

DNA: Interpretation of DNA Typing Results

Introduction The interpretation of results in casework is a matter of professional judgment and expertise. Not every situation can or should be covered by a pre-set rule; however, it is important that the laboratory develops and adheres to minimum criteria for interpretation of analytical results. These criteria are based on validation studies, scientific literature references, casework experience, SWGDAM documents, the FBI's QAS document, and input from the forensic community at large.

Purpose The purpose of these guidelines is to establish a general framework and outline minimum standards to ensure that:

- conclusions in casework reports are scientifically supported by the analytical data, including data obtained from appropriate standards and controls
- interpretations are made as objectively as possible, consistently from analyst to analyst, and within previously agreed limits

Terminology See [DNA: Glossary](#).

DNA Manager *DNA Manager* is an *Excel* based workbook used to manage samples and document the steps of DNA analysis.

The workbook is located in a limited-access shared drive.

The embedded worksheets are not reviewed or approved by the Director as they are considered components of the examination documentation for the case file.

COSTaR COSTaR is an *Excel* based workbook used to manage samples for possible upload to CODIS.

The workbook is located in a limited-access shared drive.

The embedded worksheets are not reviewed or approved by the Director as they are considered components of the examination documentation for the case file.

Continued on next page

DNA: Interpretation of DNA Typing Results, Continued

Interpretation – The steps for the interpretation of STR data amplified with Fusion 6C are listed in the table below. The steps employ a fully continuous probabilistic genotyping system, STRmix™, which uses biological modelling, statistical theory, computer algorithms, and probability distribution to calculate likelihood ratios and/or infer genotypes for the DNA typing results of forensic samples.

Promega
PowerPlex®
Fusion 6C

| Step | Action |
|------|---|
| 1 | Decide whether the profile is suitable for interpretation using STRmix™ . |
| 2 | Assign the number of contributors. |
| 3 | Identify possible assumed contributors. |
| 4 | Compose the likelihood ratio (LR) propositions appropriate for the case circumstances, if needed. |
| 5 | If suitable, interpret the profile with STRmix™ . |
| 6 | Review STRmix™ outputs. |
| 7 | Report conclusions and statistics. See DNA: Reporting DNA Typing Results. |

Continued on next page

DNA: Interpretation of DNA Typing Results, Continued

Profiles suitable for STRmix™ interpretation

STRmix™ should be used to interpret all qualifying profiles, where a qualifying profile is defined below.

- all probative samples, including single-source and 2, 3, and 4 person mixtures suitable for interpretation

STRmix™ should not be used to interpret profiles where:

- the number of contributors cannot be determined
- there are more than four contributors
- an inclusion does not provide probative value in the context of the case. For example:
 - an individual's own DNA on intimate swabs or clothing
 - alleles from a non-sperm fraction of an intimate swab are consistent with carry-over from the sperm fraction and the victim
 - alleles from a steering wheel are consistent with one or more individuals who are known to drive the car

If it has been determined that a profile is unsuitable for interpretation, an explanation as to why the DNA evidence is uninterpretable shall be given.

If a person of interest (POI) is visually excluded as a contributor to a single-source profile, a likelihood ratio does not need to be calculated, but the profile should be interpreted with STRmix™ for potential future comparisons or for upload to CODIS. The POI is reported as an exclusion.

Continued on next page

DNA: Interpretation of DNA Typing Results, Continued

Assigning the number of contributors to the DNA profile

The assignment of the number of contributors to a profile must be made by the analyst prior to starting a STRmix™ interpretation. The number of contributors chosen for the analysis should be the most likely number required to reasonably explain the observed profile. The following components of the mixture should be examined to estimate the number of contributors:

- the locus (or loci) that contain the greatest number of alleles
- the presence of peak height imbalance at more than one locus
- possible peaks that fall below the analytical threshold
- elevated stutter peaks

The presence of an assumed contributor(s) can help in assessing the number of contributors.

The presence of alleles at the YSTR markers (DYS391, DYS570, and DYS576) may provide additional information about the number of contributors.

After analysis of a STRmix™ output file, a deconvolution may not conform to qualitative expectation based on the assessment of STRmix™ diagnostics. If the assumed number of contributors is determined to be the cause, the deconvolution may be re-run under a different assumption for the contributor number. In this case, the course of action should be detailed in the case file and the electronic file of the first deconvolution should be included with the electronic data. The analyst should report the interpretation(s) that best conforms to scientific expectation.

At times, for profiles with low levels of DNA or high numbers of contributors, there may be so much allele dropout that it is difficult to estimate the number of contributors. If the number of contributors cannot be reliably estimated, the profile may be determined to be unsuitable for STRmix™ interpretation.

If a profile is determined to have five or more contributors, the profile is unsuitable for STRmix™ interpretation.

Continued on next page

DNA: Interpretation of DNA Typing Results, Continued

Identifying possible assumed contributors

In some cases, profiles contain DNA from an individual who is reasonably expected to be a contributor to the sample. Whenever possible, the genetic profile of the assumed known contributor can be used to infer the genetic profile of the remaining contributor(s) in the sample. Examples of assumed contributor scenarios include:

- Intimate samples that are collected from an identified anatomical location (ex: body orifice, skin surface, fingernail scrapings) and/or a piece of clothing (ex: panties, bra, shirt).
- Samples in which an individual may be expected to contribute to a profile (ex: bedding, vehicle swabs, bloodstains from crime scenes in which a known individual was bleeding) as long as there is reasonable expectation of association.
- Samples with an additional known contributor, such as a consensual partner.
- Samples from which DNA is isolated by a differential extraction. In such situations, a contributor from one fraction (sperm or non-sperm) can be used to deduce information from its complementary fraction.
- For the purpose of obtaining a CODIS profile, the criteria for identifying an assumed contributor can be broader to include contributors that have already been identified in the case. For example, if one or more perpetrators have left clothing behind after a burglary, a contributor identified on a cutting from the sweatshirt can be used as an assumed contributor on a cutting from the cap to help STRmix™ deconvolute a complex mixture in order to develop one or more CODIS profiles from the cap. Before making this assumption, the proposed assumed contributor must give an $LR \geq 100$ in the questioned evidence sample.

For intimate evidence samples, the analyst can make a visual comparison between the assumed contributor profile and the evidence profile to ensure the assumed contributor is present in the evidence profile. If it is unclear whether an intimate contributor is present in a profile, a STRmix™ deconvolution can be run, and an LR calculated for the potential intimate contributor can be assessed. If the LR is greater than the inconclusive range ($LR \geq 100$), the profile can be re-run in STRmix™ with the intimate contributor assumed.

Continued on next page

DNA: Interpretation of DNA Typing Results, Continued

Identifying possible assumed contributors (continued)

To identify assumed contributors from non-intimate evidence samples and samples with possible consensual partners, a STRmix™ deconvolution should be run before making the assumption, regardless of apparent contribution level. If the LR for the questioned contributor is greater than the inconclusive range ($LR \geq 100$), the profile can be re-run in STRmix™ with the contributor assumed.

If the profile for the assumed contributor is partial (with any missing loci or single alleles under 450 rfu), an attempt should be made to obtain a full profile, if possible. If only a partial profile is available for assumption, only the loci with two alleles or one allele over 450 rfu should be used. The loci that do not have full information must be ignored during the deconvolution.

There may be times when the profile needed for the assumption is part of a mixture. In these instances, a STRmix™ deconvolution should be run on the mixture containing the proposed assumed contributor. Using the Contributor Summary document in the STRmix™ results folder for that sample, obtain the profile for the contributor where full genotypes are expected >98% of the time. Loci with an obligate allele (e.g. 10,0) will not be used. This profile can then be used as an assumed contributor, ignoring the loci that were not >98%. Because there is a level of risk in using the >98% genotypes, it is preferable to use a single source sample for assumption. If a mixture contributor is used for assumption, the run diagnostics should be closely examined.

Continued on next page

DNA: Interpretation of DNA Typing Results, Continued

**Setting up the
Likelihood
Ratio (LR)
propositions
appropriate for
the case
circumstances**

The LR is represented by the following equation:

$$LR = \frac{\Pr(E \mid H^I)}{\Pr(E \mid H^2)}$$

The LR assesses the probability of the evidence given two alternate propositions or hypotheses, one that aligns with the prosecution (H_1) and one that aligns with the defense (H_2). H_1 is typically inclusionary of the person of interest (POI). H_2 is typically exclusionary of the person of interest (POI).

Depending on the case scenario, efforts are made to set propositions that will minimize the LR. This can be accomplished by minimizing the difference in the number of unknowns between H_1 and H_2 .

H_1 should contain any assumed contributors, a POI (if applicable), and any unknown contributor(s) representing the total number of contributors in the sample. H_2 should contain any assumed contributors and the appropriate number of unknown contributors representing the total number of contributors in the sample. The total number of contributors in H_1 must equal the total number of contributors in H_2 .

Examples:

- Two person mixture from intimate swab collected from C. One POI identified.
 - H_1 : DNA originates from C and POI
 - H_2 : DNA originates from C and unknown
- Three person mixture from firearm located at scene. One POI identified.
 - H_1 : DNA originates from POI and two unknowns
 - H_2 : DNA originates from three unknowns
- Three person mixture from firearm located at scene. Two POIs identified. Two LRs can be calculated.
 - LR_a:
 - H_{1a} : DNA originates from POI1 and two unknowns
 - H_{2a} : DNA originates from three unknowns
 - LR_b:
 - H_{1b} : DNA originates from POI2 and two unknowns
 - H_{2b} : DNA originates from three unknowns

Continued on next page

DNA: Interpretation of DNA Typing Results, Continued

**Setting up the
Likelihood
Ratio (LR)
propositions
appropriate for
the case
circumstances
(continued)**

This scenario represents an example where it is unavoidable to limit the difference in the number of unknowns under H_1 and H_2 . Assuming that inclusionary LR's were obtained for POI1 and POI2, a third LR should be calculated.

- LR_c:
 - H_{1c} : DNA originates from POI1, POI2, and one unknown
 - H_{2c} : DNA originates from three unknowns

This third LR will typically be more than additive if POI1 and POI2 are present in the mixture together.

Multiple propositions may be required in certain instances. This information will be documented in the case file. If multiple runs are performed in STRmix™ for multiple propositions that are evaluated, maintain all analyses in the case file. The analyst will determine which proposition(s) to report based on available case information.

If the defense hypothesis suggests that a relative of the defendant is the true contributor to the DNA results, then an attempt should be made to obtain a reference sample from the relative. If a reference sample cannot be obtained, alternate LR's may be provided. STRmix™ automatically calculates LR's for siblings, parents/children, grandparents/grandchildren, aunts/uncles, and first cousins of the compared individual, and the appropriate LR's may be included in the report.

- For example, the sibling LR would be:
 - H_1 : DNA originates from POI and one unknown
 - H_2 : DNA originates from a sibling of POI and one unknown

Continued on next page

DNA: Interpretation of DNA Typing Results, Continued

Special considerations for reference profiles

There may be times when the reference profile needed for a comparison is not a full profile. In this case, the loci that have partial information must be ignored for the STRmix™ analysis.

There may be times when the profile needed for a comparison is an evidence sample instead of a reference sample.

- If the profile is full and single-source, it can be analyzed with the stutter peaks removed and used directly in STRmix™.
- If the profile is partial and single-source, the loci that have partial information must be ignored for the STRmix™ analysis.
- If the profile is part of a mixture, a STRmix™ deconvolution should be run on the profile containing the target individual. Using the Contributor Summary document in the STRmix™ results folder for that sample, obtain the profile for that contributor where full genotypes are expected >98% of the time. Loci with an obligate allele (e.g. 10,0) will not be used. This profile can then be used for comparison, ignoring the loci that were not >98%. Because there is a level of risk in using the >98% genotypes, it is preferable to use a single source sample for comparison. If a mixture contributor is used for comparison, the diagnostics should be closely examined.

If reference samples from POIs are not available, deconvolute the sample in STRmix™ without setting up an LR. Assumed contributors can be included in the deconvolution, if appropriate. If reference samples become available at a later date, an LR can be calculated from this previous deconvolution.

Importing data into STRmix™

STRmix™ requires input data describing an electropherogram (epg) in order to run, and this data must be formatted in a specific way for it to be entered. As STRmix™ models the height of both allelic and stutter peaks, +/- one repeat stutter must not be removed from evidence samples at epg analysis. Labels must be removed from all other analysis artifacts including pull-up, spikes, dye blobs, and additional stutter (aside from back and forward stutter). Non-numeric values such as OL or </> are not permitted within the STRmix™ input files. There is no function to accommodate somatic mutations or trisomy in STRmix™ calculations. If a profile has a triallelic STR pattern, STRmix™ can still use the input data if the locus is ignored.

Continued on next page

DNA: Interpretation of DNA Typing Results, Continued

Importing data into STRmix™ (continued)

Input file format

There are two types of files that can be input into STRmix™:

- Evidence files (which require sample name, marker, allelic designations, size (i.e. molecular weight) and height information).
 - As STRmix™ models back and forward stutter, the stutter filters will need to be turned off during analysis to capture potential stutter information within your input file.
- Reference files (which require sample name, marker, allelic designations and size information).
 - As STRmix™ does not *interpret* the reference profile, stutter data should not be included in your reference input file.

The two types of input file formats that can be used for import into STRmix™ are text files (extension .txt) and common separated values files (extension .csv).

Replicates

A replicate is defined as a repeated amplification of the same extract. The replicate can be amplified with different amounts of template DNA. Replicates may improve the sensitivity and specificity of the interpretation. Replicates get imported into STRmix™ together and result in one deconvolution. Multiple injections of the same sample cannot be treated as replicates.

Replicate profiles are not required to be included in the interpretation. For example, if a full profile was obtained from one amplification and no further information was obtained from previous or subsequent amplifications, the amplification with the most information should be chosen for STRmix™ analysis. A replicate PCR may help in determining the number of contributors. If the replicate amplifications of the sample indicate the presence of a different number of contributors, the highest likely number of contributors should be used and reported.

Continued on next page

DNA: Interpretation of DNA Typing Results, Continued

Run parameters in STRmix™

The following settings should be used for the interpretation of profiles.

- DNA profiling kit used for interpretation:

STRmix - Manage DNA Profiling Kits

Manage DNA Profiling Kits

DNA Profiling Kit: Fusion6C_3500 [Delete Kit]

Kit name: Fusion6C_3500

Kit Type: Fusion6C

Stutter File: Sac county stutter file.csv [Select File] [Edit File]

Stutter Exceptions File: Sac county stutter exceptions file.csv [Select File] [Edit File]

Forward Stutter File: Sac county stutter file N+4.csv [Select File] [Edit File]

Size Regression File: Fusion6C_SizeRegression.csv [Select File] [Edit File]

Manage Loci [Add] [Delete] Number of Loci: 27 [Move Up] [Move Down]

| Locus Name | Gender Locus? | Ignore Locus? | Detection Threshold |
|------------|-------------------------------------|--------------------------|---------------------|
| AMEL | <input checked="" type="checkbox"/> | <input type="checkbox"/> | |
| D3S1358 | <input type="checkbox"/> | <input type="checkbox"/> | 100 |
| D1S1656 | <input type="checkbox"/> | <input type="checkbox"/> | 100 |
| D2S441 | <input type="checkbox"/> | <input type="checkbox"/> | 100 |
| D10S1248 | <input type="checkbox"/> | <input type="checkbox"/> | 100 |
| D13S317 | <input type="checkbox"/> | <input type="checkbox"/> | 100 |
| Penta E | <input type="checkbox"/> | <input type="checkbox"/> | 100 |
| D16S539 | <input type="checkbox"/> | <input type="checkbox"/> | 100 |
| D18S51 | <input type="checkbox"/> | <input type="checkbox"/> | 100 |

[Bulk Edit] [Ignored Loci...] [Detection Thresholds...]

Stutter max: 0.3 Drop-in cap: 150 Allelic Variance: 6.518, 1.439

Forward stutter max: 0.1 Drop-in frequency: 0.0013 Stutter Variance: 3.354, 3.882

Degradation starts at: -1.0 Drop-in parameters: 0.0, 0.0 Var > mode: 0.5

Degradation max: 0.01 Saturation: 30000 Locus Amp Variance: 0.021

[Cancel] [Save Kit]

STRmix V2.5.11 - User:

Continued on next page

DNA: Interpretation of DNA Typing Results, Continued

Run parameters in STRmix™ (continued)

- Default parameters used for interpretation:

STRmix - Configure STRmix Defaults

Default Analysis Settings

MCMC settings

8 Number of chains

100000 Burn-in accepts (per chain)

50000 Post burn-in accepts (per chain)

9.0 Post burn-in shortlist

0.005 Random Walk SD

Default Kit: Fusion6C_3500

Inputs and Outputs

☐ Extended Output

Gelman Rubin

☐ Auto continue?

1.2 GR Threshold

10000 Extra Iterations

Likelihood Ratio

1000 HPD iterations

99.0 Quantile

1 Sides

☒ Factor of N! LR

☒ Include MCMC uncertainty

Performance settings

4 Number of threads

☐ Low Memory Mode

Default Results Directory: Results

User Interface Settings

Default Text File Directory: C:\ProgramData\STRmix\Results

Default STRmix File Directory: C:\ProgramData\STRmix\Results

☐ Allow Text Sample Renaming

STRmix V2.5.11 - User:

Continued on next page

DNA: Interpretation of DNA Typing Results, Continued

Run parameters in STRmix™ (continued)

- Population data used for LR calculations (African-American shown):

STRmix - Manage Populations

Manage Populations

Population:

Population Name:

Allele Frequency File:

Population Proportion:

Default FST:

Population Size:

Children Per Family:

| | | | |
|------------|----------------------------------|--------------|----------------------------------|
| Siblings | <input type="text" value="0.0"/> | Niece/Nephew | <input type="text" value="0.0"/> |
| Parents | <input type="text" value="0.0"/> | Grandparent | <input type="text" value="0.0"/> |
| Children | <input type="text" value="0.0"/> | Grandchild | <input type="text" value="0.0"/> |
| Uncle/Aunt | <input type="text" value="0.0"/> | Cousin | <input type="text" value="0.0"/> |
| Unrelated | <input type="text" value="1.0"/> | | |

STRmix V2.5.11 - User:

Continued on next page

DNA: Interpretation of DNA Typing Results, Continued

Run
parameters in
STRmix™
(continued)

- Population settings used for LR calculations:

STRmix - Population Settings

Case Details: test - test Open Results Directory

Population Settings

NIST1036_AfAm Add Population Remove Population

| Population | Proportion | FST | Allele Freq File |
|----------------|------------|----------------|---------------------------|
| NIST1036_AfAm | 0.25 | 0.01b(1.0,1.0) | 1036-Revised-Allele-Fr... |
| NIST1036_Asian | 0.25 | 0.01b(1.0,1.0) | 1036-Revised-Allele-Fr... |
| NIST1036_Cauc | 0.25 | 0.01b(1.0,1.0) | 1036-Revised-Allele-Fr... |
| NIST1036_Hisp | 0.25 | 0.01b(1.0,1.0) | 1036-Revised-Allele-Fr... |

Factor N!

☒ Factor of N! LR

Sampling Variation

☒ Calculate HPD ☒ Include MCMC uncertainty

HPD iterations: 1000 Quantile: 99 Sides: 1

Save as Default Populations Cancel Back Start Start & Search

STRmix V2.5.11 - User:

Continued on next page

DNA: Interpretation of DNA Typing Results, Continued

Running STRmix™ - new deconvolution

This is the procedure for running a new deconvolution in STRmix™.

| Step | Action |
|------|--|
| 1 | Select Start Analysis . |
| 2 | Enter the case number into the Case Number field. |
| 3 | The Sample ID should be the item number and should also include any run specific information, such as deconvolution, assumed contributors, and comparisons. |
| 4 | Case Notes may be used for other important information, such as loci ignored. Optionally, the item description can be entered here. |
| 5 | Enter the number of assumed contributors in the Number of Contributors box. |
| 6 | The number of burn-in accepts (per chain) and post burn-in accepts (per chain) should be left at the default of 100,000 and 50,000, respectively. These values may be raised as needed during troubleshooting of STRmix™ diagnostics (see Reviewing STRmix™ outputs). |
| 7 | Click on Run Settings if Informed priors are needed (see Informed priors). |
| 8 | Select Confirm to proceed. |
| 9 | Add the appropriate evidence profile. |
| 10 | Reference profiles may be added for the purposes of comparing a POI or assuming a contributor. If assuming, ensure that the reference is included in both the H_1 and the H_2 columns. If a profile is not added to the Reference profile section, STRmix™ will deconvolute the profile, but an LR will not be calculated. |
| 11 | The DNA kit used should be Fusion6C_3500. |

Table continued on next page

DNA: Interpretation of DNA Typing Results, Continued

Running STRmix™ - new deconvolution (continued)

| Step | Action |
|------|--|
| 12 | <p>Click on Kit Settings to ignore loci or change the degradation max.</p> <ul style="list-style-type: none"> • The ignore loci function can be used to leave out loci during the deconvolution. This may be necessary if: <ul style="list-style-type: none"> – The evidence and/or reference profile has a tri-allele – The evidence and/or reference profile has an off-ladder allele that cannot be given an allele designation – The profile used for comparison has one or more loci with partial information (i.e. one allele typing below the stochastic threshold of 450 rfu). Loci with <i>no</i> alleles typing do not need to be ignored, but can be if desired. • A small subset of degraded mixtures may benefit from increasing the degradation max setting to 0.1. This may be necessary if: <ul style="list-style-type: none"> – The allelic variance parameter is higher than expected – The degradation values proposed on the STRmix™ output are low compared to the expectation given profile morphology – The log (likelihood) value is low for the type of profile being analyzed – The amplification efficiencies display as a steep “ski slope” pattern. |
| 13 | <p>Select Confirm settings.</p> <p>NOTE: Default population settings should be used if calculating an LR (NIST1036_Afam, NIST1036_Asian, NIST1036_Cauc, and NIST1036_Hisp).</p> |
| 14 | Select Start & Search to initiate sample deconvolution and a search of the employee database. |
| 15 | After each deconvolution, a report will be saved into the Results folder, along with all of the files used for the deconvolution. |
| 16 | The steps above may be achieved individually or in batch mode, which queues up the deconvolution of multiple profiles (see Running STRmix™ – LR Batch). |
| 17 | Select Finish to return to the STRmix™ home screen. |

Continued on next page

DNA: Interpretation of DNA Typing Results, Continued

Running STRmix™ - LR from previous analysis

If a profile has previously been deconvoluted, then the genotype probability distributions can be used to calculate an LR by comparing to additional references as they are obtained without the need to reanalyze the profile.

| Step | Action |
|------|--|
| 1 | Select LR from Previous . |
| 2 | Click Select File and navigate to the desired results folder for the deconvolution of interest. Select the config.xml file. As an alternative, the entire results folder can be dragged and dropped into the Previous Analysis box. |
| 3 | Edit the sample ID and case notes with appropriate run information. |
| 4 | Select Confirm to proceed. |
| 5 | Add reference profile(s) for comparison/LR calculation. NOTE: An assumed contributor cannot be edited at this stage. Changes to the assumed contributor(s) can be performed by starting a new analysis. |
| 6 | The ignore locus function can be accessed, if needed, by clicking on Kit Settings . |
| 7 | Select Confirm settings . |
| 8 | Default population settings should be used. Select Start . |
| 9 | When complete, a report will be saved into the Results folder, along with all of the files used for the run. |
| 10 | Select Finish to return to the STRmix™ home screen. |

Continued on next page

DNA: Interpretation of DNA Typing Results, Continued

Running STRmix™ - LR Batcher

The LR Batcher tool allows for one or more deconvolutions to be compared to one or more reference samples.

| Step | Action |
|------|---|
| 1 | Select Tools . Select LR Batcher . |
| 2 | Select Add Analysis and navigate to the desired Results folder for the deconvolution of interest. Select the config.xml file. As an alternative, the entire results folder can be dragged and dropped into the Deconvolutions box. Repeat this step for additional deconvolutions, if necessary. |
| 3 | Select Add Reference (or drag and drop) to add as many reference sample files as desired. An LR, using the default LR Settings, will be calculated for every reference sample compared to every deconvolution. A results folder and report will be generated for each comparison. |
| 4 | For the Seed setting, make sure the Random box is checked. |
| 5 | Click Start . |

Reviewing STRmix™ outputs

Review the results of the database search to check for contamination from an individual in the employee database. The employee database consists of current and some former employees of the laboratory. The search will return an LR value using the Caucasian database for any database individual giving an LR greater than 1000. The LR proposition is as follows:

- H_1 : DNA originates from database individual and N-1 unknowns
- H_2 : DNA originates N unknowns

If an $LR > 1000$ is obtained, the analyst should attempt to determine the reason for the association. A STRmix™ comparison of the target employee to the sample may be done, and if the associated LR is less than 100, the database “hit” can be disregarded, unless contamination is truly suspected. A STRmix™ comparison will return a lower LR than the database search because the database calculates LRs without the use of HPD (highest posterior density), and it only returns values from one population group.

If the STRmix™ LR is greater than or equal to 100, and a means of contamination exists, then contamination should be investigated. Follow up with the Technical Leader if contamination is suspected.

Continued on next page

DNA: Interpretation of DNA Typing Results, Continued

Reviewing STRmix™ outputs (continued)

In the report of the STRmix™ deconvolution, there are numerous diagnostics that may indicate that an interpretation has not completed as expected.

The primary diagnostics include:

- genotype weights
- mixture proportions
- LRs

The secondary diagnostics include:

- total accepts
 - effective sample size
 - average log (likelihood)
 - Gelman-Rubin value
 - allele and stutter variances
-

Review of the primary diagnostics

- **Mixture proportions:** STRmix™ provides DNA amounts for each contributor in the profile (expressed in rfu units). These amounts are used to calculate mixture proportions. Reporting mixture proportions replaces the previous designation of major and minor contributors to a mixture. These mixture proportions should meet qualitative expectations when compared to the observed profile.
 - **LRs:** If a review of the individual LR_s reveals an inclusion or an exclusion of a contributor that is not intuitive, or if inclusionary LR_s are obtained for every locus with one or two loci giving a low (even 0) LR, this could be due to:
 - extreme stochastic events at one or two loci, either large unexpected drop-in or large unexpected dropout
 - an error in the input file (e.g. pull-up peak left in the input file)
 - a partial genotype in a reference sample (and the locus has not been ignored)
 - an incorrect number of contributors assumed, especially one less than the true number
 - **Genotype weights:** Where possible, the weights generated by STRmix™ should be assessed at each locus for their intuitiveness.
-

Continued on next page

DNA: Interpretation of DNA Typing Results, Continued

Review of the secondary diagnostics

No single run diagnostic alone is demonstrative of a problem with the STRmix™ deconvolution. When multiple diagnostics are affected (e.g., very low acceptance rate with very high Gelman-Rubin value), that may be indicative of a problem with the STRmix™ deconvolution and warrant further action. The secondary diagnostics are found in the Run Information section of the report as shown in the example below.

POST BURN-IN SUMMARY

| | | | |
|-------------------------------------|---------|---------------------------------|-----------|
| Total iterations | 1436078 | Acceptance rate | 1 in 3.59 |
| Effective sample size | 9966.88 | log(likelihood) | 20.53 |
| Gelman-Rubin convergence diagnostic | 1.05 | | |
| Allele variance (mode = 7.94) | 6.402 | Stutter variance (mode = 9.138) | 11.469 |

- **Total iterations (Acceptance Rate):** The total iterations value is the number of post burn-in iterations run by the MCMC in order to reach the target of 500,000 accepts (100,000 burn-in plus 400,000 post burn-in accepts). The acceptance rate is 400,000 divided by the total iterations.
 - An acceptance rate of 1 in thousands or millions may indicate that the analysis needs to be re-run with additional accepts.
- **Effective sample size:** Effective sample size (ESS) is the number of independent steps the MCMC has taken from the posterior distribution of all parameters.
 - A low ESS in relation to the total number of iterations suggests that the MCMC has not moved very far with each step or has had a low acceptance rate.
 - A low absolute value of ESS (e.g. 10s or 100s) will mean that there is potential for a large difference in weights if the analysis was run again. This will be taken into account during HPD interval generation in any LR calculations unless the genotype sets are completely resolved on a single combination, in which case there will be no effect of ESS on the HPD interval.
 - A low ESS on its own is not an indication that reanalysis is required.

Continued on next page

DNA: Interpretation of DNA Typing Results, Continued

Review of the secondary diagnostics (continued)

- **Average (log) likelihood:** This value is the log of the average probability value created at each of the post burn-in MCMC iterations.
 - The larger this value, the better STRmix™ has been able to model the observed data.
 - A low or negative value may suggest that STRmix™ has not been able to model the data well, given the information provided. Reasons for this include:
 - The profile is very low level and there is little data making up the likelihood.
 - The incorrect number of contributors was entered and there are forced stochastic events in the STRmix™ run as a result (e.g. large heterozygote peak imbalances or variation in mixture proportions across the profile).
 - Data has been removed that was real, particularly stutter peaks, and must now be described by dropout.
 - Artifactual peaks have been left in the STRmix™ input and must be accounted for by drop-in.
- **Gelman-Rubin (GR) convergence diagnostic:** This diagnostic is a measure of how well the eight STRmix™ chains carrying out the MCMC have converged on a final profile deconvolution.
 - For a fully converged analysis, the GR value should be 1.
 - A GR value above 1.2 by itself should not be interpreted as indicating a problem with the run. The GR value should be reviewed in conjunction with other STRmix™ diagnostics. It has been found that profiles with a high number of contributors often produce larger GR values.

Continued on next page

DNA: Interpretation of DNA Typing Results, Continued

Review of the secondary diagnostics (continued)

- **Allele variance and Stutter variance:** These values are the average value for allele variance and stutter variance constants across the entire post burn-in MCMC analysis. These values can be used as an indication of the level of stochastic variation in peak heights present in the profile.
 - For the Fusion 6C kit, the allele variance mode is 7.940, and the stutter variance mode is 9.138.
 - If the allele or stutter variance constant values are significantly higher than the mode, this may indicate that the DNA profile is of poor quality or that an incorrect number of contributors was assumed.
 - Used in conjunction with the average log (likelihood), a large allele or stutter variance value can indicate poor PCR.
 - The deconvolution of low level samples may result in a low average log (likelihood) and average allele and stutter variance constants.
 - If some data has been inappropriately omitted, artifacts left in, or the profile was otherwise misinterpreted, the deconvolution should result in a low or negative average log (likelihood) and high allele and/or stutter variances relative to the mode.
-

Troubleshooting

If the primary or secondary diagnostics indicate that further work should be done, then a number of options are available (listed below).

NOTE: Additional work requires a new STRmix™ run.

- Re-amplify the sample.
- Re-inject or re-run the sample on the genetic analyzer.
- Edit the input file if an artifactual peak has been left in.
- Increase the number of MCMC accepts (requires documented DNA Technical Leader approval).
- Change the number of contributors.
- Introduce informed priors into the interpretation. See [Informed priors](#).

If a sample is re-run using any of the previously listed options, the reasons should be documented in the case file. The analyses of all STRmix™ runs should be maintained in the case electronic data. Based on available case information, the analyst may choose to report all interpretations or the one that best meets scientific expectations.

Continued on next page

DNA: Interpretation of DNA Typing Results, Continued

Informed priors

Informed priors are a means by which an analyst can provide STRmix™ with a starting-point for contributor ratios. See the *STRmix™ Operation Manual* for specific instructions on the use of the informed prior functionality. This function may be especially useful if the evidentiary mixture is likely to be comprised of known first-order relatives.

The use of informed priors requires documented DNA Technical Leader approval. If informed priors are used, the Run Information section of the report must be included in the case notes, indicating the mean and the variance for each contributor.

Interpretation-Y23

The following are the steps for Y23 profile interpretation.

| Step | Action |
|------|--|
| 1 | Identify the profile as being from a single contributor or being from multiple contributors. |
| 2 | Determine the assumed number of contributors. |
| 3 | Identify alleles from an assumed known contributor, if applicable. |
| 4 | Determine the approximate ratio of contributors to the mixture to evaluate if a major/minor contributor proportion exists. |
| 5 | Identify the mixture type. |
| 6 | Depending on the mixture type, identify haplotypes that must be present in the profile of a true contributor. |
| 7 | Compare the interpreted evidence profile with probative reference profiles and perform statistical analysis, if appropriate. |

Continued on next page

DNA: Interpretation of DNA Typing Results, Continued

Y23

Interpretation –
Step 1

Identify the profile as being from a single contributor or being from multiple contributors.

- **Single contributor profiles** - A profile may be considered to be from a single male contributor when the observed number of alleles at each locus does not exceed one, except for DYS385 a/b, which may have one or two alleles.
 - The stochastic threshold for DYS385 is 400 RFUs. A single peak below this threshold must not be designated as a homozygote.
 - If two alleles are detected at DYS385, the peak height ratio between the two alleles should be $\geq 55\%$ from a single contributor.
 - **Multiple contributor profiles** - A sample is generally considered to be a mixture of more than one male if two or more alleles are present at two or more single-copy loci. All loci should be evaluated in making this determination. However, it may be difficult to discern minor alleles from stutter so these mixtures should be interpreted with caution.
 - Dropout of a contributor is possible if any alleles fall below 400 RFUs.
-

Y23

Interpretation –
Step 2

Determine the assumed number of contributors.

The following components of the mixture should be examined to estimate the number of contributors:

- the locus (or loci) that contain the greatest number of alleles
- possible peaks that fall below the 150 RFU detection threshold
- elevated stutter peaks

If possible, the analyst should assume a specific number of contributors for the mixture or for a component of the mixture. If necessary, the analyst can interpret a profile using more than one assumption about the number of contributors. Any assumptions about the number of contributors should be documented in the notes and stated in the report.

At times, for samples with very little DNA, there may be so much allele dropout that it is difficult to estimate the number of contributors. If the number of contributors cannot be reliably estimated, the profile may be determined to be uninterpretable.

Continued on next page

DNA: Interpretation of DNA Typing Results, Continued

Y23
Interpretation –
Step 3

Identify alleles from an assumed known contributor, if applicable.

In some cases, mixture profiles contain DNA from a male who is reasonably expected to be a contributor to the mixture. Whenever possible, the genetic profile of the assumed known contributor can be used to infer the genetic profile of the remaining contributor(s) in the case. See *Identifying possible assumed contributors* for examples of known contributors.

Any assumptions that account for the subtraction of the profile of a known contributor should be documented in the notes and stated in the report. Statistical associations will not be provided for the profile of an individual whose DNA is reasonably expected to be on an item.

Y23
Interpretation –
Step 4

Determine the approximate ratio of contributors to the mixture to evaluate if a major/minor contributor proportion exists.

A donor ratio is the ratio of DNA proportions of multiple individuals to a mixed DNA typing result. The donor ratio is calculated using the following formula:

$$\left(\begin{array}{c} \text{RFU contribution} \\ \text{of apparent Major} \\ \text{component} \end{array} \right) : \left(\begin{array}{c} \text{RFU contribution} \\ \text{of apparent Minor} \\ \text{component} \end{array} \right)$$

To calculate the donor ratio for a 2-person mixture, identify the loci which do not appear to have allele sharing between contributors:

- For all loci except DYS385 a/b, divide the RFU value of the larger allele by the RFU value of the smaller allele.
 - Example: Alleles with RFUs of 2000 and 500.
2000 : 500 =
4 : 1 ratio
-

Continued on next page

DNA: Interpretation of DNA Typing Results, Continued

Y23

Interpretation – Step 4 (continued)

- For DYS385 a/b:
 - If there are four alleles, sum the RFUs of the two larger alleles divided by the sum of the RFUs of the two smaller alleles.
 - Example: Alleles with RFUs of 1000, 800, 250 and 200.
$$(1000 + 800) : (250 + 200) =$$
$$4 : 1 \text{ ratio}$$
 - If there are three alleles for which the major contributor appears to be homozygous and the minor contributor appears to be heterozygous, divide the RFU of the largest allele by the sum of the RFUs of the two smaller alleles.
 - Example: Alleles with RFUs of 2500, 260 and 240.
$$2500 : (260 + 240) =$$
$$5 : 1 \text{ ratio}$$

If there is an assumed known contributor to the mixture, additional loci with unshared alleles may be identified and used for donor ratio calculations.

Determining the donor ratio in a 3+ male mixture can be very difficult unless there are clear differences in donor contribution or if a known contributor is assumed to be present.

In a 3+ male mixture, a donor ratio approximation may be calculated between the major component alleles and the minor component alleles.

Profiles showing signs of degradation (especially differential degradation) may lead to unreliable donor ratio calculations.

The donor ratio can and does vary from one locus to another within one amplification. To estimate the donor ratio across the entire profile, average the individual donor ratios obtained from all applicable loci.

Continued on next page

DNA: Interpretation of DNA Typing Results, Continued

Y23
Interpretation –
Step 5

Identify the mixture type.

To determine the mixture type, consider

- the assumed number of contributors and
- the relative DNA proportions of each contributor resulting in a mixture that is either resolvable or irresolvable

A resolvable mixture is a DNA typing result from a mixed sample for which alleles can be attributed to their relative contributor(s).

An irresolvable mixture occurs when

- relatively similar amounts of DNA ($< \sim 4:1$ ratio) are contributed to the sample by two or more unknown contributors or
- the quality of the profile prevents the analyst from confidently resolving the mixture

There are three mixture types as follows:

Type 1 Mixture - A Type 1 mixture consists of DNA from two males and the profiles of the two individuals can be resolved into their individual components resulting in major/minor or known/unknown contributors.

NOTE: When resolving a 2-person mixture, the terms “major contributor” and “minor contributor” can be used for mixtures with a donor ratio average greater than $\sim 4:1$. At a $4:1$ ratio, the major contributor’s alleles are expected to be “major” across the entire profile. When resolving a mixture with a known contributor, the terms “known” and “unknown” can be used when the donor ratio average is less than $\sim 4:1$.

A Type 1 mixture must have one of the following features:

- The average donor ratio of the two contributors is $\sim 4:1$ or greater.
- A mixture with an assumed known contributor can be resolved into known/unknown contributors regardless of the donor ratio.

Continued on next page

DNA: Interpretation of DNA Typing Results, Continued

Y23
Interpretation –
Step 5
(continued)

Type 2 Mixture - A Type 2 mixture consists of DNA from three or more males and the average donor ratio allows for a major component to be separated from a minor component.

NOTE: When interpreting a Type 2 mixture, the terms “major component,” and “minor component” can be used. “Component” describes a portion of the mixture and may include one or more contributors. The terms used to describe each of the components should be defined in the case notes.

For mixtures of three or more contributors, reliable donor ratio calculations may be difficult. The analyst may use one or more of the following to determine if the mixture can be resolved into its relative components:

- donor ratio approximations
- the presence of an assumed known contributor, if applicable

In a 3+ person mixture, the analyst should make the following observations before determining if the mixture can be resolved into its relative components:

- the quality of the profile, including signs of degradation, differential degradation, inhibition, etc.
- the presence of alleles below 400 RFUs
- the effects of allele sharing

Type 3 Mixture - A Type 3 mixture consists of DNA from two or more males and the mixture *cannot* be resolved into its individual components.

- A Type 3 mixture has an average donor ratio of ~ 4:1 or less with no assumed known contributors.

Continued on next page

DNA: Interpretation of DNA Typing Results, Continued

Y23

Interpretation –
Step 6

Depending on the mixture type, identify haplotypes that must be present in the profile of a true contributor.

Type 1 Mixture Interpretation - A Type 1 mixture consists of DNA from two males and the profiles of the two individuals can be resolved into their individual components resulting in major/minor or known/unknown contributors.

- If the average donor ratio is $\sim 4:1$ or greater, the haplotypes of the major and minor contributors can be interpreted separately.
- Single copy loci with a single allele typing must be interpreted with caution for the minor contributor because dropout must be considered. If any minor-contributor alleles are equal to or below a peak height of 400 RFU, loci containing a single allele are inconclusive for the minor contributor.
- If a known contributor can reasonably be assumed, subtract his allelic contribution to determine the contribution of the unknown contributor.

At times, a profile as a whole may be considered resolvable, but there may be individual loci which are irresolvable. If a locus is irresolvable, the analyst may mark the locus inconclusive. Consequently, the locus cannot be used for comparison with reference profiles.

Type 2 Mixture Interpretation - A Type 2 mixture consists of DNA from three or more males and the average donor ratio allows for a major component to be separated from a minor component. Listed below are some examples of Type 2 mixtures:

- one major contributor, multiple minor contributors (minor component)
- multiple major contributors (major component), one minor contributor
- multiple major contributors (major component), multiple minor contributors (minor component)

The interpretation approach will vary depending on

- the total number of assumed contributors
- the assumed number of contributors within each major/minor component
- the presence of an assumed known contributor

Continued on next page

DNA: Interpretation of DNA Typing Results, Continued

Y23

Interpretation – Step 6

(continued)

If the major or minor component can be assumed to be from two individuals, then the interpretation of that component can follow the Type 1 mixture interpretation approach.

Dropout of a contributor within a component is possible if any alleles are below 400 RFUs.

Type 3 Mixture Interpretation - A Type 3 mixture consists of DNA from two or more males and the mixture *cannot* be resolved into its individual components. These mixtures can only be used for exclusionary purposes at loci with no dropout.

Y23

Terminology

The following terminology will be used for the interpretation of Y23 profiles and in the results table:

| Terminology | Explanation | Example | Included haplotypes |
|---|--|-----------------------------------|---------------------|
| allele ₁ | Single copy haplotype | 11 | 11 |
| allele ₁ , allele ₂ | Duplicate copy haplotype (DYS385) | 9, 10 | 9, 10 |
| allele ₁ + | Obligate allele paired with anything at a duplicate copy locus | 10+ | 10, anything |
| inc | Not suitable for comparison with reference profiles | inc | - |
| - | No signal above 150 RFU | - | - |
| * | All detected alleles | List all alleles at a given locus | All alleles listed |

Y23

Interpretation – Step 7

Compare the interpreted evidence profile with probative reference profiles and perform statistical analysis, if appropriate.

Refer to [DNA: Reporting DNA Typing Results](#) for the appropriate reported conclusions and information on how to perform the statistical analysis based on these comparisons.