SDPD Forensic Science Section – Forensic Biology Unit

Modification to the Genemapper ID-X v1.4 Software

Comparison to GMID-X v1.1.1

Introduction

GeneMapper ID-X Software version 1.4 can process data from .fsa files generated on 3130 Genetic Analyzers as well as .hid data files generated on 3500 Genetic Analyzers. It is capable of processing six-dye fragment analysis, which is necessary for the Applied Biosystems GlobalFiler PCR Amplification kit.

A new feature includes a Y-marker check box in the Panel Manager. This checkbox is defaulted for DYS391 and the Yindel in the GlobalFiler kit. There is a new Amelogenin Cross Check PQV (ACC) that functions to cross reference Amelogenin and any Y-maker results and flag concordance. Refer to the Life Technologies User Bulletin (Part Number 4477684 Rev. A) for additional features such as disabling spike detection for samples and controls, modified allele number PQV, modified mixture analysis functionality, and modified duplicate homozygous allele label functionality.

Materials and Methods

To determine whether GeneMapper ID-X Software version 1.4 produces comparable results to GeneMapper ID-X Software version 1.1.1, which has been in use in the Forensic Biology Unit since 2009, a comparison of .fsa raw data was interpreted with the two versions. Raw data from 20 (prior case work) samples was imported into each version and an evaluation of allele call, peak height, base pair size, and data point size was performed.

Results

Raw data from 20 (prior case work) samples (previously analyzed with Identifiler Plus) was imported into each version of the software. The data was analyzed with the default Identifiler Plus Panels and Bins. An evaluation of allele call, peak height, base pair size, and data point size was performed. This information was equivalent between version 1.1.1 and 1.4. Results can be observed in attached table.

Conclusions

The algorithms and coding used in the two versions of the software were demonstrated to produce equivalent data. Given the equivalence of the software versions, GMID-X v1.4 should be considered as a valid software package for analyzing all previous data from the 3130s.

Analysts

Marybeth Sciarretta
Criminalist

Shawn Montpetit
DNA Technical Manager

10/5/15
Supervisor Review **Date**

10/7/15
Supervisor Review **Date**

10/7/15
Quality Assurance Review **Date**

GeneMapper ID v3.2 vs. GeneMapper ID-X Analysis

Introduction

Applied Biosystems has recently developed a new expert system analysis software called GeneMapper ID-X, which replaces and improves on all the functions of GeneMapper v3.2, and includes new features. Our goal is to implement the use of GeneMapper ID-X in casework. The purpose of this validation was to compare data previously analyzed with GeneMapper v3.2 with the same data analyzed using GeneMapper ID-X to determine whether or not the final DNA testing results were equivalent, and to assess additional features of GeneMapper ID-X (ladder, control, and sizing quality flags, sample/lab reference comparison tool, spike labeling, allele table export, and Yfiler stutter filtering) for accuracy. GeneMapper ID-X also incorporates a new mixture analysis tool to assist the analyst with interpreting mixed profiles. The mixture software can calculate the most likely possibilities for major and minor profiles, extract known profiles, and calculate statistics. The main objective of implementing the mixture tool is to assist the analyst with deciphering mixtures, particularly for determining major/minor profiles for entry into CODIS.

Materials and Method

Genotypes and Sizing

Data from three Identifiler casework run folders was analyzed using GeneMapper ID-X with parameters similar to those used with GeneMapper v3.2. The projects were randomly selected; one project used the 310 Genetic Analyzer while the other two projects used the 3130 Genetic Analyzer. Results for up to five alleles per DNA locus were tabulated and exported to Excel spreadsheets. The GeneMapper ID-X analyzed data was compared to the previous GeneMapper v3.2 analyzed data with respect to genotyping, base pair sizing, and peak heights.

Quality Flags

Additional sample data was analyzed in GeneMapper ID-X to assess the accuracy of the quality flags (ladder, positive, negative, and sizing quality). For the ladder quality assessment, a random full profile sample was incorrectly designated as an allelic ladder prior to performing analysis. An actual failed ladder was also analyzed. For the positive control quality assessment, a random full profile sample was incorrectly designated as a positive control prior to performing analysis. Additionally, a true positive control sample (amped with 1uL instead of 10uL) with one peak missing was analyzed. For the negative control quality assessment, a random full profile sample was incorrectly designated as a negative control prior to performing analysis. Additionally, a low profile sample with only one peak called was analyzed. For the sizing quality assessment, four failed or low quality size standard samples were analyzed.

Lab Reference/Sample Comparison Tool

The Lab Reference (Staff Profile) Comparison tool was examined by analyzing two separate projects known to contain staff contamination. The staff member's profile was added to the

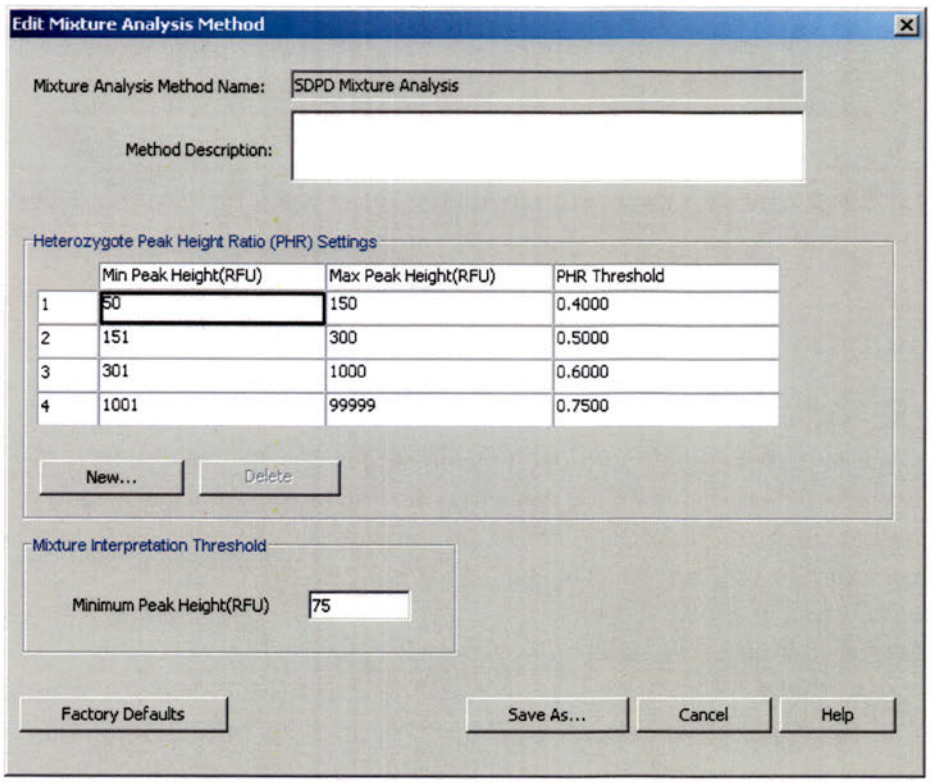
profile manager so it would be available for comparison. The Sample Comparison Tool was examined by analyzing two projects previously shown to have matching profiles.

Spike Labeling Feature

The automatic spike labeling feature was assessed by analyzing three samples known to contain spikes.

Mixture Analysis Tool and Stats

The mixture analysis tool was evaluated based on the features that the lab plans to potentially utilize. The following capabilities were examined for two person mixtures: ability of the software to accurately separate major and minor contributors; the extraction of a known contributor from a mixture; and statistical calculations. In order to assess these features, mixture training samples with known contributors and proportions were analyzed. The statistics generated by GeneMapper ID-X were compared to those from PopStats. Three or more person mixtures were not reviewed because the software does not interpret these, it only calculates statistics. A new mixture analysis method, named "SDPD Mixture Analysis", was used for our analysis. The method has the following parameters:



Allele Table Export

GeneMapper ID-X provides the option to export the allele table in the traditional horizontal format typically used. To determine if the software correctly exports the proper allele calls, two projects were randomly selected for export and the data was manually reviewed.

Y-STR Stutter

To assess the Y-STR stutter filtering, a Yfiler project was analyzed in GeneMapper ID-X and the following stutter positions were examined: DYS19 n-2 bp and DYS392 n+3 bp.

Results

Genotypes

The three Identifiler casework projects [redacted] contained data from approximately 94 samples, allelic ladders, and controls.

In all samples, the major allele assignments were identical whether analyzed using GeneMapper ID-X or GeneMapper v3.2. Any "no match" conclusion was manually reviewed and confirmed to be identical (see green notes in comparison spread sheets - tab 4). The reasons for any false "no match" calls were attributed to one of the following factors:

1) Artifact peaks (minus A, pullup, etc.) had been manually reviewed and labeled by an analyst for the completed GeneMapper v3.2 projects, while the GeneMapper ID-X projects were left as OL, etc. and not labeled by an analyst. Had the GeneMapper ID-X peaks been reviewed and manually labeled, the data would have resulted in matches. The peaks were confirmed to be identical and labeled as "OK" (in green font) in the comparison spreadsheets.
2) A spike was labeled OL by GeneMapper v3.2 software and this peak was included in the genotype table. GeneMapper ID-X has a new spike labeling feature which detected this spike and automatically labeled it "spike". These software labeled spikes are not included in the genotype table. The OL peak from GeneMapper v3.2 and the software labeled spike from GeneMapper ID-X were confirmed to be the identical and labeled as "OK" (in green font) in the comparison spreadsheets.

Allele Base Pair Size and Peak Heights

All of the allele base pair sizing and peak heights for the three projects were identical, except for the one instance of a spike described in number two of the above Genotypes Results. GeneMapper v3.2 included the spike in the genotype table while GeneMapper ID-X excluded it, which resulted in a discrepancy in the comparison spreadsheet. The peaks were confirmed to be identical and labeled as "OK" (in green font) in the comparison spreadsheets.

Allelic Ladder Quality Flag

To assess the accuracy of the allelic ladder quality flag, a random sample [redacted] was designated as an allelic ladder before GeneMapper ID-X analysis was preformed. This resulted in the sample being flagged red, indicating that it did not satify the requirements for an allelic ladder. The quality value details for the sample (see Figures 1 and 2) explain why the sample was flagged. All loci have to be in concordance (green) for the sample to be flagged green overall.

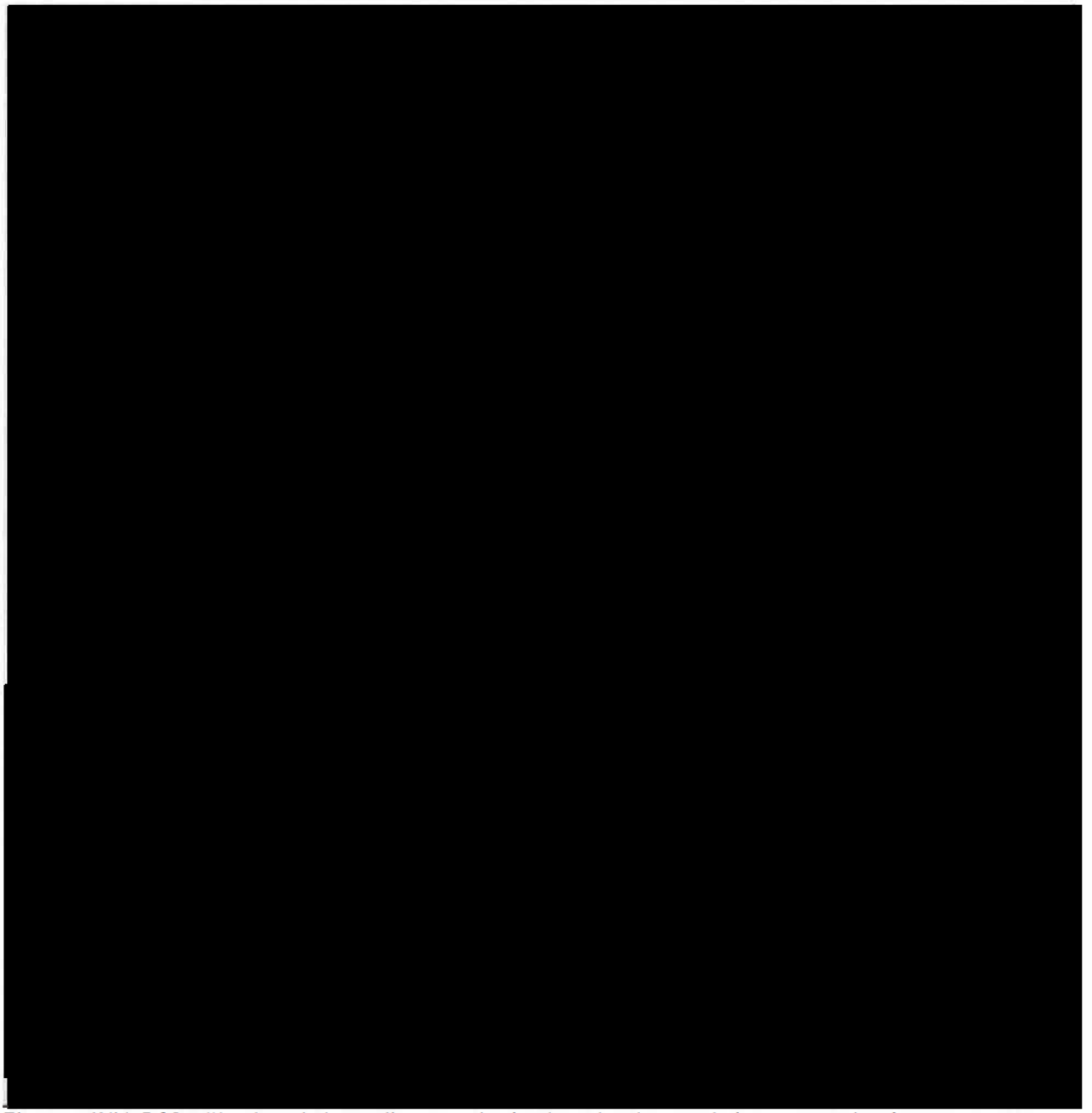


Figure 1: With D8S1179 selected, the quality value details show that the sample is non-conordant because one or more peaks were not detected. Notice that the locus is marked red.

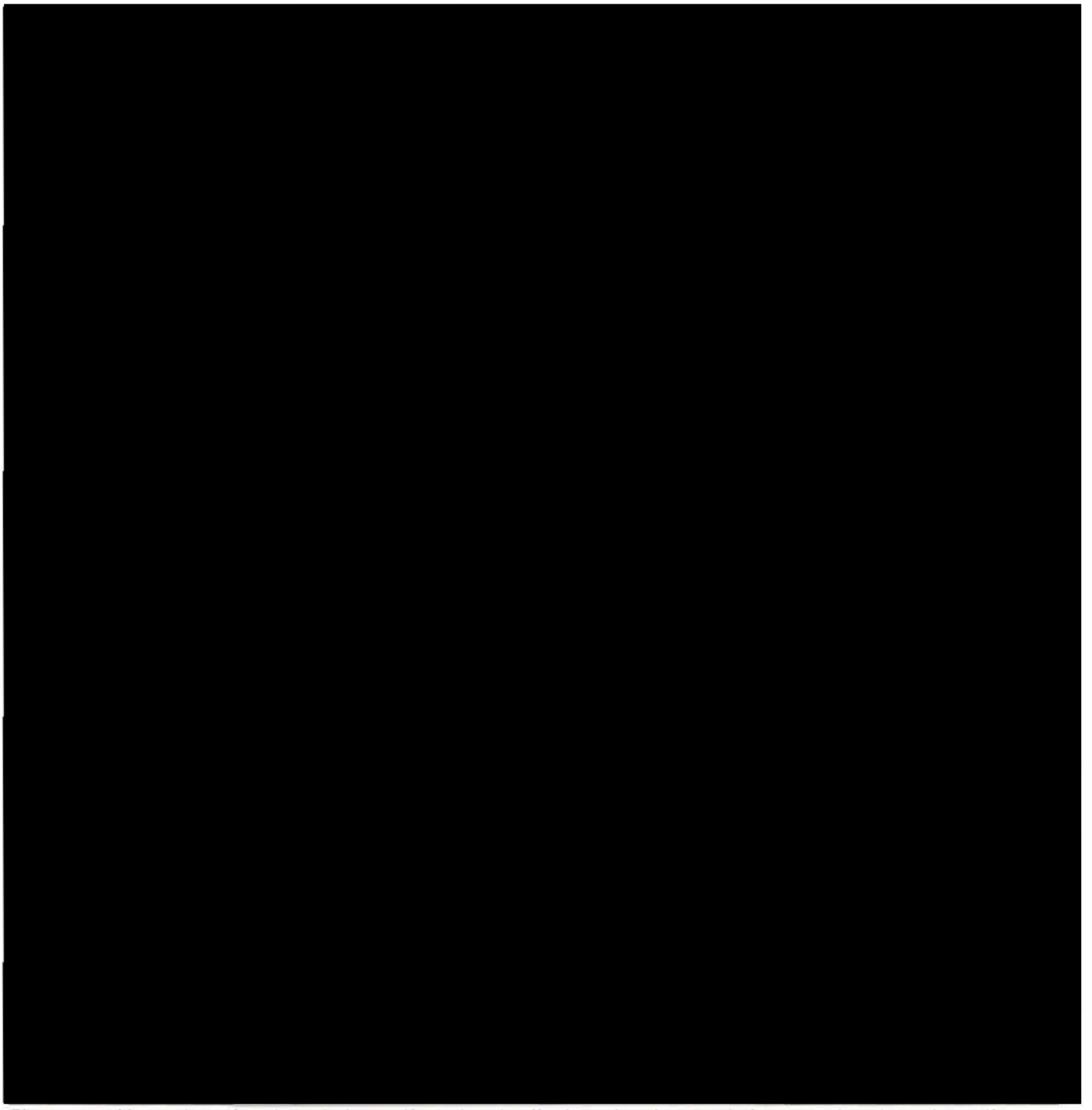


Figure 2: With amelogenin selected, the quality value details show that the sample is concordant (X and Y calls match that of a true ladder). Notice that the locus is marked green. However, the sample was still flagged red overall because all of the loci have to be in concordance to be flagged green.

In addition, an actual failed ladder was analyzed (sample *LAD0804100* courtesy of Trinity DNA Solutions in Milton, FL). The ladder was above thershold, however the capillary was going bad and resulted in poor peak resolution. The TH01 locus was flagged red due to poor resolution between the 9.3 and 10 allele (see Figures 3 and 4). As a result of the ladder not having the expected allele calls (due to 9.3 and 10 being indistinguishable), the software flagged the CGQ and this ladder could not be used.

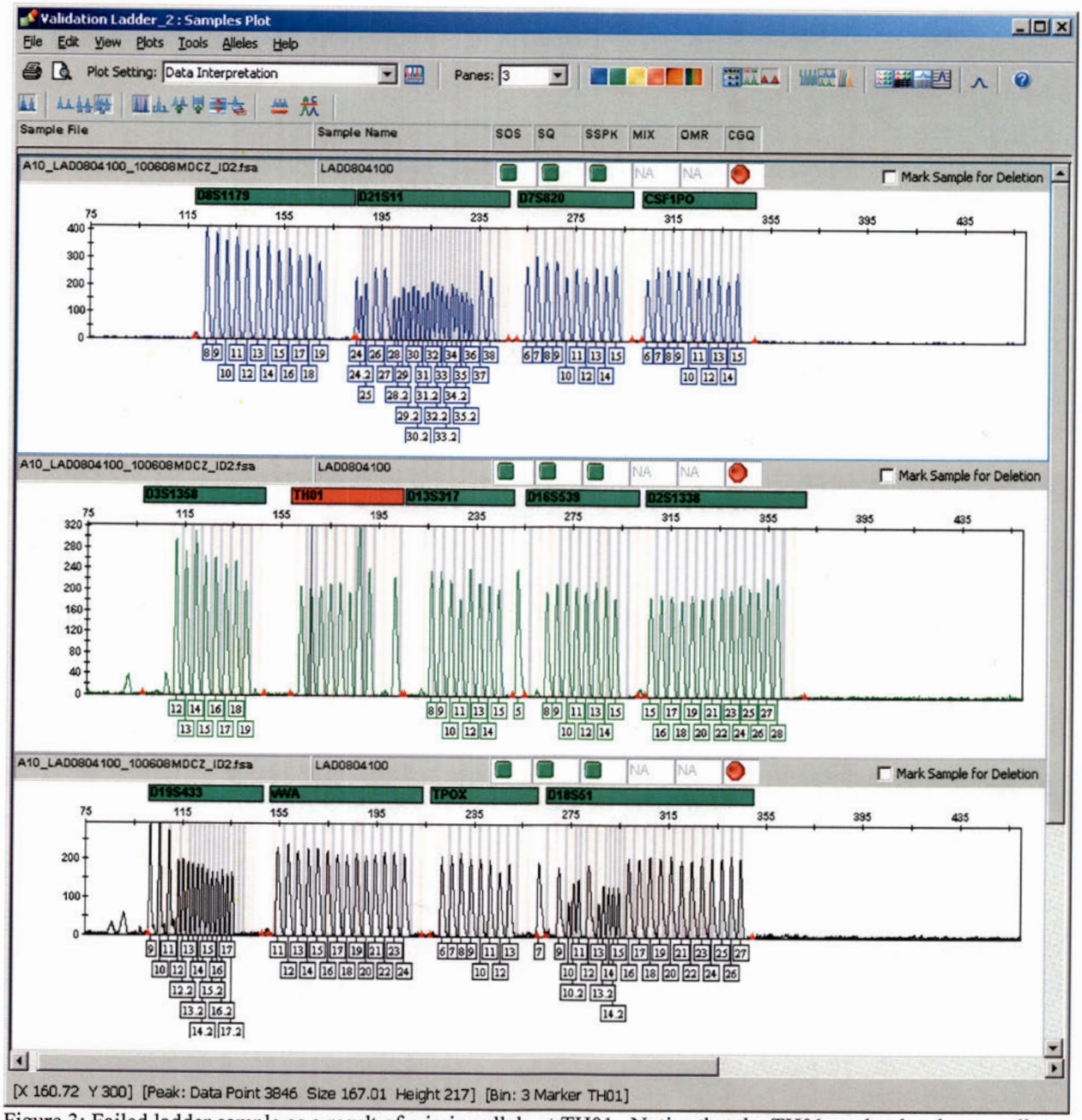


Figure 3: Failed ladder sample as a result of missing allele at TH01. Notice that the TH01 marker header as well as the sample CGQ are flagged red.

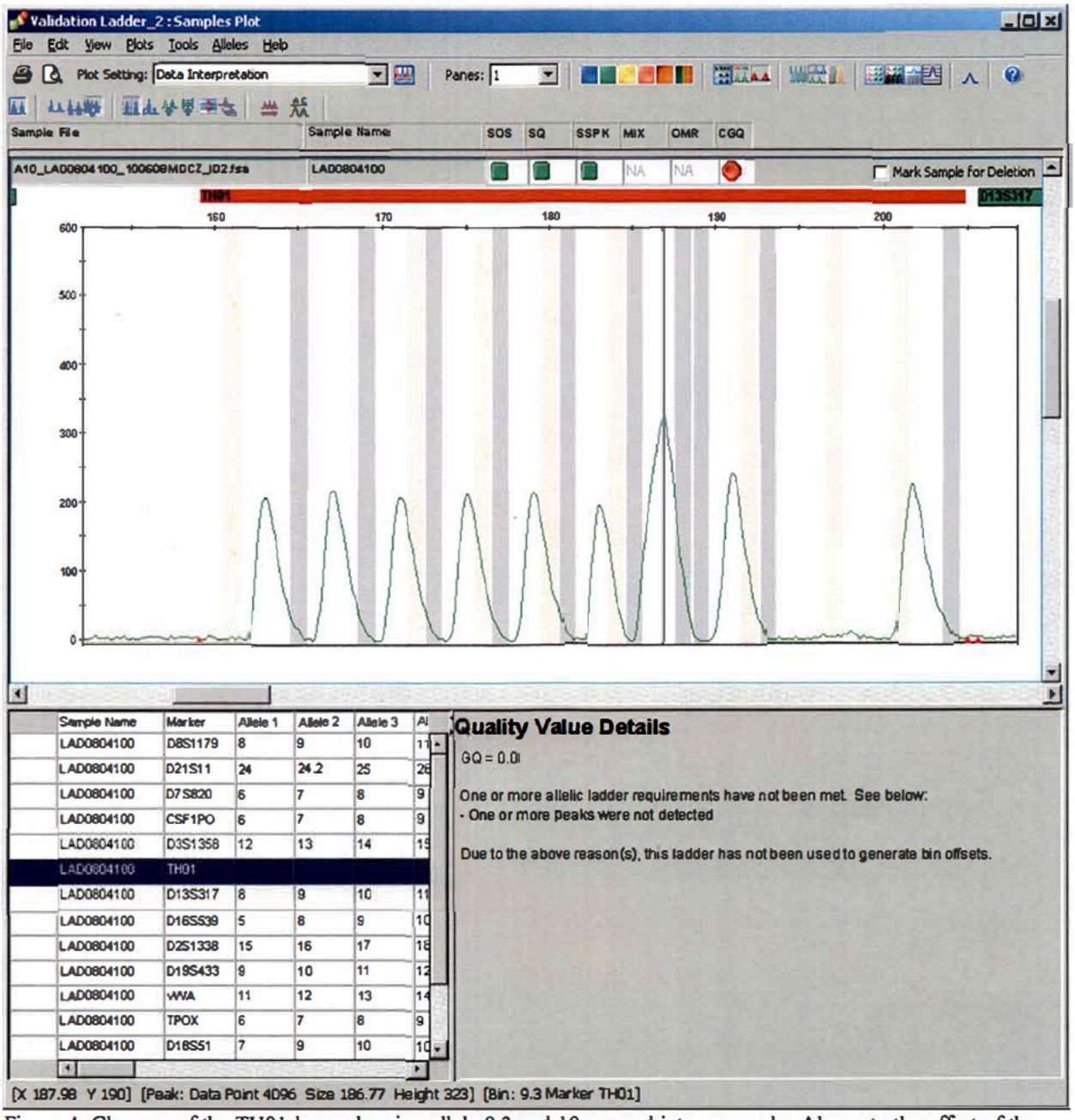


Figure 4: Close up of the TH01 locus showing allele 9.3 and 10 merged into one peak. Also note the offset of the peaks from the expected bin positions.

Positive Control Quality Flag

To assess the accuracy of the positive control quality flag, a random sample [redacted] was designated as a positive control before GeneMapper ID-X analysis was performed. This resulted in the sample being flagged red, indicating that it did not satify the requirements for a positive control. The quality value details for the sample (see Figure 5) explain why the sample was flagged. All loci have to be in concordance (green) for the sample to be flagged green overall.

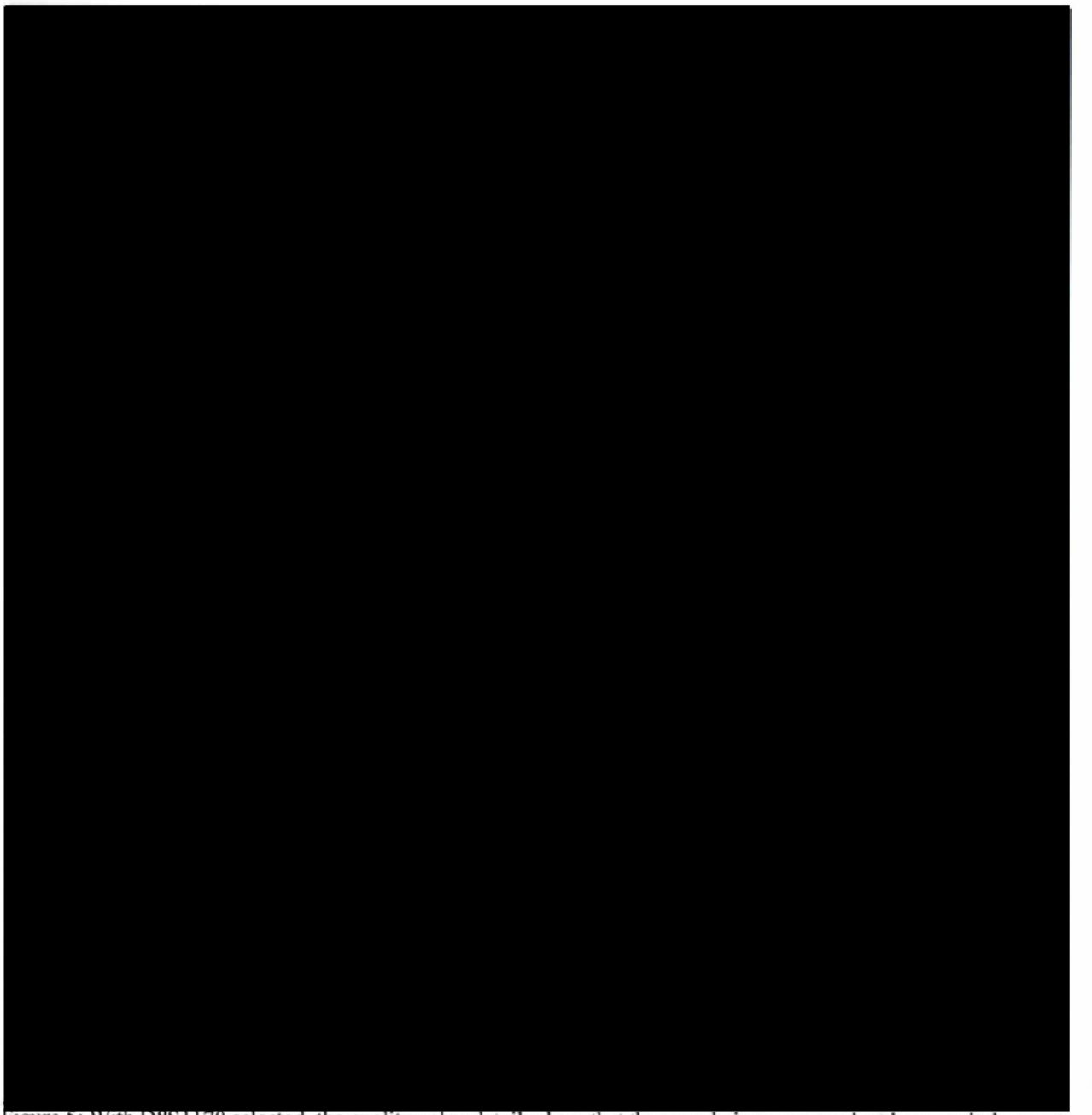


Figure 5: With D8S1179 selected, the quality value details show that the sample is non-conordant because it does not match the values of a true positive control. Notice that the locus is marked red.

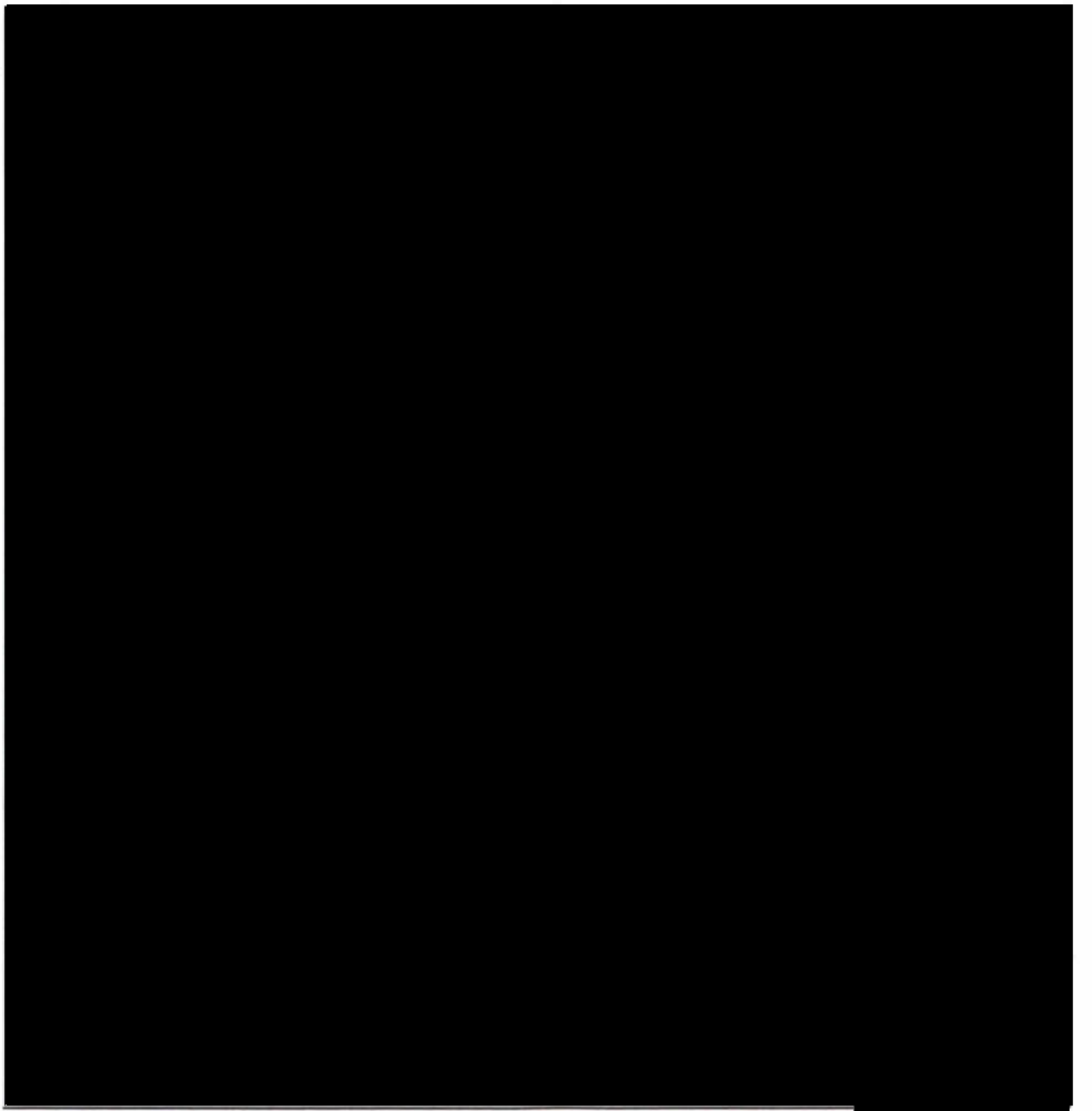


Figure 6: With D7S820 selected, the quality value details show that the sample is concordant [redacted] coincidently match that of a true positive). Notice that the locus is marked green. However, the sample was still flagged red overall because all of the loci have to be in concordance to be flagged green.

In addition, an actual positive control sample was analyzed [redacted] that had one peak drop out at the TH01 locus. The missing peak was recognized by the software, and the locus, as well as the overall sample, were flagged red (see Figure 7).

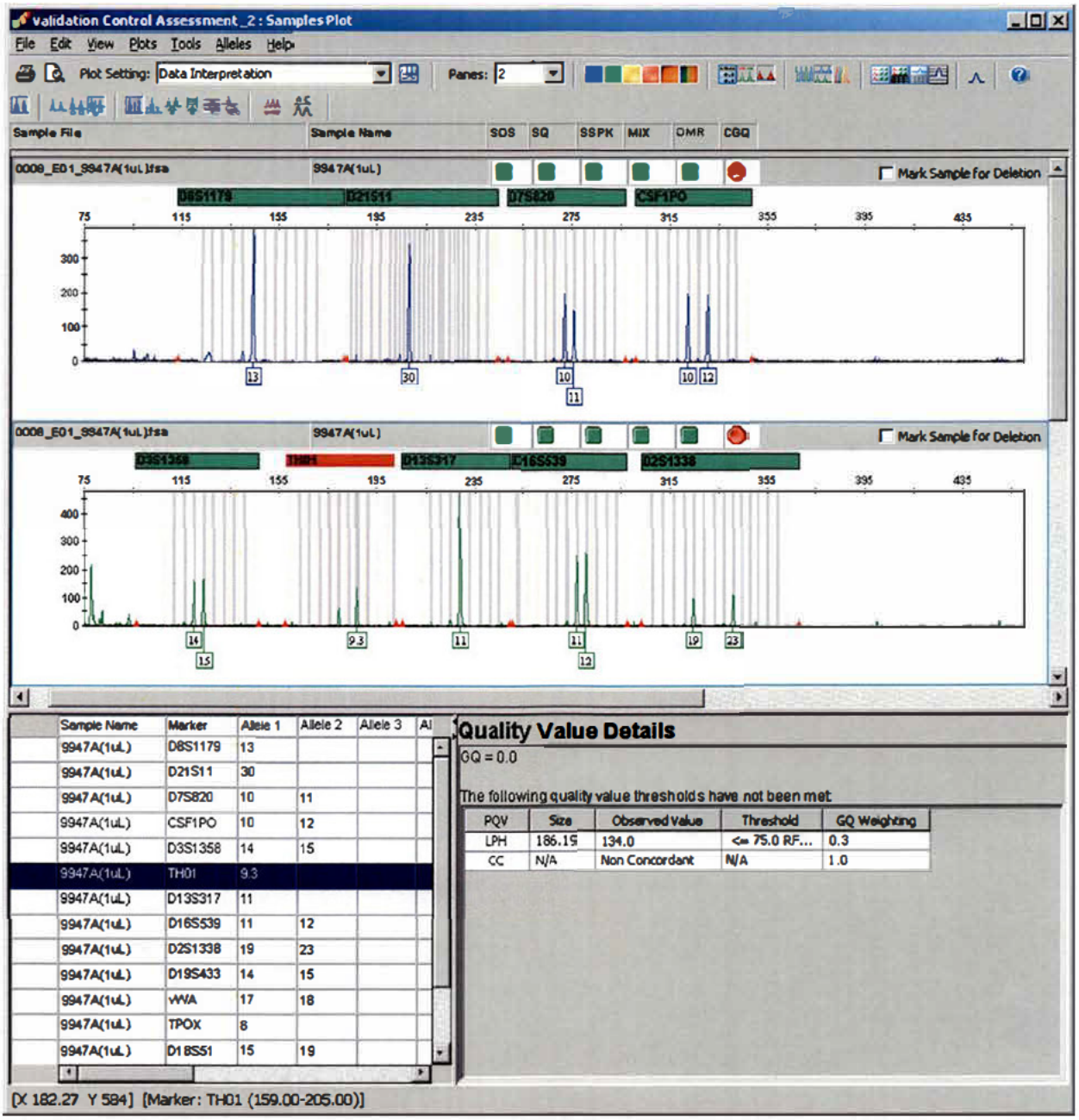


Figure 7: With TH01 selected, the quality value details show that the sample is non-conordant (missing one allele call), and there is a peak below threshold. Notice that the locus is marked red. As a result of the one peak non-concordance, the entire sample is flagged red.

Negative Control Quality Flag

To assess the accuracy of the negative control quality flag, a random sample ([redacted] *from projec*[redacted] was designated as a negative control before GeneMapper ID-X analysis was preformed. This resulted in the sample being flagged red, indicating that it did not satisfy the requirements for a negative control. The quality value details for the sample (see Figure 8) explain why the sample was flagged. All loci have to be in concordance (green) for the sample to be flagged green overall.

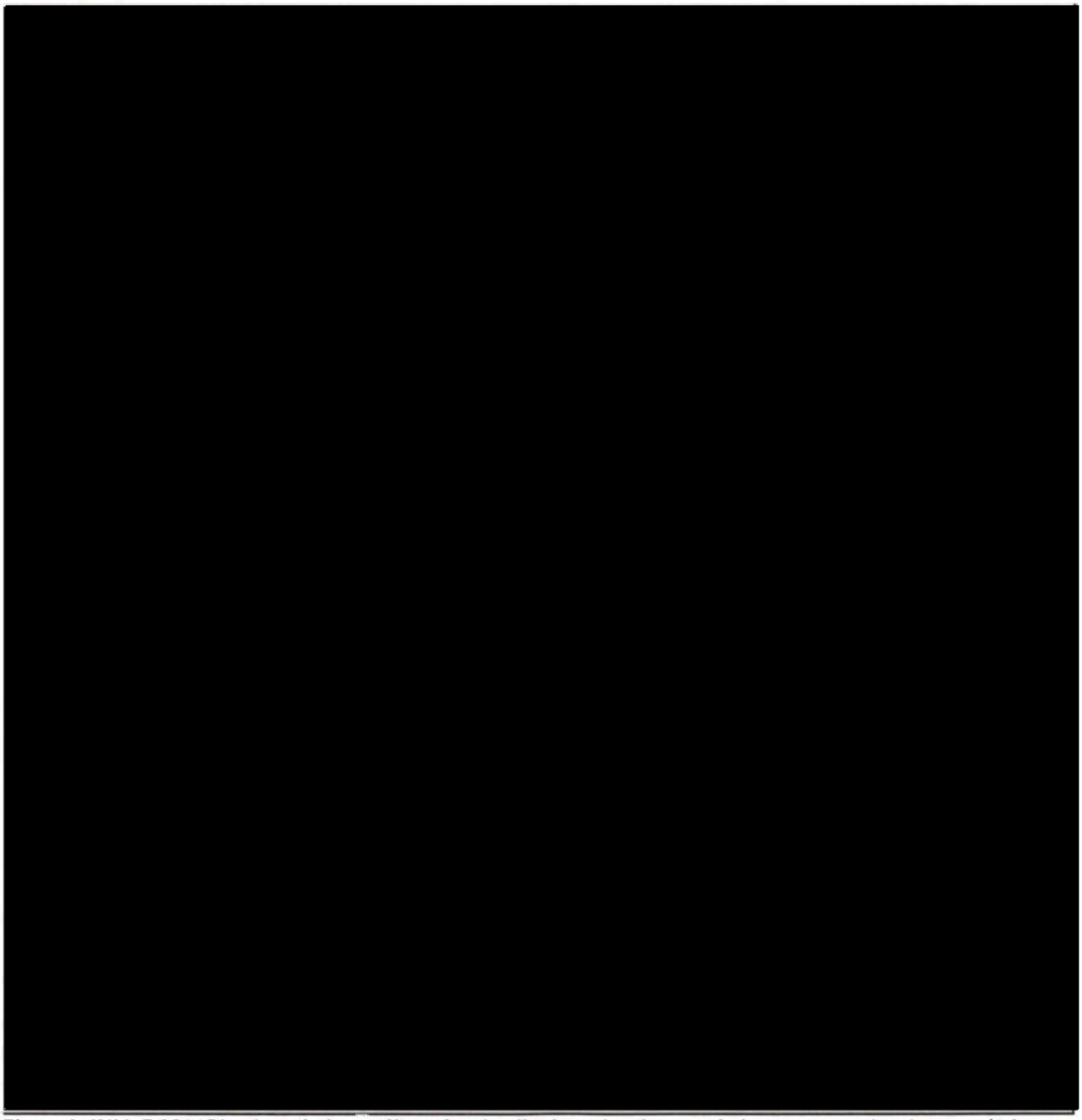


Figure 8: With D8S1179 selected, the quality value details show that the sample is non-concordant because it does not match the values of a true negative control (blank). Notice that the locus is marked red.

In addition, a low profile sample [redacted] with only one peak called was analyzed. The peak was detected by the software and the loucs, as well as the overall sample, was flagged red (see Figure 9). The software will not flag a negative control for peaks emerging below threshold.

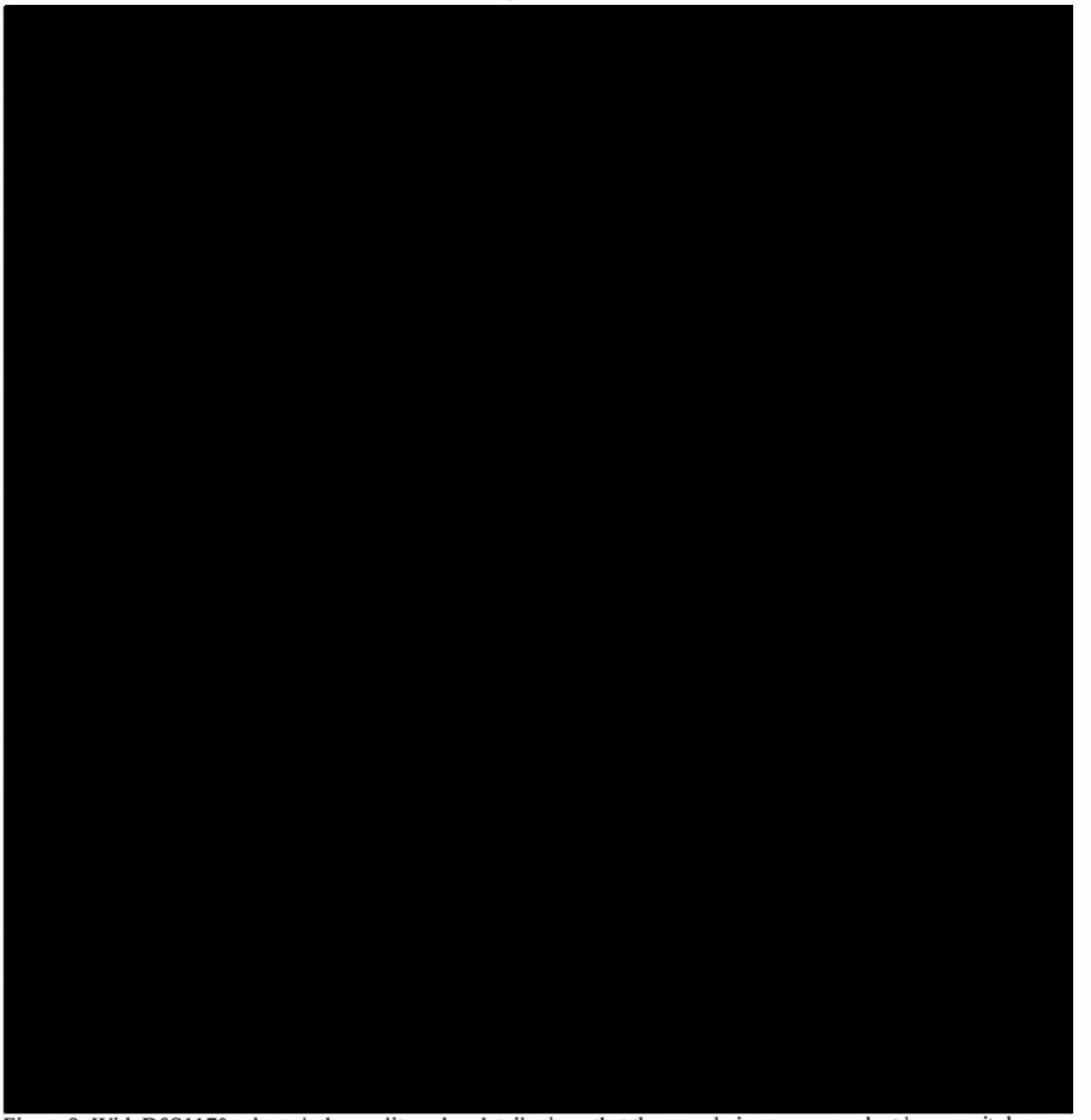


Figure 9: With D8S1179 selected, the quality value details show that the sample is non-concordant because it does not match the values of a true negative control (blank). Notice that the locus is marked red. The software does not flag loci with peaks below threshold.

Sizing Standard Quality Flag

To assess the accuracy of the sizing quality flag, a project [redacted] with several failed sizing standards was analyzed with GeneMapper ID-X. This resulted in the sample SQ being flagged red or yellow, indicating that there is a problem with the size standard peaks. The default SQ values are 1.0 – 0.75 for green, 0.74 – 0.26 for yellow, and 0.25 – 0 for red. A new feature affecting sizing quality in GeneMapper ID-X is the broad peak SQ weighting. If the user defined broad peak threshold (default 1.5 bp) is exceeded, the SQ value will be affected by the user-defined broad peak weight (default 0.5). To have the software ignore resolution when

determing sizing quality, set the size standard Broad Peak value to 0. To use a low quality sizing standard, click "Override SQ" in the Size Match Editor. Figures 10 – 13 show examples of flagged sizing standards.

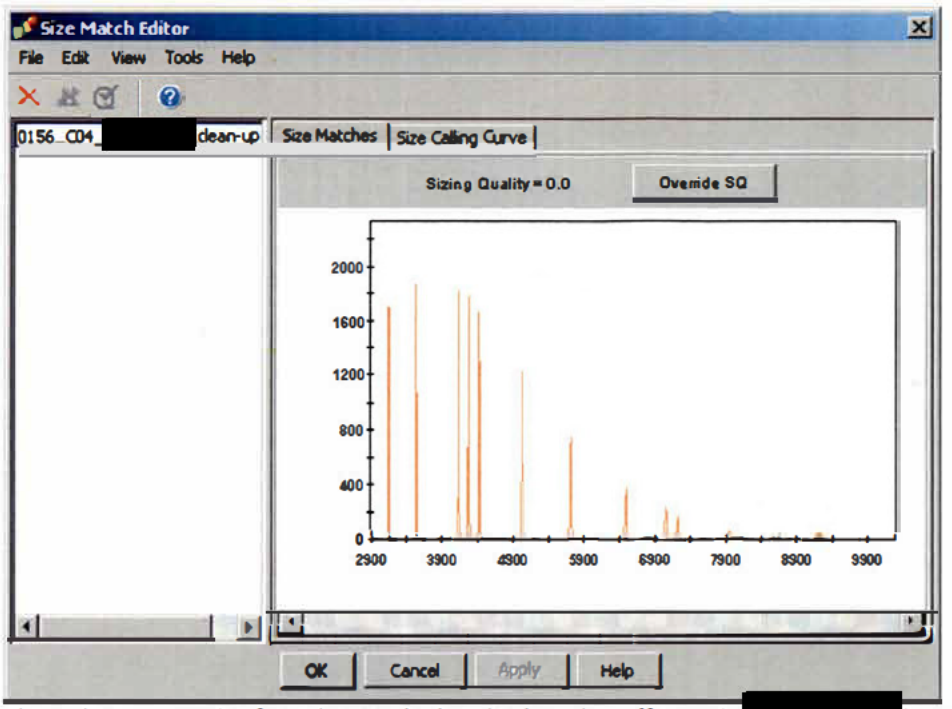


Figure 10: As a result of the size standard peaks dropping off, sample [redacted] clean-up had a red flag for SQ. No sizing data was generated for this sample.

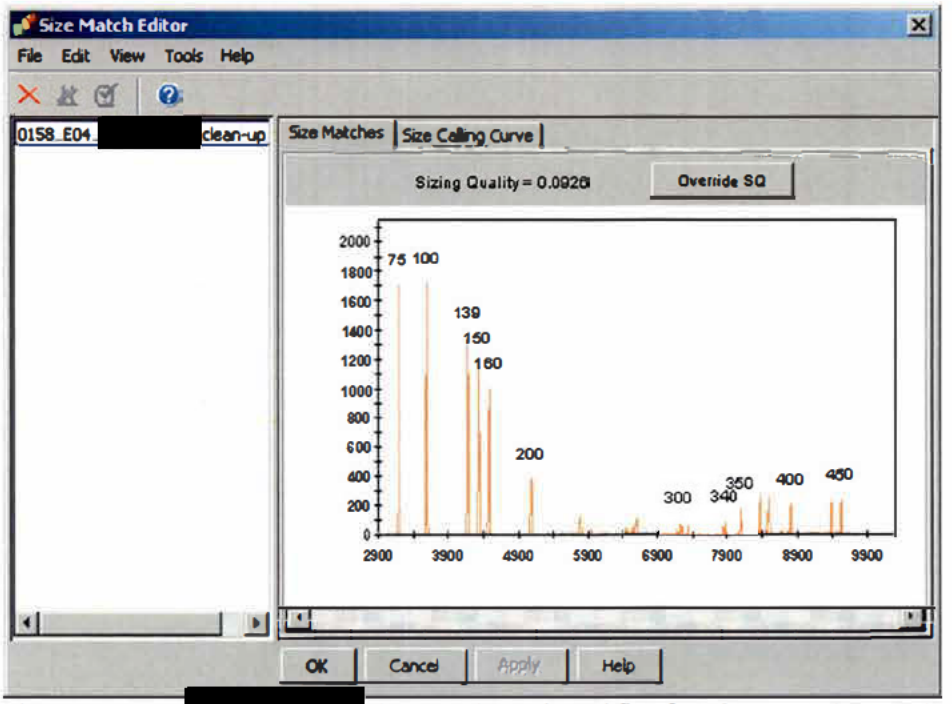


Figure 11: Sampl[redacted] clean-up also had a red flag for SQ as a result of the low sizing quality. No sizing data was generate e.

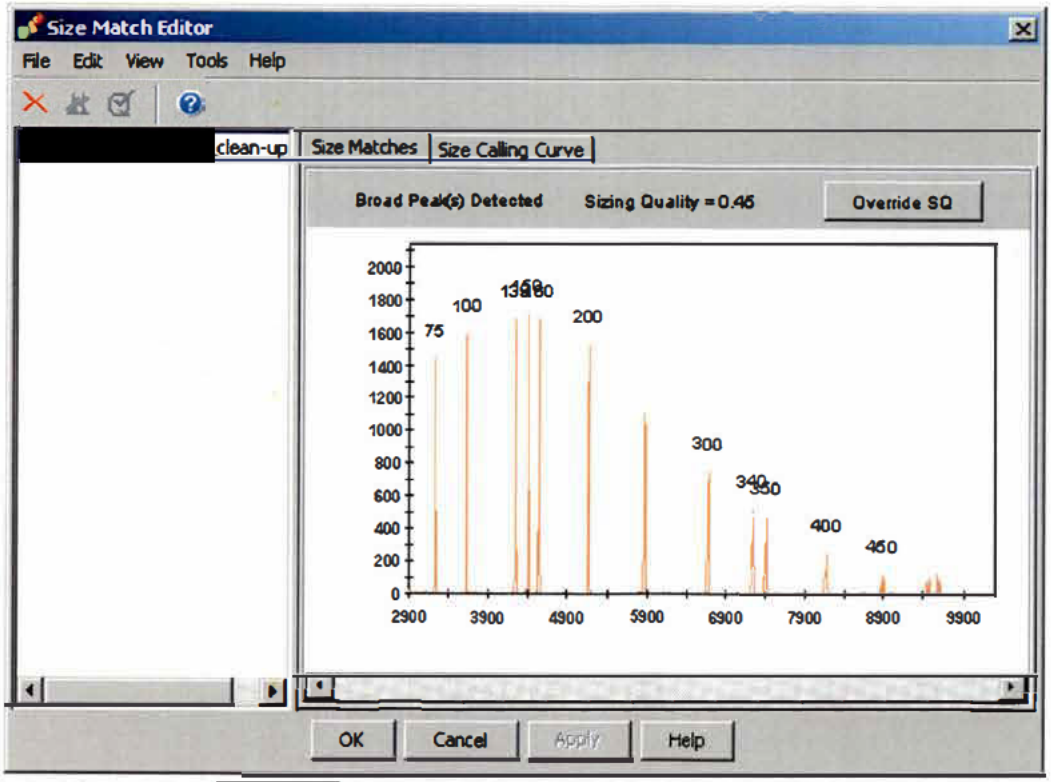


Figure 12: Sample [redacted]_clean-up had a yellow SQ flag. Sizing data is generated for samples flagged yellow.

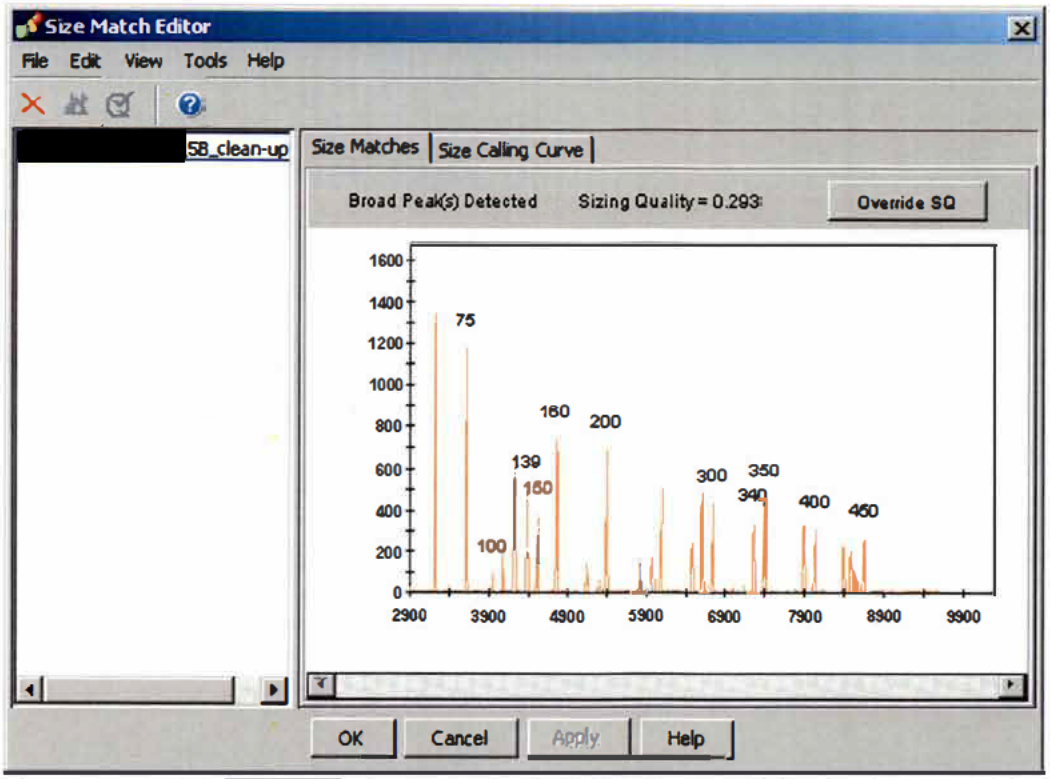


Figure 13: Sample [redacted]_clean-up also had a yellow SQ flag. Sizing data is generated for samples flagged yellow.

Lab Reference Comparison Tool

Two separate projects ([redacted]8) [redacted]ach contained samples containing DNA [redacted]ember [redacted]. A[redacted]er analyzing the projects with GeneMapper ID-X, the Lab Reference Comparison Tool was used (80% match threshold) to see if [redacted]as detected in the samples. [redacted]as shown to be a contributor to three samples from the two projects, which is consistent with the lab's past finding of contamination (when the same threshold is considered). Figures 14 and 15 show a portion of the profile comparison windows.

[redacted]

entire profile was detected within the sample.

[redacted]

Figure 15: Staff member sample [redacted]is a 100% match to case work sample [redacted]and a 92.6% match (25 of 27 alleles were present) to case work s[redacted].

Sample Comparison Tool

Two separate projects [redacted]ach contained samples with matching [redacted]rofiles. After analyzing the projects with GeneMapper ID-X, the Sample Comparison Tool was used (80% match threshold) to see which samples were

labeled as matching. A variety of matches were detected by the comparison tool. The data was manually reviewed and the matches were determined to be correct. A window similar to the "Lab Reference Comparison" seen above (Figures 14 and 15) is used to show sample comparison.

Spike Labeling Feature

Three samples with spikes [redacted] were analyzed with GeneMapper ID-X to determine if the spike labeling feature would call the spikes appropriately. The peak raw data was reviewed to confirm whether or not the peaks were actual spikes. The spike in sample pos alp_ was correctly labeled by the software (see Figure 16).

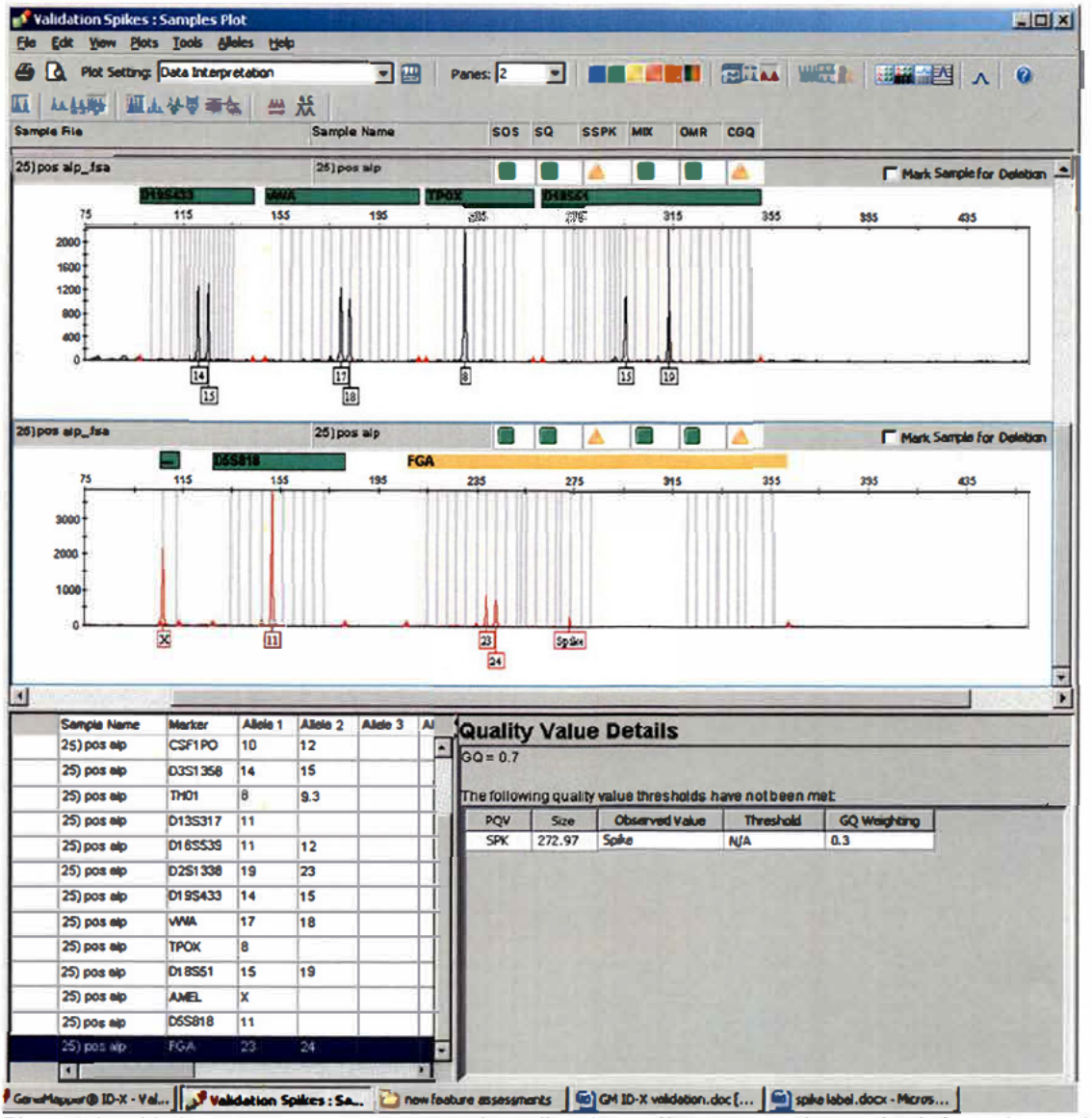


Figure 16: With the FGA locus selected (contains spike) the quality value details provide information on the spike. Notice that the locus, SSPK, and CGQ are flagged yellow due to the spike.

The spike in sample RB-E2 was correctly recognized by the software (see Figure 17). The SSPK was correctly flagged yellow; however, the D2S1338 locus and CGQ were still green. These should have been flagged yellow as well to alert the analyst to the presence of the spike and provide quality value details about the spike. ABI was contacted and they determined that this is a bug in the software and it will be corrected in a future upgrade. Refer to the May 15, 2009 email from April Orbison of ABI for more information (tab 8).

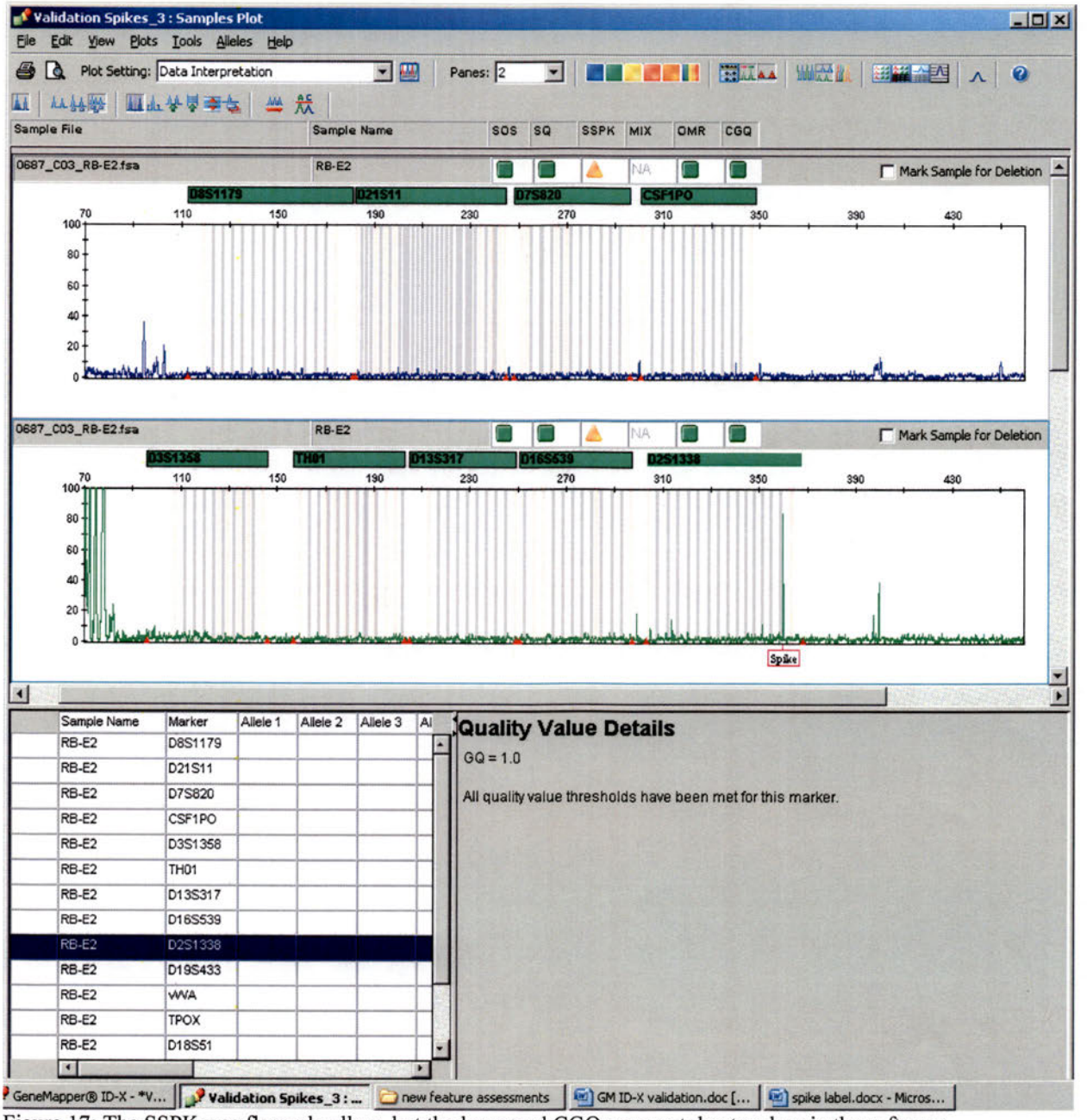


Figure 17: The SSPK was flagged yellow, but the locus and CGQ were not due to a bug in the software.

The spike in sample [redacted] was not labeled by the software (see Figure 18). The software has a specific spike calling algorithm that requires the peak to be within at least three dye colors and all must have a minimum peak half width of 0.5 bp or less. After reviewing the spike in sample [redacted] it was determined that the spike was not called because it was too broad (minimum peak half width greater than 0.5 bp).

peaks are too broad.

Mixture Analysis Tool and Stats

-Extraction of known contributor:
Three mixture samples ([redacted] ith known contributors ([redacted]) were [redacted] o determine if the s[redacted]ftware was accurately extracting the known contributor and assigning the correct profile to the remaining contributor.

[redacted] e M[redacted] w[redacted] MS[redacted] [redacted]xtracted. The [redacted] xtracted the [redacted] rofile of sample MS[redacted] and correctly determined the remaining contributor profile of sample [redacted] In addition, when the known

contributor was switched to sampl[redacted] M[redacted] the software correctly extracted it and determined the remaining contributor profile of sampl[redacted] MS[redacted] no figure shown).

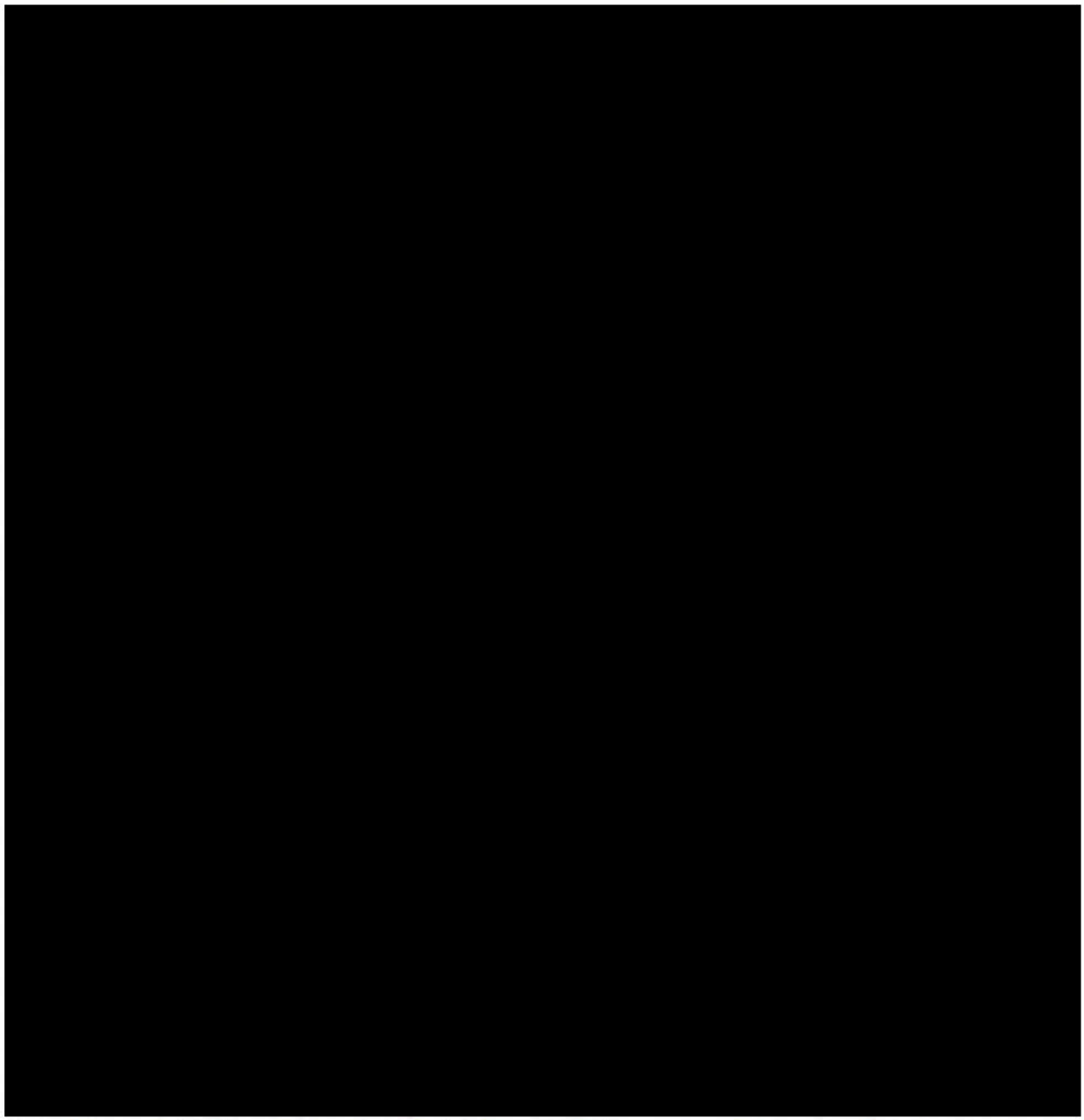


Figure 19: Sample [redacted] is a mixture of samples [redacted] (approximately 1:1). See printouts of samples [redacted] in validation binder.

Figure 20: The selected genotype combinations for sample [redacted] when known contributor [redacted] is extracted. The profiles are highlighted in bright yellow and green, signifying that all markers are present and can be attributed to the known profile. Note from figure 19 that the mixture is about a 1:1 ratio. When this is the case, the software assigns the known profile to the major contributor.

Figures 21-26 show an example of mixture sample [redacted] with known contributor [redacted] extracted. The software correctly assigned the known profile to the minor contributor at 13 of the 16 loci. For one locus (TPOX), the software includes the possibility of allelic dropout for the minor contributor, so there are two possibilities listed (see Figure 22, rows 10 and 11). The first profile listed (the most probable) is the correct genotype combination for the contributors. The other two loci (D21S11 and TH01) are missing from the selected genotype combinations table due to PHR status. For D21S11, the PHR for the minor contributor is about 59%. The threshold set in the SDPD mixture analysis method for that particular peak height range (300-1000) is 60%. For TH01, the PHR for the minor contributor is about 52%, with a threshold of 60%. As a result of not meeting the PHR threshold set in the analysis method, these loci are flagged red for PHR status and included in the unselected genotype combinations table with a yellow IQ (inclusion quality) (see Figure 24 and Figure 25). The analyst will have to manually review these loci to determine which profile possibilities should be moved to the selected genotype combinations table (highlight the profile and press the "Select" button). After manually reviewing the unselected genotype combinations table, the appropriate profiles were moved to the selected genotype combinations table (see Figure 26). The unknown major contributor profile is now complete and correct ([redacted]). When the extracted known contributor is switched

[redacted] a similar trend in the results is seen (no Figures). The same two markers (D21S11 and TH01) are missing due to PHR status and must be manually reviewed. Once reviewed, the major profile matches the extracted known profile. A few of the unknown minor contributor loci have multiple combinations listed due to the possibility of allelic dropout. For all of these instances, the first profile listed (most probable) was the correct profile.

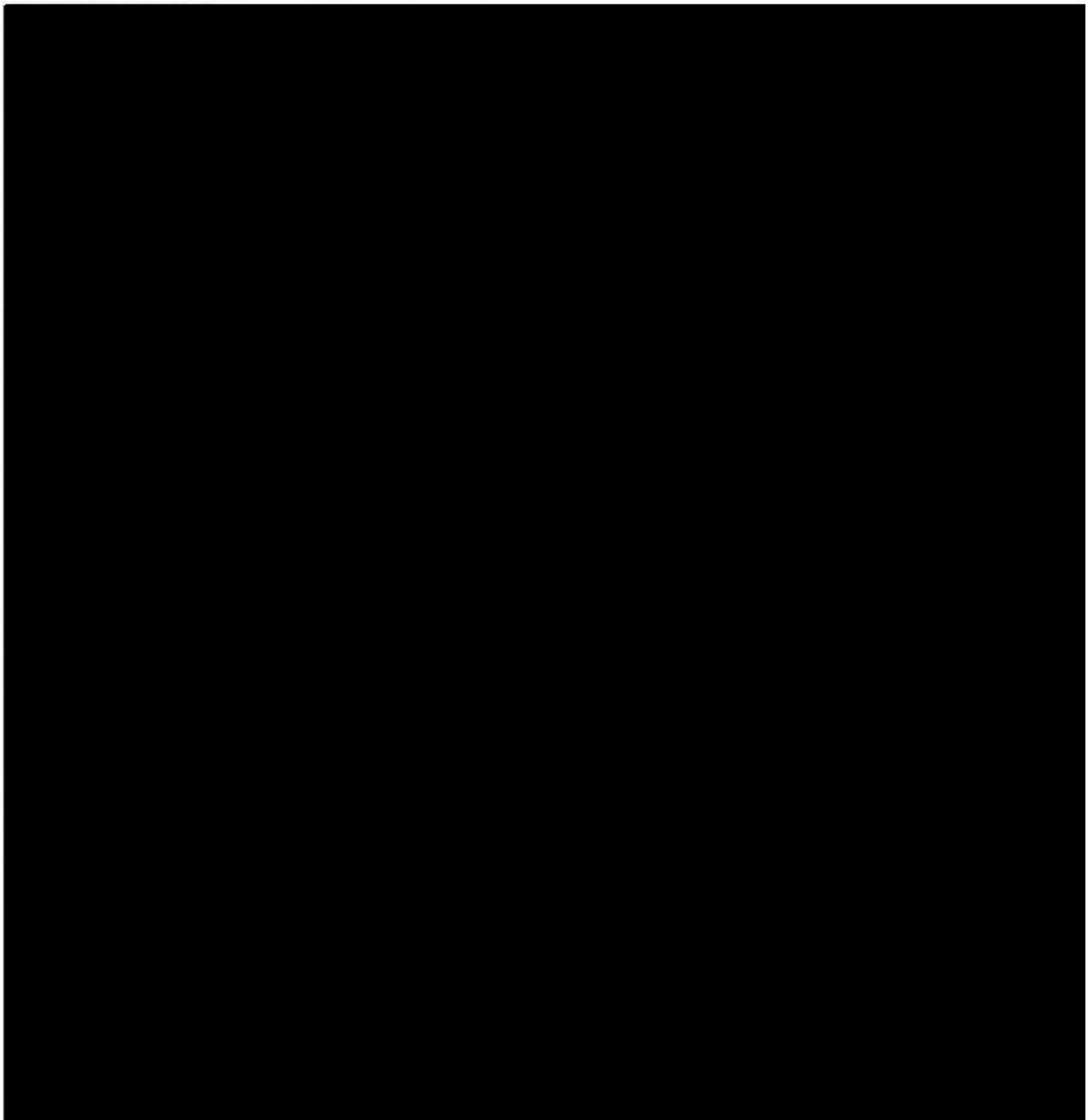


Figure 21: Sample [redacted] present at a much lower proportion.

Figure 22: The selected genotype combinations for sample [redacted] is extracted. The profiles are highlighted in light yellow and green, signifying that markers are missing.

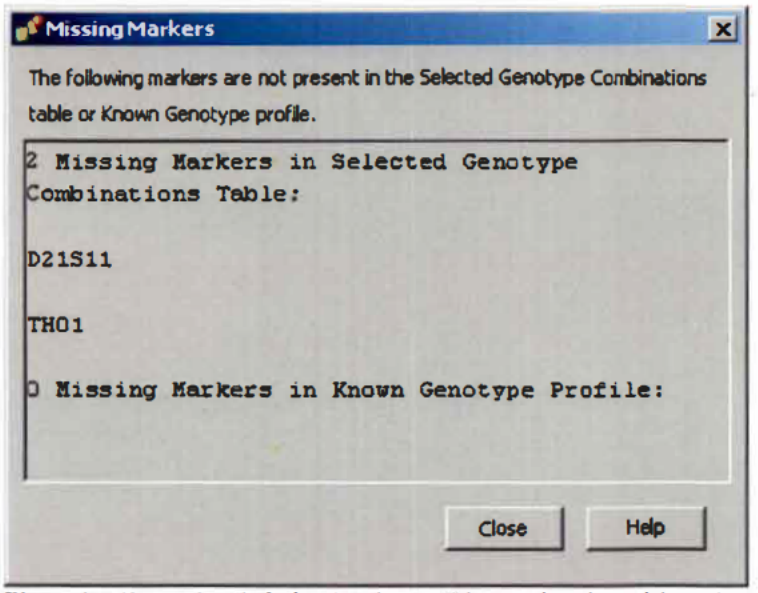


Figure 23: When the "Missing Markers..." button is selected (see right side of figure 22), the above information is given. The analyst will have to view the unselected genotypes combination table and determine what profile possibilities to include for these markers.

Figure 24: A portion of the unselected genotype combinations for sample [redacted] showing marker D21S11. Notice row 6 (1st row in figure) was not included in the selected genotypes table due to PHR status (red flag) making the IQ yellow.

Figure 25: A portion of the unselected genotype combinations for sample [redacted] showing marker TH01. Notice row 63 (1st row in figure) was not included in the selected genotypes table due to PHR status (red flag) making the IQ yellow.

Figure 26: The selected genotype combinations for sample [redacted] after manually reviewing and moving the missing profiles for D21S11 and TH01. Notice that D21S11 and TH01 still retain their red PHR flag and yellow IQ flag. The profiles are now highlighted in bright yellow and green, signifying that all markers are present.

Figures 27-29 show an example of mixture sample [redacted] with known contributor [redacted] extracted. The software correctly assigned the known profile to the major contributor at all loci. For the remaining minor contributor, 10 loci were correctly assigned by the software. The remaining 6 loci all had multiple possible profiles listed. For 4 of these 6 loci, the first profile listed (most probable) was correct. For one locus, the second profile listed was correct. For the final locus (D2S1338), the correct minor profile was missing from both the selected and unselected genotype combinations tables. This was a result of one of the minor alleles being masked by stutter from a major allele (Figure 29). The unselected genotype combinations table did have [redacted] as the first possibility for the minor contributor [redacted]. However, this is not entirely accurate because the [redacted] did not dropout (signified by F1), but was instead masked by stutter. When the extracted known contributor is switched to [redacted] (no figures), the unknown major contributor is determined at 15 of the 16 loci. As mentioned above, locus D2S1338 was not accurately deciphered due to the [redacted] being masked by stutter. As a result, D2S1338 was a missing marker from the selected genotype combinations table because it could not be matched to the known minor profile.

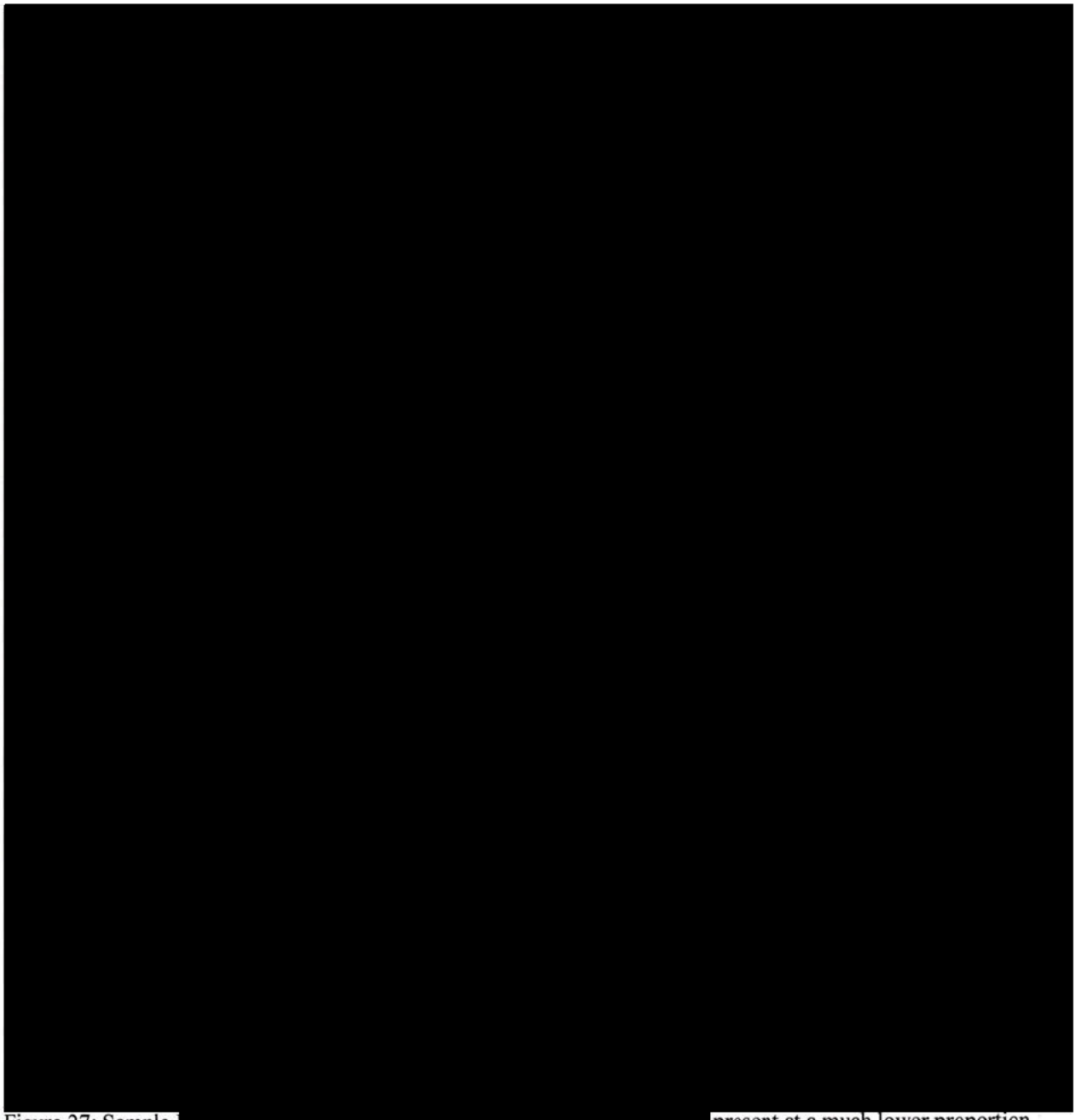


Figure 27: Sample [redacted] present at a much lower proportion.

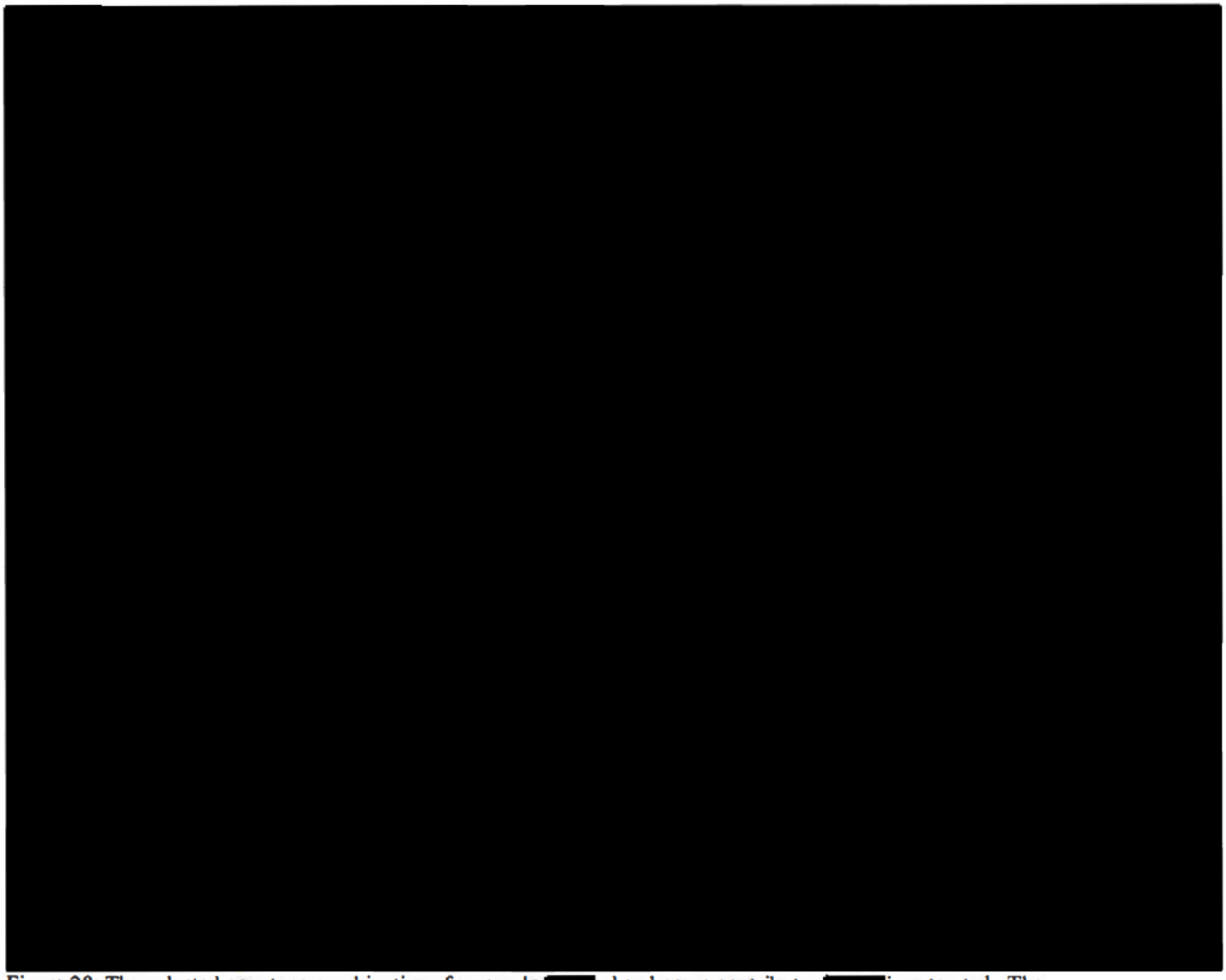


Figure 28: The selected genotype combinations for sample [redacted] when known contributor [redacted] is extracted. The profiles are highlighted in bright yellow and green, signifying that all markers are present. Notice that several markers for the minor contributor have multiple possibilities listed.

determine the correct minor contributor due to the [redacted] (highlighted) being masked by [redacted] stutter. The [redacted] was within the acceptable stutter for this locus (11.1%), and therefore it was not called by the software.

-Major/minor determination:

Two mixture samples ([redacted] with known contributors were analyzed to determine if the software was accurately determining the major and minor contributors.

Figures 30-32 show an example of mixture sample 1A when analyzed with the mixture analysis tool. The software correctly assigned the major and minor profile combinations for 10 of the loci. The remaining 6 loci had multiple possibilities listed. For 4 of these 6 loci, the first profile listed (most probable) was correct. For one locus, the second profile listed was correct. The last locus (CSF1PO) is missing from the selected genotype combinations table due to PHR status. The PHR for the major contributor's peaks is about 66%. The threshold set in the SDPD mixture analysis method for that particular peak height range (above 1000 RFU) is 75%. As a result, the locus is flagged red for PHR and included in the unselected genotypes combination table with a yellow IQ (see Figure 32). The analyst will have to manually review this locus to determine which profile possibilities should be moved to the selected genotype combinations table (highlight the profile and press the "Select" button).

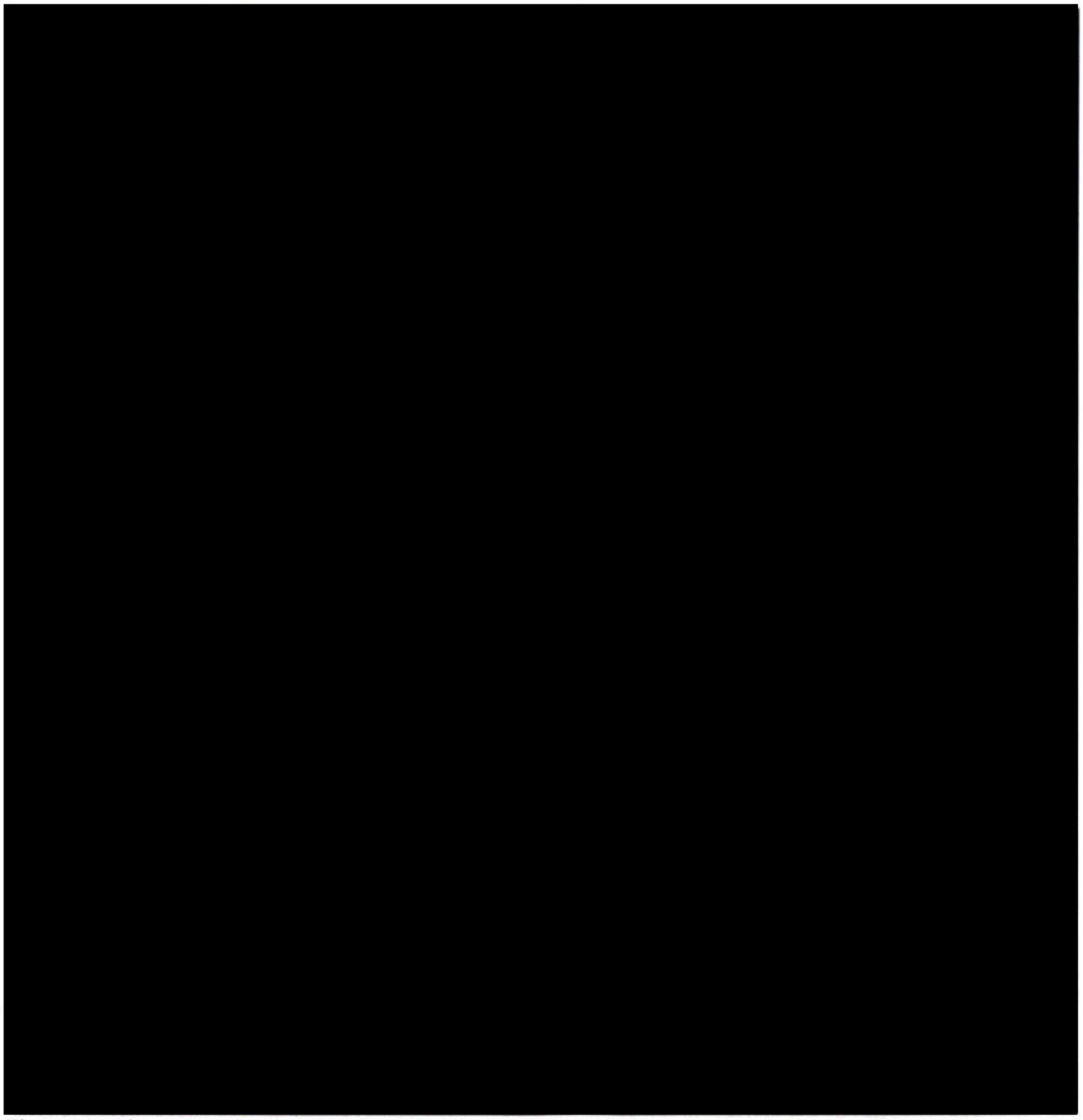


Figure 30: Mixture sample 1A.

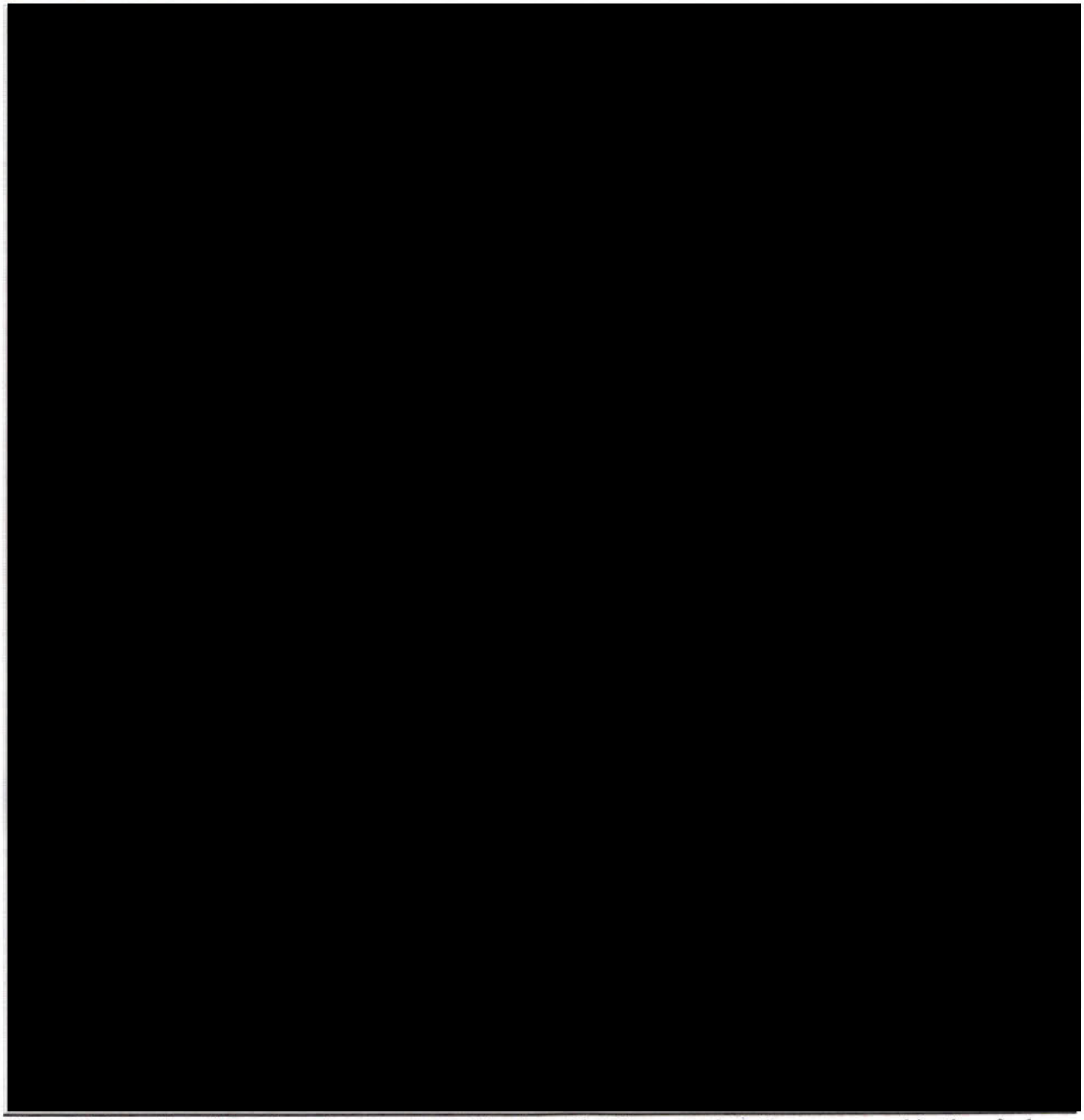


Figure 31: A portion of the selected genotypes combination table. The two selected genotype combinations for locus D21S11 are shown highlighted in blue. For this locus, the second choice (row 3) is the correct profile combination. Rows 9-13 show the five possibilities for locus TH01. For this locus, the first choice (row 9) is the correct profile combination.

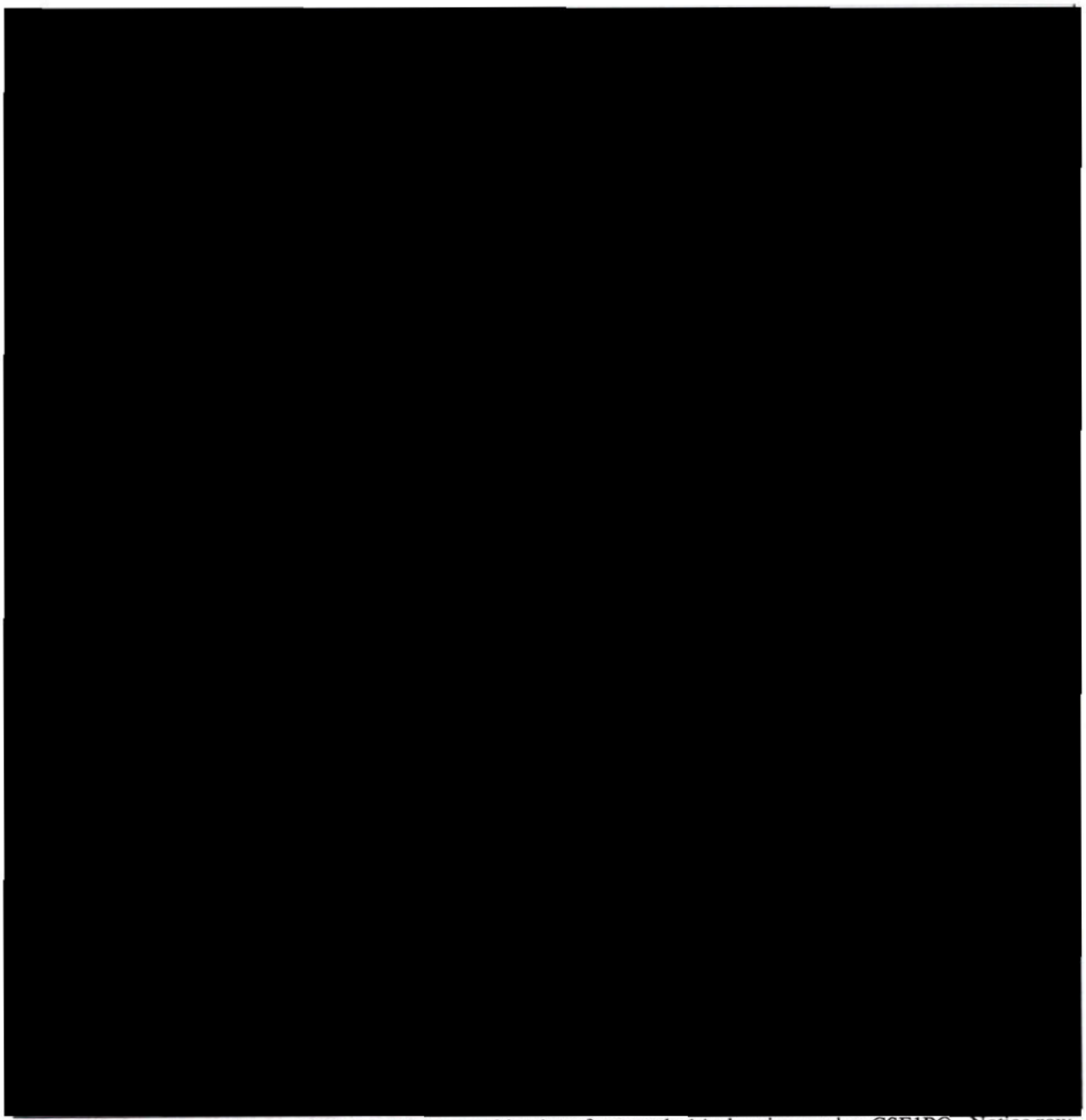


Figure 32: A portion of the unselected genotype combinations for sample 1A showing marker CSF1PO. Notice row 33 (the correct profile combination) was not included in the selected genotypes table due to PHR status (red flag) making the IQ yellow.

Figure 33 shows an example of mixture sample 5A. When analyzed with the mixture analysis tool, the software correctly assigned the major and minor profile combinations for 6 of the loci. The remaining 10 loci had multiple possibilities listed. For 9 of these 10 loci, the first profile listed (most probable) was correct. For one locus, the second profile listed was correct.

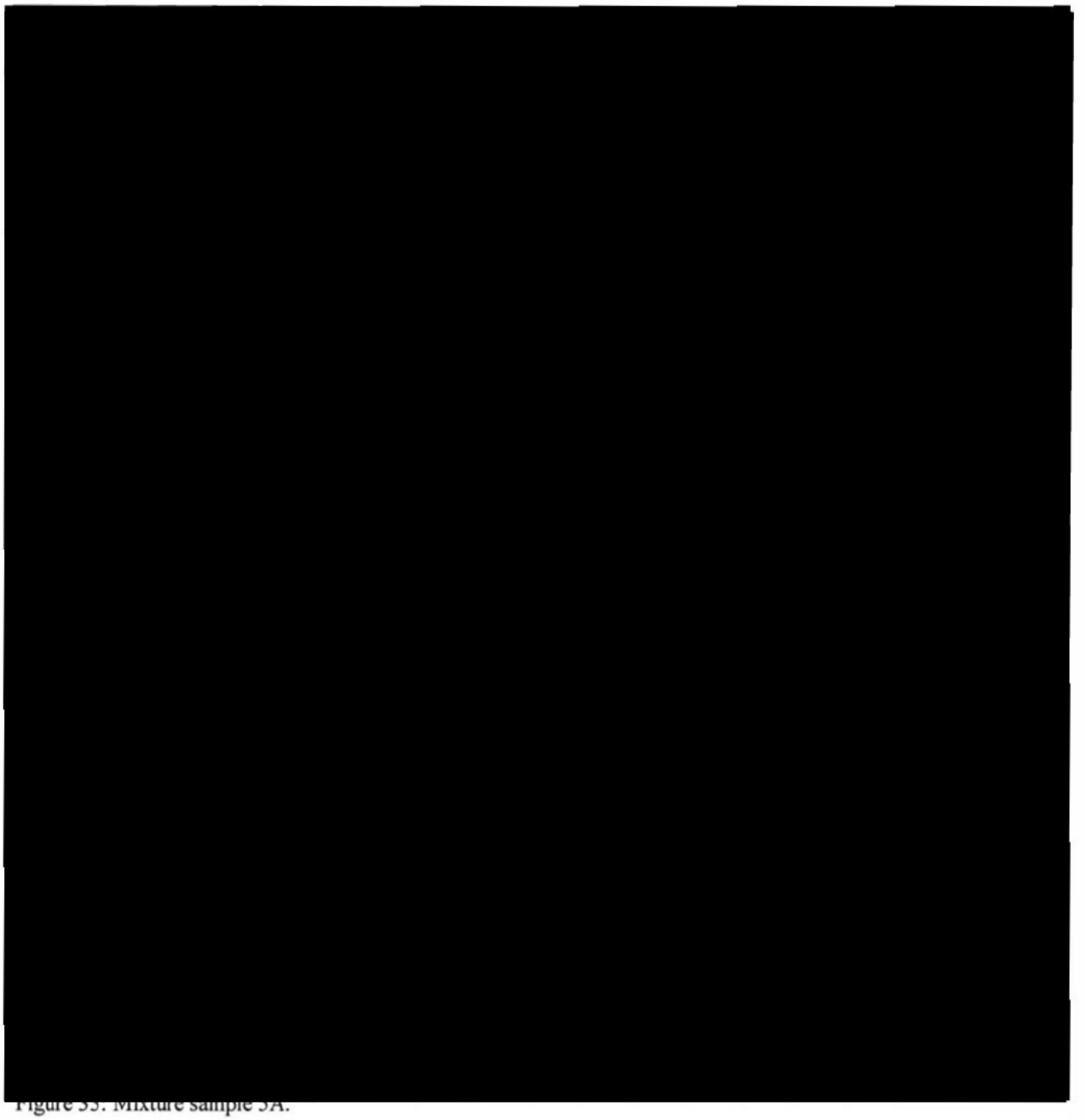
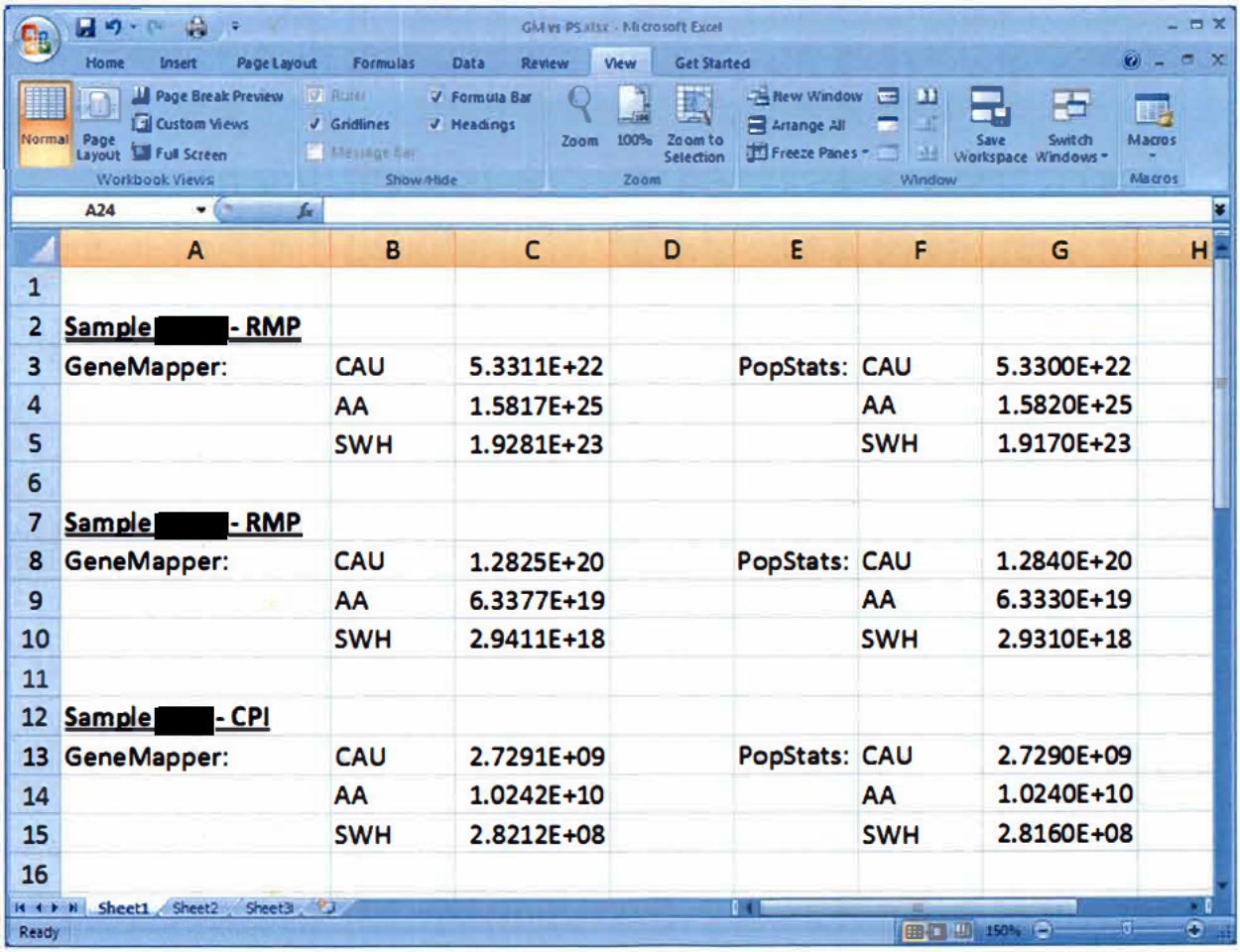


Figure 33: Mixture sample 5A.

It appears as though the same type of issues seen when extracting known contributors can be encountered when the software is determining major/minor contributors without known contributors. The analyst must be aware of these types of situations when analyzing mixture data.

-Statistics:
Mixture sample [redacted] as analyzed with the mixture [redacted] o generate a major and minor [redacted] rofile consistent with single source samples [redacted] The statistics generated in GeneMapper ID-X for these samples were

compared to the statistics from PopStats (see Figure 34). The numbers were not an exact match, which was determined to be a result of rounding differences within the software programs.



	A	B	C	D	E	F	G	H
1								
2	**Sample [redacted]- RMP**							
3	**GeneMapper:**	**CAU**	**5.3311E+22**		**PopStats:**	**CAU**	**5.3300E+22**	
4		**AA**	**1.5817E+25**			**AA**	**1.5820E+25**	
5		**SWH**	**1.9281E+23**			**SWH**	**1.9170E+23**	
6								
7	**Sample [redacted]- RMP**							
8	**GeneMapper:**	**CAU**	**1.2825E+20**		**PopStats:**	**CAU**	**1.2840E+20**	
9		**AA**	**6.3377E+19**			**AA**	**6.3330E+19**	
10		**SWH**	**2.9411E+18**			**SWH**	**2.9310E+18**	
11								
12	**Sample [redacted]- CPI**							
13	**GeneMapper:**	**CAU**	**2.7291E+09**		**PopStats:**	**CAU**	**2.7290E+09**	
14		**AA**	**1.0242E+10**			**AA**	**1.0240E+10**	
15		**SWH**	**2.8212E+08**			**SWH**	**2.8160E+08**	
16								

Figure 34: The Random Match Probabilities for single source contributors [redacted] and the Combined Probability of Inclusion for mixture sampl[redacted] The GeneMapper ID-X numbers are slightly different than the PopStats numbers due to rounding differences.

Allele Table Export

Sample profiles (excluding ladders, blanks, controls, etc.) for projects [redacted] were exported to an allele table using the GeneMapper ID-X Report Manager. The profiles exported to the allele table were manually reviewed using the electropherograms. No discrepancies were found in the exported allele table.

Y-STR Stutter

Project s[redacted] was analyzed in GeneMapper ID-X to show that the appropriate stutter positions would be filtered (not called) by the software. Figure 35 shows that the stutter was properly filtered. Specifically, the DYS19 n-2 bp and DYS392 n+3 bp positions were not called.

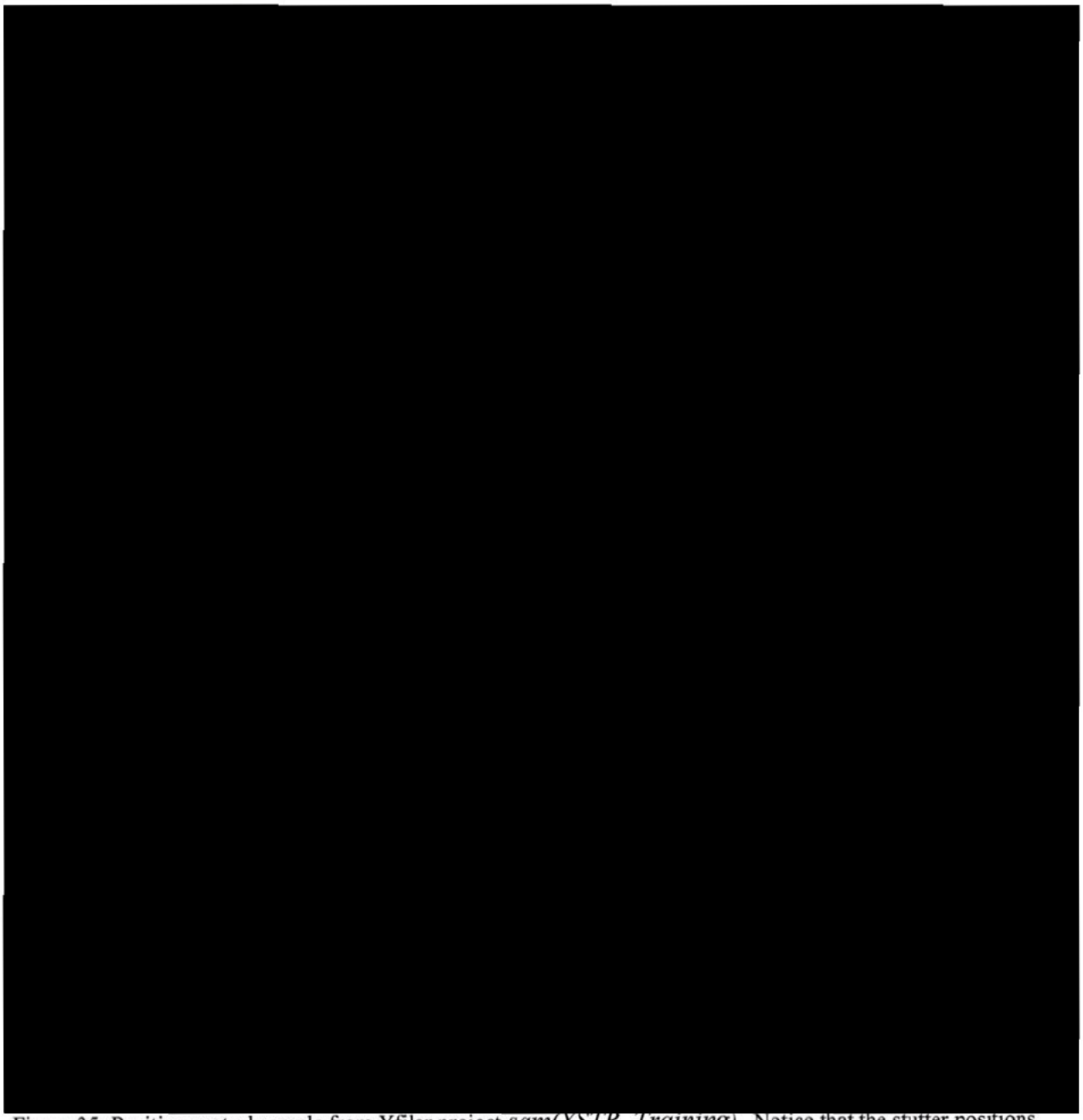


Figure 35: Positive control sample from Yfiler project *sam(YSTR_Training)*. Notice that the stutter positions have been appropriately filtered by the software.

Conclusions

Genotypes, Base Pair Size, and Peak Heights

For all three Identifiler projects, the final DNA results were equivalent in allele assignments, base-pair sizing, and peak heights whether the data was analyzed using GeneMapper v3.2 or GeneMapper ID-X.

Quality Flags

The quality flags assessed (ladder, positive, negative, and sizing quality) were determined to be accurate. For the negative control quality flag, the analyst should be aware that the software does not give any indication for the possibility of low level contamination with peaks emerging below threshold.

Lab Reference/Sample Comparison

The Lab Reference and Sample Comparison Tools were determined to be accurate. The percent match is the percent of the reference profile alleles detected in the comparison profile. This is calculated by dividing the number of matching alleles by the total number of alleles in the reference sample (including amelogenin). While this feature would be useful, there is no way to directly import staff profiles. The software requires you to add all of the old files to a project, analyze them, and assign them a profile ID. This is a problem for our lab due to the number of profiles we have and the fact that many of them were run/analyzed many years ago. In order to use this feature, all of the old files would need to be located or all of the samples re-run from the beginning. ABI has been contacted about a more efficient way to import existing staff profiles, but they have not come up with a solution as of yet.

Spike Labeling Feature

Due to the specific spike calling requirements of the software (spikes must be within three dye colors, within one data point of each other, and have a minimum peak half width of 0.5 bp or less), it is possible to have spikes that do not meet the calling requirements. Analysts will have to manually confirm spikes that are not called by the software, and should not assume that spikes are not present just because they are not automatically labeled. However, due to the specific calling algorithm of the software, it is accurate to conclude that spikes that are labeled are true spikes. Currently, the software has a bug that prevents it from correctly flagging negative controls and blank samples (no alleles called), even though it does recognize and label spikes within these samples. It will label the spike and flag the SSPK yellow, but the locus and CGQ will incorrectly remain green. ABI claims that the bug will be fixed in a future upgrade. Until then, analysts should be aware of this issue and interpret spikes within negative controls and blank samples accordingly.

Mixture Analysis Tool and Stats

The GeneMapper ID-X mixture tool is designed to assist the analyst in mixture interpretation, not replace their analysis. There are situations where the analyst may not agree with the software's findings and will have to determine a more accurate conclusion. In many instances, the software will provide a list of profile combinations that the analyst will have to review for the most likely possibility.

-Major/minor determination and Extraction of known contributor:

As demonstrated in the results section, there are issues that arise during the software's interpretation of mixtures that require the analyst's review. These include, but may not be limited to the following: multiple possible profile combinations listed, especially when allelic dropout is a possibility; correct profile moved to unselected genotype combinations table due to PHR being below limits; correct profile excluded from both tables (not listed) due to a minor

allele peak being masked by stutter. The analyst will have to recognize these issues and interpret them appropriately. That being said, the mixture tool is still a very useful aid to the analyst during mixture interpretation. It can be especially useful in helping to determine the most likely profiles from contributors for entry into CODIS. For those samples with multiple profile combinations listed, the major and minor contributor interpretation becomes ambiguous. For these types of samples, it is probably best to search all alleles in CODIS.

-Statistics:
When the statistics were compared between GeneMapper ID-X and PopStats, the numbers were found to be slightly different. ABI was contacted, and it was determined that these minor differences were a result of rounding differences within the two software programs. The differences are not significant and either software is considered accurate when calculating statistics. When using the mixture analysis statistics for a sample with a known contributor extracted, it should be noted that the software only calculates statistics for the unknown contributor. Another factor to consider when using GeneMapper ID-X for statistical analysis is allelic dropout. For RMP calculations, if the software determines that allelic dropout is possible (designated by an "F" allele in the profile) it will use $2p$ for calculations. For CPI (when allelic dropout is possible) the software will substitute a value of "1.0" for the genotype frequency, essentially making that locus inconclusive and not including it in the calculation. The analyst may need to review the software's conclusions for allelic dropout and confirm their findings before using the statistics.

Allele Table Export
The allele table export feature of GeneMapper ID-X was determined to be accurate. It is extremely important that the proper allele number be set in the project genotypes table settings *and* the report manager settings. If they are set to show too few alleles (i.e. sample has 6 alleles per locus but settings say to show 4 alleles), the report, and therefore the exported allele table, may omit some of the alleles at a locus.

The exported table will still require some editing to obtain the desired format that is currently used. However, this is reduced by importing into an existing template. This feature is expected to save ample time and decrease input errors when compared to the manual entry method. A macro is planned for future implementation to insert parentheses for lower intensity alleles and possibly delete the extra ending comma that remains after export.

Y-STR Stutter
The software was successful in filtering the Y-STR stutter. The software uses the following new stutter percentages for filtering: DYS19 n-2 bp = 10.21% and DYS392 n+3 bp = 7.90%. These values can be adjusted in the Panel Manger in each marker's individual stutter ratio chart if needed.

7/22/09
Meghan Didier, Criminalist (Intern)

08/01/2009 (APPROXIMATELY) /SM 11/18/09
Shawn Montpetit, DNA Technical Manager

8/3/09
Patrick O'Donnell, Supervising Criminalist

JMS 8-7-09