Forensic Biology Unit Legacy Kit Interpretation Guidelines Manual

Forensic Biology

SAN DIEGO POLICE DEPARTMENT

June 9, 2025
Approved by: Adam Dutra, DNA Technical Manager

There may be instances where further work may be required in a case where data were originally generated using an STR typing kit (i.e., Profiler Plus, COfiler, Identifiler, Identifiler Plus, Identifiler Direct, MiniFiler, or Yfiler) and/or electrophoresis instrument (i.e., 3130 or 310) which is not currently employed by the Forensic Biology Unit. *Reinterpretation* (reevaluation of any of the allele calls or genotype calls [to include potential allelic drop-out], removal of alleles (or entire loci) from statistical estimates, or a change in the assumptions (including number of contributors) may not be necessary and/or possible in all circumstances. There are six broad categories of scenarios:

1. **The legacy data are from a reference sample to be compared to data generated with the current STR typing kit and electrophoresis instrument.**

   The legacy data may be used by an analyst authorized in the current STR typing kit and instrument for comparisons to the current STR typing kit data following the current interpretation guidelines.

2. **Comparisons of additional reference samples to previously interpreted legacy data.**

   If the interpretation of the original sample(s) included documentation of acceptable genotype combinations for comparison and *reinterpretation* of the legacy data is not required, comparisons of the data from the original interpretation(s) to additional reference samples may be performed by an analyst who was previously qualified in the STR typing kit(s) or who is currently authorized to interpret legacy data for the STR typing kit(s).

   - For samples previously analyzed using STRmix, likelihood ratio calculations/comparisons should be performed using the STRmix likelihood ratio procedures in the current Forensic Biology Technical Manual.

   - Calculations/comparisons to autosomal STR samples not previously analyzed using STRmix should be performed using the relevant sections of the Autosomal STR Interpretation Guidelines (Binary Interpretation) section of this document. Calculations/comparisons to Yfiler data should be performed using the relevant sections of the AmpF*l*STR Yfiler™ Y-STR Interpretation Guidelines section of this document.

- Although genotype combinations may not have been explicitly documented for samples previously described as single source or as having a single predominant contributor, these descriptions implied the genotype combinations at the detected loci and can be treated as having the genotype combinations documented.

3. **STRmix reinterpretation of legacy data.**

   Legacy data from some combinations of autosomal STR kits, electrophoresis instruments, and injection times can be reinterpreted using STRmix. The legacy data typically will need to be reanalyzed in GeneMapper prior to STRmix deconvolution. Legacy data from the following combinations can be reinterpreted using STRmix:

   - MiniFiler data from the 3500 Genetic Analyzer
   - Identifiler data from 310 Genetic Analyzer (up to 5s injections)
   - Identifiler data from 3130 Genetic Analyzer (up to 5s injections)
   - Identifiler Plus data from 3130 Genetic Analyzer (up to 5s injections)

   SDPD has two STRmix kits for reinterpretation of legacy data: SDPD Identifiler Plus and SDPD MiniFiler. The SDPD Identifiler Plus kit can be used in the reinterpretation of Identifiler or Identifiler Plus data collected under the conditions described above. STRmix reinterpretation of legacy data from the combinations of STR kits and genetic analyzers described above may only be performed by analysts authorized to perform legacy data reinterpretation for the applicable combinations of kits and genetic analyzers.

4. **Legacy data from DNA mixtures amplified with Profiler Plus and/or COfiler kits.**

   The original interpretations of these data did not incorporate an explicit stochastic threshold. Stochastic thresholds for these kits were developed in 2017. Reinterpretation of Profiler Plus and/or COfiler data can be performed using the applicable stochastic threshold and the Autosomal STR Interpretation Guidelines (Binary Interpretation) section of this document.

5. **MiniFiler data collected on the 3130 Genetic Analyzer and Identifiler and/or Identifiler Plus data from samples with injection times greater than 5 seconds.**

   Reinterpretation of these data can be performed using the Autosomal STR Interpretation Guidelines (Binary Interpretation) section of this document.

6. **Legacy samples for which the electronic and printed copies of the electrophoresis data cannot be located.**

   Interpretation of samples requires electrophoresis data. These samples cannot be reinterpreted.

Forensic Biology

Method: **GeneMapper ID-X™ Reanalysis of MiniFiler Data (from 3500 genetic Analyzers)**

Date: 06/01/2006 Revision Date: 05/28/2024 Approved by: ARD

INTRODUCTION:

GeneMapper™ ID-X software is an automated genotyping software package that will genotype samples based on allelic bin definitions stored within the Panel Manager. Custom SDPD Analysis Methods have also been created in the GeneMapper Manager for use in re-analyzing raw data files for MiniFiler amplifications. An analysis method defines the analysis parameters to be used for the analysis of samples imported into the Project Window.

MATERIALS:

Applied Biosystems GeneMapper™ ID-X Software

PROCEDURE:

IMPORTANT: STRmix™ analysis requires the use of different analysis methods to analyze evidence and reference samples. To have multiple analysis methods within one project, each requires at least one ladder associated with it. To accomplish this, the .hid files of the ladders can be duplicated within the run folder and imported into GeneMapper ID-X if two ladders were not run with the original data.

1. Open the GeneMapper™ ID-X software by clicking on the icon located on the computer's desktop and log-in. Type in or select your user ID from the drop-down menu as the User Name. The password is case sensitive. Upon successful log-in the **Project Window** will open. The project window is the primary window for the user interface.

2. To add samples to the Project Window. Select **Edit>Add Samples to Project** or click the  icon.
   For samples run on the 3130:
   - Navigate to Forensic Biology on FB server H:\3130-310 Run Folders to locate the desired run folder.

- Select the desired run folder and click **Add To List**. The selected folder will appear in the **Samples To Add** box on the right. Note: Multiple run folders can be added to the same project. Click **Add** to import the files into the project and close the dialog box.

3. For each sample within the **Project Window** select the appropriate **Sample Type**, **Analysis Method**, **Panel**, **Size Standard** (see below). Unknowns should be designated with sample type "Sample", Allelic Ladders should be designated with sample type "Allelic Ladder", the positive PCR control should be designated with sample type "Positive Control", and any reagent blanks, amplification blanks, or formamide blanks should be designated with sample type "Negative Control".

Kit	Analysis Method	Panel	Size Standard
MiniFiler	MF_reference	MiniFiler_reference 2017	GS600_LIZ_(80-400)
MiniFiler	STRmix_MF_evidence	MiniFiler_evidence 2017	GS600_LIZ_(80-400)

Note: These MiniFiler panels are newly created panels and should only be used when reanalysis of the original data is to be re-done. In general, analysis parameters will not change with the exception of the analysis range which will be dependent on the electrophoresis of each run.

4. Click ► (Analyze), and the Save project dialog box opens. Save the project with a uniquely identifying name [e.g. initials(Case #)]. After analysis, the project is automatically saved, and the Genotypes Tab becomes available. Analyzing the data sizes and genotypes the samples in the project. There must be at least one ladder in each run folder analyzed for the analysis to be successful.

5. Accurate application of the size standard to the samples can be verified in a variety of ways. One method uses the Process Component Quality Values (PQVs) **Sizing Quality** flag as an efficient method of verifying the sizing of individual samples. In the Samples tab of the Project Window the sizing quality of each sample is displayed as a symbol (■ = Pass, ▲ = Check sizing for problems, or ● = Low Quality, sizing failed).

6. The individual sample files can be reviewed and printed for the analytical record. If needed, labels can be edited within the samples. To edit labels, select the peak to be edited, right-click and select the applicable option of changing the artifact or allele label. An Edit Allele Comment dialog box will

open and a reason for the change can be input. Sample plots with edited labels should be printed for the analytical record with the edit displayed in the plot. Labels may be changed but must not be deleted.

7. PQVs can be used to assess the quality of a variety of aspects of each sample. For a full list of the PQVs please use the GeneMapper™ ID-X help function within the software.

8. When the analysis of the GeneMapper ID-X project is completed the project file should be exported to the appropriate folder on the Forensic Biology network. To export the project: highlight the project in the **Projects** tab in **GeneMapper Manager** and click the **Export** button. Browse for the appropriate folder on the network (**Forensic Biology on FB server (H:)>GeneMapper ID Projects>Analysts folder**) and save the exported project using the Project name. Use of analysts initials at the start of the file name will be useful in locating projects as the number of files increases and is highly recommended [e.g. initials(XX-XXXXXX;etc)].

9. The original project should be deleted from the GeneMapper ID-X database once the project is exported. To delete the original project: highlight the project in the **Projects** tab in **GeneMapper Manager** and click the **Delete** button. Prior to deleting the project, it is recommended to check that the file was exported to the appropriate folder.

10. To view exported GeneMapper ID-X projects. Open **GeneMapper Manager** and from the **Projects** tab click **Import**. Browse to **Forensic Biology on FB server (H:) > GeneMapper ID Projects** and select the desired project. Click **OK**. The project will have been imported to the GeneMapper ID-X database and can be opened as usual. When the imported project is no longer needed it should be deleted from the database (see step 9).

REFERENCES:
AmpF*l*STR MiniFiler™ PCR Amplification Kit User's Manual
GeneMapper™ ID-X Software version 1.4 User Manual.
GeneMapper™ ID-X Software Tutorial.

Forensic Biology

Method: **GeneMapper ID-X™ Reanalysis of Identifiler™ (310 or 3130 Genetic Analyzers) or Identifiler™ Plus Data (3130 Genetic Analyzers)**

Date: 06/01/2006 Revision Date: 05/28/2024 Approved by: ARD

INTRODUCTION:

GeneMapper™ ID-X software is an automated genotyping software package that will genotype samples based on allelic bin definitions stored within the Panel Manager. The AmpFlSTR panels and bins are pre-installed in the Panel Manger within the software. Custom SDPD Analysis Methods have also been created in the GeneMapper Manager for use in re-analyzing raw data files for Identifiler™ and Identifiler™ Plus amplifications for the purpose of STRmix™ analysis. An analysis method defines the analysis parameters to be used for the analysis of samples imported into the Project Window.

MATERIALS:

Applied Biosystems GeneMapper™ ID-X Software

PROCEDURE:

IMPORTANT: STRmix™ analysis requires the use of different analysis methods to analyze evidence and reference samples. To have multiple analysis methods within one project, each requires at least one ladder associated with it. To accomplish this, the .fsa files of the ladders can be duplicated within the run folder and imported into GeneMapper ID-X if two ladders were not run with the original data.

1. Open the GeneMapper™ ID-X software by clicking on the icon located on the computer's desktop and log-in. Type in or select your user ID from the drop-down menu as the User Name. The password is case sensitive. Upon successful log-in the **Project Window** will open. The project window is the primary window for the user interface.

2. To add samples to the Project Window. Select **Edit>Add Samples to Project** or click the  icon.
   For samples run on the 310 or 3130:

- Navigate to Forensic Biology on FB server G:\3130-310 Run Folders to locate the desired run folder.
    - If the desired run folder is from prior to 2006, the original .fsa files may be stored on a CD-ROM stored in the laboratory supply room. Once located, navigate to the appropriate folder on the CD-ROM.
- Select the desired run folder and click **Add To List**. The selected folder will appear in the **Samples To Add** box on the right. Note: Multiple run folders can be added to the same project. Click **Add** to import the files into the project and close the dialog box.

3. For each sample within the **Project Window** select the appropriate **Sample Type**, **Analysis Method**, **Panel**, **Size Standard** (see below). Unknowns should be designated with sample type "Sample", Allelic Ladders should be designated Controls with sample type "Allelic Ladder", the positive PCR control should be designated with sample type "Positive Control", and any reagent blanks, amplification blanks, or formamide blanks should be designated with sample type "Negative Control".

Kit	Analysis Method	Panel	Size Standard
Identifiler	ID-X Identifiler	Identifiler_v1.2X	CE_G5_HID_GS500
Identifiler	STRmix_ID_evidence	Identifiler_v1.2X	CE_G5_HID_GS500
Identifiler Plus	ID-X Identifiler Plus	Identifiler_Plus_Panels_v1X	CE_G5_Identifiler_Plus_GS500
Identifiler Plus	STRmix_IDP_evidence	Identifiler_Plus_Panels_v1X	CE_G5_Identifiler_Plus_GS500

Note: These STRmix evidence panels are newly created panels and should only be used when reanalysis of the original data is to be re-done. In general, analysis parameters will not change with the exception of the analysis range which will be dependent on the electrophoresis of each run.

4. Click ► (Analyze), and the Save project dialog box opens. Save the project with a uniquely identifying name [e.g. initials(Case #)]. After analysis, the project is automatically saved, and the Genotypes Tab becomes available. Analyzing the data sizes and genotypes the samples in the project. There must be at least one ladder in each run folder analyzed for the analysis to be successful.

5. Accurate application of the size standard to the samples can be verified in a variety of ways. One method uses the Process Component Quality Values (PQVs) **Sizing Quality** flag as an efficient method of verifying the sizing of individual samples. In the Samples tab of the Project Window the sizing quality of each sample is displayed as a symbol (■ = Pass, ▲ = Check sizing

for problems, or ● = Low Quality, sizing failed).

6. The individual sample files can be reviewed and printed for the analytical record. If needed, labels can be edited within the samples. To edit labels, select the peak to be edited, right-click and select the applicable option of changing the artifact or allele label. An Edit Allele Comment dialog box will open and a reason for the change can be input. Sample plots with edited labels should be printed for the analytical record with the edit displayed in the plot. Labels may be changed but must not be deleted.

7. PQVs can be used to assess the quality of a variety of aspects of each sample. For a full list of the PQVs please use the GeneMapper™ ID-X help function within the software.

8. When the analysis of the GeneMapper™ ID-X project is completed the project file should be exported to the appropriate folder on the Forensic Biology network. To export the project: highlight the project in the **Projects** tab in **GeneMapper Manager** and click the **Export** button. Browse for the appropriate folder on the network (**Forensic Biology on FB server (H:)>GeneMapper ID Projects>Analysts folder**) and save the exported project using the Project name. Use of analysts initials at the start of the file name will be useful in locating projects as the number of files increases and is highly recommended [e.g. initials(XX-XXXXXX;etc)].

11. The original project should be deleted from the GeneMapper™ ID-X database once the project is exported. To delete the original project: highlight the project in the **Projects** tab in **GeneMapper Manager** and click the **Delete** button. Prior to deleting the project, it is recommended to check that the file was exported to the appropriate folder.

12. To view exported GeneMapper™ ID-X projects. Open **GeneMapper Manager** and from the **Projects** tab click **Import**. Browse to **Forensic Biology on FB server (H:) > GeneMapper ID Projects** and select the desired project. Click **OK**. The project will have been imported to the GeneMapper™ ID-X database and can be opened as usual. When the imported project is no longer needed it should be deleted from the database (see step 9).

<u>REFERENCES</u>:

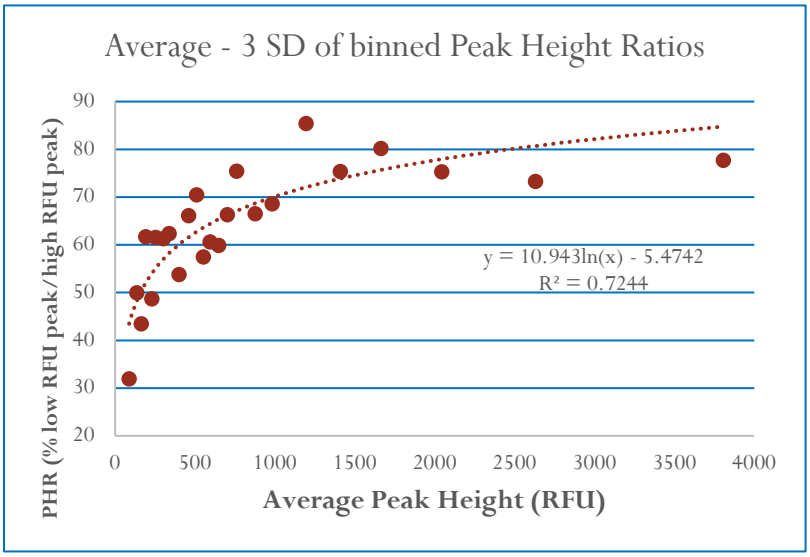
AmpF*l*STR Identifiler™ PCR Amplification Kit User's Manual
AmpF*l*STR Identifiler™ Plus PCR Amplification Kit User's Manual

GeneMapper™ ID-X Software version v1.4 User Manual.
GeneMapper™ ID-X v1.4 Software Tutorial.

Method: **AmpF*l*STR Profiler Plus Reinterpretation Thresholds**

Date: 10/10/2017 Revision Date: 06/02/2025 Approved by: ARD

- Peak detection thresholds of 50rfu or 75rfu were used for analysis of AmpF*l*STR Profiler Plus. Peaks below the analytical threshold are not reported, but can be used qualitatively to determine the possible presence of a mixture.
- Homozygous allele peaks are often approximately twice the height of heterozygous allele peaks as a direct result of allele dosage.
- At heterozygous loci, the ratio of the height of the shorter peak to the height of the higher peak is generally above 60% (see Figure below), however, peak height ratios tend towards larger imbalances as the height of the tallest peak decreases toward the stochastic threshold.



- n-1 repeat stutter peaks will be filtered by the software when they are below the percent cutoff value for a locus. The stutter cutoff values used by the software are presented below:

Marker	Stutter Cut-off Value
D3S1358	12.36%
vWA	12.36%
FGA	12.36%
D8S1179	13.64%
D21S11	14.95%
D18S51	19.05%
D5S818	11.11%
D13S317	11.11%
D7S820	9.89%

- n+1 repeat stutter was not observed in this kit due to the different reaction mix formulation than those of more advanced kits.

- New stutter expectation charts have been created for assistance in reinterpreting Profiler Plus data. The new stutter files are located on the Forensic Biology network in the following location:

  H:\QA-QC files\Validations, etc\Validations\Validations for former and legacy methods\Profiler Plus and Cofiler Reinterpretation Data

- Validation efforts have established the stochastic thresholds (or homozygote peak height threshold) above which both alleles of a heterozygote pair will most likely be detected. When evaluating an evidence profile, the following table details the stochastic thresholds to be used when evaluating data obtained using the Profiler Plus PCR amplification kit.

Analytical Threshold	Stochastic Threshold
50 RFU	150 RFU
75 RFU	235 RFU

Probability of Identity using the Applied Biosystems Profiler Plus kit.

The Probability of Identity (P_I) [Sensabaugh 1982] is the probability that two individuals selected at random will have an identical Profiler Plus kit genotype. For the Profiler Plus kit the P_I are as follows:

1 in 96 trillion in the U.S. Caucasian population,
1 in 68 trillion in the African American population,

The P_I expresses the average 9 locus DNA profile frequencies using the Applied Biosystems Profiler Plus kit. This data was taken from the Identifiler user's manual page 13-9.

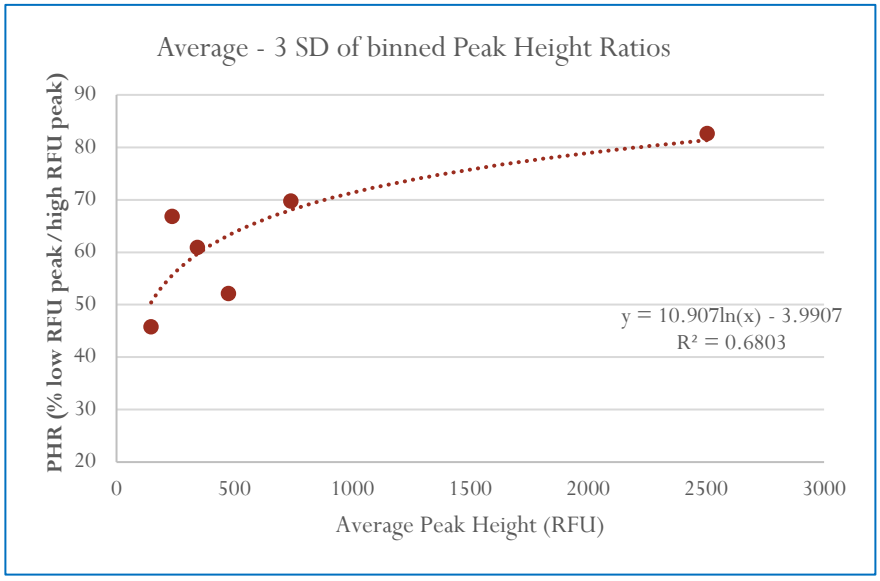
REFERENCES:

SDPD Profiler Plus Validation studies
AmpF*l*STR® Profiler™ Plus User's Manual

Method: **AmpF*l*STR COfiler Reinterpretation Thresholds**

Date: 10/10/2017 Revision Date: 06/02/2025 Approved by: ARD

- Peak detection thresholds of 50rfu or 75rfu were used for analysis of AmpF*l*STR COfiler. Peaks below the analytical threshold are not reported, but can be used qualitatively to determine the possible presence of a mixture.
- Homozygous allele peaks are often approximately twice the height of heterozygous allele peaks as a direct result of allele dosage.
- At heterozygous loci, the ratio of the height of the shorter peak to the height of the higher peak is generally above 60% (see Figure below), however, peak height ratios tend towards larger imbalances as the height of the tallest peak decreases toward the stochastic threshold.



- n-1 repeat stutter peaks will be filtered by the software when they are below the percent cutoff value for a locus. The stutter cutoff values used by the software are presented below:

Marker	Stutter Cut-off Value
D3S1358	12.36%
D16S539	14.95%
TH01	6.38%
TPOX	6.38%
CSF1PO	9.89%
D7S820	9.89%

- n+1 repeat stutter was not observed in this kit due to the different reaction mix formulation than those of more advanced kits.

- New stutter expectation charts have been created for assistance in reinterpreting COfiler data. The new stutter files are located on the Forensic Biology network in the following location:

  H:\QA-QC files\Validations, etc\Validations\Validations for former and legacy methods\Profiler Plus and Cofiler Reinterpretation Data

- Validation efforts have established the stochastic thresholds (or homozygote peak height threshold) above which both alleles of a heterozygote pair will most likely be detected. When evaluating an evidence profile, the following table details the stochastic thresholds to be used when evaluating data obtained using the COfiler PCR amplification kit.

Analytical Threshold	Stochastic Threshold
50 RFU	150 RFU
75 RFU	175 RFU

Probability of Identity using the Applied Biosystems Profiler Plus kit.

The Probability of Identity (P_I) [Sensabaugh 1982] is the probability that two individuals selected at random will have an identical Profiler Plus kit genotype. For the Profiler Plus kit the P_I are as follows:

1 in 480 thousand in the U.S. Caucasian population,
1 in 3.2 million in the African American population,

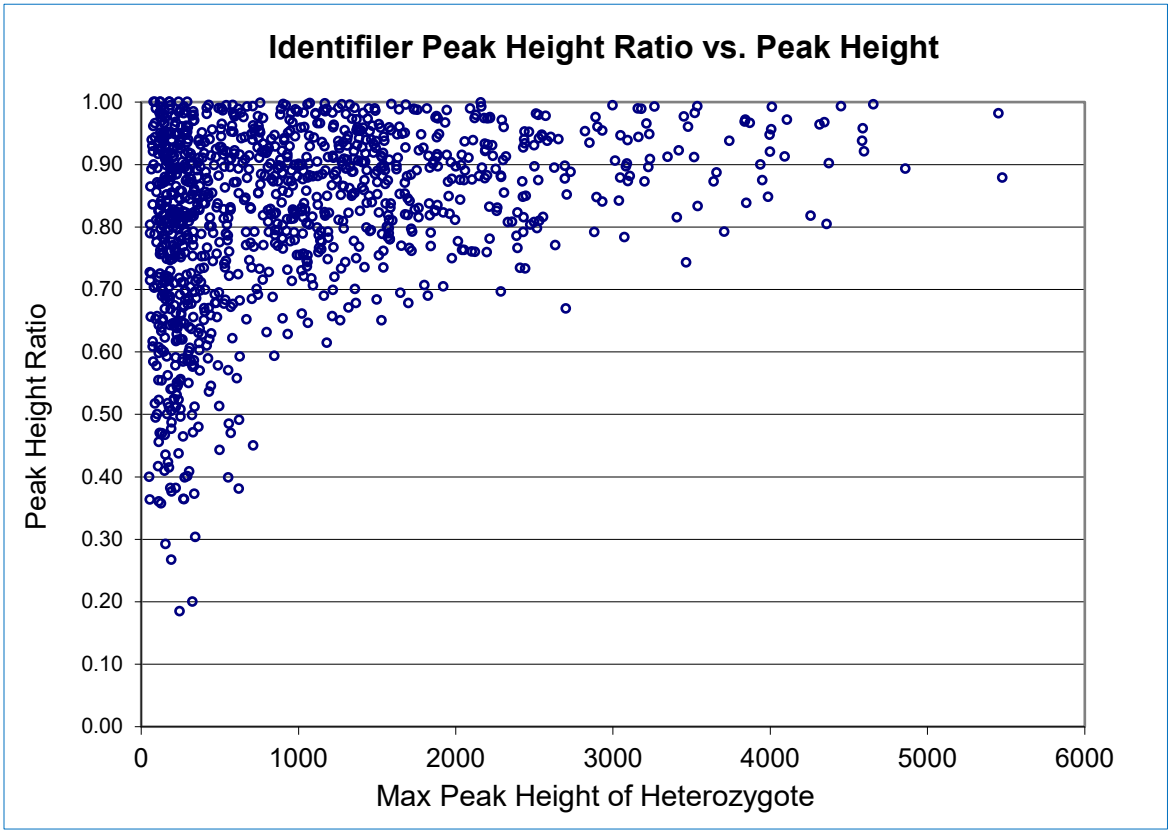
The P_I expresses the average 6 locus DNA profile frequencies using the Applied Biosystems COfiler kit. This data was taken from the COfiler user's manual page 4-7.

REFERENCES:
SDPD COfiler Validation studies
AmpF*l*STR® COfiler™ User's Manual

Method: **AmpF*l*STR Identifiler Reinterpretation Thresholds**

Date: 10/10/2017 Revision Date: 06/02/2025 Approved by: ARD

- Peak detection thresholds of 50rfu or 75rfu were used for analysis of AmpF*l*STR Identifiler. Peaks below the analytical threshold are not reported, but can be used qualitatively to determine the possible presence of a mixture.
- Homozygous allele peaks are often approximately twice the height of heterozygous allele peaks as a direct result of allele dosage.
- At heterozygous loci, the ratio of the height of the shorter peak to the height of the higher peak is generally above 60% (see Figure below), however, peak height ratios tend towards larger imbalances as the height of the tallest peak decreases toward the stochastic threshold.



- n-1 repeat stutter peaks were filtered by the software when they were below the percent cutoff value for a locus. The stutter cutoff values used by the software are presented below:

Marker	Kazam Stutter Filter	GMID Stutter Filter
D8S1179	8.93%	8.2%
D21S11	10.37%	9.4%
D7S820	8.93%	8.2%
CSF1PO	10.13%	9.2%
D3S1358	11.98%	10.7%
TH01	5.37%	5.1%
D13S317	8.70%	8.0%
D16S539	11.60%	10.4%
D2S1338	12.48%	11.1%
D19S433	15.34%	13.3%
vWA	14.41%	12.6%
TPOX	5.04%	4.8%
D18S51	20.49%	17%
D5S818	7.29%	6.8%
FGA	17.24%	14.7%

- n+1 repeat stutter was not observed in this kit due to the different reaction mix formulation than those of more advanced kits.

- New stutter expectation charts have been created for assistance in reinterpreting Identifiler data. The new stutter files are located on the Forensic Biology network in the following location:

  H:\QA-QC files\Validations, etc\Validations\Validations for former and legacy methods\Identifiler validation\Identifiler\Identifiler Kit Stutter Graphs 2017.xls

- Validation efforts have established the stochastic thresholds (or homozygote peak height threshold) above which both alleles of a heterozygote pair will most likely be detected. When evaluating an evidence profile, the following table details the stochastic thresholds to be used when evaluating data obtained using the Identifiler PCR amplification kit. STRmix parameters have not been established for evidence profiles where increased injections times were used. If re-interpretation of samples with increased injection times are required, binary interpretation should be performed with consideration that increased injection times will lead to higher stochastic

thresholds than those listed below.

Analytical Threshold	Stochastic Threshold
50 RFU	200 RFU
75 RFU	250 RFU

Probability of Identity using the Applied Biosystems Identifiler kit.

The Probability of Identity (P_I) [Sensabaugh 1982] is the probability that two individuals selected at random will have an identical Identifiler kit genotype. For the Identifiler kit the P_I are as follows:

1 in 200 quadrillion in the U.S. Caucasian population,
1 in 760 quadrillion in the African American population,
1 in 130 quadrillion in the Hispanic population.

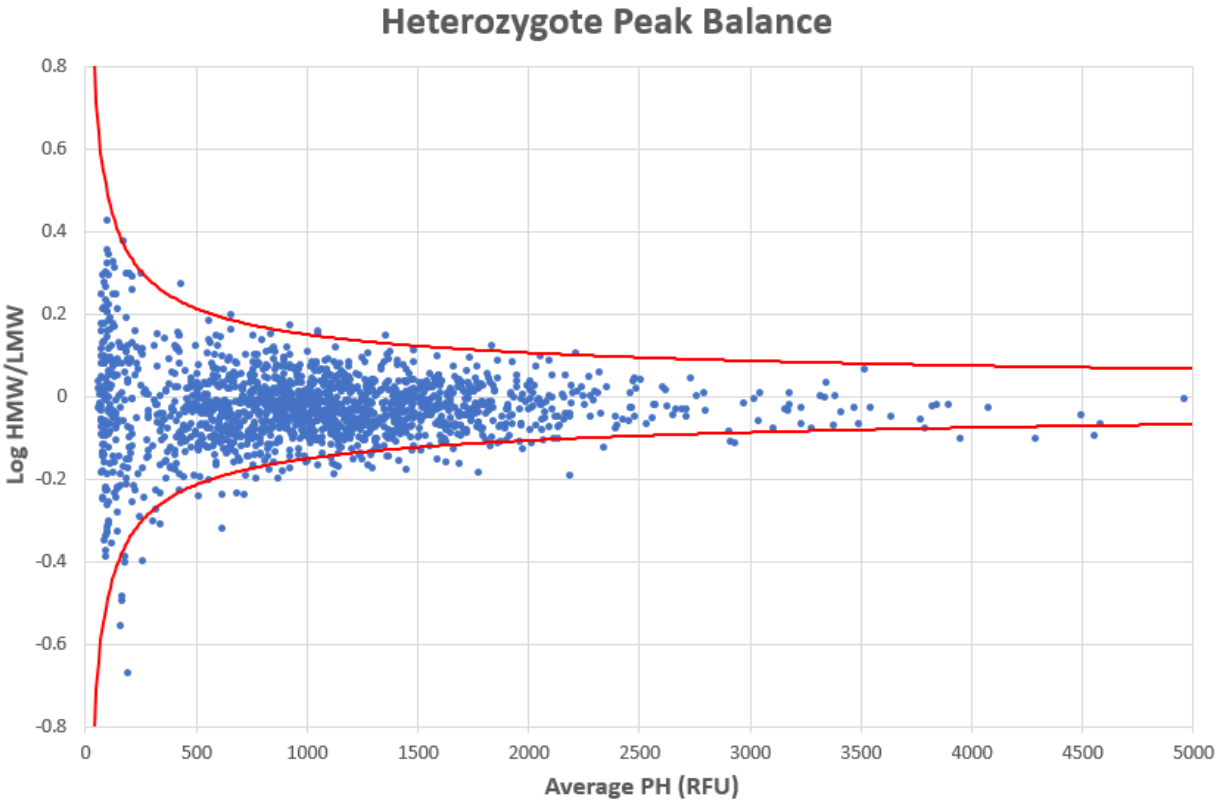
The P_I expresses the average 15 locus DNA profile frequencies using the Applied Biosystems Identifiler kit. This data was taken from the Identifiler user's manual page 4-55.

REFERENCES:
SDPD Identifiler Validation studies
AmpF*l*STR® Identifiler™ User's Manual

Method: **AmpF*l*STR Identifiler™ Plus Interpretation Thresholds**

Date: 04/01/2011 Revision Date: 06/02/2025 Approved by: ARD

- A peak detection threshold of 50rfu will be used for analysis of AmpF*l*STR Identifiler™ Plus data. Peaks below 50rfu are not reported, but can be used qualitatively to determine the possible presence of a mixture.
- Homozygous allele peaks are often approximately twice the height of heterozygous allele peaks as a direct result of allele dosage.
- At heterozygous loci, the ratio of the height of the shorter peak to the height of the higher peak is generally above 60% (see Table below), however, peak height ratios tend towards larger imbalances as the height of the tallest peak decreases toward the stochastic threshold.



- n-4bp stutter peaks will be filtered by the software when they are below the percent cutoff value for a locus. The stutter cutoff values used by the software are presented below:

Marker	Stutter Cut-off Value
D8S1179	10.32%
D21S11	10.67%
D7S820	9.69%
CSF1PO	9.2%
D3S1358	12.27%
TH01	4.08%
D13S317	9.93%
D16S539	10.39%
D2S1338	12.44%
D19S433	11.21%
vWA	12.45%
TPOX	6.38%
D18S51	13.68%
D5S818	10.06%
FGA	13.03%

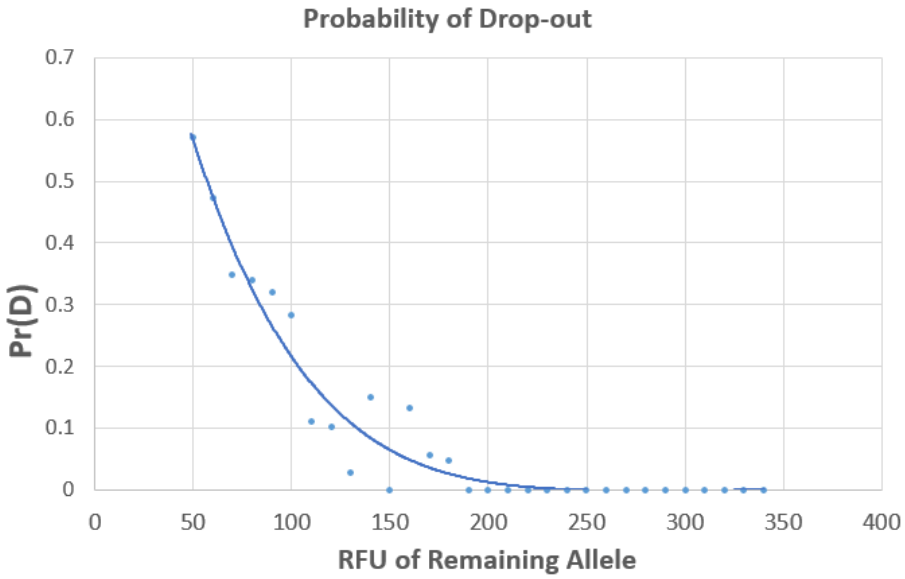
For any reinterpretation of Identifiler™ Plus data, detailed stutter information can be obtained by consulting the stutter data tables created in conjunction with the validation stutter study conducted. The electronic files are located in the following location on the network:

H:\QA-QC files\Validations, etc\Validations\Validations for former and legacy methods\Identifiler Plus\ID+ STRmix_Legacy Interpretation

- n+4nt Stutter does occur more frequently than in past AB kits due to the different reaction mix formulation. On average, n+4 stutter products are less than 2% of the parent peak, however n+4 stutter peaks as high as 3% have been observed at several loci. n+4nt stutter expectations are included in the stutter data tables referenced above.

- Validation efforts have established the stochastic thresholds (or homozygote peak height threshold) above which both alleles of a heterozygote pair will most likely be detected. Validation data have been used by STRmix™ to establish a probability of drop-out for alleles of various peak heights. This probability of drop-out is used in lieu of a traditional stochastic threshold (or homozygote peak height threshold). Analysts should be aware of what a theoretical stochastic threshold may be when evaluating MiniFiler DNA profiles. As such,

the following graphical representation (histogram) is an approximation of the STRmix™ probability of drop-out graph for the ≤5s injection time.



- STRmix parameters have not been established for evidence profiles where increased injections times were used. If re-interpretation of samples with increased injection times are required, binary interpretation should be performed, and the following stochastic threshold values should be taken into account when interpreting DNA peaks, as allelic drop-out may have occurred if the peaks heights are less than the values below.

7 second injection	280rfu
10 second injection	400rfu

Probability of Identity using the Applied Biosystems Identifiler™ Plus kit.

The Probability of Identity (P_I) [Sensabaugh 1982] is the probability that two individuals selected at random will have an identical Identifiler kit genotype. For the Identifiler™ Plus kit the P_I are as follows:

1 in 200 quadrillion in the U.S. Caucasian population,
1 in 760 quadrillion in the African American population,
1 in 130 quadrillion in the Hispanic population.

The P_I expresses the average 15 locus DNA profile frequencies using the Applied Biosystems Identifiler™ Plus kit. This data was taken from the Identifiler user's manual page 4-55.

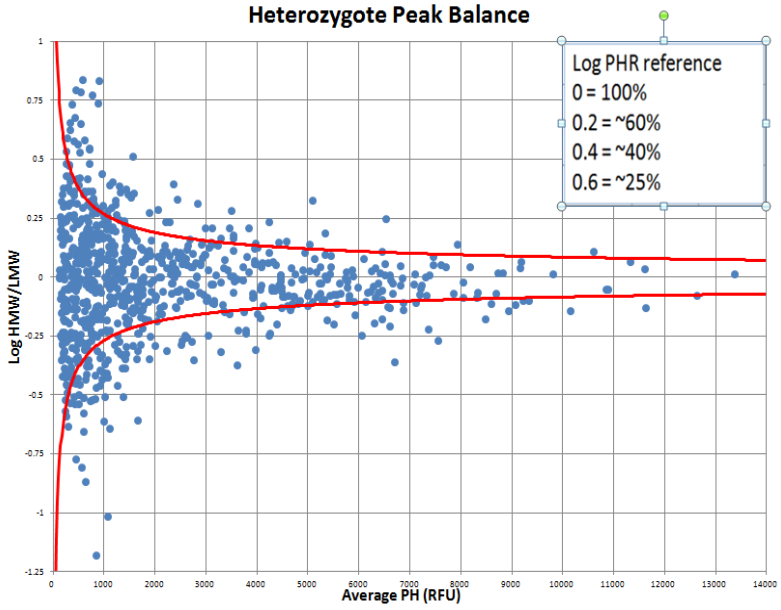
REFERENCES:

SDPD Identifiler™ Plus Validation studies
AmpF*l*STR® Identifiler™ Plus User's Manual

Method: **AmpF*l*STR MiniFiler™ Interpretation Thresholds**

Date: 07/25/2007 Revision Date: 06/02/2025 Approved by: ARD

- Analytical Threshold (Prior to October 2015 for data collected on a 3130 Genetic Analyzer): a peak detection threshold of 75rfu was used for analysis of MiniFiler™ data. Peaks below 75rfu were not reported but could be used qualitatively to determine the possible presence of a mixture or to determine the number of contributors.
- Analytical Threshold (Between October 2015 and September 28, 2017 for data collected on a 3500 Genetic Analyzer): a peak detection threshold of 100rfu was used for analysis of MiniFiler™ data. Peaks below 100rfu could be used qualitatively to determine the possible presence of a mixture or to determine the number of contributors.
- Homozygous allele peaks are approximately twice the height of heterozygous allele peaks as a direct result of allele dosage.
- Peak Height Ratios: At heterozygous loci, the ratio of the height of the lower rfu peak to the height of the higher rfu peak is generally above 50% (see Figure below where stutter ratios are presented in log format to also show the relationship between molecular weight), however, peak height ratios tend towards larger imbalances as the height of the peaks decreases toward the analytical threshold. Red lines are 95% boundaries.



- Stutter Threshold: n-1 repeat stutter peaks will be filtered by the GMID-X software for reference samples when they are below the percent cutoff value for a locus. For data analyzed prior to October 2015, the following were the stutter thresholds employed in GMID-X:

Marker	n-1 Stutter Threshold
D13S317	14.0%
D7S820	11.0%
D2S1338	18.0%
D21S11	16.0%
D16S539	15.0%
D18S51	18.0%
CSF1PO	14.0%
FGA	15.0%

For samples analyzed between October 2015 and September 28, 2017, evidence samples were analyzed in GMID-X without filtering n-1 repeat stutter, however; n+1 repeat stutter still required editing. For any re-interpretations of MiniFiler™ data obtained from the 3500, re-analysis with updated stutter filters is possible (see the GeneMapper™ ID-X Re-Analysis of MiniFiler™ Data (from 3500 genetic Analyzers) procedure in this manual). Re-analysis with the new panels will filter n+1 stutter. With the following thresholds:

Marker	n+1 Stutter Threshold
D13S317	5.2%
D7S820	n/a
D2S1338	4.9%
D21S11	5.7%
D16S539	2.9%
D18S51	6.4%
CSF1PO	6.5%
FGA	4.9%

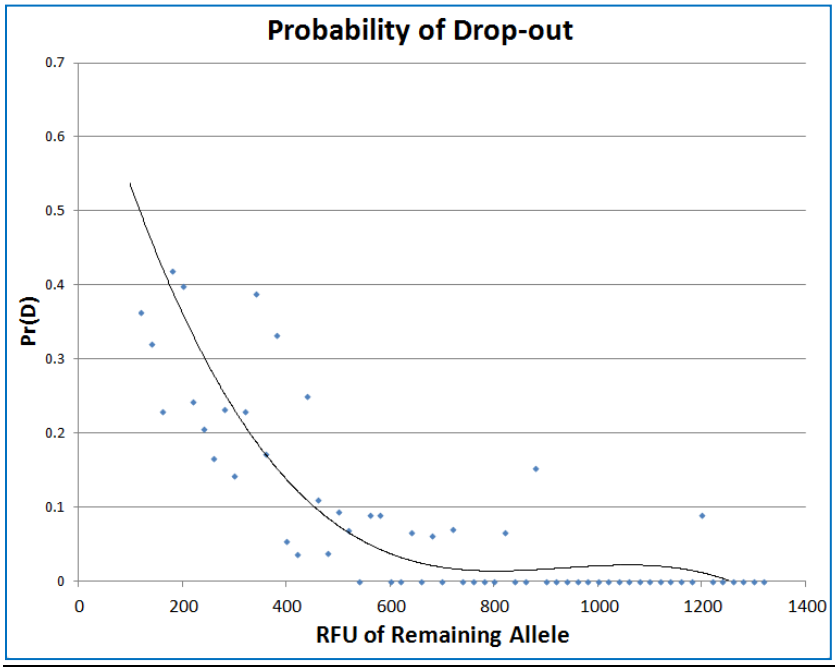
For any reinterpretation of MiniFiler™ data, updated and more detailed stutter information can be obtained by consulting the stutter data tables created in conjunction with the supplemental stutter validation conducted in 2017. The electronic files are located in the following location on the network:

H:\QA-QC files\Validations, etc\Validations\Validations for former and legacy methods\3500 MiniFiler YFiler\Minifiler\Supp Stutter 2017

Please see the autosomal interpretation guidelines for more information regarding stutter.

- Stochastic Threshold (prior to October 2015 for data collected on a 3130 Genetic Analyzer): Validation indicated that both alleles of a heterozygote pair will be detected when one of the alleles reaches at least 450rfu (>99% of the time, however allelic drop-out was observed when the larger peak was approximately 680rfu). When evaluating an apparent single source evidence profile, care should be taken when interpreting a single DNA peak below 450rfu as allelic drop-out may have occurred.

- Stochastic Threshold (between October 2015 and September 28, 2017 for data collected on a 3500 Genetic Analyzer): Validation data have been used by STRmix™ to establish a probability of drop-out for alleles of various peak heights. This probability of drop-out is used in lieu of a traditional stochastic threshold (or homozygote peak height threshold). Analysts should be aware of what a theoretical stochastic threshold may be when evaluating MiniFiler™ DNA profiles. As such, the following graphical representation (histogram) is an approximation of the STRmix™ probability of drop-out graph.



Probability of Identity using the Applied Biosystems MiniFiler™ kit.

The Probability of Identity (P_I) [Sensabaugh 1982] is the probability that two individuals selected at random will have an identical MiniFiler™ kit genotype. For the MiniFiler™ kit the P_I are as follows:

1 in 12 billion in the U.S. Caucasian population,
1 in 15 billion in the African American population,
1 in 4.8 billion in the Hispanic population.

The P_I expresses the average 8 locus DNA profile frequencies using the Applied Biosystems MiniFiler™ kit. This data was taken from the MiniFiler™ user's manual page 5-46.

REFERENCES:

SDPD MiniFiler™ Validation studies
AmpF*l*STR® MiniFiler™ User's Manual

Method: **AmpF*l*STR® Identifiler Direct® Interpretation Thresholds**

Date: 04/01/2012 **Approved by: ARD**

- A peak detection threshold of 50rfu will be used for analysis of AmpF*l*STR® Identifiler Direct® data. Peaks below 50rfu are not reported.
- Homozygous allele peaks are often approximately twice the height of heterozygous allele peaks as a direct result of allele dosage.
- At heterozygous loci, the ratio of the height of the shorter peak to the height of the higher peak is generally above 70% (see Peak Height validation study), however, peak height ratios tend towards larger imbalances as the height of the tallest peak decreases toward the stochastic threshold.
- A global filter of 20.00% will be applied to data as opposed to marker specific stutter percentages.
- Validation efforts have established the stochastic thresholds (or homozygote peak height threshold) above which both alleles of a heterozygote pair will most likely be detected.

≤5 second injection	150rfu
7 second injection	210rfu
10 second injection	300rfu

<u>Probability of Identity using the Applied Biosystems Identifiler Direct® kit.</u>

The Probability of Identity (P_I) [Sensabaugh 1982] is the probability that two individuals selected at random will have an identical Identifiler Direct® kit genotype. For the Identifiler Direct® kit the P_I are as follows:

1 in 200 quadrillion in the U.S. Caucasian population,
1 in 760 quadrillion in the African American population,
1 in 130 quadrillion in the Hispanic population.

The P_I expresses the average 15 locus DNA profile frequencies using the Applied Biosystems Identifiler Direct® kit. This data was taken from the Identifiler Direct® user's manual page 105.

REFERENCES:

SDPD Identifiler Direct Validation studies
AmpF*l*STR® Identifiler Direct® User's Manual

Forensic Biology

METHOD: **STRmix™ Re-analysis Settings for MiniFiler™ Data (3500 Genetic Analyzers), Identifiler™ Data (310 Genetic Analyzer), or Identifiler™ Plus Data (310 or 3130 Genetic Analyzers)**

Date: 10/10/2017 Revision Date: 06/02/2025 Approved by: ARD

Reanalysis of Identifiler™ or Identifiler™ Plus data from samples with four or fewer contributors is possible with STRmix. STRmix reanalysis of samples amplified with the MiniFiler™ amplification kit is limited to those with data collected on the 3500 Genetic Analyzer.

To perform STRmix analysis for any of the compatible data listed in the paragraph above, follow the procedures for STRmix analysis and interpretation in the current version of the Forensic Biology Technical Manual.

If the original .fsa or .hid files cannot be located, it may be possible to create a .txt file for STRmix import using the data from the GeneMapper ID project. If a .txt file is created outside of GeneMapper ID-X, all potential allelic and stutter peaks (n-1 and n+1 repeat) are required for STRmix analysis. Artifact peaks (e.g., pull-up and spikes) must be omitted from the .txt file as STRmix is unable to process mixtures that contain artifact peaks.

In addition, staff elimination, contaminant, and random profile database files are not maintained for the legacy kits. In order to perform the quality assurance check of the staff database file, one should be created from the latest version of the GlobalFiler elimination/QC database. This can be accomplished by opening the database .txt file in Excel and deleting the markers that are *not* contained in the legacy kit being analyzed, then reorganizing the remaining loci into the same order as the kit.

The following are STRmix settings and parameters that should be used in the STRmix analysis:

For MiniFiler:

Kit Name	SDPD MiniFiler
Size regression file	MiniFiler_Size Regression.csv
Analytical Threshold	100
Allele Variance	6.677, 3.921
Locus Amplification Variance	0.01996
Back Stutter Variance	2.639,10.041
Maximum Back Stutter Ratio	0.32
Inversely Proportional To	Observed Height of Parent Allele
Stutter Regression File	SDPD_MF_Stutter.txt
Stutter Exceptions File	SDPD_MF_StutterExceptions.csv
Minimum Variance Factor	0.5
Variance Minimization Parameter	1,000
Drop-In Cap	612
Drop-In Frequency	0.006 (uniform)
Maximum Degradation	0.01
Degradation Start Point	Use Smallest Peak
Saturation Threshold	30,000

For Identifiler™ and Identifiler™ Plus:

Kit Name	SDPD Identifiler Plus
Size regression file	Identifiler_Size Regression.csv
Analytical Threshold	50
Allele Variance	4.195, 0.923
Locus Amplification Variance	0.013
Back Stutter Variance	2.740, 1.854
Maximum Back Stutter Ratio	0.3
Inversely Proportional To	Observed Height of Parent Allele
Back Stutter Regression File	SDPD_IdentStutter.txt
Back Stutter Exceptions File	SDPD_ID_stutter exceptions.csv
Forward Stutter Variance	4.616, 1.522
Maximum Forward Stutter Ratio	0.07
Inversely Proportional To	Expected Height of Stutter Peak
Forward Stutter Regression File	SDPD_Ident_FwdStutter.txt
Minimum Variance Factor	0.5
Variance Minimization Parameter	1,000
Drop-In Cap	100
Drop-In Frequency	0.000365 (uniform)
Maximum Degradation	0.01
Degradation Start Point	Use Smallest Peak
Saturation Threshold	10,000

The following are the settings for the NIST population databases:

STRmix uses the same population data for each kit. The allele frequencies used for the populations are the NIST 1036 allele frequencies. There is also one additional population created for use in VarNOC deconvolutions and database searches. For each of the populations, the Population proportion = 1 and Unrelated Proportion = 1, with no details about population size, number of children, or family relationships.

Below lists each population, the allele frequency file name, and the default FST. The allele frequency file has GF in the name to indicate that it contains all the markers in GlobalFiler, but can be used for other kits with fewer markers, as well. July 2017 indicates that the allele frequencies incorporate the corrections issued in the 2017 release of data.

NIST_SDPD_AfricanAmerican = NIST1036_GF_AfAm_July2017.csv, F_{ST} = 0.01b(1.0,1.0)
NIST_SDPD_Asian = NIST1036_GF_Asian_July2017.csv, F_{ST} = 0.02b(1.0,1.0)
NIST_SDPD_Caucasian = NIST1036_GF_Cauc_July2017.csv, F_{ST} = 0.01b(1.0,1.0)
NIST_SDPD_Hispanic = NIST1036_GF_Hisp_July2017.csv, F_{ST} = 0.01b(1.0,1.0)
NIST_SDPD_All = NIST1036_GF_All_July2017.csv, F_{ST} = 0.01b(1.0,1.0)

Forensic Biology

METHOD: **Autosomal STR Interpretation Guidelines (Binary Interpretation)**

Date: 06/01/2004 Revision Date: 04/20/2020 Approved by: ARD

The following guidelines are meant to be used for legacy kits for which probabilistic genotyping parameters have not been established. These include the Profiler Plus, COfiler, MiniFiler amplification kits (specifically MiniFiler data collected on the 3130 Genetic Analyzers), and increased injection times for Identifiler and Identifiler Plus data.

Introduction

The Scientific Working Group on DNA Analysis Methods (SWGDAM) states that:

> "The interpretation of DNA typing results for human identification purposes requires professional judgment and expertise. Additionally, laboratories that analyze DNA samples for forensic casework purposes are required by the Quality Assurance Standards for Forensic DNA Testing Laboratories (effective July 1, 2009) to establish and follow documented procedures for the interpretation of DNA typing results and reporting. Due to the multiplicity of forensic sample types and the potential complexity of DNA typing results, it is impractical and infeasible to cover every aspect of DNA interpretation by a preset rule. However, the laboratory should utilize written procedures for interpretation of analytical results with the understanding that specificity in the standard operating protocols will enable greater consistency and accuracy among analysts within a laboratory. It is recommended that standard operating procedures for the interpretation of DNA typing results be sufficiently detailed that other forensic DNA analysts can review, understand in full, and assess the laboratory's policies and practices. The laboratory's interpretation guidelines should be based upon validation studies, scientific literature, and experience."

Here we present a number of general interpretation guidelines for interpreting samples in forensic DNA casework. The rationale behind these guidelines is to establish a set of standards to ensure that conclusions are supported by reliable

data, derived using a scientific approach, and that interpretation of results is as objective as possible and is consistent from one analyst to another.

The interpretation of DNA results should take into account the DNA results obtained, the nature of the sample, the condition of the profile, any forensically valid assumptions made, and should be scientifically justifiable and based on the collected experience and knowledge of the laboratory and the scientific community. These guidelines are designed to provide a general foundation for interpretation.

Please refer to the complete Autosomal STR Interpretation Guidelines document if more information is desired on any section within this summary document. See Appendix A for a glossary of terms.

DNA interpretation should take place in the following order:

Preliminary Evaluation of Data and Allele Designation (section 1)
Interpretation of DNA Typing Results (section 2)
Comparison of DNA Typing Results (section 3)
Statistical Analysis of DNA Typing Results (section 4)
Reporting of DNA Typing Results and Conclusions (section 5)

Section 1 – Preliminary Evaluation of Data and Allele Designation

1.1 Detection of peaks

The detection of peaks in the data is performed by the GeneMapper ID-X software (Profiler Plus, COfiler, and some Identifiler data may have had peak detection and analysis performed in GeneScan and GenoTyper software packages). The analytical thresholds (peak detection thresholds) for the autosomal STR kits are listed within documents specific to each individual kit.

1.2 Evaluation of internal standards

The internal size standard should be evaluated to ensure that for each sample it has produced the expected pattern of peaks. These peaks should generally be between 1000 and 2000rfu for a 5 second injection (310 and 3130 data) and not show signs of poor resolution. If a sample has issues in the size standard that prevent proper sizing, or has signs of poor

resolution that affect sizing or genotyping, the sample will not be interpreted.

1.3 In general, internal size standards that use "Local Southern" sizing should contain at least two size standard peaks above and below the size of the smallest and largest peaks within the sample. For MiniFiler analysis, the use of 3rd order least squares sizing method requires only a single size standard peaks above and below the size of the smallest and largest peaks within the sample.

1.4 Evaluation of allelic ladders

All allelic ladders used in genotyping samples must be evaluated for proper genotyping. The GeneMapper ID-X software compares the ladder injections in the run with the expected sizes. Analysis of at least one of the injected allelic ladder must produce the known allele peaks associated with the ladder in order for the software to type the other samples of the run. If a ladder injection passes the evaluation by the GeneMapper ID-X software, the ladder can be relied upon without further manual evaluation or reinjection.

1.5 Designation of alleles

Allele assignments are made by comparing the base pair size of the peaks in the amplified samples with the base pair size of the alleles within the allelic ladder(s). Allele assignments are performed automatically by the GenoTyper or GeneMapper ID-X software. Analysts may on occasion need to manually edit "off-ladder" peaks to designate them as true alleles, or to re-designate peaks originally called as alleles by the software as artifacts.

1.5.1 Occasionally an ***allele*** will be detected that is ***outside the ladder range*** for a given locus. Alleles that fall above or below the ladder alleles (even those in virtual bins) will be designated as greater than (>) or less than (<) the respective ladder allele.

Occasionally an ***allele*** will fall ***outside the ladder range between two loci***. An analyst should consider the size of the apparent allele and the expected zygosity in an attempt to determine to which locus the allele belongs. A list of variant alleles is maintained by NIST on the

STRBase website (https://strbase.nist.gov). This list can be consulted to assist in assist in determining to which locus the allele belongs. The allele will then be renamed with the appropriate (>) or (<) allele designation for the appropriate locus. Supporting documentation should be included in the administrative documents for the case.

1.5.2 Off ladder alleles that occur within the ladder ranges (i.e. ***microvariants*** that contain an incomplete repeat motif) should be designated by the number of complete repeats and, separated by a decimal point, the number of base pairs in the incomplete repeat (e.g. D21S11 29.1 allele). ***Microvariants*** will be confirmed through a re-injection of the sample as a means of demonstrating the reproducibility of the retention time and sizing of the peak. If a microvariant is observed in multiple samples from the same case, re-injection for confirmation purposes is unnecessary. ***Tri-allelic patterns*** should have been confirmed through reamplification. If re-amplification is not possible, the locus containing the suspected tri-allelic pattern will be declared inconclusive.

1.6 Designation of artifacts

Some data contained within the electropherograms may not represent actual alleles that originate in the sample. Non-allelic peaks such as stutter, non-template dependent nucleotide addition (minus-A), nonspecific amplification products, spikes, raised baseline, pull-up, or disassociated primer dyes should be identified prior to the comparisons of reference samples to the data. Generally, non-allelic data such as stutter, non-template dependent nucleotide addition, disassociated dye, and incomplete spectral separation are reproducible; spikes and raised baseline are generally non-reproducible.

1.6.1 ***Off –scale data*** is often the result of over-amplification and can cause several other interpretational difficulties, such as: “minus-A” products, increased stutter, and pull-up (see below). If off-scale data is present in a sample, the sample may be re-injected with a decreased injection time, reamplified using less input DNA, or the amplified product may be diluted for reinjection.

1.6.2 ***Pull-up*** results from poor spectral separation and is characterized by signal from a peak in one fluorescent dye color producing a smaller, artifactual peak in another, often adjacent, dye color. Amplification of less sample, injection of a sample for less time, dilution of amplified product, or running the samples after creating a new spectral for the instrument can reduce pull-up peaks.

1.6.3 ***Minus-A*** (-A) is a form of PCR product that does not possess an extra nucleotide at the 3' end. Over-amplification can result in prominent "minus-A" products. Re-amplifying using less input DNA will likely reduce minus-A peaks. Incubating amplified products at 60°C for an additional time period may also reduce minus A.

1.6.4 ***Stutter*** peaks are minor peaks appearing one repeat unit smaller (e.g. n-4 in tetranucleotide loci) or larger (n+4 in tetranucleotide loci) than a primary STR allele. If a called peak is suspected to be stutter, the validation study should be consulted to determine the expected stutter percentage for that allele. If the observed percentage is within the expected stutter percentage range for that allele, but above the filter threshold, an analyst may determine the peak to be stutter. If a filtered allele falls outside 3 standard deviations from the mean for a peak, the analyst should interpret that filtered peak as an actual allele for interpretation purposes. For alleles that are at the larger base pair range of a locus, if the suspected stutter peak is above the stutter filter, but within three standard deviations of the mean for that allele based on the validation an analyst may determine the peak to be stutter. Additional factors such as spectral overlap or the presence of a potential stutter peak between two true alleles may enhance stutter peaks above the expected values. In reference samples, unfiltered stutter peaks that are less than three standard deviations from the mean observed for an allele during the validation study can be interpreted as stutter peaks, provided that there are no additional indications that the sample is a mixture. An examination of the baseline may provide additional indication that the reference sample is single source. For rarer alleles that do not have enough data to establish a mean or meaningful standard deviation, the value of the putative stutter peak must be shown to be below the extrapolated trend line for three standard deviations to be

considered as stutter. Care should be taken in assessing these instances as tri-allelic patterns, although rare, are known to occur.

1.6.5 ***Spikes*** are artifacts of electrophoresis, which may resemble DNA peaks. Often spikes will be narrower than true DNA peaks and will be filtered by the GeneMapper ID-X software. Multi-color spikes can be identified by the presence of peaks occurring in several colors at the same data point location. Single-color spikes can be identified by the presence of a sharp peak occurring in a single color in the raw data, with no spectral overlap into the other colors. Since spikes are generally random, reinjection of a sample should eliminate a questioned spike.

1.7 Evaluation of analytical controls

Controls must be assessed prior to interpretation of any associated samples. If the analytical controls are properly designated, the GeneMapper ID-X software will flag controls that fail to meet expectations. Controls that have passed GeneMapper ID-X verification can be relied upon without further evaluation.

1.7.1 The ***reagent blank*** is a check for possible contamination of the sample preparation reagents by extraneous DNA. If DNA types are detected in the reagent blank this is an indication of a possible contamination event or that extraneous DNA was introduced into the sample from some unintended source. Based on an evaluation of the DNA types in the reagent blank, the results of associated samples may be deemed inconclusive. The analyst and the DNA Technical Manager will confer on an appropriate action based on the particular circumstance.

Samples extracted in a batch where the reagent blank shows DNA types greater than the interpretation threshold may need to be re-extracted. Not all contamination events will necessitate reanalysis and will be judged on a case-by-case basis.

1.7.2 The ***negative amplification control*** (Amplification Blank, NAC, or No DNA Control) is a check for possible contamination of samples during the amplification process. The appearance of DNA types in this control indicates that problems similar to those described for the reagent blank may have occurred during the amplification.

The interpretation of detectable DNA in the amplification blank should take into account the results obtained from any samples and reagent blanks amplified in the same batch. A batch where both the amplification blank and the reagent blank give the same typing result, or where only the amplification blank gives a result, may be caused by a problem limited to the amplification process. In this case, re-amplification of the samples may resolve the problem. A batch that shows a typing result for the reagent blank, but not for the amplification blank is more likely to be caused by an extraction-related problem.

If the reagent blank and amplification blank give different typing results, both the extraction and amplification processes are potentially suspect and should be repeated.

1.7.3 The ***positive amplification control*** (PAC) is a check for effectiveness of the amplification and the Short Tandem Repeat (STR) analytical processes. This control should always show interpretable results consistent with the known genotype published in the DNA kit literature. No other interpretable peaks should be present. Peaks that are determined to be artifacts (elevated stutter, pull-up, etc) are not considered interpretable peaks.

Results from samples associated with a positive amplification control for which these criteria are NOT met must be jointly evaluated by the DNA analyst and the DNA Technical Manager (or his/her designee). On a case-by-case basis, the DNA Technical Manager will evaluate the results obtained and determine the appropriate interpretation, taking the observed control results into account. If the observed control results bring the reliability of a sample result into question (as determined by the DNA Technical Manager), the sample result will not be interpreted or reported. The evaluation of the control results and the DNA Technical Manager's approval (or his/her designee's) of the process must be documented in the analyst's case notes and, in an Quality Incident Summary (see FB policy manual section 1.8). If the results from samples associated with the failed positive control are deemed usable, the failure of the positive control must be included in the final report.

If the cause of the positive amplification control cannot be determined, samples associated with the failed positive amplification control should be reamplified. Samples associated with the failed positive control that were consumed for the original amplification would require re-examination (re-extraction) of the original evidence. If no evidence remains, then the samples will be reported as inconclusive.

If the cause of the positive amplification control cannot be determined, the only time the results from samples associated with a failed positive amplification control that were consumed for the original amplification will be reported is when they form the basis of an exclusion. In this instance, troubleshooting of the problem must lead to the determination that the amplification for that sample is reliable and full disclosure of the unexpected result (failed positive amplification control) as well as the exclusion will be documented in the final report.

1.8 Evaluation of samples that have been reinjected, reamplified, and/or reextracted

1.8.1 Analysts should evaluate ***multiple data sets for an item amplified with the same autosomal STR kit*** to determine if there are indications of differences in peak proportions between the injections/amplifications for the sample. Analysts must indicate in the table of results if the represented result is a composite of multiple amplifications/or injections (see section 5). When generating a ***composite DNA profile***, the analyst must use the highest RFU value of a reproducible peak when determinations of contribution level are done. If significant differences exist between the data generated from reamplification of the same DNA extract the analyst may choose to consider the data as separate results instead of generating a composite profile.

1.8.2 Analysts should evaluate data for an ***item amplified with multiple autosomal STR kits*** to verify that no unexplainable discordances exist between loci common to the kits. Possible reasons for discordances include; but are not limited to; degradation or inhibition affecting one of the amplifications to a greater degree than the other, variation in primer sequence, stochastic effects of low level samples, and the

slight differences in peak proportions typically encountered during multiple amplifications of mixed samples.

Section 2 – Re-interpretation of DNA Typing Results

In general, interpretation of DNA samples will be completed prior to comparison of reference samples. All interpretation of samples will be performed using the actual electrophoresis data and not the table of results. DNA interpretation will include the following steps:

Determine whether a sample is a mixture (section 2.1)
Estimate the number of contributors to a sample (section 2.2)
Determine whether stochastic effects are likely (section 2.3)
Identify whether a sample is degraded and/or inhibited (section 2.4)
Determine contribution level in mixed DNA samples (section 2.5)
Document relevant assumptions (section 2.6)
Determine which loci and combinations of DNA types are useful for statistical support of inclusions (section 2.7).

These steps will be documented in the analytical record on an interpretation worksheet prior to comparison to reference samples.

2.1 Determination of whether a sample is a mixture

A mixed DNA profile possesses certain common attributes that will tend to indicate the presence of more than one contributor to the observed results. The attributes may include:

- Greater than two alleles present at a locus, unattributable to a mutation or a tri-allelic pattern.
- Imbalance in peak height between peaks at a locus greater than expected based on validation data for the specific kit (see the interpretation thresholds specific kit).
- Generally, evidence of a mixed DNA sample exists at multiple loci.

To determine whether a mixed DNA profile exists, the DNA profile must be evaluated in its entirety. Possible peaks observed below the analytical threshold may be used in the determination of the number of contributors. The conclusion in the Forensic Biology report should state that the minimum number or assumed number of contributors for any evidence

sample, and may be stated independently of the conclusion(s) surrounding potential individual contributors (see section 5).

2.2 Estimation of the number of contributors to a sample

Conclusions with respect to the minimum number of contributors must be included in a report. Generally, an estimate of the minimum number of contributors is based on the locus that exhibits the greatest number of allelic peaks and the fact that each contributor would contribute at most two alleles to the locus (although tri-allelic patterns could be present, but are extremely rare). Additional information such as expected peak height ratios can be used to determine minimum number of contributors. While the minimum number of contributors to a mixture can be established with a high probability, an absolute determination of the number of contributors cannot be ascertained. Assumptions as to the number of contributors can be made if the data appears consistent with a given number of contributors (see section 2.6)

2.3 Determination of whether stochastic effects are likely

Validation data has been used to determined stochastic thresholds to be used with each specific kit for the injection parameters used. Please refer to the specific kit interpretation thresholds sections for the stochastic thresholds to be used for interpretations.

If a table of results is to be included with the Forensic Biology Unit report, any DNA types below the stochastic threshold will be designated with an asterisk (e.g. 16*) (see section 5).

2.4 Determination of whether inhibition and/or degradation is present

2.4.1 ***Inhibition*** of the PCR process can be caused by a variety of substances and is generally described as a reduction in efficiency of the PCR process. Examples of the effects of inhibition can be observed in the SDPD Inhibition presentations on the FB network in the presentations folder.

Evaluation of the IPC in the quantification assay(s), if applicable, can be helpful to determine whether inhibitors may be present. Because a larger amount of DNA may be amplified than used for

quantitation, a sample with an IPC result within expectations may still show evidence of inhibition when amplified.

2.4.2 DNA is generally stable if stored properly; however, ***degradation*** of DNA can occur over time, through exposure to the elements, or through the action of bacteria. Degradation can be identified by a classic pattern in samples where rather robust results are obtained at the smaller (base pair) loci while little to no results are obtained at the larger (base pair) loci. The overall results have been described as being in a ski slope pattern.

In general, inhibition will affect all contributors to a DNA mixture equally and could raise the concern of actually detecting all alleles that should be present in the mixture. Degradation, however, may not affect all contributors equally (***differential degradation***) and mixtures containing degraded DNA can pose a challenge to the analyst attempting to interpret them. Analysts are encouraged to exercise caution when interpreting a mixture that is thought to contain, or is known to contain, degraded DNA.

2.5 Mixture Terminology in Binary Interpretations.

2.5.1 In general, a ***major DNA contributor*** is an individual who appears to contribute a significant portion of the DNA to a mixture. A DNA contributor should be contributing more DNA than at least one of the other contributors to the mixture to be considered a major contributor. Major contributors to a DNA mixture do not possess DNA types that are designated as minor alleles unless a significant amount of stacking of alleles is likely.

2.5.2 In general, a ***minor DNA contributor*** is an individual who appears to contribute a small portion of the DNA to a mixture. Minor contributors will possess multiple DNA types that are designated as minor alleles. Alleles from possible minor DNA contributors may be masked by alleles from major contributors. If no alleles from minor contributors are identified in the mixed DNA sample, then the mixture should be considered unresolvable in terms of determining major and minor DNA contributors.

2.5.3 When a DNA profile presents as a mixture of three or more individuals, with at least one major DNA contributor, at least one minor DNA contributor, and a contributor at a level somewhere between the major and minor contributor(s), that contributor can be termed a ***mid-level contributor*** or ***second significant contributor***. Additional terms such as "significant minor" may be used to further clarify assumptions about the contribution levels in the mixture provided they are clearly defined in the case notes (see section 2.9.1.4 c). The interpretation of the presence of a significant second contributor to a mixture should be based on a discernable difference in peak intensities between the second significant contributor and any other minor alleles present in the mixture. Alleles designated as likely originating from the significant second contributor should be documented in the analytical record.

2.6 Determination of contribution levels in mixed DNA samples obtained with Profiler Plus, COfiler, Identifiler, and MiniFiler (from the 3130 Genetic Analyzer) amplification kits.

The validations studies conducted for the Profiler Plus, COfiler, Identifiler, or MiniFiler kits were not comprehensive enough to support the use of the Identifiler Plus amplification kit partitioning model (see section 2.7).

As such, analysts seeking to re-interpret mixtures obtained when using kits other than Identifiler Plus should rely heavily on observations regarding the number of contributors to a sample, any assumptions regarding the number of contributors or reasonably expected contributors, observed mixture ratios, expected peak height and stutter ratios when interpreting mixed DNA samples. The thresholds for each kit are detailed in the interpretations thresholds documents for each kit.

To assist in the interpretation of mixed DNA samples of samples tested with the Profiler Plus, COfiler, Identifiler, or MiniFiler kits an attempt to make an assumption as to the number of contributors is recommended. If no such assumption can be made as to the total number of DNA contributors in the sample, analysts may make an assumption as to the number of major contributors (if possible) and an interpretation of the major contributors predicated on that assumption may be performed without an interpretation as to minor DNA contributors.

2.7 Determination of contribution levels in mixed DNA samples obtained with the Identifiler Plus amplification kit.

Validation studies conducted on 2-, 3-, and 4-person mixtures indicates that the partitioning model presented below can be useful in the interpretation of complex mixtures. This model will not apply to all mixtures and consideration should be taken in applying the partitioning model, especially for complex mixtures outside the scope of the validation (i.e. mixtures of 5- or more people). In mixtures known to contain at least 5– or more people, for instance, if the loci with the most DNA types detected contain data indicating that any DNA type could pair with any other at the marker (based on validated peak height ratio expectations), there may not be any contributors contributing more DNA than another. Not applying the partitioning model would be the best approach to interpreting mixtures in these instances.

The individual peak heights of alleles within a mixture can provide information regarding the relative contribution level (major/minor) to the overall mixture for people possessing those types. This assessment begins by determining whether there appears to be multiple significant contributors (section 2.7.1). Once balance has been evaluated, the mechanism for identifying DNA types more likely from minor DNA contributors can be determined. The determination of the appropriate method for assigning DNA types likely from minor contributors is critical in the interpretation of complex mixtures.

The two methods for assigning minor DNA types in a mixture are by using 1/3 the tallest peak at a marker or using a threshold of <10% the total rfu detected at a locus (these are explained further in sections 2.7.1.1 and 2.7.1.2). Using a threshold of ⅓ the tallest peak at a marker will generally provide a subset of DNA types that includes the genotype of the single strongest DNA contributor in a mixture and is used when there is an indication that the mixture is imbalanced in contributions levels between the contributors. Using the threshold of <10% the total rfu detected a marker will provide the subset of DNA types that include the genotypes for the two or three strongest contributors to the mixture.

DNA types identified as more likely from minor DNA contributors will be designated by brackets (e.g. [16]) in the table of results (see section 5).

2.7.1 In general, mixtures having multiple major contributors (a balanced mixture) in a mixed DNA profile can be determined by examining the results using the following criteria:

> One locus with at least 5 alleles greater in height than ⅓ the tallest peak at that locus (five including the tallest peak) and/or three loci with four alleles within 50% of the tallest allele at those loci (four including the tallest peak).

If the above indicators are not met, but the overall information in the profile indicates that the situation of at least two significant contributors is likely, analysts should adopt the <10% the total rfu at a marker for the designation of minor contributor alleles. Some two person mixtures may not have 3 loci with 4 DNA types and therefore peak proportions at the other loci should be evaluated for indications of balance. Loci with two alleles that have close to 1:1 or 3:1 ratio, or three alleles with some combination close to a 2:1:1 ratio is indicative of balance between significant contributors.

Additional factors or phenomena such as inhibition or degradation within the sample may lead an analyst to adopt a more conservative interpretation and statistical assessment of any potential inclusions. Analysts should be able to articulate the basis for their interpretation, such as low level mixtures that may have peaks below the detection threshold that suggest balance in the mixture.

2.7.1.1 Minor contributor types in *mixtures with a **single major contributor** indicated*

> In mixtures with a single major DNA contributor indicated (imbalanced mixtures) a DNA type is considered to be from a minor contributor if it is less than ⅓ the intensity of the tallest peak at a marker.
>
> Validation data has shown in mixtures that have varying contribution levels from the DNA donors, individuals that generally contribute greater than 50% of the DNA in a sample will not possess any alleles designated as “minor” in the mixture. In these situations, “minor” alleles can be used

as an interpretational tool for determining major vs. minor contributors to a mixture.

In some circumstances the imbalance in a sample will be so extreme, that the genotype from a single major contributor (predominant) see 2.9.1.4a can clearly be deduced.

2.7.1.2 Minor contributor types in *mixtures with* ***multiple major contributors***

In mixtures with multiple major contributors (balanced mixtures), minor contributor alleles will generally be less than 10% of the total RFU at a locus. 10% is meant as a guideline and phenomena such as the overall intensity of the types in the mixture, stutter and baseline artifacts may affect the percentage from locus to locus. If the percentage of an allele is such that it is suggestive that it may be from a more significant contributor it should not be considered minor. When the overall peak heights in a mixture are low, it is more likely to have stochastic issues affect expected peak height ratios of the contributors, as such, 7% will be used as the threshold for loci with less than 1600 total rfu in mixtures with multiple major contributors (balanced mixtures).

2.7.2 When identifying DNA types in a mixture that are more likely from minor contributors, the robustness of the DNA profile should be taken into consideration in evaluating and interpreting a DNA profile. As results approach the stochastic range (e.g. 200 RFU for Identifiler Plus), peak height ratios of less than 30% may be observed. Based on the increased imbalance expected with lower levels of DNA, loci where the tallest peak is less than or equal to 350rfu should not be used to determine major/minor contributor DNA types. In some mixtures where the majority of the types detected are stochastic level and only a few DNA types are above 350rfu, analysts may determine that assigning major vs minor alleles should not be done due to the low level of the data.

2.8 Documenting Assumptions

In certain circumstances, ***assumptions regarding DNA profiles*** may be made to assist the analyst in drawing conclusions and to put the evidence in the proper context. All assumptions regarding any interpretation will be presented in the report and should also be in the notes (see section 5).

2.8.1 While the absolute ***number of contributors*** to a DNA mixture is difficult to determine, analysts may make ***assumptions as to the number of contributors*** detected in samples. The assumptions as to the number of contributors should take into account phenomena such as allelic dropout, degradation, differential degradation, inhibition, and possible masking of alleles by stutter filters. The assumption as to the number of contributors must be supported by the data.

2.8.2 In certain circumstances ***assumptions regarding the presence of a contributor*** in a DNA mixture is permissible (i.e. intimate samples, or reasonably expected contributors). In these circumstances the DNA profile from the assumed contributor(s) can be used to assist interpretation of the additional contributor(s) DNA profile/types (section 2.9).

The assumption of the presence of one or more contributors may require the analyst to change the minimum number of contributors to a mixture. It also may be possible that the assumption of the presence of a contributor may change the confidence with which an assumed number of contributors can be made.

2.9 Determining the Loci to be used for Statistics/Comparisons

Prior to making comparisons, the profile should be evaluated to ***determine what loci and/or combinations of DNA types are acceptable to support a conclusion statistically***. Much of the evaluation to determine what parts of the DNA results can be used for statistical weight takes into account the presence of DNA types below the stochastic threshold and whether assumptions regarding the number of contributors have been made. If all DNA types are above the stochastic threshold, all DNA types can be used for comparison, but it may useful to further refine the genotypes for comparison if a limited number of combinations are likely.

If an analyst believes that ***filtered stutter peaks*** should be evaluated for comparison, the mean stutter percentage observed for the allele during validation and other studies should be used to assess whether excess stutter is likely. Peaks in stutter positions that are more than 2 standard deviations above the mean observed stutter (for Identifiler Plus amplification kit), and 3 standard deviations above the mean observed stutter (for the Profiler Plus, COfiler, Identifiler, and MiniFiler amplification kits), indicate the presence of excess stutter and that DNA contribution from a source other than stutter is likely. If a peak in a stutter position is deemed to have originated from a separate contributor to the sample, then all stutter peaks in the mixture should be considered as possible allelic peaks from the additional contributor(s) (within the limitations of the number of contributors) and can be evaluated in the same manner described above. Subsequent to an evaluation of stutter peaks in the mixture it may be necessary to reassess the minimum number of DNA contributors to the sample.

2.9.1 Samples with an assumed number of contributors.

2.9.1.1 An apparent ***single source sample*** consists of no more than two relatively balanced alleles at each locus. Single source samples with three allele peaks at a single locus and/or partial null alleles have been reported, though are extremely rare. For single source samples, allele number and peak height ratios should be appropriate for the level of the detected DNA.

Alleles used for comparison:
When two alleles, or one allele detected above the stochastic threshold, are detected at a locus, allelic dropout is not a concern. These alleles are used for comparison. When one allele is detected and the allele is below the stochastic threshold, allelic dropout is possible. The allele may be used for comparison, but allelic drop-out must be considered in the interpretation.

2.9.1.2 When the data supports it, an analyst may make an ***assumption as to the number of DNA contributors in a mixture***. It is not recommended that analysts make assumptions on mixtures of four or more individuals as the data in those mixtures becomes much harder to interpret.

When the number of contributors to a DNA mixture is assumed to be two or three people, there are restrictions on the genotypes that could have contributed to the profile. Given the number of assumed contributors, the number of detected alleles, and the ratio of the contributors, it is possible to remove certain allele combinations from consideration. Any criteria used for the interpretation of a mixture, or to select acceptable genotype combinations for any contributors, must be listed in the analytical record.

Alleles used for comparison:
All alleles at loci where the analyst is confident that all DNA types are detected will be used for comparison for mixtures of two or three people.

Loci with DNA types below the stochastic threshold can be used in some circumstances in mixtures of two or three people:

a) If all DNA types from all contributors appear to be present at a locus, all alleles at the locus, including those below the stochastic threshold, will be used for comparison.

b) If DNA types below the stochastic threshold are observed and at least one allele is present from each contributor (three alleles in a two-person mixture or five alleles in a three-person mixture), all alleles can be used for comparison if the analyst considers the possibility that the DNA type(s) below the stochastic threshold can be paired with an undetected allele.

c) If the analyst does not have confidence that at least one allele has been detected from each contributor, the locus should not be used for comparison if DNA types below the stochastic threshold are detected unless a major contributor or major contributors are present (see section 2.8.1.4a).

2.9.1.3 Mixtures of two or three contributors where a forensically valid assumption regarding a contributor can be made

In addition to the guidelines in section 2.9.1.2, when the number of contributors is assumed and a forensically valid assumption regarding a contributor is made, further refinement of the mixture is possible. In these circumstances, possible combinations representing the source(s) of the DNA foreign to the assumed contributor(s) should be considered for comparison as described below. DNA types foreign to the assumed contributor(s) are referred to as obligate foreign alleles.

Loci used for comparison:
Loci where the analyst is confident that all DNA types are detected can be used for comparison. Loci with DNA types below the stochastic threshold can be used if all of the types below the stochastic threshold are due to the assumed contributor and/or if at least one DNA type from the foreign contributor(s) is present, in a similar fashion to 2.9.1.2a and 2.9.1.2b.

2.9.1.4 Some samples may not allow for confident assumption as to the number of total contributors, but do allow for an ***assumption as to the number of major contributors***.

a) ***Assumption of one major contributor***. If the overwhelming contribution to the DNA in a mixture appears to originate from a single individual, then the term predominant DNA contributor will apply. A ***predominant DNA profile*** will be one where a single source genotype from the single strongest contributor to the mixture is able to be deduced from all of the DNA markers tested. In general, all alleles from predominant DNA contributors will be designated as major. Factors that will affect deducing a predominant DNA profile such as the genotype of an assumed additional contributor or other issues commonly observed in PCR systems (stutter, peak height ratios, preferential amplification, etc.) should all be considered in the interpretation of a predominant genotype.

In some circumstances a single major contributor will be apparent, but some ambiguity may exist for the genotype of the major contributor at a few loci due to stacking of minor contributor alleles. Peak height ratio expectations may limit the ambiguity, but if uncertainty still exists, an analyst may choose all combinations of types that are not designated as minor at the ambiguous loci or all loci.

Alleles/loci used for comparison:
The DNA types for a predominant contributor will be considered for comparison in a similar fashion as single source samples (section 2.9.1.1).

Minor contributors to a mixture with a single major contributor, but no assumption as to the number of total contributors will be evaluated in a similar fashion as in section 2.9.2.2

b) ***Assumption of two major contributors***. It is possible that two DNA contributors will contribute significantly more DNA than any other contributors to a mixed DNA sample. In this circumstance, an analyst can treat the two major contributors in a fashion similar to section 2.9.1.2.

Minor contributors to a mixture with two major contributors, but no assumption as to the number of total contributors will be evaluated in a similar fashion as in section 2.9.2.2

c) ***Assumption of one major contributor, one mid-level DNA contributor, and at least one minor DNA contributor***. In certain mixtures, it is possible to identify a "mid-level" contributor that is contributing more than at least one other contributor but less than another. In these circumstances, the assumptions of the analyst should be documented in the case packet. Additional terms such as "mid-level" or "significant" can be used provided they are clearly defined in the case notes. Alleles designated as

likely originating from the significant second contributor should be documented in the analytical record.

Alleles/loci used for comparison:
The interpretation of the presence of a significant second contributor to a mixture should be based on a discernable difference in peak intensities between the second significant contributor and any other minor alleles present in the mixture. Alleles designated as likely originating from the significant second contributor should be documented in the analytical record. Care should be taken with any alleles thought to be from a second significant contributor if those alleles are in the stochastic range.

2.9.2 Even when the ***number of total contributors and/or the number of major contributors cannot be confidently assumed***, the sample may be evaluated for the presence of possible major contributors (section 2.9.2.1) and minor contributors (section 2.9.2.2). In some circumstances, the mixture cannot be resolved into major and minor components (section 2.9.2.3) or is not useful for comparison (section 2.9.2.4).

2.9.2.1 ***Major contributors with no assumption of the number of contributors.*** Major contributors should not have types that are designated as minor contributor alleles (see section 2.5); the subset of alleles in a DNA mixture above the level of minor contributor alleles are considered possible major contributor DNA types. It may be possible that the number of major contributors can be assumed even if the number of total contributors cannot be confidently determined (see section 2.8.1.4).

Loci used for comparison:
Loci with all possible major contributor DNA types above the stochastic threshold will be used for comparison to the major contributor DNA types. If the mixture contains multiple major contributors (i.e. a balanced mixture), caution should be exercised in using loci for comparison where the total signal detected is less than 700rfu. At this

level it is possible that major contributor alleles may be below the detection threshold of 50rfu (using the 7% threshold). Variables including the peak heights, number of detected alleles, and the minimum number of contributors will all play a part in determining whether a locus should be used for comparisons.

If possible major contributor DNA types at a locus are detected below the stochastic threshold, in a mixture evaluated without assumptions; the locus will not be used for comparison purposes to possible major contributors. The locus will be not be used for comparison due to the concern that one or more major contributors may have DNA types that were not detected.

In mixtures where only a subset of the loci are deemed usable for comparison/statistical calculations, any inclusions should be supported by the information contained within the remaining loci. If the remaining loci contain instances that conflict with the stated inclusion, the analyst may deem that exclusion is the more appropriate conclusion given the totality of the information in the results. If an analyst opts for exclusion under the previously described scenario, the reasons underlying the exclusion must be articulated in the notes.

2.9.2.2 ***Minor contributors with no assumption of the number of contributors.*** Minor DNA contributors should show proportional contribution across the DNA mixture. There can be more than one minor DNA contributor to a mixture. In general, minor DNA contributors will possess weaker DNA types across several loci (see section 2.5). Masking and allelic drop-out are more likely with minor DNA contributors. Caution should be taken when interpreting a mixture with only one or two minor DNA types in a *complex* mixture as allele sharing may have caused the allele designation to be a "minor" type(s). The validation data indicates that it is possible to have alleles less than 1/3 the height of the tallest in mixtures with relatively equivalent contributors due to allele stacking effects.

If no assumptions are made about the number of minor DNA contributors, then all combinations of alleles at a locus should be considered as possibly originating from a minor contributor and not just genotypes containing the designated minor alleles.

Loci used for comparison:
Loci with minor contributor alleles present and where the analyst is confident that all DNA types are detected will be used for comparison to possible minor contributors. Analysts should be able to articulate their reasoning for the use of loci for comparison to potential minor contributors. Peak heights, number of detected peaks, minimum number of potential DNA contributors may all play a role in an analyst's decision use a locus for comparison purposes. Loci with DNA types below the stochastic threshold will not be used for comparison purposes to possible minor contributors due to concern that one or more minor contributors may have DNA types that were not detected. If fewer than four loci are useful for comparison to possible minor contributors, the profile will be considered ***uninterpretable*** with respect to minor contributors.

2.9.2.3 ***Mixtures with no major contributors and no assumption of the number of contributors.*** Samples with no DNA types designated as minor contributor alleles (see section 2.5) are not resolvable into major or minor contributors. Mixtures, such as those of more than four contributors, may have DNA types designated as minor contributor alleles, but ambiguity may still exist as to whether one or more contributors is donating more DNA than others in the mixture. Such samples will be evaluated for possible ***DNA contributors***.

Loci used for comparison:
Loci where the analyst is confident that all DNA types are detected will use all alleles for comparison to possible DNA contributors. Loci with DNA types below the stochastic threshold will not be used for comparison purposes to possible DNA contributors due to concern that one or more

DNA contributors may have DNA types that were not detected. If fewer than four loci are useful for comparison to possible DNA contributors, the profile will be considered uninterpretable.

2.9.2.4 ***Uninterpretable mixtures*** are mixtures where no assumptions can be confidently made and where all observed alleles are below the stochastic threshold are not useful for comparison and are considered uninterpretable. A mixture may be useful for comparison to possible major contributors and may be uninterpretable with respect to possible minor contributors.

3. Comparison of DNA Typing Results

3.1 ***Inclusion.*** If a person of interest has no unexplainable differences from the alleles or allele combinations determined to be useful for statistical support of inclusions, and thus comparison, they are included as a source of the DNA or possible contributor to the DNA mixture.

3.2 ***Exclusion.*** If a person of interest has unexplainable differences from the alleles or allele combinations determined to be useful for comparisons, they are excluded as a source of the DNA or possible contributor to the DNA mixture.

3.3 ***Inconclusive.*** An analyst may determine that due to a combination of variables including (but not limited to) the number of DNA contributors as well as a lack of information in the results that no comparisons can be made to the profile (or comparison for possible minor contributors).

3.4 ***Unsuitable for comparisons***. Some mixtures such as those with five or more contributors may not be suitable for comparison, unless a minimal number of major contributors exist. Comparisons to these mixtures will be considered inconclusive.

Mixtures with no assumptions made to the number of contributors, that have less than 4 loci available for comparison, will be deemed unsuitable for comparisons.

If the positive amplification control does not provide the expected results, or if DNA types are detected in the negative amplification control or the

relevant reagent blank, the sample may not be useful for comparison (see section 1.6) and would thus be uninterpretable.

4. Statistical Analysis of DNA Typing Results

When a comparison leads to a conclusion that an individual is included as a possible source of the DNA or a possible contributor to a DNA mixture, an assessment of the significance of that inclusion should be performed by calculating the rarity of such an inclusion. If the inclusion has little probative value a calculation may not be necessary. If an individual is included as a possible contributor to the DNA from multiple items in a case, calculations may only be necessary for the most probative item(s) and/or the item(s) with the highest discrimination potential.

Calculations will only be performed on loci and/or allele combinations that have been deemed acceptable for comparison purposes (section 2.9).

All calculations will be generated for the Caucasian, African American, Hispanic, and Asian populations using the DNA allele frequency data taken from Forensic Sci. Int. Genet. 7 (2013) e82–e83 and Forensic Sci. Int. Genet. 31 (2017) e36–e40 unless noted. If a different database is used for calculations, it should be included in the administrative documents with the case file and noted in the report.

The product rule will be used for calculations by determining the probability at each locus and multiplying the results for each locus together.

In general, Popstats is used for statistical calculations. There is also the SDPD RMP calculator, which can be used for mixtures, that allows for the exclusion of certain genotypes from the calculation.

4.1 For ***single source samples*** and samples with a ***predominant DNA profile*** the significance of a match should use the random match probability (RMP) formulae for single source samples. Loci with two alleles (heterozygotes) or with a single allele above the stochastic threshold (homozygotes) can be calculated using Popstats with the "Forensic Single Sample" calculation selected. For loci with a single allele below the stochastic threshold, the analyst can use Popstats Forensic Single Sample calculation with "Rec 4.1" selected for the locus and "$2p-p^2$" selected as the "Rec 4.1 formula" in the NRC '96 box of the Probability Formula options in

Popstats. The analyst may alternatively decide not to use the locus with a single allele below the stochastic threshold for statistical weight if the added value would be insignificant to the overall conclusion. The formulae for the Random Match Probability for single source samples are shown below:

$f = 2pq$	(heterozygote)	
$f = p^2 + p(1 - p)\theta$	(homozygote)	(θ=0.01 in most instances)
$f = 2p - p^2$	(single peak below stochastic threshold)	

4.2 There are two methods used for calculations when an ***assumption is made regarding the number of contributors or the number of major contributors*** to a mixture: the likelihood ratio (section 4.2.1) and the random match probability (section 4.2.2). Both methods can be restricted (taking into account peak height ratios and mixture ratios) or unrestricted. The RMP can also be modified based on an assumed number of contributors to exclude unreasonable genotypes from consideration. Both methods can take into account the presence of an assumed contributor.

4.2.1 If the number of contributors can be assumed with confidence, the ***likelihood ratio*** takes into account all of the information in the sample. The likelihood ratio is the ratio of two probabilities of the same event under two mutually exclusive hypotheses. Given the same data, it is possible to calculate more than one likelihood ratio under a different set of assumptions and hypotheses.

In general, the likelihood ratio formula is:

$$L = \frac{P(E|C_x)}{P(E|C_y)}$$

Typically, the calculation in the numerator is the probability of obtaining the evidence under the hypothesis that a person of interest is one of the contributors to the DNA mixture along with additional known or unknown people. The calculation in the denominator is the probability of obtaining the evidence under the hypothesis that a randomly selected individual is contributing to the DNA mixture along with additional known (generally the same as

the numerator) or unknown contributors. If the result of the ratio is a number greater than one, the DNA results are more probable if the person of interest is a contributor to the mixture. If the result of the ratio is less than one, the DNA results are more probable if the evidence is a mixture of DNA from randomly selected individuals.

In the SDPD lab, the likelihood ratio has been more commonly calculated when a contributor is assumed and the number of contributors is also assumed, but it can also be calculated without the assumption of a contributor. Likelihood ratios are also used in parentage and kinship calculations, but these will be dealt with later (sections 4.4 and 4.5).

Please see the complete document for additional scenarios where likelihood ratios could be useful. Popstats can calculate likelihood ratios; however, it does not employ formulas to deal with restricted likelihood ratios or scenarios where allelic dropout is possible.

4.2.2 Although commonly used for single source samples, ***random match probabilities (RMP)*** calculations can also be performed for mixtures with an assumed number of contributors or for major contributors to a mixture with an assumed number of major contributors. The RMP calculates the probability of each genotype that could reasonably represent a contributor. The RMP differs from a combined probability of inclusion (see section 4.3) in that fewer possible genotypes exist due to the assumption regarding the number of contributors.

When a mixture is encountered that appears to have a single major contributor, but where some ambiguity exists in the genotype at one or more loci (see section 2.8.1.4a for an example), an RMP incorporating multiple genotypes at the ambiguous loci can be performed.

In balanced mixtures of DNA from two people, the RMP could be calculated by using only the possible combinations that could have contributed to the mixture (e.g. heterozygote combinations where four alleles are detected).

In circumstances where DNA types below the stochastic threshold are detected and where dropout has possibly occurred, a modified version of the $2p - p^2$ calculation in section 4.1 can be performed using the following formula for the alleles below the stochastic threshold (p, q, r, etc):

$$2(p + q + r + \ldots) - (p + q + r + \ldots)^2$$

In mixtures of DNA where the presence of a specific contributor is assumed, an RMP could be calculated for the combinations of DNA types foreign to the assumed contributor. Such calculations focus on the obligate alleles in the mixture.

4.3 Combined Probability of Inclusion

When no assumptions regarding the number of contributors can be made with confidence, the ***combined probability of inclusion (CPI)*** is used to express the significance of an inclusion. The probabilities of all possible combinations at each locus are calculated and summed to generate a probability of inclusion or PI. The PI for each of the loci acceptable for comparison are multiplied together to generate the CPI. The PI for each marker and the CPI are calculated as:

$$PI = (p + q + r + \ldots)^2$$
$$CPI = PI_1 * PI_2 * PI_3 * \ldots$$

A CPI using the possible major contributor DNA types (e.g. types above minor) can be calculated to express the significance of an inclusion as a major contributor. A separate CPI calculation using all DNA types can be performed to express the weight of an inclusion of a minor contributor. For samples without major and minor contributors, a CPI using all alleles can be performed to express the significance of an inclusion as a possible DNA contributor.

4.4 Parentage calculations involving legacy kit loci will be handled using current interpretation guidelines.

5. Reporting of DNA Typing Results and Conclusions

5.1 Reporting of STRmix DNA typing results from legacy data will be handled using current reporting guidelines.

5.2 Reporting of legacy data not interpreted using STRmix.

5.2.1 Results from legacy data will include the comparison typing results for each person of interest (see section 3).

5.2.2 If an evidence profile or one or more components of an evidence profile (i.e., a predominant contributor, minor contributor, etc.) are searchable in, and uploadable to CODIS, a statement as to whether DNA types or a subset of the DNA types from an item will be searched in CODIS and what level of CODIS they are being searched at must be included.

5.2.3 If the comparison typing results for one or more persons of interest would be reported as "included", the applicable statistical calculation for the racial group giving the lowest statistic will be reported. If the results are only applicable to a component of an evidence profile (i.e., a predominant contributor, minor contributor, etc.), the interpretation should add such a specification.

5.2.3.1 Popstats likelihood ratio results from legacy data will be handled similar to current reporting guidelines, omitting reference to the estimated percent contribution of each contributor and the contributor order to which the POI best fits.

5.2.3.2 Example of reporting of random match probability or combined probability of inclusion results from legacy data:

The approximate probability that a person chosen at random would be included as (major) DNA contributor to the mixture is 1 in 2.3 septillion

Forensic Biology Unit Legacy Kit Interpretation Guidelines Manual

Forensic Biology

METHOD: **AmpF*l*STR Yfiler™ Y-STR Interpretation Guidelines**

Date: 06/01/2004 Revision Date: 05/28/2024 Approved by: ARD

Introduction

The Scientific Working Group on DNA Analysis Methods (SWGDAM) states that:

> "The interpretation of DNA typing results for human identification purposes requires professional judgment and expertise. Additionally, laboratories that analyze DNA samples for forensic casework purposes are required by the Quality Assurance Standards for Forensic DNA Testing Laboratories (effective July 1, 2009) to establish and follow documented procedures for the interpretation of DNA typing results and reporting. Due to the multiplicity of forensic sample types and the potential complexity of DNA typing results, it is impractical and infeasible to cover every aspect of DNA interpretation by a preset rule. However, the laboratory should utilize written procedures for interpretation of analytical results with the understanding that specificity in the standard operating protocols will enable greater consistency and accuracy among analysts within a laboratory. It is recommended that standard operating procedures for the interpretation of DNA typing results be sufficiently detailed that other forensic DNA analysts can review, understand in full, and assess the laboratory's policies and practices. The laboratory's interpretation guidelines should be based upon validation studies, scientific literature, and experience."

Here we present a number of general interpretation guidelines for interpreting samples in forensic Y-STR casework. The rationale behind these guidelines is to establish a set of standards to ensure that conclusions are supported by reliable data, derived using a scientific approach, and that interpretation of results is as objective as possible and is consistent from one analyst to another.

The interpretation of Y-STR results should take into account the DNA results obtained, the nature of the sample, the condition of the profile, any forensically valid assumptions made, and should be scientifically justifiable and based on the collected experience and knowledge of the laboratory and the

scientific community. These guidelines are designed to provide a general foundation for interpretation.

DNA interpretation should take place in the following order:

Preliminary Evaluation of Data and Allele Designation (section 1)
Interpretation of DNA Typing Results (section 2)
Comparison of DNA Typing Results (section 3)
Statistical Analysis of DNA Typing Results (section 4)
Reporting of DNA Typing Results and Conclusions (section 5)

This document will reference the Autosomal STR Interpretation Guidelines document where information applicable to Y-STR interpretation is available.

Section 1 – Preliminary Evaluation of Data and Allele Designation

Please refer to section 1 of the Autosomal STR Interpretation Guidelines document as all the information therein is relevant to Y-STR interpretation.

Section 2 – Interpretation of DNA Typing Results

In general, interpretation of Y-STR data will be completed prior to comparison of reference samples. All interpretation of samples will be performed using the actual electrophoresis data and not the table of results. DNA interpretation will include the following steps:

Determine whether a sample is a mixture (section 2.1)
Estimate the number of contributors to a sample (section 2.2)
Identify whether a sample is degraded and/or inhibited (section 2.3)
Determine contribution level in mixed DNA samples (section 2.4)
Document relevant assumptions (section 2.5)
Determine which loci and combinations of DNA types are useful for statistical support of inclusions (section 2.6).

2.1 Determination of whether a sample is a mixture

A mixed Y-STR profile possesses certain common attributes that will tend to indicate the presence of more than one contributor to the observed results. The attributes may include:

- Greater than one allele is present at a locus other than DYS385 where a gene duplication results in a bi-allelic pattern.
- Significant imbalance in peak height between peaks at a DYS385.
- Generally, evidence of a mixed DNA sample exists at multiple loci.

To determine whether a mixed Y-STR profile exists, the DNA profile must be evaluated in its entirety. If a DNA mixture exists, a conclusion to that fact should be stated in the Forensic Biology report, and may be stated independently of the conclusion(s) surrounding potential individual contributors (see section 5).

2.2 Estimation of the number of contributors to a sample

Conclusions with respect to the minimum number of contributors must be included in a report. Generally, an estimate of the minimum number of contributors is based on the locus that exhibits the greatest number of allelic peaks and the fact that each contributor would contribute at most one allele to the locus (except DYS385). While the minimum number of contributors to a mixture can be established with a high probability, an absolute determination of the number of contributors may be more difficult.

2.3 Determination of whether inhibition and/or degradation is present

Please refer to section 2.4 of the Autosomal STR Interpretation Guidelines document for more information on inhibition and degradation. The SDPD does not currently have good examples of degraded or inhibited Y-STR profiles, however, the same basic tenets apply to both autosomal and Y-chromosome analysis.

2.4 Determination of contribution levels in mixed DNA samples

The relative peak intensity differences between detected alleles can be used to identify differences in contribution levels between contributors to relatively simple (i.e. 2-person) mixtures. In general, a 3:1 difference in peak intensity will be used to identify major vs. minor alleles. In more complex mixtures, such as mixtures with more than two contributors or mixtures that may exhibit inhibition or degradation, it may be difficult to determine major vs. minor alleles and analysts are encouraged to exercise caution when interpreting Y-STR mixtures.

2.5 Documentation of Assumptions

In certain circumstances, ***assumptions regarding DNA profiles*** may be made to assist the analyst in drawing conclusions and to put the evidence in the proper context. All assumptions regarding any interpretation will be presented in the report and should also be in the notes (see section 5).

In certain circumstances ***assumptions regarding the presence of a contributor*** in a DNA mixture is permissible (i.e. intimate samples). In these circumstances the DNA profile from the assumed contributor(s) can be used to assist interpretation of the additional contributor(s) DNA profile/types (section 2.7).

Prior to making comparisons, the profile should be evaluated to ***determine what loci and/or combinations of DNA types are acceptable to support a conclusion statistically***. Since there is technically no stochastic threshold to be concerned with (with the exception of DYS385) much of the evaluation to determine what parts of the DNA results can be used for statistical weight takes into account the possible presence of DNA types below the analytical (or detection). If, based on the totality of the information in the Y-STR result, it is reasonable to assume that all information is accounted for in the results, all DNA types can be used for comparison, but it may useful to further refine the genotypes for comparison if a limited number of combinations are likely.

If an analyst believes that ***filtered stutter peaks*** should be evaluated for comparison, the mean stutter percentage observed for the allele during validation and other studies should be used to assess whether excess stutter is likely. Peaks in stutter positions that are more than 2 standard deviations above the mean observed stutter indicate the presence of excess stutter and that DNA contribution from a source other than stutter is likely. If a peak in a stutter position is deemed to have originated from a separate contributor to the sample, then all stutter peaks in the mixture should be evaluated in the same manner and included in the assessment of the mixture. After the evaluation of stutter peaks in the mixture it may be necessary to reassess the minimum number of DNA contributors to the sample.

Section 3 - Comparison of DNA Typing Results

3.1 ***Inclusion.*** If a person of interest has no unexplainable differences from the alleles or allele combinations determined to be useful for statistical support of inclusions, and thus comparison, they are included as a source of the DNA or possible contributor to the DNA mixture.

3.2 ***Exclusion.*** If a person of interest has unexplainable differences from the alleles or allele combinations determined to be useful for comparisons, they are excluded as a source of the DNA or possible contributor to the DNA mixture.

3.3 ***Inconclusive.*** Some mixtures such as those with five or more contributors may not be suitable for comparison, unless a minimal number of major contributors exist. Comparisons to these mixtures will be considered inconclusive.

An analyst may determine that due to a combination of variables including (but not limited to) the number of DNA contributors as well as a lack of information in the results that no comparisons can be made to the profile (or comparison for possible minor contributors).

Section 4 - Statistical Analysis of DNA Typing Results

Y-STR loci are linked on the Y-chromosome and do not undergo linkage disequilibrium. Therefore, neither Recommendation 4.1 (NRC 1996) nor the Product Rule can be applied to Y-STR haplotypes. The significance of a haplotype match in casework must be expressed in terms of the observed frequency of the particular haplotype in the population (the counting method). Numerous Y-STR haplotype databases are available on-line.

4.1 Apparent Single Source profiles

For the purposes of expressing the significance of a Y-STR haplotype match all alleles for loci deemed acceptable for comparison will be searched in the database. More DNA profiles in the US YSTR database have a subset of the Yfiler loci than have the full complement of the Yfiler loci. As a result, an analyst may elect to search a subset of the Y-STR results in the database for the purposes of obtaining a more discriminating search.

4.2 Mixed Y-STR DNA profiles

4.2.1 If a significant difference in allele peak height exists in a male DNA mixture, a predominant Y-STR DNA haplotype may be deduced. If a predominant haplotype is deduced, all applicable alleles from the predominant haplotype will be searched in the database (as in section 4.1). For the purposes of expressing the significance of a Y-STR haplotype match all applicable alleles will be searched in the database.

4.2.2 If a predominant Y-STR DNA profile cannot be deduced, the frequencies in a given database may be determined for each possible haplotype and the sum of all combinations of possible haplotypes may be reported.

4.3 Y-STR haplotype frequencies can be calculated from a number of established databases. The laboratory has incorporated the CADOJY-Mix Tool v. 4.2 R60(US)1 for the calculation of Y-STR haplotype frequencies. This haplotype calculator can be found on the forensic biology computer network:

H:\Worksheets\Statistical Tool\Y-Mix -version 4.2 –R60(US)1

This tool will not only allow for the frequency estimates of single source haplotypes, but will also allow for calculations to be done on Y-STR mixture profiles. Instructions are located within the Excel spreadsheet on the "Instructions" tab. The database that this spreadsheet uses for its haplotype and Y-STR mixture calculations is release 60 of the Y-chromosome Haplotype Reference Database (www.yhrd.org).

To search for a haplotype, enter the alleles of interest for each marker by using the drop-down menu, or typing the alleles into each appropriate cell in the worksheet (only one allele per cell). To specify a value not listed in the drop-down menu, enter the variant allele into the appropriate new variant cell. www.yhrd.org is a useful resource for Y-STR information. If mutation rates are required for any Y-STR calculations, analysts must use the mutation rates listed on this website.

4.3.1 Using Y-Mix version 4.2 R60(US)1

Open the Y-Mix tool. Entered haplotype information for all available loci in alpha-numerical order with each individual DNA type listed in a single cell within the appropriate locus column. The Y-Mix tool allows for haplotype mixtures to be searched.

The Y-Mix tool has an option to either limit the search to only database samples with the same set of data as the entered haplotype or to search against all samples within the database. Selecting "Yes" for *Limit database to samples with all the loci entered above?* results in searching against only samples within the database than contain all the loci entered. Selecting "No" results in a search against all haplotypes in the database. Selecting "No" will always result in a more discriminating search and should be the default search parameter when searching haplotypes.

Y-Mix Filter Tool v4.2 also has tool allowing drop-out to be considered at the duplicated markers during the database search of single source samples. When "Yes" is selected for *Treat this profile as a single source sample?*, an option appears: *Allow for dropped alleles at multicopy loci?*. Selecting "No" searches for the exact haplotype entered. When only one type is entered at a duplicated locus (i.e., DYS385), selecting "Yes" allows for any haplotypes with the entered type as a homoallele, and the entered allele paired with any other type (heteroallelic haplotype) to be returned in the search. This feature is useful when single types below 570RFU are present at either of the duplicated markers.

The search and results should be saved to an analyst folder as a separate Excel file.

4.3.2 YHRD may be used on occasions when a search of a broader spectrum of the haplotypes are required to be searched. YHRD contains data from over 340,000 haplotypes from around the world. All samples searched against YHRD are only searched against samples with the same complement of loci that are entered for the haplotype to be searched. YHRD does not accommodate searching mixtures.

To use the YHRD, access www.yhrd.org, then click on the *Estimate Frequency* tab. Click on the *Manually enter haplotype(s)* button (alternatively, the YSTR haplotype may be uploaded via an Excel spreadsheet, .csv file, or GeneMapper ID-X export file). Select *Yfiler Plus*, enter the all available loci as appropriate, and click the "Search" button. Click on the *Add feature to the report* and select "National Database (with subpopulations, 2014 SWGDAM – compliant)". Ensure the drop down lists *United States* to obtain the data for the US population dataset. The statistical calculations of interest will be listed under the *Observed* title.

To achieve more discriminating searches, users may alternately reduce the number of loci by selected the YFiler, or Minimal, buttons to reduce the number of loci searched (without having to re-enter the haplotype).

Searches using YHRD can be printed, but cannot be saved.

Section 5 - Reporting of DNA Typing Results and Conclusions

The following are meant as guidelines for reporting conclusions for DNA testing in Forensic Biology Unit reports. There are many ways to state a conclusion in a DNA report and the suggestions listed below do not cover every possible scenario. Alternative wording may be used to express a conclusion so long as the conclusion is correct, supported by the data, and is appropriate based on the statistical probability limits employed by the Unit.

5.1 Reporting of DNA typing results will be handled using current reporting guidelines.

5.1.1 Please refer to the Autosomal STR Interpretation Guidelines document for additional information.

REFERENCES:
1. AmpF*l*STR Yfiler™ PCR Amplification Kit User's Manual
2. SDPD AmpF*l*STR Yfiler™ validation studies.
3. SDPD verification of the CADOJY-Mix Tool