13 Analysis and Interpretation of DNA Results

13.1 Autosomal Analysis and Interpretation Guidelines

The AmpF/STR® Identifiler® and AmpF/STR® Identifiler® Plus PCR Amplification Kits are short tandem repeat (STR) multiplex assays that amplify 15 tetranucleotide repeat loci and the Amelogenin gender determining marker in a single PCR amplification. The AmpF/STR® Identifiler® Plus PCR Amplification Kit uses the same primer sequences as the earlier generation AmpF/STR® Identifiler® PCR Amplification Kit. The Identifiler® Plus kit uses modified PCR cycling conditions for enhanced sensitivity, a new buffer formulation to improve performance with inhibited samples, and an improved process for DNA synthesis and purification of the amplification primers to deliver a much cleaner electrophoretic background. With these modifications, the AmpF/STR® Identifiler® Plus PCR Amplification Kit delivers the same power of discrimination as, better sensitivity than, and better robustness than the earlier generation of the AmpF/STR Identifiler® Kit. The AmpF/STR Identifiler® Plus kit employs the latest improvements in primer synthesis and purification techniques to minimize the presence of dye-labeled artifacts. These improvements result in a much cleaner electropherogram background that enhances the assay's signal-to-noise ratio and simplifies the interpretation of results.

Loci Targeted by AmpF/STR® Identifiler® and Identifiler® Plus PCR Amplification Kits

Locus designation	Chromosome location	Alleles included in Identifiler® Plus Allelic Ladder	Dye label	Control DNA 9947A
D8S1179	8	8, 9 10, 11, 12, 13, 14, 15, 16, 17, 18, 19	6-FAM™	13‡
D21S11	21q11.2-q21	24, 24.2, 25, 26, 27, 28, 28.2, 29, 29.2, 30, 30.2, 31, 31.2, 32, 32.2, 33, 33.2, 34, 34.2, 35, 35.2, 36, 37, 38		30§
D7S820	7q11.21-22	6, 7, 8, 9, 10, 11, 12, 13, 14, 15		10, 11
CSF1PO	5q33.3-34	6, 7, 8, 9, 10, 11, 12, 13, 14, 15		10, 12
D3S1358	3p	12, 13, 14, 15, 16, 17, 18, 19	VIC®	14, 15
TH01	11p15.5	4, 5, 6, 7, 8, 9, 9.3, 10, 11, 13.3		8, 9.3
D13S317	13q22-31	8, 9, 10, 11, 12, 13, 14, 15		11#
D16S539	16q24-qter	5, 8, 9, 10, 11, 12,13, 14, 15		11, 12
D2S1338	2q35-37.1	15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28		19, 23
D19S433	19q12-13.1	9, 10, 11, 12, 12.2, 13, 13.2, 14, 14.2, 15, 15.2, 16, 16.2, 17, 17.2	NED™	14, 15
vWA	12p12-pter	11,12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24		17, 18
TPOX	2p23-2per	6, 7, 8, 9, 10, 11, 12, 13		8‡‡
D18S51	18q21.3	7, 9, 10, 10.2, 11, 12, 13, 13.2, 14, 14.2, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27		15, 19
Amelogenin	X: p22.1-22.3 Y: p11.2	X, Y	PET®	X
D5S818	5q21-31	7, 8, 9, 10, 11, 12, 13, 14, 15, 16		11§§
FGA	4q28	17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 26.2, 27, 28, 29, 30, 30.2, 31.2, 32.2, 33.2, 42.2, 43.2, 44.2, 45.2, 46.2, 47.2, 48.2, 50.2, 51.2		23, 24

13.1.1 Analysis

13.1.1.1 GeneMapper Analysis

Allele assignment occurs through a three-step process:

1. **Spectral separation:** The 4 dyes that Identifiler uses in the STR amplification are 6-FAM, VIC, NED, and PET. A fifth dye, LIZ, is used in the internal size standard. Although each dye emits its maximum fluorescence at a different wavelength, there is overlap in the emission spectra. Multi-component analysis is the process that separates the five different fluorescent dye colors into distinct spectral components. This analysis occurs automatically after each instrument run.

   Note: The precise spectral overlap is determined by separately analyzing DNA fragments labeled with each of the dyes (spectral standards). A spectral should be run on each 3100 at least every six months, following maintenance/repairs (other than routine cleaning), and anytime poor spectral separation is suspected.
2. **Peak BP sizing:** The internal size standard (GS500 LIZ, for the Identifiler, Identifiler Plus, and Yfiler multiplexes) is used to calculate precise peak bp sizing. The internal size standard is used to normalize injection-to-injection variations. The Local Southern method is used to compare allele peaks with the LIZ peaks and calculate the bp size. See the GeneMapper User Manual for explanation of the Local Southern method.
3. **Allele designations**: Allele calls for each peak are made by comparing the bp size of the sample or control peak to the bp sizes of peaks in the allelic ladder. Thus for allele designations to be determined, an allelic ladder must be present and interpretable in each run, and each ladder must have the appropriate alleles present for each locus when analyzed.

13.1.1.2 GeneMapper Procedure

1. Login to GeneMapper using the appropriate username and password.

2. Add Samples to Project: Browse to appropriate data folders. Select folder, select all samples, click Add to List, and add to the right of window.

3. Check that the Sample Type, Analysis Method, Panel, and Size Standard are correctly labeled.
   1. For Identifiler:
      i. Sample types: Select the sample type; allelic ladder samples must be labeled as “Allelic Ladder”
      ii. Analysis method: HPD 3130
         1. Verify the analysis range is set to exclude the primer peak
         2. Verify the minimum peak threshold is 100 RFU
      iii. Panel: Identifiler_v2
      iv. Size standard: CE_G5_HID_GS500

2. For Identifiler Plus:
    i. Sample types: Select the sample type; allelic ladder samples must be labeled as "Allelic Ladder"
    ii. Analysis method: Identifiler_Plus_AnalysisMethod_v1
        1. Verify the analysis range is set to exclude the primer peak
        2. Verify the minimum peak threshold is 50 RFU
    iii. Panel: Identifiler_Plus_Panels_v1
    iv. Size standard: CE_G5_HID_GS500

4. Analysis settings are GeneMapper Default settings unless otherwise documented in this SOP. During analysis of Identifiler samples, the threshold may be lowered to 50 RFU to investigate any potential alleles below the 100 RFU threshold. Note the presence of possible additional alleles with an asterisk at that locus in the allele chart, but do not use these in the interpretation of results. Highlight a column heading and hit the green arrow on the toolbar.

4.1 The following is recommended for naming project files:

4.1.1 If a single ladder can be used for all samples in a run:
1. Choose one ladder.
2. Uniquely name the project, such as with the initials and the date the run was set up. For example, a run on May 19, 2009 will be named JD051909. Multiple runs on the same day can be put into sequential order utilizing either numbers or letters (e.g. JD051909, JD051909B, PA051909, and PA051909#2). Alternatively, the LIMS-generated 3130 plate number may be used to identify the GeneMapper project.

4.1.2 If multiple ladders are required for samples in a run:
1. Multiple projects will need to be created. Samples not being analyzed do not need to be deleted from the project.
2. Choose one ladder. Print out the injection grid, ladder, positive control, and all samples associated with that particular ladder.
3. The project name will be adjusted to indicate the appropriate ladder used for analysis. For example, the same run on May 19, 2009 using two different ladders would be named JD051909_1_ladder_D3 (run name_injection#_"ladder"_well location).

5. At this point, the plots can be analyzed. Hit the Display Plots icon on the toolbar. A separate window will open containing all system plots. Click OFF the all colors except LIZ. The internal size standard for each lane appears in the window. Verify the following for each lane:
    5.1 All of following peaks are present in LIZ: 75, 100, 139, 150, 160, 200, floater (245-250), 300, 340, 350, 400, and 450 bps.
    5.2 Sample floater peaks must be within +/- 0.5 bp of the allelic ladder floater peak. Choose one ladder. A new project will have to be created for samples that do not fall within +/- 0.5 bp of the selected ladder, using a different ladder from the

same plate record. Make sure each project contains only one ladder used for analysis.

5.3 Only one print out of the negative control is required per run, but each GeneMapper project must include at least one positive control.

6. Examine ladder and all controls and confirm:
   6.1 Allelic ladder was correctly labeled (see amplification kit User's Manual).
   6.2 All positive control peaks are present and labeled correctly. No extra peaks indicating contamination are present. Each amplification set and each analyzer run must have one acceptable positive control. If the amplification positive control fails to amplify/type correctly, all associated samples must be re-amplified.
   6.3 All negative controls (reagent blanks and amplification negative control) show no labeled peaks indicating contamination. Each reagent blank must have acceptable results for data from the corresponding extraction set to be used. The amplification negative control must have acceptable results for data from the corresponding amplification set to be used. Additionally, each analyzer run must have at least one acceptable negative control. Given there is no data for sizing, negative controls with less than ideal LIZ peaks may be used if it is clear that an injection with stronger LIZ data would not have resulted in allelic designations (e.g., peak morphology weakens towards the end of the injection). However, the Technical Leader must approve the use of such injections.

7. Examine all electropherograms for data quality and allele calls. Refer to the discussion below for guidance on evaluation data and editing artifacts and microvariant calls. When possible, use the "track changes" function of GeneMapper to electronically document changes made to the GeneMapper calls. Then if necessary, to change or delete the allele type or call, click on the peak or allele call to be edited to select it. Choose either to Delete Allele or Rename Allele. After making selection, a dialog box will appear asking for sample comment. You must type in explanation of edit.

8. Upon completion of analysis, save the project. Print plots for all case samples, associated controls, and ladders to be included in the case file. Place in each case folder a printout of the complete GeneMapper project list for each run and the electropherograms (all 5 colors) from GeneMapper ID for the ladder, positive control, negative control, and all samples and reagent blanks for that case. If an entire run is not used, it is not necessary to print the plots; however, the injection list must be included in the case file with an explanation as to why data was not used (e.g., no acceptable positive control). Document each time a sample is manipulated including longer or shorter injection times.

9. A second analyst must perform the GeneMapper analysis independently and agree with the allele calls reported by the first analyst on the allele chart by noting their "second read" on the injection list.

10. If LIMS was used to create the 3100 Plate map, you must export the project for entry into LIMS. After saving the project, select the GenoTypes tab and then choose HPD Table. Select File>Export Table and choose the location to which you will be exporting, such as the Removable Drive (jump drive). From the 3100 Plate Map screen in LIMS, click the Data Review button at the bottom. The Allele Call Table Worksheet will be displayed. Click on the "1st Read" button at the bottom of the screen. Browse to the location of your saved GeneMapper® *ID* file. Select your file and click Open to bring your allele calls for each of the samples in the batch into the table.

11. The analyst performing your second read will also create a GeneMapper® *ID* project and import their project back into LIMS. However, he/she will choose the "2nd Read" button in the Allele Call Table Worksheet to import his/her allele calls. Any mismatches will be highlighted in red and the run that should be imported will have to be selected. Every effort should be made to resolve any discrepancies between the analysts. If an agreement cannot be reached, the Technical Leader must be involved.

12. If LIMS was not used to create the 3100 Plate Map prior to the genetic analyzer run, the allele calls can be manually entered into LIMS in the DNA Analysis Results Chart section of the Report Writing tab in each case's DNA assignment.

13.1.1.3 Artifacts and unusual results

True alleles are defined as any peak that meets established threshold values, which is clearly visible above baseline noise, is of a size that falls within a defined category as determined by the GeneMapper program, and is not an artifact. Peaks other than true alleles may be detected on the electropherogram and labeled by GeneMapper. The source of these artifacts should be determined where possible. If such a peak can be definitely identified as an artifact and the peak does not interfere with interpretation of the data, then the artifact can be disregarded (See 13.1.1.2.7) and the data can be used for interpretation. The presence of the artifact will be documented on the plot. A technical reviewer must agree with the analyst's decision and shall initial the plot to indicate his/her agreement. Disagreements between the analyst and technical reviewer must be brought to the technical leader for resolution. If an artifact or suspected artifact may interfere with interpretation, the data cannot not be used; the locus may be called inconclusive or the sample may be re-analyzed (re-inject or re-amplify) to resolve the issue. Commonly observed artifacts include:

1. **Spikes:** Spikes are sharp, narrow peaks generally present in all colors and occur at the same location. This is often caused by electrical anomalies. This can also be caused by air bubbles and urea crystals.

2. **Stutter:** A stutter peak is a reproducible minor product peak usually four bases (1 repeat) shorter than the corresponding main peak allele peak. Stutter products are the most common source of additional peaks in an STR sample, caused by a slipped-strand mispairing during PCR. Rarely, stutter may be observed at two or three repeats shorter that the true allele, or one repeat greater than the true allele. Peaks in the minus 1 repeat position that fall below the maximum published % stutter values (see the

amplification kit's User's Manual) may be assumed to be stutter, and GeneMapper does not label these peaks. If an apparent stutter peak exceeds the published maximum stutter percent, the sample must be carefully evaluated to determine if the peak may be a true peak or stutter. In general, for each AmpF/STR® Identifiler® Plus Kit locus, the percent stutter increases with allele length. Forward stutter peaks generally exhibit a much lower percentage difference from the true allele than minus one repeat stutter peaks. Both minus 1 repeat and plus 1 repeat stutter seem to be more common when excessive template DNA is amplified (when one or more of the critical components of the amplification master mix becomes limited or depleted, causing a loss of processivity).

For mixtures in which minor contributors are determined to be present, a peak in stutter position (generally n-4) may be determined to be 1) a stutter peak, 2) an allelic peak, or 3) indistinguishable as being either an allele or stutter peak. This determination is based principally on the height of the peak in the stutter position and its relationship to the stutter percentage expectations established by the laboratory. Generally, when the height of a peak in stutter position exceeds the laboratory's stutter expectation for a given locus, that peak is consistent with being of allelic origin and should be designated as an allele. If a peak is at or below this expectation, it is generally designated as a stutter peak. However, it should also be considered as a possible allelic peak, particularly if the peak height of the potential stutter peak(s) is consistent with (or greater than) the heights observed for any allelic peaks that are conclusively attributed (i.e., peaks in non-stutter positions) to the minor contributor(s).

A phenomenon known as over-subtraction could cause elevated pull-up, stutter and minus A peaks when the analysis software normalizes the true off-scale peak(s) but does not adjust the artifacts (because they are not off-scale). This may also result in a loss of linearity in RFU increase when injection times are extended.

3. **Incomplete 3'(+A) nucleotide addition/Minus A**: PCR amplification results in the addition of a single "A" nucleotide at the 3' end of double stranded PCR products, resulting in a product that is one bp longer than the actual target DNA sequence. PCR reactions have been optimized to favor this "A" addition, but incomplete "A" addition may occur when excessive amount of target DNA is present, or in other conditions that are less than optimal for the PCR reaction. Incomplete "A" addition, or "minus A" appears as a peak one bp shorter, and typically at a smaller peak height, than the true allele. See the stutter discussion (13.1.1.3.1) regarding over-subtraction.

4. **Pull-up/Matrix Failure:** Pull-up is the result of the instrument's inability to completely separate the spectral components. Pull-up is identified as a smaller peak of the same location as a true allele but in another color. It is the result of either excessive DNA or a faulty spectral. Complex pull-up occurs when a locus is heterozygous and the two peaks are within one repeat unit of each other. This can cause a single bridge-like peak to be observed in another color channel. See the stutter discussion (13.1.1.3.1) regarding over-subtraction.

5. **Dye Blobs:** Dye blobs occur when fluorescent dyes come off their respective primers and migrate independently through the capillary. Dye blob morphology is generally characteristically broad, and not very "peak-like".

13.1.1.4 Microvariants

Microvariants are true alleles that vary by fewer than 4 bp from the typical repeating unit. The designation of alleles containing an incomplete repeat unit, falling within the range spanned by the ladder alleles, should include the number of complete repeats, a decimal point, then the number of base pairs in the incomplete repeat (e.g. 9.3 for a TH01 allele with 9 full repeats plus three more bps). The determination of the number of additional bp present in a microvariant is made by comparing the bp size of the off ladder peak with the bp size of the flanking alleles in the allelic ladder. If the locus is heterozygous, the sister allele should be evaluated to help establish whether the off-ladder allele is truly a microvariant, or simply off-ladder due to electrophoretic drift.

If an allele falls greater than four bp away from the largest or smallest allele at a locus, it will be designated as greater than or less than the appropriate ladder allele (i.e. > 11 for TH01).
Any allele designated as off-ladder by GeneMapper and not determined to be an artifact is verified by reinjection or re-amplification. If the microvariant appears in multiple evidence samples from the same case otherwise having the same profile, it is not necessary to re-inject. Because every measure should be taken to interpret evidence samples prior to reference samples, the presence of a microvariant in an associated known sample will not suffice for confirmation, as the evidence should be interpreted before and separate from any associated knowns in the case.

It is recommended that analysts visit http://www.cstl.nist.gov/strbase/var_tab.htm (Variant Allele Reports) for a given off-ladder allele. A printout of the previously observed off-ladder allele may be retained in the case file.

13.1.1.5 Re-analysis and Additional Analysis

Any step in the process with unacceptable controls (as defined in this SOP) must be rerun. Refer to SOP 2 (Quality Assurance) for instructions on investigation and reanalyzing samples where contamination is detected or is suspected. Data from analysis with unacceptable controls may not be used for interpretation.

If low levels of DNA provide insufficient data upon initial analysis, the analyst may choose any of the following (as long as relevant maximums specified in the SOP's are not exceeded):

1. Re-amplify the sample with more template DNA (or less template DNA, if inhibition is suspected).
2. Re-inject for a longer amount of time (Identifiler samples may be extended to 20 seconds).

The associated reagent blank and amplification negative control should also be subjected to the extended injection time.

Samples with excessive RFUs (>~4000 RFUs) must be interpreted with caution. Excessive template DNA can make data interpretation difficult with a greater propensity for pull-up/matrix failure, as well as possible non-specific amplification. A phenomenon known as over-subtraction could cause elevated pull-up, stutter and minus A peaks when the analysis software normalizes the true off-scale peak(s) but does not adjust the artifacts (because they are not off-scale). This may also result in a loss of linearity in RFU increase when injection times are extended. The analyst should re-evaluate the quantification data to ensure the correct amount of template was amplified. The sample may be re-analyzed to obtain suitable data through any of the following:

1. Re-inject the sample for a shorter time.
2. Dilute the amplified product in TE buffer, add 1 µl to the formamide/LIZ mixture, and re-inject (the associated reagent blank must also be diluted to demonstrate the TE buffer is DNA-free).
3. Re-amplify the sample with less template DNA.

Results with artifacts that interfere with interpretation may be re-analyzed as needed to resolve the issue:

1. Artifacts resulting from poor electrophoresis may simply be re-injected.
2. Artifacts resulting from excessive DNA may be reinjected for a shorter time, diluted and re-injected, or re-amplified with less DNA.
3. Artifacts that are a result of poor amplification, such as excessive –A, are typically resolved best by re-amplification.
4. Pull-up due to a poor spectral can only be resolved by running a new spectral or performing other instrument maintenance.

13.1.2 Interpretation Guidelines

Results and conclusions from DNA analysis must be scientifically supported by the analytical data with appropriate standards and controls. Interpretations are made as objectively as possible and consistently from analyst to analyst. Not every situation can nor should be covered by a specific rule, and situations may occur that require an analyst to deviate from stated guidelines. Deviation from stated guidelines is allowable; however such deviations must be documented in the case file and approved by the Technical Leader. Such documentation should be sufficient for another experienced analyst to identify and understand the deviation from stated guidelines.

Whenever possible, evidentiary samples should be completely evaluated prior to the evaluation of any reference samples that are to be compared. The decision to attempt to develop a profile further through re-injection or re-amplification should be made without knowledge of the known profile for comparison, other than those of assumed contributors.

13.1.2.1 Types of Conclusions

Three types of conclusions are generally possible when both evidence (questioned) and reference (known) samples are tested and compared:

1. inclusion: the individual could have contributed to/been a source of the questioned profile; for a single-source sample, the same genotypes were obtained by an evidentiary sample and a reference sample, with no unexplained differences; for a mixture profile, all of the alleles from the reference sample are accounted for in the mixture profile of an evidentiary item (the loss of an allele due to incomplete (preferential) amplification, stochastic effects, mutation, or other factors must be considered and does not necessarily indicate an exclusion)
2. exclusion: the individual could not have contributed to/been a source of the questioned profile; the genotype comparison shows profile differences that can only be explained by the two samples originating from different sources
3. inconclusive: the data does not support a conclusion; insufficient information exists to support any conclusion. The reason for inconclusive results must be clearly stated.

Inclusion or exclusion is determined by qualitative and quantitative evaluation of the entire DNA profile produced at the various loci tested. Inconclusive results, or an uninterpretable profile, may result from, but are not limited to:

1. Insufficient amounts of template DNA
2. Degradation
3. Inhibitors
4. Mixtures of DNA from multiple donors

It should be noted that it may be acceptable for an inclusion or exclusion to be determined when one or more loci yield inconclusive results. An inclusion statement, and any resulting statistical calculations, will be based only on loci that yield conclusive results. An exclusion statement can be determined if even a single locus produces exclusionary results. However an exclusion will not be determined if technical issues such as the loss of an allele due to incomplete (preferential) amplification, stochastic effects, mutation, or other factors may have caused the non-match.

It also should be noted that it may be possible to obtain a conclusive result comparing one reference to a questioned sample, but an inconclusive result when comparing a second reference sample to the same questioned sample. This is most commonly observed when interpreting complex mixtures.

13.1.2.2 Types of Profiles

A number of different types of profiles (or a combination of these) may be obtained from evidence samples. The comparisons with references that can be made are determined by the type of profile.

1. **No profile:** No DNA results obtained at all. No comparison can be made to reference samples.

2. **Uninterpretable:** Data at too few loci or only possible activity below threshold is obtained, or an indistinguishable mixture contains excessive contributors. No comparison can be made to reference samples.

3. **Full single source profile:** A sample may be considered to be from a single person if the number of observed alleles at each locus is no more than two and the signal peak heights are balanced (the lower peak >60% of the higher peak) for heterozygous alleles. All loci must be evaluated when making this determination. If a sample has a third peak at just a single locus, with no other indication of a mixture, this may indicate an unusual mutation present in that individual; such a profile may be still considered single-source but must be interpreted with caution. If a sample has unbalanced peak heights with no other indication of a mixture, the sample may also be considered single-source but must be interpreted with caution.

Inclusions with full profiles can be made with respective statistics calculated and reported on probative samples. Exclusions with full profiles can be made if the individual being compared differs from the evidentiary profile at any one locus. If one or few non-matches are observed, the profiles should be evaluated carefully for evidence of dropout or artifacts that may have resulted in the non-match.

4. **Partial profiles:** Partial profiles exhibit allelic dropout in some, but not all, loci tested and can result from insufficient, degraded, or inhibited DNA. Typically, smaller loci amplify better under these conditions, and larger loci tend to drop out. Extremely low levels of template DNA may also lead to stochastic effects which may under represent one allele of a heterozygous locus. Exclusions with partial profiles may be made; again care must be taken to ensure that non-matches are not a result of drop-out.

5. **Mixtures:** Evidence samples may contain DNA from more than one individual. A mixture can consist of full and/or partial profiles from multiple individuals. One or more of the following may indicate the presence of a mixture:
   a. Greater than two alleles at a locus.
   b. A peak at the stutter position of significantly greater RFUs than is typically observed.
   c. Significantly unbalanced alleles at a heterozygous locus.

   Two types of mixtures may be observed:

   5.1 Mixture with Major/Minor Contributors

   Some mixed source profiles may be clearly differentiated into major and minor components. For STR analysis, if the major component fulfills certain criteria, it may be treated the same way that a single source profile is treated for estimating match significance. These criteria are: 1) interpretable loci show a major component and 2) no minor peak is greater than ~30% (≤30.5%) of the height of

the shortest major peak in the same locus for Identifiler amplifications and ~ 35% (≤35.5%) for Identifiler Plus amplifications. Heterozygous alleles considered major must also meet the 60% PHR requirement. Instances of homozygosity of the major donor or the minor donor, along with allelic sharing between contributors, may influence the balance such that the 30% or 35% parameter may be exceeded; the Technical Leader must approve instances where a major contributor is declared, despite a >30% or >35% PHR. If a clear major donor cannot be identified at the majority of the loci examined, the sample will be reported and treated statistically as an indistinguishable mixture. As with the single source evaluation criteria, no prior adjustment for stutter is assumed in the 30% or 35% maximum. It is permissible to determine a major profile at most loci, even if at some loci the major component cannot be determined. In this case, only the loci for which the major component can be determined may be used in the single source significance determination.

It should be noted that not all alleles of a minor contributor to a mixture may be determined, since minor contributor alleles may be masked by the presence of the same allele in the major contributor. The minor component of a mixture is treated in the same manner as a mixture with indistinguishable contributors for assessing inclusion or exclusion and for significance calculations. For a distinguishable mixture, a major contributor profile may be suitable for statistical analysis even in the presence of inconclusive minor contributor results.

5.2 Mixtures with indistinguishable contributors (unresolved mixture)

A mixture should be considered indistinguishable when the major and minor contributors cannot be distinguished because of signal intensities. Individuals may still be included or excluded as possible contributors to the mixture. All interpretable loci should be used in evaluating whether a person is included or excluded as a possible contributor. When evaluating whether a person should be excluded as a possible contributor, if an allele is not present at a locus, care must be taken to consider whether the allele may be missing due to drop-out. This determination can be difficult to make, and consultation with a more experienced analyst or supervisor may be helpful.

5.3 Minimum Number of Contributors

Generally, the minimum number of contributors to a mixed sample can be determined based on the locus that exhibits the greatest number of allelic peaks. For example, if at most five alleles are observed at a locus, then the DNA results are consistent with having arisen from at least three individuals. Proceed with caution when only one allele in the entire sample would lead to an increased number of possible contributors, as stutter and imbalance can complicate mixture interpretation. Alleles that do not meet the stochastic threshold may be used to determine the number of contributors.

If the minimum number of contributors equals the number of individuals being compared, but there is activity foreign to the individuals being compared, the sample interpretation should denote the foreign allelic activity.

Indistinguishable mixtures of four or more individuals will not be interpreted. Indistinguishable mixtures of three or more individuals will be interpreted at the analyst's discretion.

13.2 Interpretation of Forensic Parentage and Relationship Cases

For parentage cases, follows guidelines established by the AABB Relationship Testing Standards. Greater details of these standards and recommendations can be found in their published Standards for Relationship Testing Laboratories and Guidance for Standards for Relationship Testing Laboratories. These references may also be used to calculate significance of other relationship testing cases. The Popstats software (which used these AABB-recommended formulae) is used to calculate parentage statistics.

Paternity analysis can be performed with a reference sample from the child and the alleged father (AF), with or without the biological mother. However the statistical results of a "not excluded" case will be more significant if the mother is included in the testing. Therefore a reference sample from the mother should be obtained when possible. A reference from the mother is required in the following two types of cases:

a. The child's sample is fetal tissue. For fetal tissue samples, the mother must be typed in order to confirm that the tissue sample is of fetal, not maternal, origin. Failure to test the mother could result in a false exclusion. Note: if the mother is unavailable and the tissue is determined to be from a male, it may be assumed that the tissue is of fetal origin and the results may be used in the paternity analysis.

b. The alleged father is a close biological relative of the biological mother. In this situation, the child is likely to share some alleles with the AF because of the biological relationship between the AF and biological mother. Therefore testing the mother is required in order to take into account this shared relationship.

13.2.1 Definitions

1. Likelihood Ratio: a baysian statistical calculation that estimates the likelihood of seeing the evidence in question under two competing hypotheses.
2. Paternity Index (PI): a specialized likelihood ratio estimating the likelihood of seeing the child's profile if (1) the AF is the biological father versus (2) the AF is unrelated (or another competing hypothesis)
3. Combined Paternity Index (CPI): The product of all individual Paternity Indexes
4. Probability of Paternity: The probability, expressed as a percentage, that the AF is the biological father of the child. This calculation is dependent on the CPI and prior probability assumption.

5. Mutation: A change in DNA resulting from a copying error during DNA replication. In STR analysis, this can result in a non-matching allele between a biological parent and child.
6. Obligate paternal allele: the child's allele(s) at a locus that must have been inherited from the biological father. If the child is heterozygous at a locus and it cannot be determined which of the two alleles came from the biological father, then both alleles must be considered obligate.

13.2.2 Exclusion

An AF will be excluded as the biological father if in more than two loci, the AF does not share an obligate paternal allele with the child. In this case, no statistical calculations are required.

In the event of an exclusion, the analyst should evaluate the data for a possible sample switch between the biological mother and female child, or alleged father and male child. A biological relationship will be evident between the mother and female child, even if the samples were switched (for example, during collection). A biological relationship will not be apparent between the mother and male child if the samples were switched, with the exception of cases where the mother and alleged father are biologically related.

13.2.3 Not Excluded, or Included

If the AF shares an obligate paternal allele at all loci, he cannot be excluded as being the biological father of the child. In this case the PI at each locus will be calculated by Popstats using the formulas in SOP #14 (Statistics).

13.2.4 Inconclusive

A comparison may be inconclusive when at two loci the AF does not share an obligate paternal allele with the child. Additionally, a comparison may be inconclusive when at only one locus the AF does not share an obligate paternal allele with the child and a mutation is not suspected.

13.2.5 Mutations

Mutations in STR loci typically result in an allele one repeat shorter or longer than the parent allele. These mutations occur at different frequencies in different loci, and typically at a higher frequency in the larger loci. If a non-match is observed in only one or two loci, the possibility of mutational events must be evaluated. Two or more repeat mutations are seen at a lower frequency than single repeat mutations.

If a non-match is observed at only one locus, the possibility of a mutation must be assumed, and taken into account in the CPI calculation. At the locus with the suspected mutation, a mutational PI is calculated for that locus, following AABB recommendations. This mutational PI is incorporated into the CPI calculation.

If two non-matches are observed after standard laboratory STR (Identifiler) testing, the results should be considered inconclusive. Supplemental testing, such as additional autosomal STR loci or Y STR analysis, may yield additional information in these cases.

When a mutation is assumed and taken into account statistically, the DNA report must indicate which references were used. An example of wording for the DNA report is:

Note: The AABB Standards for Relationship Testing Laboratories, 10th Edition/Guidance Document was referenced for the mutation rate of locus D21S11. The AmpFlSTR® Identifiler™ PCR Amplification Kit User's Manual (2005, PN 4323291) was referenced for the probability of paternity exclusion for locus D21S11.

13.3 Y-STR Analysis and Interpretation Guidelines

13.3.1 Y-STR GeneMapper® *ID* Analysis

Loci Targeted by AmpF*l*STR® Yfiler PCR Amplification Kit

The following table shows the loci amplified by the Yfiler kit and the corresponding dyes used. The AmpF*l*STR Yfiler Kit Allelic Ladder is used to genotype the analyzed samples. The alleles contained in the allelic ladder and the genotype of the Control DNA 007 are listed in the table.

Locus Designation	Alleles Included in Yfiler Kit Allelic Ladder	Dye Label	DNA 007 Genotype
DYS456	13-18	6-FAM™	15
DYS389I	10-15		13
DYS390	18-27		24
DYS389II	24-34		29
DYS458	14-20	VIC©	17
DYS19	10-19		15
DYS385 a/b	7-25		11,14
DYS393	8-16	NED™	13
DYS391	7-13		11
DYS439	8-15		12
DYS635	20-26		24
DYS392	7-18		13
Y GATA H4	8-13	PET©	13
DYS437	13-17		15
DYS438	8-13		12
DYS448	17-24		19

The AmpF*l*STR® Yfiler® PCR Amplification Kit is a STR multiplex assay that amplifies 17 Y-STR loci in a single PCR amplification reaction. The loci amplified are in the "European minimal haplotype" (DSY19, DYS385a/b, DYS389I/II, DYS390, DYS391, DYS392, and DYS393), the Scientific Working Group-DNA Analysis Methods (SWGDAM) recommended Y-STR panel (European minimal haplotype plus DYS438 and DYS439), and additional highly polymorphic loci (DYS437, DYS448, DYS456, DYS458, DYS635, and GATA H4). DYS392 is a trinucleotide repeat, DYS438 is a pentanucleotide repeat, and DYS448 is a hexanucleotide repeat, while the remaining loci in the AmpF*l*STR® Yfiler® PCR Amplification Kit are tetranucleotide repeats. Given that the AmpF*l*STR® Yfiler® PCR Amplification Kit uses the same dye chemistry as the AmpF*l*STR® Identifiler® PCR Amplification Kit, the GeneMapper® *ID* Analysis outlined in section 13.1 will be similar to the GeneMapper® *ID* Analysis employed for Y-STR data analysis.

Allele assignment occurs through the same 3-step process:

1. Spectral separation will still employ the 4 dyes incorporated during amplification (6-FAM™, VIC®, NED™, and PET®), as well as the GeneScan™ 500 LIZ® Size Standard (PN 4322682) as the internal size standard. Please refer to section 13.1.1 for more detail.
2. Peak BP Sizing will still use GeneScan™ 500 LIZ® Size Standard (PN 4322682) as the internal size standard and the Local Southern Method for sizing of fragments of unknown length. Please see section 13.1.2 for more detail. The size standard chosen for GeneMapper® *ID* analysis should be CE_G5_HID_GS500.
3. Allele designations will still be determined through comparison of peaks of unknown length to peaks of known lengths in the manufacturer-provided allelic ladder, in the corresponding dye channel. An allelic ladder must be present and interpretable in each run for comparison. Because the Houston Police Department post-amplification room does not generally experience significant temperature fluctuations during a single plate run (which may consist of several injections), it is not necessary at this time to analyze using the average of multiple ladders. If multiple ladders were injected in a single run, only one should be designated the allelic ladder prior to GeneMapper® *ID* data analysis.

13.3.1.1 GeneMapper Procedure

1. Exceptions to the GeneMapper® *ID* analysis procedure outlined in section 13.1.1.2 include:
    a. Analysis Method: Yfiler
        i. Bin Set: AmpFlSTR_Yfiler_Binset_v2
        ii. The "use marker-specific stutter ratio if available" option should be selected
        iii. Minus A Ratio (under tetra): 0.1021 (to address the minus 2 stutter that occurs at DYS19)
        iv. Minus A Distance (under tetra): from 1.5 to 2.5 (to address the minus 2 stutter that occurs at DYS19)
        v. Minus Stutter Distances:
            1. Trimeric repeats: from 2.25 to 3.75

2. Tetrameric repeats: from 3.25 to 4.75
3. Pentameric Repeats: from 4.25 to 5.75
4. Hexameric Repeats: from 5.25 to 6.75

vi. Plus Stutter Ratio (under tri): 0.079 (to address the plus 3 stutter that occurs at DYS392)
vii. Plus Stutter Distance (under tri): from 2.25 to 3.75 (to address the plus 3 stutter that occurs at DYS392)
viii. Analysis range: Partial – the lower end of range should be adjusted to accommodate run conditions, but will generally range from about 2,400 to 10,000
ix. Sizing: Partial: 75 – 400 (450 is optional; largest peak approximately 335 bp at DYS392, so the 350 and 400 bp LIZ peaks will permit the Local Southern Size Calling Method)
x. Peak Amplitude (Analytical) Threshold: 65 RFUs (approximately 3 times the maximum observed peak height in the validation's minimum threshold study for 20 second injections (21 RFU in yellow dye channel; validation data show that the yellow and red dye channels tend to be the noisiest)).

b. Panel: Yfiler_v2

After analysis, the data should be examined for the following:

2. All of the following peaks are present, of good morphology, and of similar height in LIZ: 75, 100, 139, 150, 160, 200, floater (approximately 245-250), 300, 340, 350, and 400 (the 450 peak is optional)

3. The sample floater peaks of all samples being analyzed must be within +/-0.5 bp of the selected allelic ladder's floater peak. If not, the analyst may choose to analyze with another allelic ladder. Should a sample not fall within +/- 0.5bp of any of the allelic ladders included on this set of injections, and if the sample contains allelic activity that requires sizing, the sample would need to be re-injected. It may be acceptable, with the approval of the Technical Leader, to use data for a sample in which the floater peak is beyond +/-0.5 bp of the ladder's floater peak, if the sample does not contain any allelic activity that would require sizing (e.g. negative amplification control or reagent blank sample)

4. Correct and complete allelic designations on the allelic ladder chosen for analysis (see the AmpF*l*STR® Yfiler® PCR Amplification Kit user manual)

5. All positive control peaks are present and labeled correctly. Any extraneous peaks must be attributable to a known biology-related or technology-related artifact. Each amplification set and genetic analyzer run must have at least one acceptable positive control. See the table in 13.3.1 for the GenoType of the male positive control (007).

6. All negative controls (amplification negative control and regent blanks) are free of allelic activity that cannot be attributed to a known biology-related or technology-

related artifact. Each reagent blank must have acceptable results for data from the corresponding extraction set to be used for interpretation. The amplification negative control must have acceptable results for data from the corresponding amplification set to be used. Additionally, each analyzer run must have at least one acceptable negative control.

7. Projects should be named in the manner defined in section 13.1.1.2.4.1, or by using the LIMS generated 3100 plate number.

8. Eletropherograms should be examined for data quality and allele calls. See the discussion below (section 13.6) for guidance on evaluating data and editing artifacts and microvariant calls. Verify that the allele edit comment is selected in the alleles tab of the plot window. To change or delete the allele call, if necessary, click on the peak or allele call to be edited to select it. Right click and either choose Delete Allele or Rename Allele. Explain your edit in the comment box.

9. Upon completion of analysis, save the project. With labels to include allele peak calls, allele peak heights, allele peak sizes, and allele edit comments (if applicable), print plots for all case samples, associated controls, and ladders to be included in the case file. Place in each case folder a printout of the complete GeneMapper® *ID* project list for each run, along with the plots for allelic ladder, associated controls, and samples associated with that case. If an entire run is not used, it is not necessary to print the plots; however, the injection list must be included in the case file with an explanation as to why data was not used (e.g., no acceptable positive control). Document each time a sample is manipulated, such as longer or shorter injection times or if amplification product is diluted.

10. A second analyst must perform the GeneMapper analysis independently and agree with the allele calls reported by the first analyst on the allele chart by noting their "second read" on the injection list.

11. If LIMS was used to create the 3100 Plate map, you must export the project for entry into LIMS. After saving the project, select the GenoTypes tab and then choose HPD Table. Select File>Export Table and choose the location to which you will be exporting, such as the Removable Drive (jump drive). From the 3100 Plate Map screen in LIMS, click the Data Review button at the bottom. The Allele Call Table Worksheet will be displayed. Click on the "1st Read" button at the bottom of the screen. Browse to the location of your saved GeneMapper® *ID* file. Select your file and click Open to bring your allele calls for each of the samples in the batch into the table.

12. The analyst performing your second read will also create a GeneMapper® *ID* project and import their project back into LIMS. However, he/she will choose the "2nd Read" button in the Allele Call Table Worksheet to import his/her allele calls. Any mismatches will be highlighted in red and the run that should be imported will have to

be selected. Every effort should be made to resolve any discrepancies between the analysts. If an agreement cannot be reached, the Technical Leader must be involved.

13. If LIMS was not used to create the 3100 Plate Map prior to the genetic analyzer run, the allele calls can be manually entered into LIMS in the DNA Analysis Results Chart section of the Report Writing tab in each case's DNA assignment.

13.3.1.2 Y-STR Artifacts and Unusual Results

Technology-related artifacts that are observed in Identifiler®/Identifiler® Plus data analysis can be expected in Yfiler® data as well, given the same method of capillary electrophoresis used to obtain Identifiler®/Identifiler® Plus data is used to obtain Yfiler® data. Please see section 13.1.1.3 for more detail artifacts commonly observed in STR analysis.

Biology-related artifacts, or those that result from the PCR, include stutter products and incomplete 3' (+A) nucleotide addition. Please refer to section 13.1.1.3 for more detail. In addition to the -4 stutter often observed in tetranucleotide repeats, when performing Yfiler® data analysis, one must consider the possibly of -3 stutter of trinucleotide repeats, -5 stutter of pentanucleotide repeats, and -6 stutter of hexanucleotide repeats, though the increasing number of base pairs in a repeat unit is expected to correlate to a smaller ratio between the artifactual peak and the true peak from which it was created. In addition to minus stutter, plus stutter, or forward stutter, should be considered a possibility when "extra" peaks are observed in an otherwise single-source sample that occur in the plus stutter position (typically one repeat unit larger than the true allele). Plus stutter occurs when polymerase slippage occurs during amplification that may lead to a loop forming in the synthesized strand (as opposed to in the template strand, as is suggested for minus stutter). This loss of processivity may be attributable to critical components of PCR becoming limited or depleted in the later amplification cycles. The HPD internal Yfiler® validation data show that most plus stutter was less than 4% of the main allele peak height. Samples that seemingly exhibit forward stutter should be interpreted with caution, so as not to misinterpret an actual mixture of male DNA.

In addition to minus and plus stutter that is one full repeat unit smaller or larger than the true peak, DYS19 has demonstrated, both in the developmental validation and in the HPD internal validation, the propensity for incomplete, or partial, stutter activity in both the -2 and +2 positions. This may be attributable to the higher degree of secondary structure causing Taq to stall more frequently. This locus was probably retained in the European minimal haplotype for historical reasons, as was TH01 in the core loci used in autosomal analysis.

Stutter

Marker	Developmental Validation % Stutter	% bp stutter (plus or minus)	Highest Observed % in Internal Validation, if Greater than Developmental Validation Data	Stutter % Filters to be used at HPD
DYS456	13.21	-	-	13.21
DYS389I	11.79	-	-	11.79
DYS390	10.4	-	10.67	10.67
DYS389II	13.85	-	-	13.85
DYS458	12.2	-	12.72*	12.2
DYS19	11.4	10.21 (-2 bp)	-	11.4
DYS385a/b	13.9	-	-	13.9
DYS393	12.58	-	-	12.58
DYS391	11.62	-	14.26*	11.62
DYS439	11.18	-	-	11.18
DYS635	10.75	-	-	10.75
DYS392	16.22	7.9 (+3 bp)	-	16.22
Y GATA H4	11.08	-	13.86*	11.08
DYS437	8.59	-	-	8.59
DYS438	4.28	-	-	4.28
DYS448	4.96	-	-	4.96

*Observed in a sample for which the RFU activity far exceeded the “sweet spot” of 1000-3000 RFUs; most alleles were >8000 RFUs (Non-probative #19, 10 seconds)

13.3.1.3 Other artifacts, published or observed during internal validation:

13.3.1.3.1 Published (observed in developmental validation):

a. ~88 bp in VIC® (green channel)
b. ~80 bp in NED™ (yellow channel)
c. ~95 bp in NED™ (yellow channel)
d. ~80 bp in PET® (red channel)

13.3.1.3.2 Observed in internal validation:

a. ~95 bp in VIC® (green channel)
b. ~118 bp in VIC® (green channel)
c. ~98-100 bp in NED™ (yellow channel) (95 bp published artifact)
d. ~205 bp in PET® (red channel)
e. ~215 bp in PET® (red channel)

13.3.1.4 Microvariants

Please see section 13.1.1.4 for additional information on and naming instructions of microvariants. Given the inclusion of trinucleotide repeats, pentanucleotide repeats, and hexanucleotide repeats, along with tetranucleotide repeats in Yfiler®, one must consider the size of the repeat unit when assessing the possibility to microvariants (not just 4 bps, as indicated in section 13.1.1.4). To ensure that a microvariant is truly off-ladder and not just the result of electrophoretic drift, suspected microvariants must be re-injected for confirmation, with one exception. If the microvariant appears in multiple evidence samples from the same case otherwise having the same Y-STR profile, it is not necessary to re-inject. Because every measure should be taken to interpret evidence samples prior to reference samples, the presence of a microvariant in an associated known sample will not suffice for confirmation, as the evidence should be interpreted before and separate from any associated knowns in the case.

13.3.1.5 Y-STR Re-analysis and Additional Analysis

Please see section 13.1.1.5 for more detail. The steps taken to investigate unacceptable controls, to optimize data, and to resolve artifacts that may interfere with allelic designations when performing Identifiler®/Identifiler® Plus data analysis should also be employed accordingly with Yfiler® data analysis.

13.3.2 Y-STR Interpretation Guidelines

Please see section 13.1.2 for more detail.

The types of conclusions and types of profiles described in section 13.1.2 can be applied to Yfiler® data analysis, with the exceptions noted below.

Given Y chromosome markers are passed down from generation to generation without changing (except for mutational events), a match between an evidentiary item and a known sample only suggests that the individual could have contributed to the evidentiary sample, along with any relative from his paternal lineage.

13.3.2.1 Mixtures

Unlike with autosomal analysis, because of the haplotype (combination of allelic states of a set of polymorphic markers lying on the same DNA molecule) nature of Y-STR analysis, a sample is considered single-source when not more than one peak (as opposed to two in autosomal analysis) is observed at any of the loci, except for DYS385 a/b.

DYS385 a/b is a duplicated (or multi-copy) marker in which one primer binds at two separate locations on the Y chromosome. The duplicated regions are 40,775 bp apart and facing away from one another. If the locus is duplicated exactly, only one allele will be observed at this marker. However, it is very common to see two alleles at this locus. This makes this marker very informative, which made it an ideal selection for inclusion in the European minimal haplotype. Two alleles at DYS385 a/b but only one allele at all other loci should not be interpreted as a mixture of DNA. Duplicate repeats should be typed as a genotype (e.g. DYS385 – 11, 14; 16, 16).

Unlike DYS385 a/b which is two separate locations, DYS389 I and DYS389 II (DYS389 I/II) are actually a single region, but produce two PCR products because the forward primer binds twice. DYS389 II represents the fragment that results from the forward primer binding further away from the reverse primer. It is longer than 389 I, occurring from about 250 bp to 295 bp, in the blue channel. The fragment that is assigned to the DYS389 I locus results from the forward primer that binds nearer the reverse primer. This fragment is generally about 140 bp to 165 bp, also occurring in the blue channel.

It is possible to observe multiple peaks at loci other than DYS385 a/b in a single-source sample. This is likely the result of the entire region of the Y-chromosome being duplicated and then diverging. Peak heights are expected to be similar, and peaks are expected to be no more than one repeat unit apart. However, this is not expected to be a common event. As with any sample exhibiting more than one peak at a locus other than DYS385 a/b, extreme caution should be exercised during interpretation to ensure actual mixtures of male DNA are not mistaken for singles-source male DNA.

Like autosomal analysis, allelic balance and excessive stutter may be used for consideration when assessing the number of contributors to a DNA profile. Also like autosomal analysis, Yfiler® mixtures can be of two general types: mixtures with major/minor contributors and mixtures with indistinguishable contributors (unresolved).

13.3.2.1.1 Mixtures with major/minor contributors

The internal validation demonstrated that as the contribution of DNA in a mixture from the donors becomes more similar (e.g. approaches 1:2) it becomes more difficult to distinguish the profiles of the major and minor contributors. Conversely, as the amount of DNA contributed becomes more dissimilar (e.g. approaches 1:19), the ability to distinguish contributors is improved. Validation data indicate that at about a 1:4 ratio the individual contributors can be readily discerned from one another. Minor peak heights must be 35% or less than the height of the major peak.

13.3.2.1.2 Mixtures with indistinguishable contributors (unresolved)

Given the mode of inheritance of DNA specific to the Y chromosome, the product rule cannot be employed in statistical analysis. Y-STR data IS completely linked. For this reason, the counting method is used to provide an estimated frequency of a particular haplotype. A database is observed for the presence of the haplotype and the number of times it is observed is *counted*. At this time, there is no consensus of Y-STR mixture interpretation in the forensic community. Without the ability to provide significance or weight (statistics) to a Y-STR mixture, possible inclusions in a Y-STR mixture in which contributors cannot be distinguished will not be reported. Individuals can be excluded from an indistinguishable mixture.

13.3.2.2 Female Control DNA

The female control DNA sample provided in the AmpF*l*STR® Yfiler® PCR Amplification Kit may be used in an amplification, but it is not necessary. Both developmental and internal validations established that not only is female DNA not amplified when primers from the AmpF*l*STR® Yfiler® PCR Amplification Kit are applied, but moreover, even in the presence of excessive female DNA (100-fold excess), male DNA can still be successfully amplified, electrophoresed, and typed using the AmpF*l*STR® Yfiler® PCR Amplification Kit. Successful typing of the male control DNA (007) will demonstrate successful PCR and subsequently successful allelic labeling by the GeneMapper® *ID* software. The amplification negative control, when it fails to produce allelic activity that cannot be attributed to biology-related or technology-related artifacts, will demonstrate that amplification kit components were free of DNA prior to use on the associated samples.

As the HPD laboratory gains experience and data by using Yfiler® on casework, or should additional validation data be produced, it may be appropriate to consider modifying the guidelines put in place at this time regarding such things as analytical threshold, stochastic threshold, ratios for major/minor distinction in mixtures, etc.