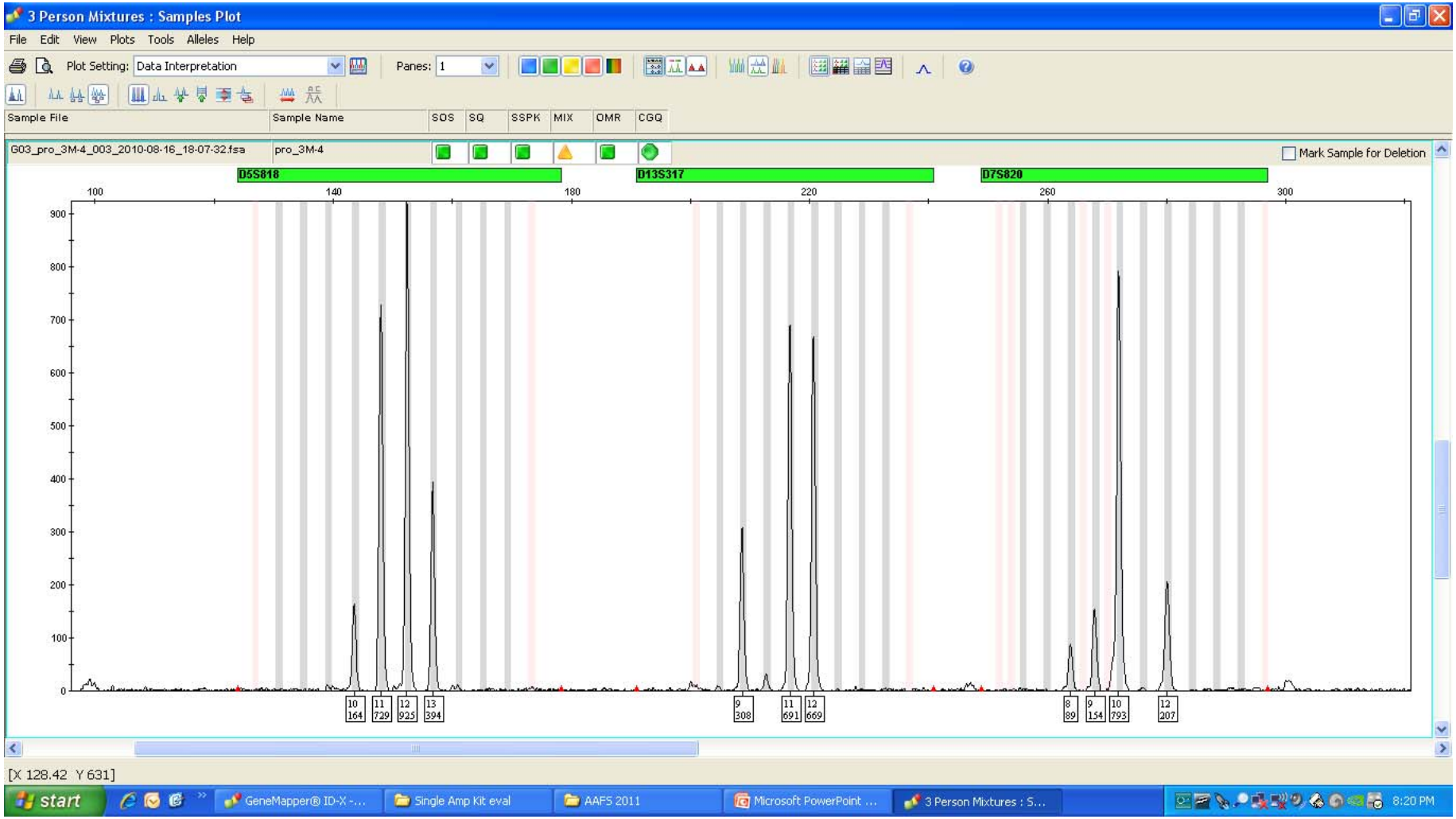
Example - 3 person mixture with no major profile provided



Example - 3 person mixture with no major profile provided



Example - 3 person mixture with no major profile provided



- Drop loci that may have dropout by working out the lowest contributor proportion
- Factor in stutter
- 7034_Mixture_Analysis 3 person mix 4.xlsx
- With K's
- 7034_Mixture_Analysis 3 person mix 4 w Ks.xlsx
  - Correctly predicted the loci with dropout

Then do a regular CPI calculation with the suitable loci

2. Some sort of statistical approach?

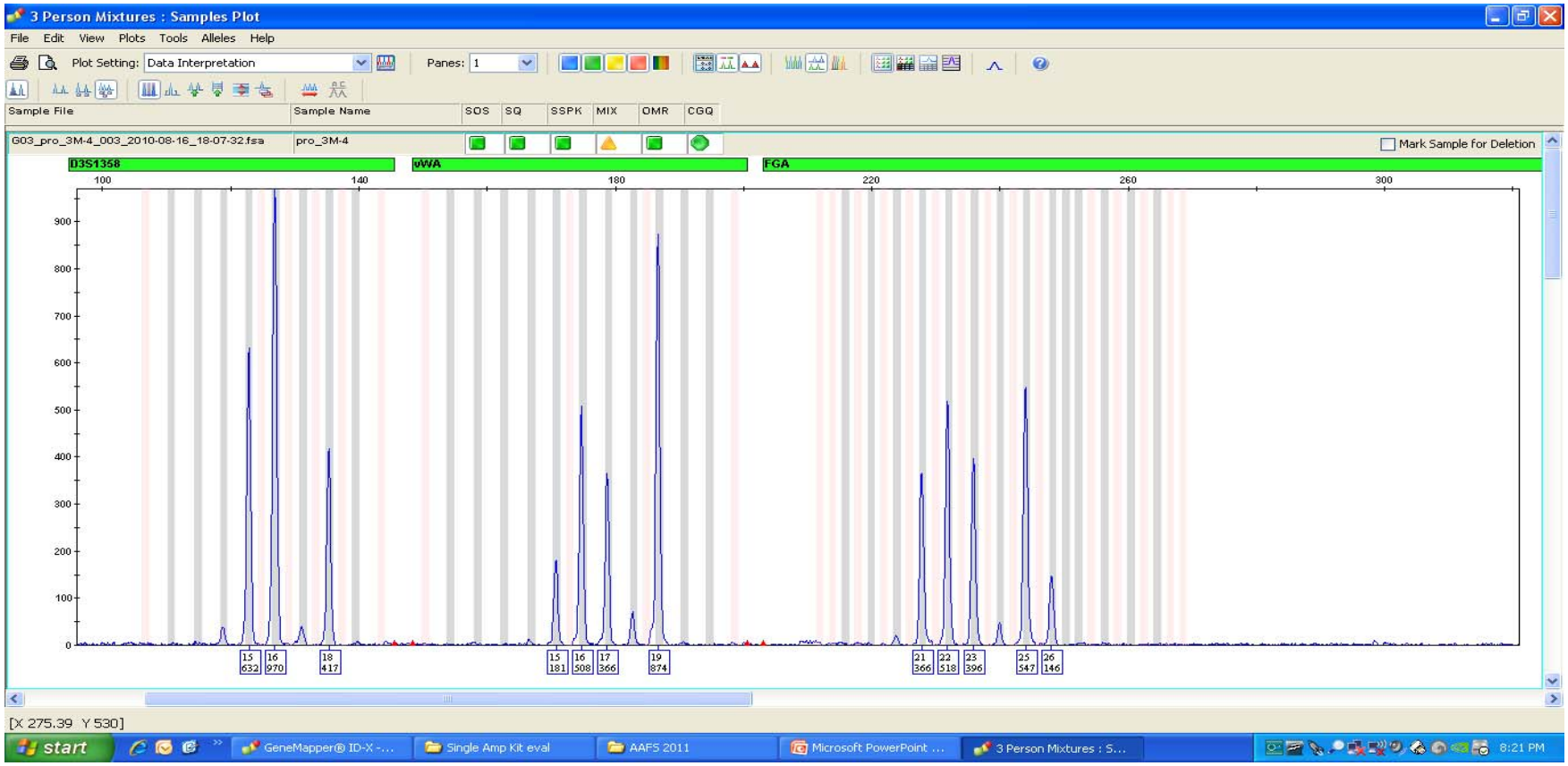
- The SWGDAM interpretation guidelines allow for a variety of unrestricted and restricted approaches, for example:
  - Compensate statistically for dropout and use assumptions
    - Could develop restricted approach LRs with assumed number of contributors and dropout probability as per Gill et al FSI 2006
    - Or alternatively use the modified RMP approach (i.e. Cal DoJ – Steve Myers) with an assumed # of contributors and using 2p for dropout.

What the heck is a modified RMP?

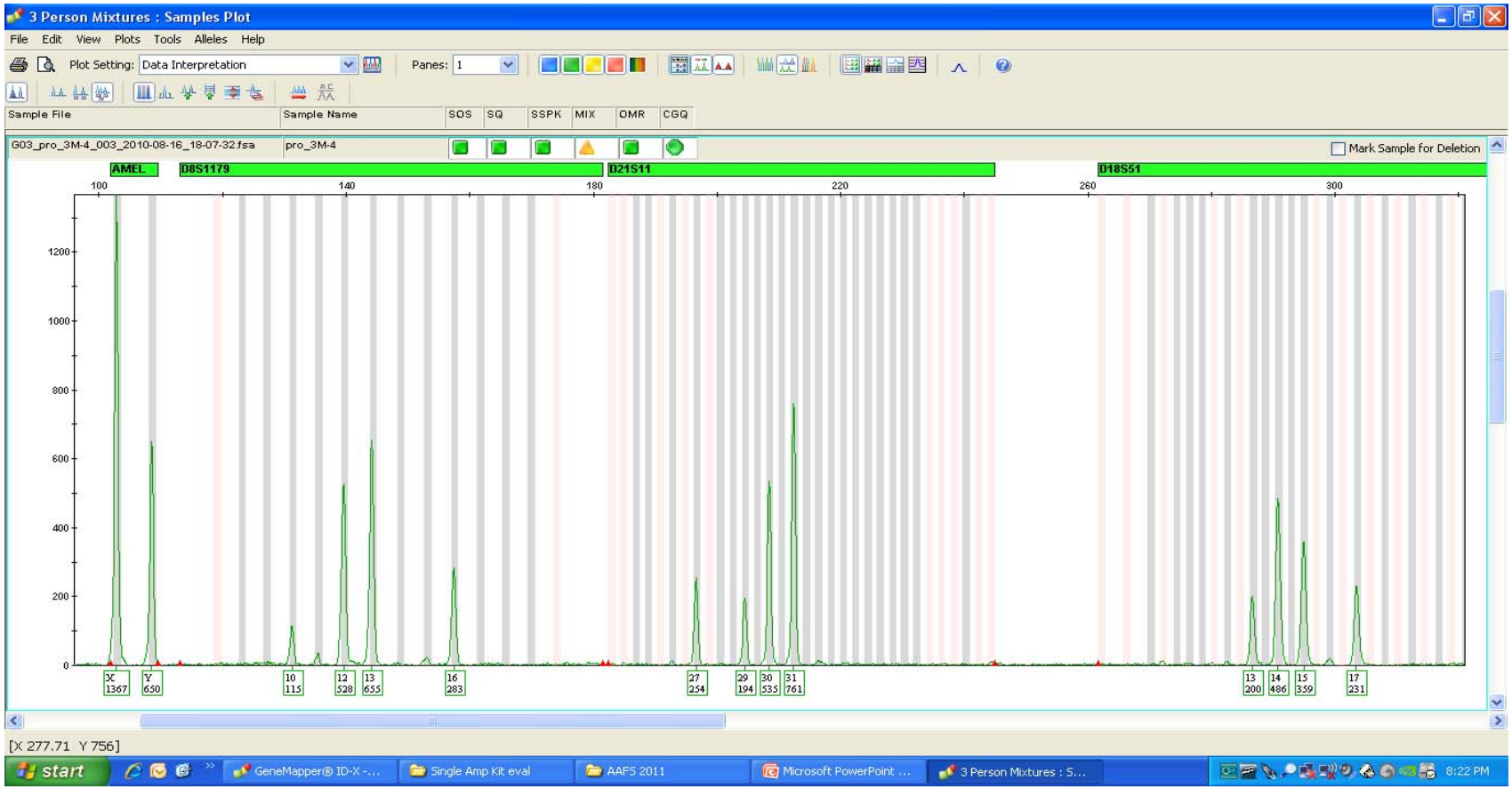
- Normally it's all 2pq and p^2, however in a mixture the RMP can be based on an assumed number of contributors which allows the summing of allele frequencies when there is more than 1 possible genotype (see section 5.2 in SWGDAM Interpretation Guidelines)

Example of modified RMP approach

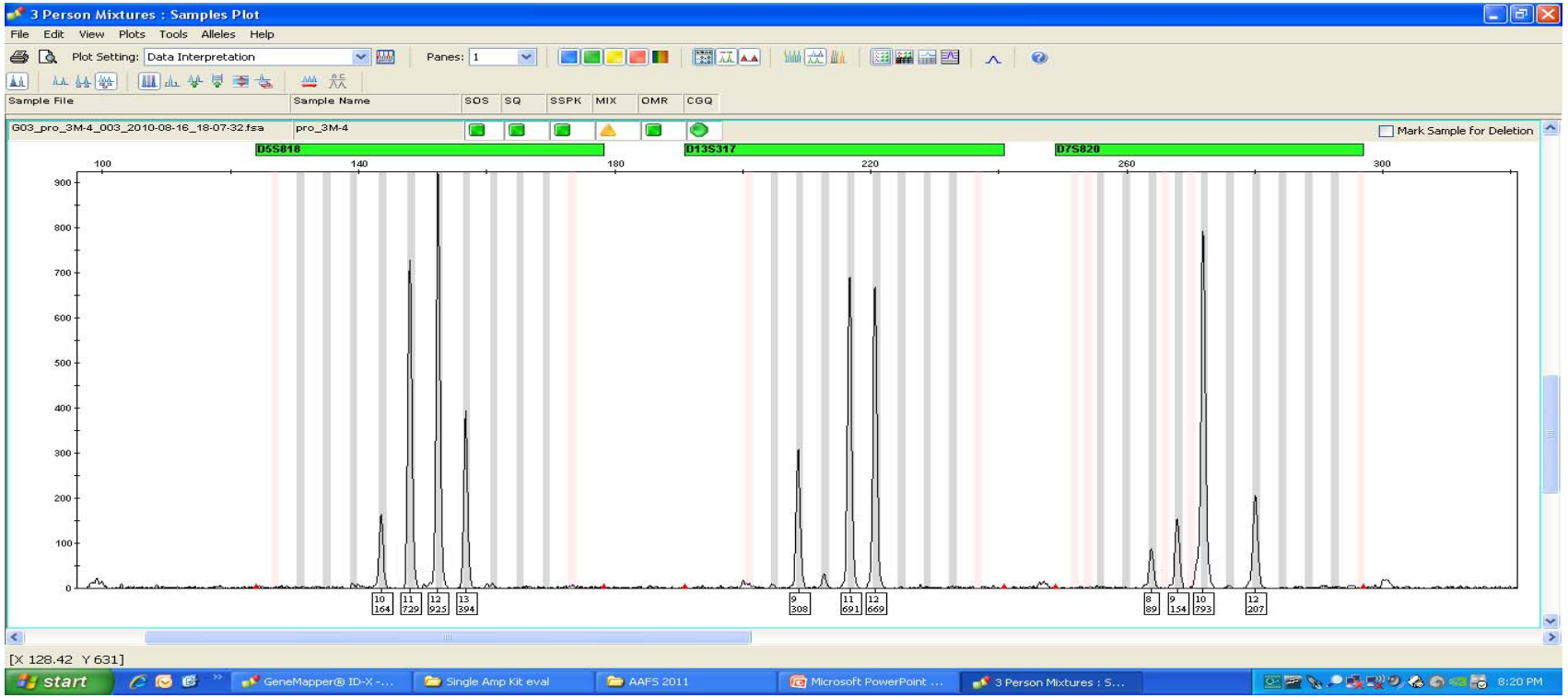
- For 3 person mixture:
  - Use loci with 5 and 6 alleles,
  - Also loci with 4 alleles provided PHI indicative of genetic info from more than 2 contributors at locus
  - Idea is to capture enough genetic info to use 2p to compensate for dropout.



D3	**vWA**	**FGA**
Only 3 alleles so not used	**16,19 PHI – 58% ok to use**	**5 alleles – ok to use (below stochastic threshold is still ok)**



D8	D21	D18
4 alleles	4 alleles (>70%)	4 alleles (>70%)
10,16 PHI = 41%	insufficient PHI	insufficient PHI
Ok to use	Not used	Not used



D5

4 alleles
10,13 PHI = 42%
Ok to use

D13

Only 3 alleles
- Not used

D7

4 alleles detected
10,12 PHI = 26%
Ok to use

- The RMP formulae that can be adapted for this approach are described in section 5.2.2 of the guidelines
- Determine all the genotypes that are included
  - "Assuming 3 contributors to this mixture, the chance that a randomly selected person would be included as a possible contributor is…"

Any Questions?

AAFS DNA Mixture Workshop : Complex mixture strategies talk