# STRmix™ v 2.7 Probabilistic Genotyping Software Operating Instructions

## 1 Guiding Principles and Scope

1.1 This procedure describes the use of STRmix™ v2.7 for the interpretation of PowerPlex® Fusion DNA profiles run on 3500xL or 3130xL Genetic Analyzers within the NYC OCME Department of Forensic Biology. Readers are also referred to the STRmix™ v2.7 Users and Operation manuals for additional information.

1.2 For STRMix™ set-up instructions please refer to QC702a – STRmixTM v2.7 Set-Up Instructions.

## 2 Preparing Data for a STRmix™ Analysis

2.1 Before performing your STRmix™ analysis, the following actions must be taken:

2.1.1 Verify that the sample is suitable for STRmix™ analysis (refer to Interpretation of PowerPlex® Fusion data run on 3500xL).

2.1.2 Evaluate your replicates, if applicable. If there are drastic inconsistencies with the alleles present between replicates and/or one has little information, only the amplification with the most information should be used, or a third amplification may be warranted.

2.1.3 Determine the best described Number of Contributors to the sample (NOC). Refer to Interpretation of PowerPlex® Fusion data run on 3500xL regarding the procedure for determining the number of contributors.

2.1.4 Create folder(s) with the FB (or FBS) case number for the STRmix™ runs within the STRmix Data folder: M:\STR_Data\STRmix Data.

- e.g. FB16-01234 or FBS16-05678

2.1.5 Confirm that the STR data is prepared correctly for STRmix™ analysis:

2.1.5.1 Evidence samples must only be amplified in PowerPlex® Fusion in order to undergo STRmix™ analysis. This procedure is specifically for evidence amplified in PowerPlex® Fusion data run on the 3500xL or 3130xL Genetic Analyzers at the NYC OCME Department of Forensic Biology.

2.1.5.2 Sample data must be assembled into the appropriate format for STRmix™ input. The standard input for STRmix™ is a .txt file. See the GeneMarker v3.0 Operation Manual for instructions on exporting data for STRmix™ input.

2.1.5.3 Evidence samples must be edited to remove all artifacts, including pull-ups, spikes, dye blobs, etc. before being input into STRmix™. **Back and forward stutter in 3130xL** data should not be removed, and **back, forward, half back, and double back stutter in 3500xL** data should not be removed before importing into STRmix™. Refer to STR Results Interpretation - PowerPlex® Fusion & STRmix™, Interpretation of PowerPlex® Fusion data run on 3500xL, and the Appendix for PowerPlex® Fusion Stutter.

2.1.5.4 Reference samples must be edited to remove all artifacts including stutter. Incomplete or tri-allelic loci must not be imported into STRmix™ for a reference sample – remove all allele(s) for that locus within the text file. If a possible drop-in peak is present in a reference sample, remove this peak from the text file before STRmix™ import.

2.1.5.5 Non-numeric values such as OL or OB, < or > are not permitted within the STRmix™ input files. Unambiguous alleles including those that are rare should appear in the corresponding input file as their actual allelic size designation, for example D21: 30.1. If an actual allelic size designation cannot be determined, the data for this locus should be removed completely from the text file and the locus should be ignored.

2.1.5.6 To modify a STRmix™ input text file: open the STRmix™ .txt file associated with the appropriate STR project (e.g. in Notepad or Microsoft Excel®). Locate the sample and locus containing the non-numeric value within the .txt file and manually replace the value with the appropriate actual allelic size designation. Save the .txt file **replacing the original file**.

2.1.5.7 An attempt should be made to amplify reference samples in PowerPlex® Fusion (see Case Management). If unavailable, STRmix™ allows the user to calculate a likelihood ratio when the evidence and reference samples are analyzed in different autosomal typing kits. LRs will only be calculated for those loci in common between the two kits.

2.1.5.7.1 If a reference sample was not amplified in Fusion, the data must be converted to a .txt file for import into STRmix™ with the locus order matching that of the evidence (PowerPlex® Fusion order). The reference sample data can be converted to the proper PowerPlex® Fusion order and appropriate .txt file format using the following macro: Identifiler to Fusion Exemplar STRmix Input Creation.

2.1.5.8 If a DNA donor is being used from one sample to condition or compare to another, use the following macro: Reference profile for STRmix Input Creation.

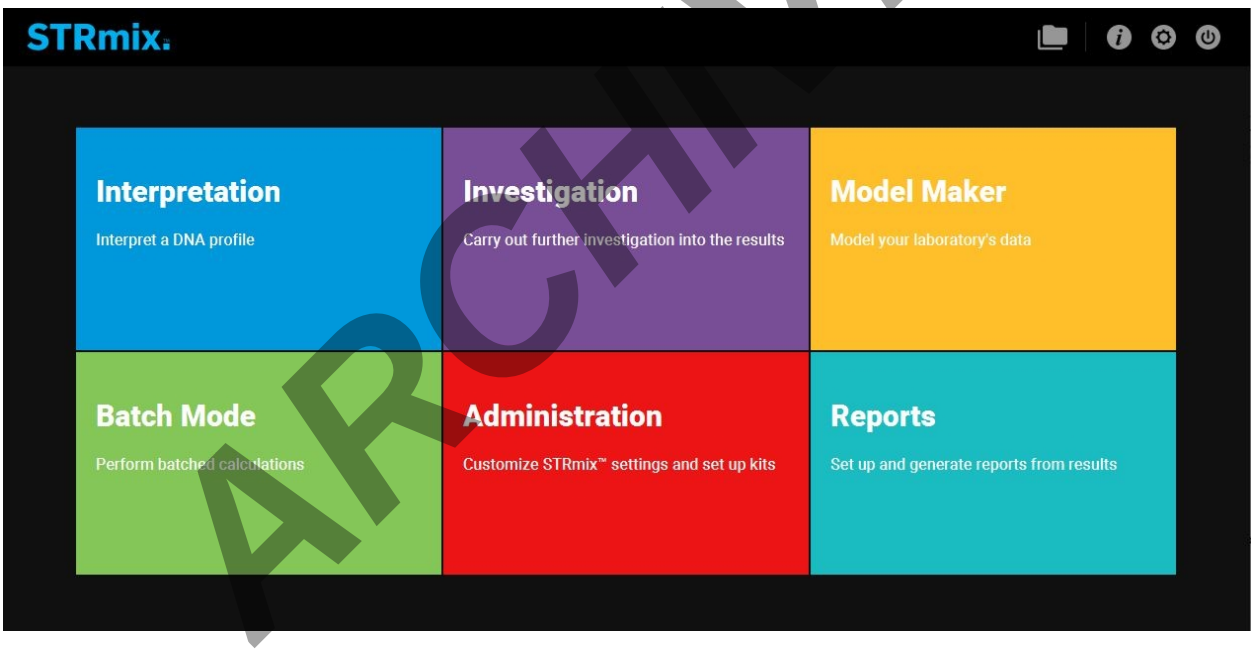
2.1.6 Evaluate evidence samples to determine if a locus needs to be ignored before STRmix™ deconvolution is performed. A comment should be added to the **Case Notes** field within the STRmix™ analysis to indicate why the locus was ignored for that run.

2.1.6.1 The following is a list of reasons data may need to be ignored at a particular locus. For any situation not covered in the list below, the technical leader should be consulted.

- Tri-allelic pattern
- Unresolved allelic or stutter peak that is visible above the AT
- OB/OL allele or stutter peak that cannot be assigned a correct allelic designation
- Stutter or allelic peak for an allele belonging to locus A is being called in a neighboring locus B (ignore both loci)
- Where a conditioning sample does not have data at a locus that is present in the evidence sample

2.1.7 Loci should not be ignored for likelihood ratio calculations due to a partial comparison sample. Ensure that your reference sample text file has been updated appropriately as described above in 2.1.5.

2.2 Launch the STRmix™ application and prepare the scenarios to be run in STRmix™. Open the STRmix™ software by locating STRmix™ in the task bar or by double clicking on the STRmix™ icon on the desktop. The main menu is shown below:





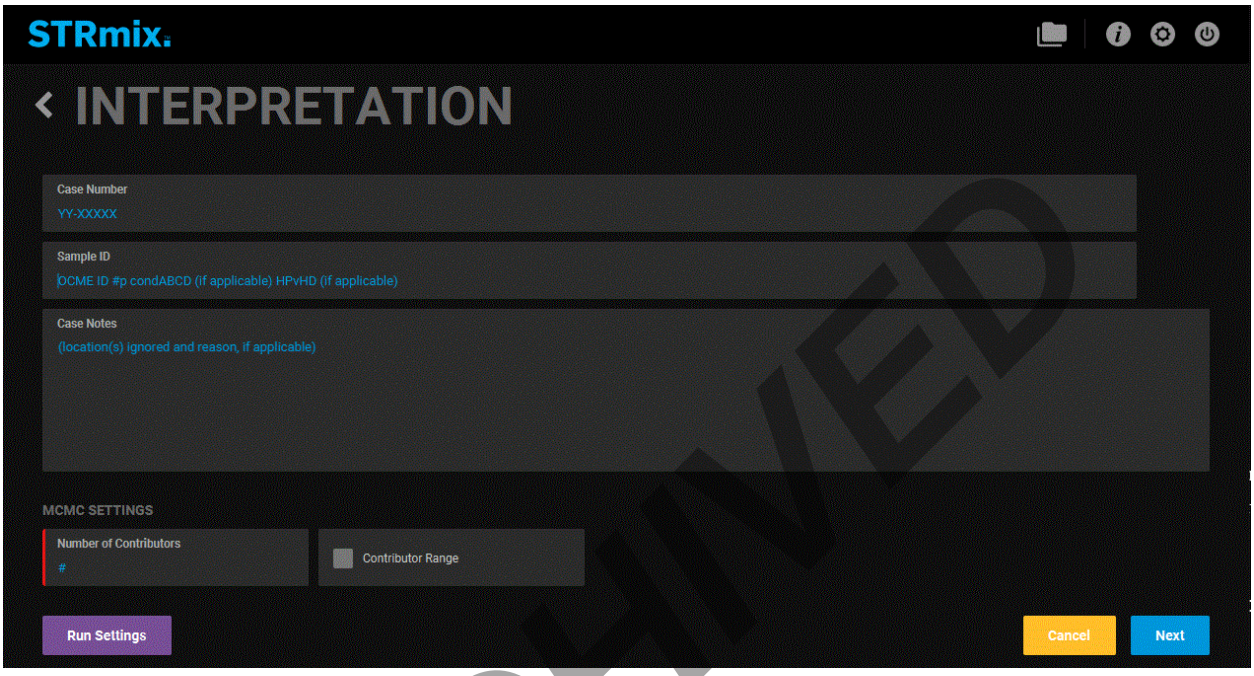

2.2.1 For deconvolutions of evidence profiles (Interpretation), go to Section 3.

2.2.2 For comparisons to previously deconvoluted profiles (Investigation: LR From Previous), go to Section 4.

2.2.3 To set up multiple STRmix™ runs to run consecutively in Batch Mode, go to Section 6.

# 3 Deconvolutions with STRmix™ (Interpretation)

3.1 Select **Interpretation** from the **Main Menu.** This will open the Interpretation Setup screen:






3.2 The STRmix™ output folder and file names are created by stringing together the values in the **Case Number** and **Sample ID** fields in the software followed by the date and time of the analysis run. The information in the **file name** is separated by dashes. Therefore, if other characters are entered, such as a comma, underscore, period, etc., the software will convert them into dashes.

3.3 Refer to Interpretation of PowerPlex® Fusion data run on 3500xL for guidance on when a **conditioned contributor** may be applied. A deconvolution of the evidence sample without conditioning and an LR against a potential conditioned contributor may need to be run first.

3.4 An LR may be run in conjunction with the deconvolution in the following scenarios:

- Single source evidence sample that did not require a STRmix™ deconvolution for determination of a profile, needing an LR to a matching comparison sample.
- Evidence mixture sample undergoing deconvolution and an LR check to determine if a reasonably expected reference sample can be used for further conditioning (ex. car owner on a swab from the steering wheel).
- To assess a probative comparison sample's presence within a mixture where that reference sample data is contained within the same evidence file (ex. victim's blood on suspect's clothing.)

3.5 The following naming conventions should be used for deconvolutions:

3.5.1 Evidence Files

- **Case Number** = YY-XXXXX (do not include "FB")
- **Sample ID** = remainder of evidence sample OCME ID* #p (NOC) condElimInitials (if applicable) scenario for LR (if applicable)
- **Case notes** = a comment should be added here if a locus is ignored, indicating the reason: e.g. "D2S441 was ignored due to an unresolved allelic peak"
- *Suffixes such as '_mcon' or '_reamp' should not be included in the OCME ID

3.5.2 Suspect Files

- **Case Number** = SYY-XXXXX (do not include "FB")
- **Sample ID** = evidence sample OCME ID* (include evidence file FB# without the "FB") #p (NOC) condElimInitials (if applicable) scenario for LR
- **Case Notes** = a comment should be added here if a locus is ignored, indicating the reason: e.g. "D2S441 was ignored due to an unresolved allelic peak"
- Suffixes such as '_mcon' or '_reamp' should not be included in the OCME ID

3.5.3 For LR scenarios, the naming format should start with the comparison sample's initials, followed by any conditioned samples' initials, and then the number and "U" for unknowns, followed by a "v" to separate the numerator hypothesis from the denominator hypothesis.

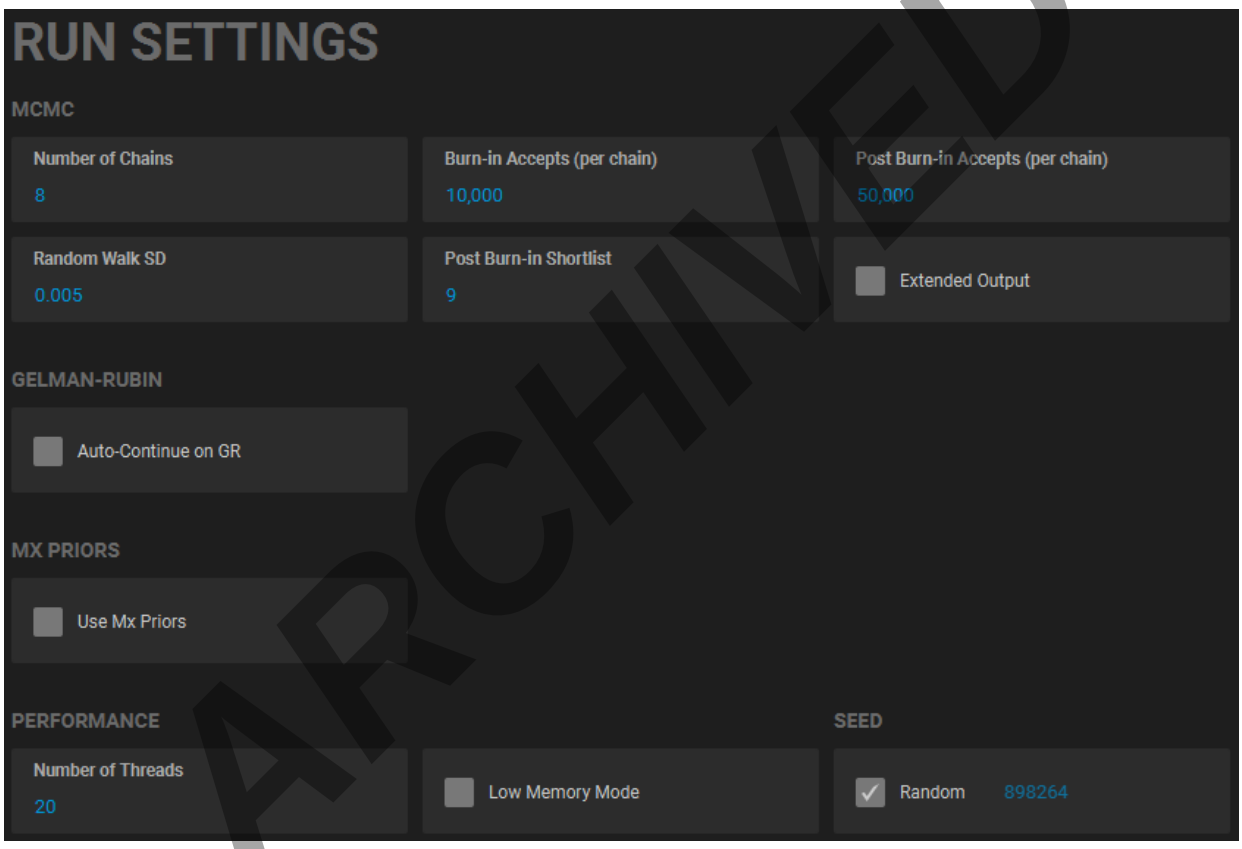
| Examples | Resulting STRmix™ file name |
|---|---|
| ***Evidence File*** | |
| 3-person, deconvolution, no conditioning, no comparisons | 16-01234-567-1-1-1-trig-GS-3p |
| 4-person deconvolution, no conditioning, comparing elimAB | 16-01234-567-1-1-1-shirt-BL-4p-AB3Uv4U |
| 3-person deconvolution, conditioning vic CD, comparing elimAB | 16-01234-567-1-1-1-shirt-BL-3p-condCD-ABCD1UvCD2U |
| 4-person deconvolution, conditioning elims CD, EF and GH, no comparisons | 16-01234-567-1-1-1-trig-GS-4p-condCDEFGH |
| ***Suspect File*** | |
| 1-person deconvolution, no conditioning, comparing suspTS | S16-05678-16-01234-567-2-1-1-slide-GS-1p-TSv1U |

3.6 Set the number of contributors and select Next.

3.7 Select Run Settings, and confirm the settings against the following screenshot. They should always be the same for every STRmix™ analysis unless an exception is listed below. Any changes that are made will appear in bold on the run report.

- **MCMC Settings**: Burn-in Accepts (per chain) and Post Burn-in Accepts (per chain) must not be modified without documented approval from the technical leader (or his/her designee).

- **Mx Priors**: this will not be used without documented approval from the technical leader (or his/her designee). See Section 3.8 for more information and set-up instructions for Mx priors.
- **Performance, Number of Threads**: it is okay to proceed if the Number of Threads does not match the screenshot below; this is specific to the computer being used.
- **Performance, Low Memory Mode**: This setting allows the computer to minimize the memory used for the run and can be turned on if a run fails to finish due to computer memory.





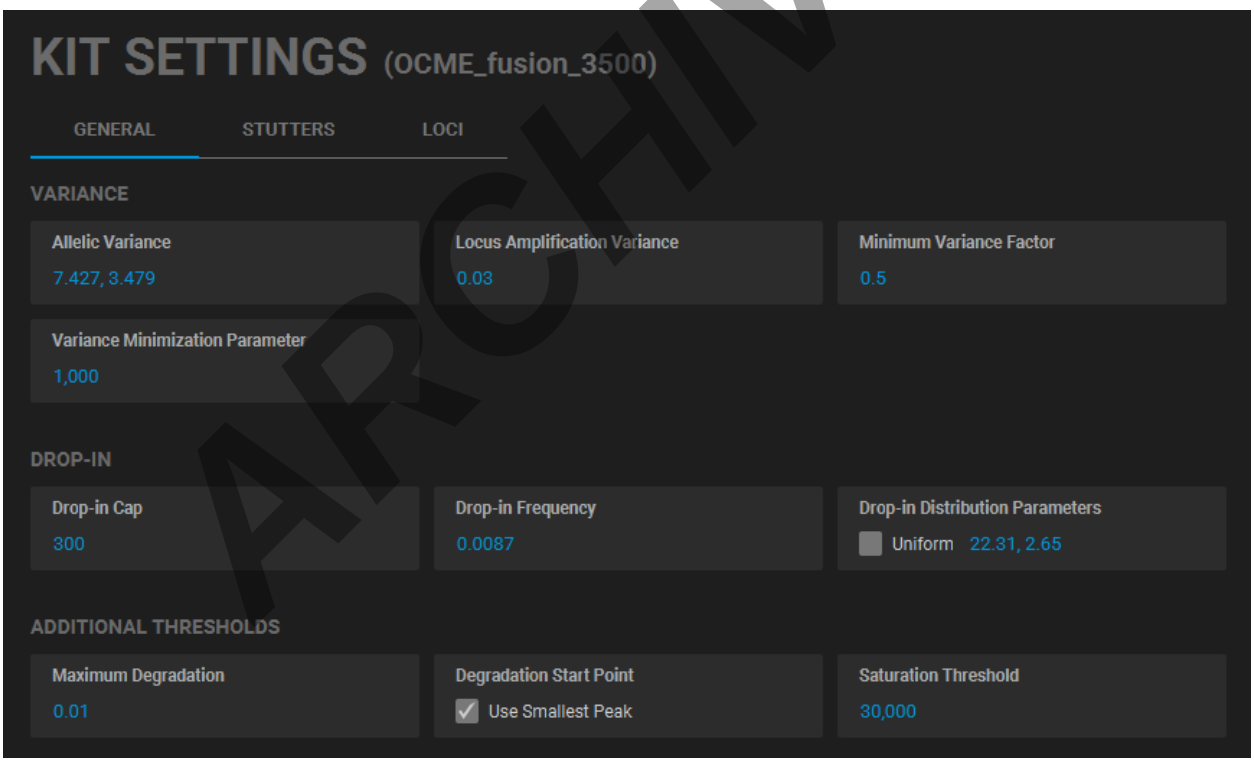

3.8 If you are not using Mx Priors, skip to 3.9; Mx Priors is used only with documented approval from the technical leader (or his/her designee). User Informed Mixture Proportion Priors (Mx Priors) is a function within STRmix™ that allows users to set approximate mixture proportion percentages for each contributor. If a proposed genotype does not fit the proportion percentage and variance set by the user, a penalty is applied to that iteration indicating an overall poor fit to the observed profile.

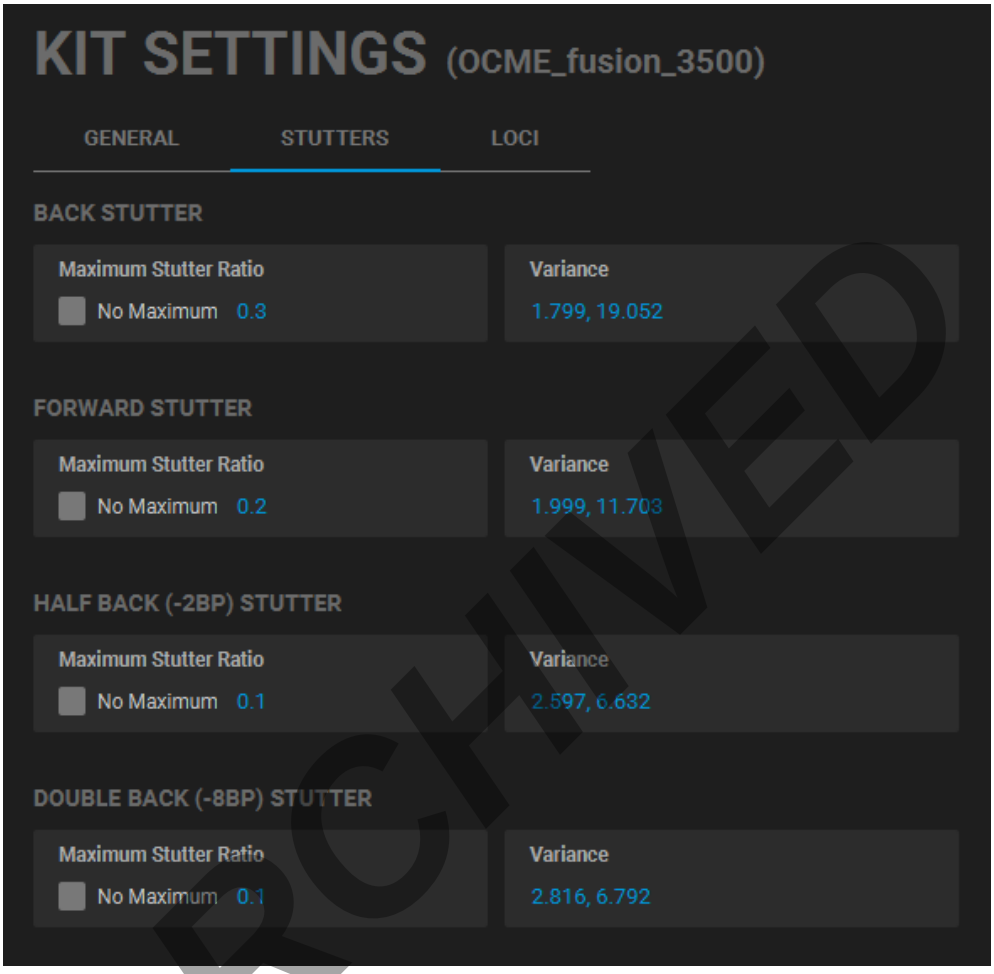
3.9 Select Cancel if all run settings are correct. If approved changes were made, Select Apply.

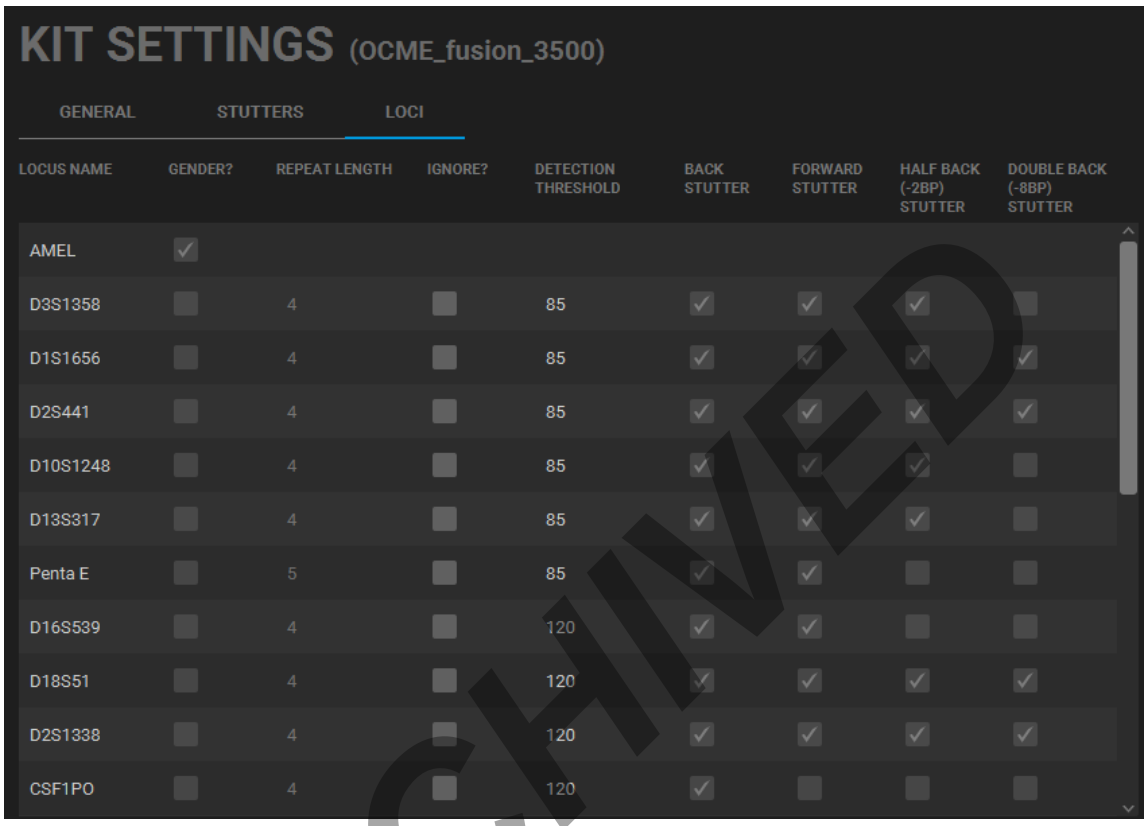
3.10 Check that the **Profiling Kit** selected is **OCME_fusion_3500** or **OCME_fusion_3130 depending on the data being analyzed**. The STRmix™ output folder and file name (based on the combined Case Number and Sample ID you previously entered) are located at the top right.

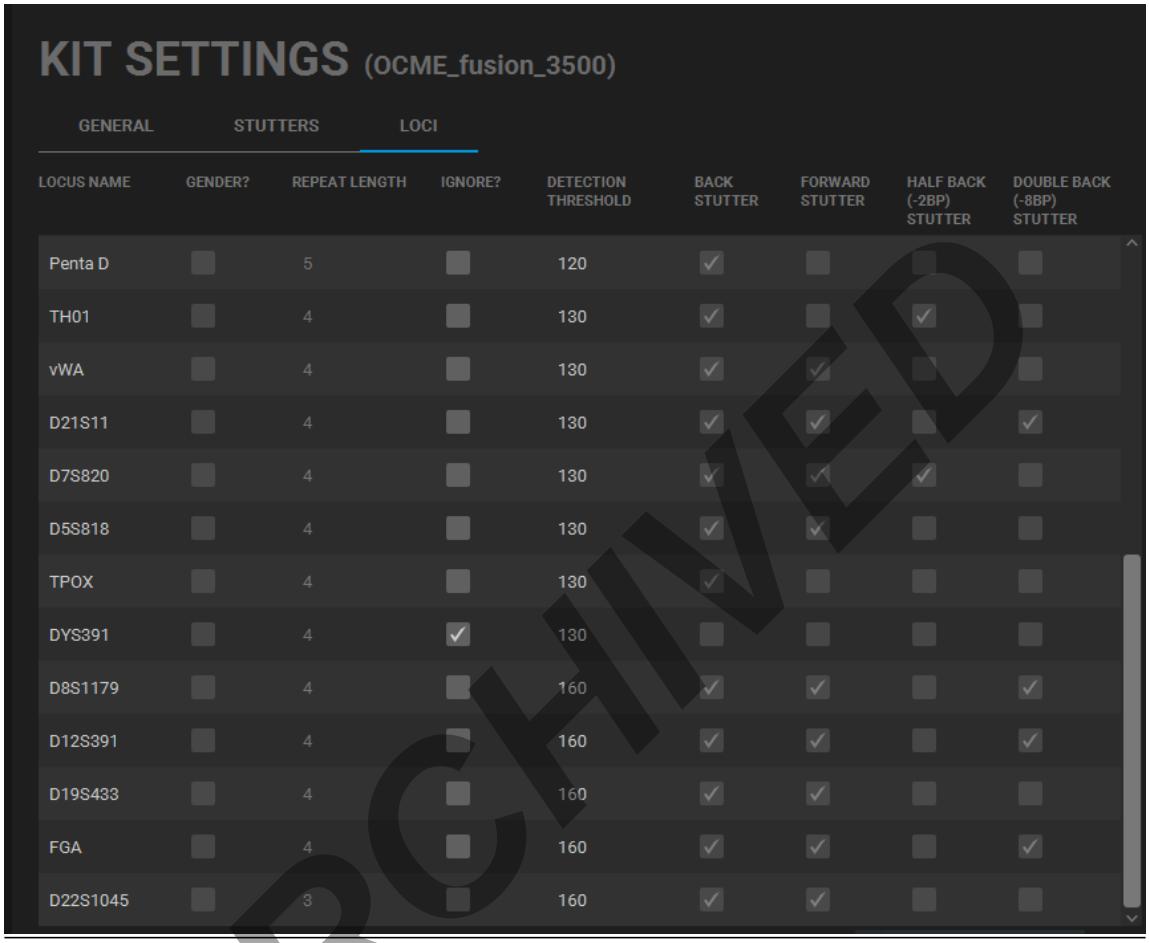

3.11 Click on Kit Settings. There are three tabs of settings to verify against the following screenshots in this window - **General**, **Stutters**, and **Loci**. If a locus needs to be ignored for a deconvolution, this is where you will be able to do that. See 3.12 for further instruction on ignoring a locus.

3.11.1 For the **OCME_fusion_3500 Profiling Kit**, ensure settings match the following:

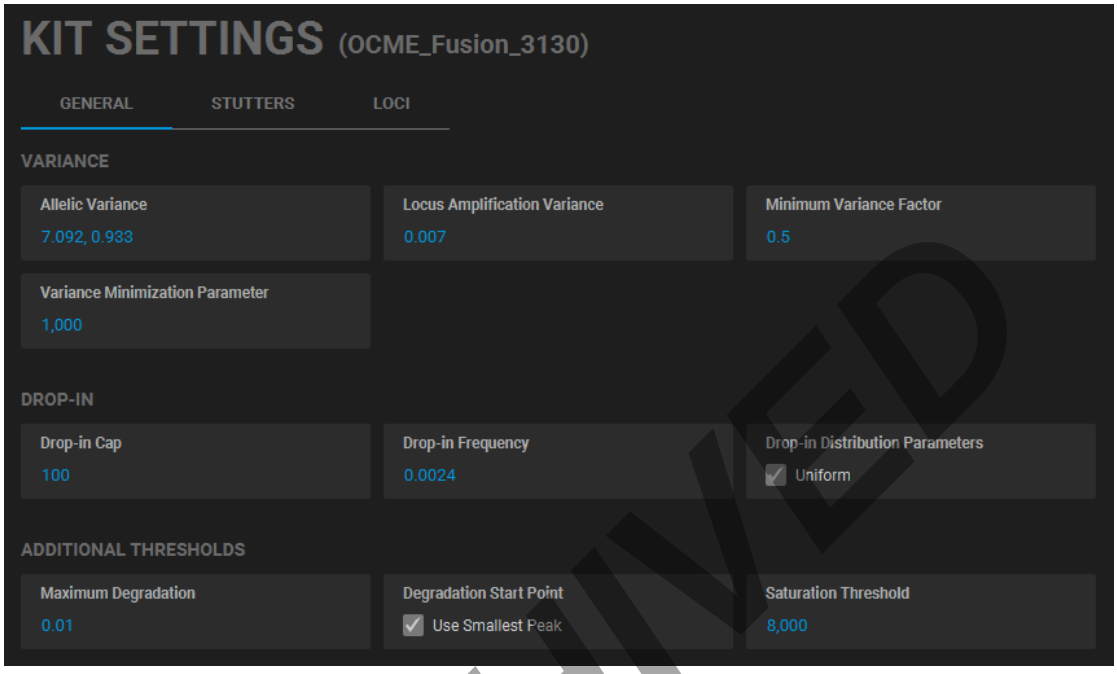


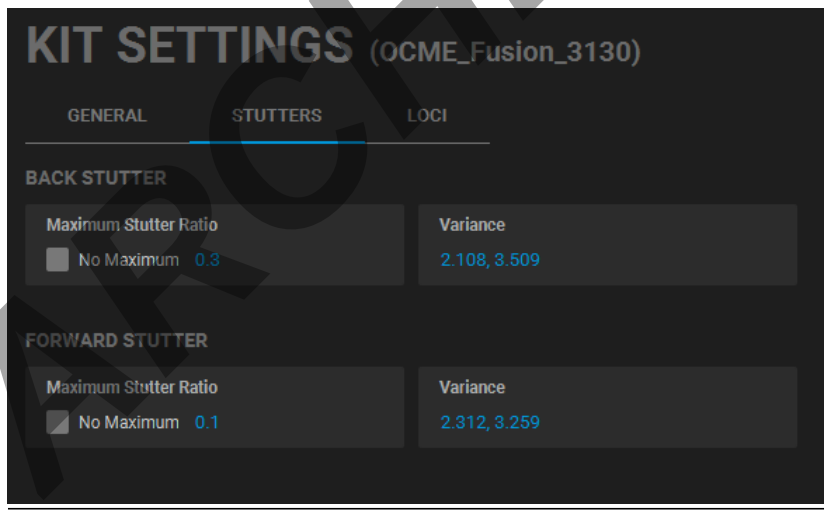


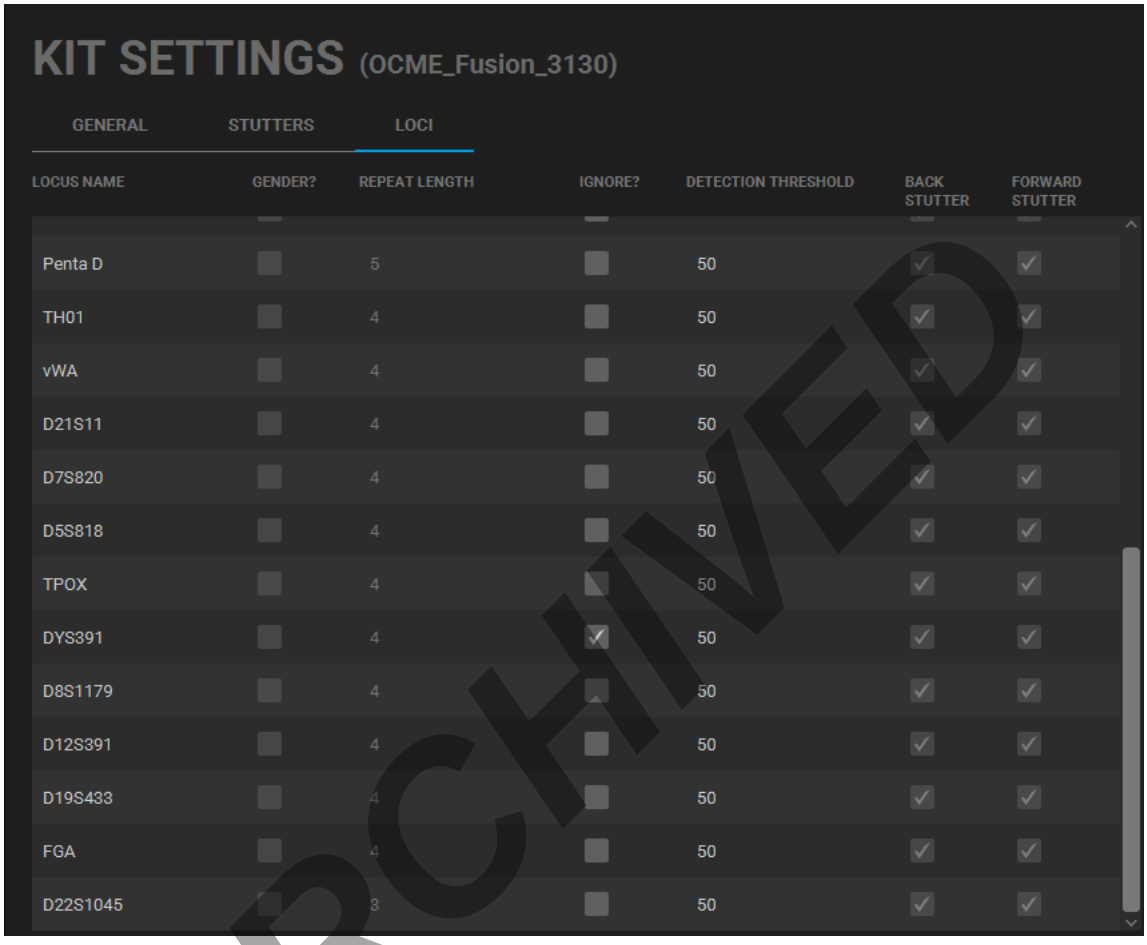





3.11.2 For the **OCME_fusion_3130 Profiling Kit**, ensure settings match the following:



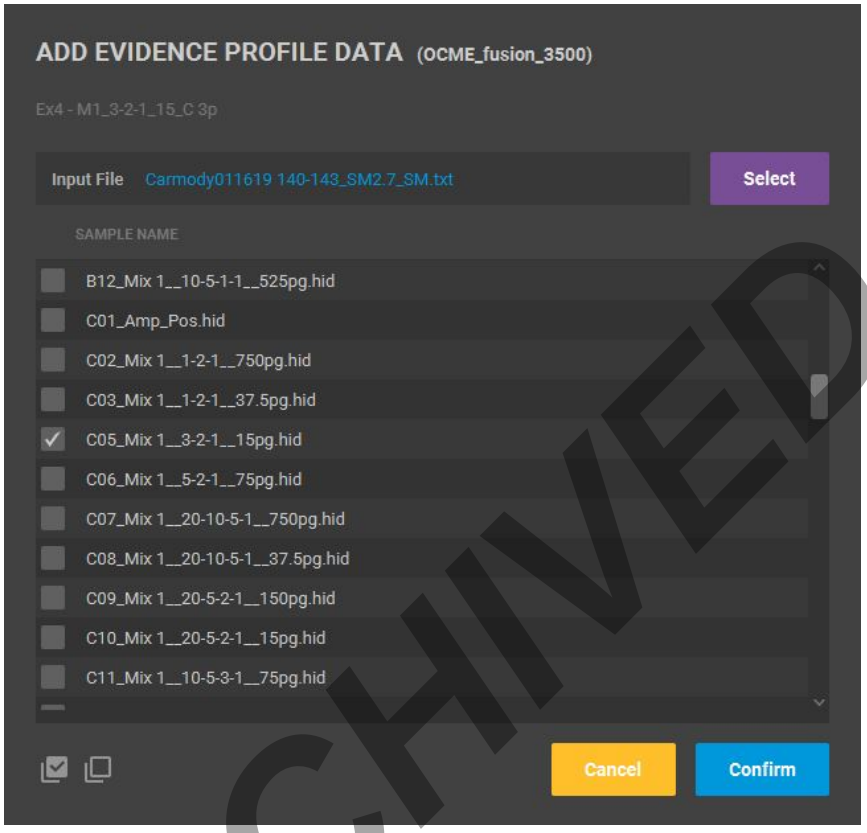





3.12 Once all settings have been verified, click Cancel to return to the previous screen. For any **locations that need to be ignored**, click the checkbox in the **Ignore** column before clicking Apply and make sure a note is made in the **Case Notes** section of the interpretation.

3.13 Add your **Evidence Profile Data** by either dragging and dropping the .txt file into the box or clicking on the + to find your .txt file. The following window will pop up and you can select your evidence profile(s). If multiple replicates are in the same run, you can select multiple. Click Confirm to return to the file import screen.

3.13.1 Repeat for any replicate(s) from other .txt files.




3.14 If you are performing a deconvolution without a conditioned contributor, proceed to step 3.15. **To add a conditioned contributor**, add your **Reference Profile Data** by either dragging and dropping the .txt file into the box or clicking on the + to find your .txt file.

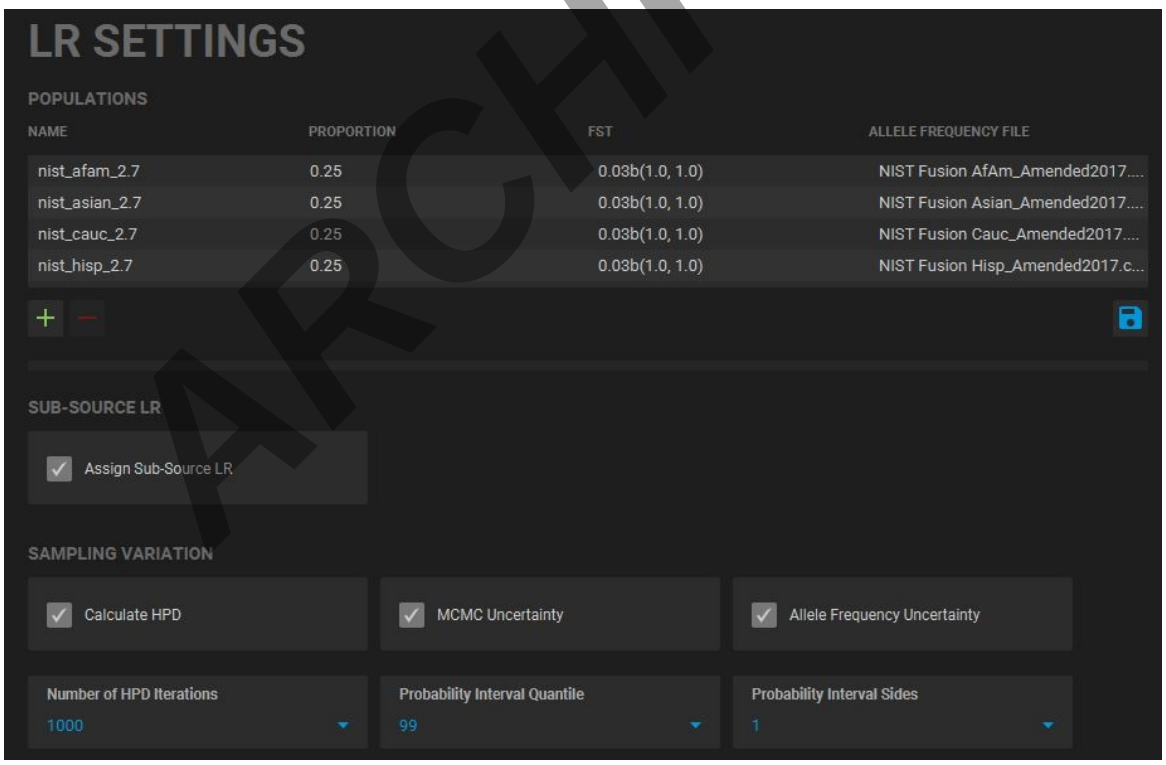
3.14.1 Select your reference profile(s) for conditioning using the checkboxes and click Confirm (similar to the screen for selecting your evidence profile).

3.14.2 Once the profile data is added, the box for HP will be checked on the right side. To condition, you must also check the box for HD. Conditioned contributors are considered true donors in Hp and Hd.


3.15 If you are performing a deconvolution without an LR, proceed to step 3.16. To perform an **LR in conjunction with the deconvolution**, add your **Reference Profile Data** by either dragging and dropping the .txt file into the box or clicking on the + to find your .txt file.

3.15.1 Select your reference profile for LRs using the checkboxes and click Confirm (similar to the screen for selecting your evidence profile). The reference will be added within the numerator of the LR (i.e. assigned to HP only).

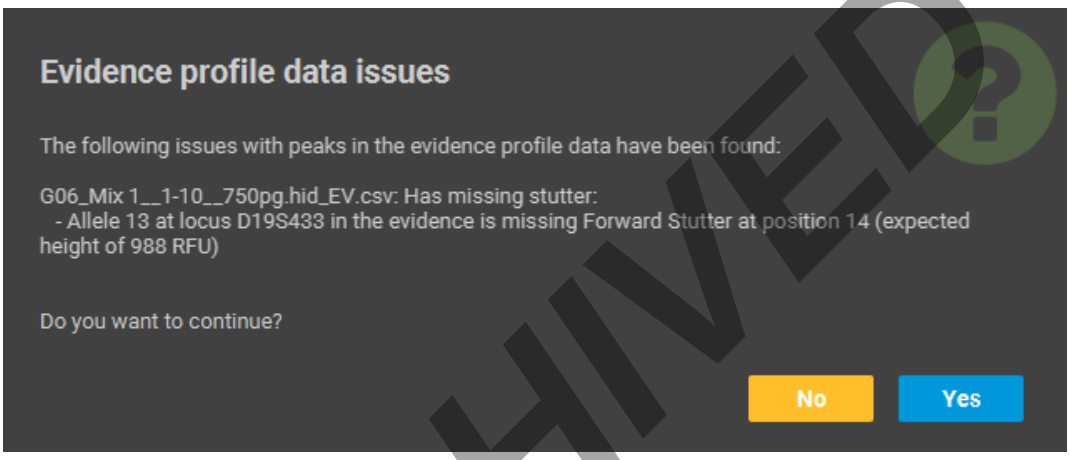
3.15.2 Once the reference profile is added to your HP, you can then select LR Settings, and check them against the following screenshot:




3.15.3 Select Cancel to return to the file import screen.

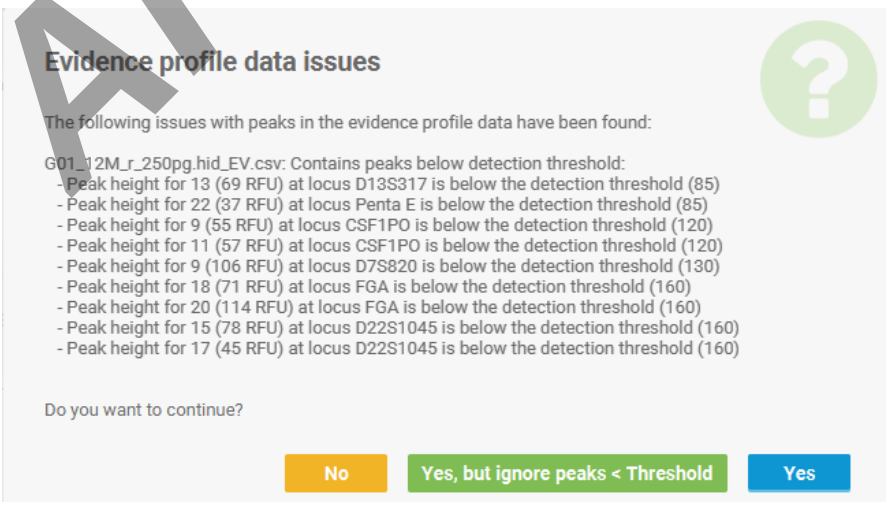
3.16 Select Start (or Queue for Batch Mode, Section 6). The < PROGRESS window will open showing the Pre-Burnin, Burnin progress, and MCMC Progress (or for Batch Mode, you will return to the batch screen to continue to add to the batch).

3.16.1 If a flag fires about missing an expected stutter (see below), review the input data and confirm that an allele label was not inadvertently deleted during analysis and that the input .txt file is correct before proceeding.




3.16.1.1 If no editing changes need to be made, you can click **Yes** to continue. If changes need to be made, select **No**, correct the labeling in the electropherogram and see Section 2.1.5.6 for editing the input file. Return to the beginning of setting up the run.

3.16.2 If a flag fires about peaks below analytical threshold (see below). Click No and do not continue. Review the input data and the GeneMarker run settings used for analysis.



3.17 Once complete, you will be at the RESULTS screen.

3.17.1 You can click on the single file folder icon on the right side to be directed to that run's **Results** folder. Select Finish to return to the Main Menu.

3.17.2 Alternatively, you can manually navigate to your STRmix™ Run Folder(s) within the **OCME_STRmix_Fileshare** folder (\\csc\ocme\OCME_STRmix_Fileshare).

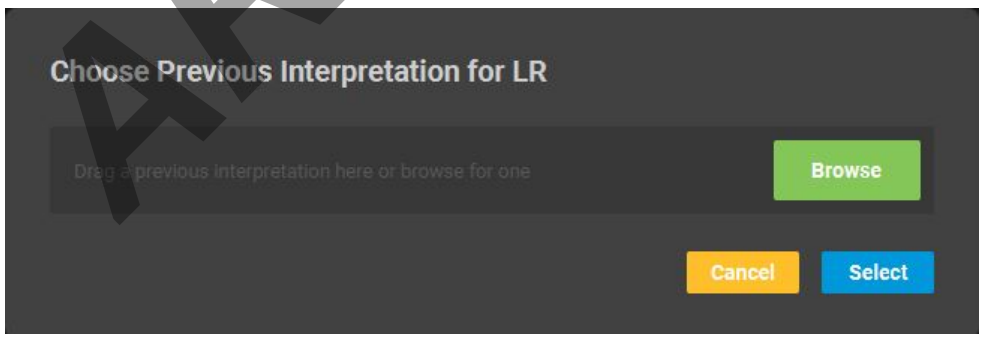
3.18 **COPY** your run folder(s) into the previously created FB sample folder within the **STRmix Data** folder.

3.19 Once you have copied the folder, **CONFIRM that all files for that run have transferred over correctly** to the **STRmix Data** folder. After confirmation, the copy of the STRmix™ Run folder located in the **OCME_STRmix_Fileshare** folder should be **deleted**.

# 4 Likelihood Ratio Calculations with STRmix™ (Investigation: LR from Previous)

4.1 Note: This section is specific for running LRs using a previously run deconvolution. Samples must undergo deconvolution prior to (or in conjunction with) running an LR for a comparison sample. Refer to Section 3 for setting up deconvolutions in conjunction with the LR and the Interpretation of PowerPlex® Fusion data run on 3500xL manual for further guidance on running deconvolutions and LRs.

4.2 From the **Main Menu**, select **Investigate**. Within the **Investigate** screen, select **LR from Previous**. You will then select the deconvolution file for which you would like to calculate an LR for a comparison sample:




4.3 Drag and drop the entire decon run folder or select Browse to locate **the deconvolution file,** labeled as config.xml or results.xml **(**or settings.ini for older versions of STRmix™) within your run folder in the STRmix Data folder in the M drive. Click **Select** to navigate to the **LR from Previous** screen to name the run.

4.4 When naming STRmix™ Likelihood Ratio runs, the Case Number and Sample ID will auto-populate from the deconvolution file. These should be updated to the appropriate naming convention for an LR run before proceeding.

4.5 The following naming convention should be used for likelihood ratios:

4.5.1 Evidence Files

- **Case Number** = YY-XXXXX (leave off "FB")
- **Sample ID** = remainder of evidence sample OCME ID* #p (NOC) condElimInitials (if applicable) scenario for LR
- **Case notes** = a comment should be added here if a locus is ignored, indicating the reason: e.g. "D2S441 was ignored due to an unresolved allelic peak"
- *Suffixes such as '_mcon' or '_reamp' should not be included in the OCME ID

4.5.2 Suspect Files

- **Case Number** = SYY-XXXXX (leave off "FB")
- **Sample ID** = evidence sample OCME ID* (include evidence file FB# without the "FB") condElimInitials (if applicable) scenario for LR
- **Case Notes** = a comment should be added here if a locus is ignored, indicating the reason: e.g. "D2S441 was ignored due to an unresolved allelic peak"
- *Suffixes such as '_mcon' or '_reamp' should not be included in the OCME ID

4.5.3 For LR scenarios, the naming format should start with the comparison sample's initials, followed by any conditioned samples' initials, and then the number and "U" for unknowns, followed by a "v" to separate the numerator hypothesis from the denominator hypothesis.

| Examples | Resulting STRmix file name |
|---|---|
| ***Evidence File*** | |
| comparing elimAB to 4p decon, no cond | *16-01234-567-1-1-1-shirt-BL*-AB3Uv4U |
| comparing elimAB to 2p decon, cond vicCD | *16-01234-567-1-1-1-shirt-BL*-ABCDvCD1U |
| comparing elimAB to 3p decon, cond vicCD and elimEF | *16-01234-567-1-1-1-shirt-BL*-ABCDEFvCDEF1U |
| ***Suspect File*** | |
| comparing suspTS to 1p decon, no cond | S16-05678-*16-01234-567-2-1-1-slide-GS*-TSv1U |
| comparing suspTS to 2p decon, no cond | S16-05678-*16-01234-567-2-1-1-slide-GS*-TS1Uv2U |
| comparing suspTS to 3p decon, cond vicCD | S16-05678-*16-01234-567-2-1-1-slide-GS*-TSCD1UvCD2U |
| comparing suspTS to 4p decon, cond vicCD and elimEF | S16-05678-*16-01234-567-2-1-1-slide-GS*-TSCDEF1UvCDEF2U |

4.6 The run settings will be pulled from the deconvolution and should not be changed. To check, click on Run Settings and see 3.7 for Run Settings screenshot. Click Cancel if all settings are correct.

- **Performance, Number of Threads**: it is okay to proceed if the Number of Threads is different; this is specific to the computer being used.
- **Performance, Low Memory Mode**: This setting allows the computer to minimize the memory used for the run and can be turned on if a run fails to finish due to computer memory.
- **Performance, Seed**: This number may be different; there is a Seed for each run, deconvolution or LR.

4.7 The kit settings will also be pulled from the deconvolution. To check, click on Kit Settings and see 3.11 for the appropriate kit settings for 3500 or 3130 data. Select Cancel to return to the previous screen **unless a locus needs to be ignored.**

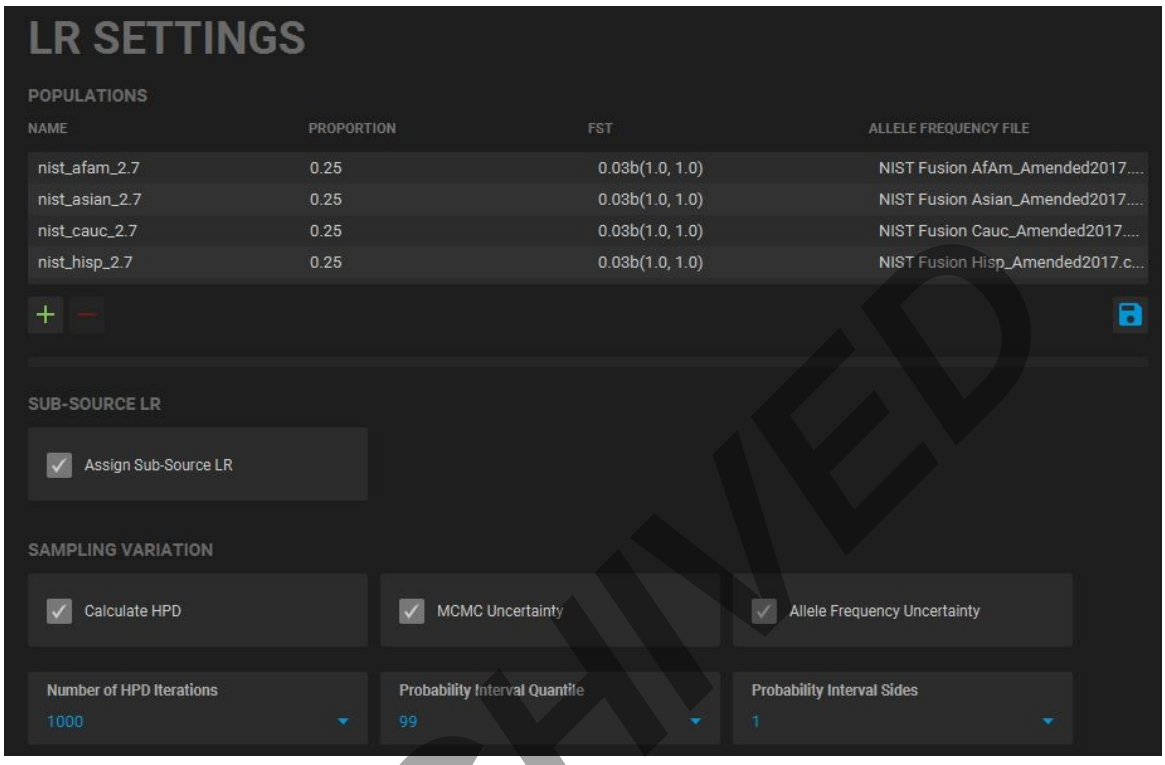
4.7.1 Rarely, a locus may be ignored at this step. For example, in the case of a tri-allelic pattern that matches your reference sample, which was not recognized at the deconvolution stage. For any **locations that need to be ignored**, click the checkbox in the **Ignore** column before clicking Apply.

4.8 Your Evidence Profile Data and Reference Profile Data for any conditioned contributors (HP & HD selected) will be populated from the deconvolution.

4.9 Add the **profile(s) for comparison** to the **Reference Profile Data** by either dragging and dropping the .txt file into the box or clicking on the + to find your .txt file(s).

4.9.1 Select your reference profile(s) for the LR using the checkboxes and click Confirm. The reference will be added within the numerator of the LR (i.e. assigned to HP only).

4.9.2 Once the reference profile(s) is added to your HP, you can then select LR Settings, and check them against the following screenshot:




4.10 Select Start (or Queue for Batch Mode, Section 6). The PROGRESS window will open (or for Batch Mode, you will return to the batch screen to continue to add to the batch).

4.11 Once complete, you will be at the RESULTS screen.

4.11.1 You can click on the single file folder icon on the right side to be directed to that run's **Results** folder. Select Finish to return to the Investigation Menu.

4.11.2 Alternatively, you can manually navigate to your STRmix Run Folder within the **OCME_STRmix_Fileshare** folder (\\csc\ocme\OCME_STRmix_Fileshare).

4.12 **COPY** your run folder into the previously created FB sample folder within the **STRmix Data** folder.

4.13 Once you have copied the folder, **CONFIRM that all files for that run have transferred over correctly** to the **STRmix Data** folder. After confirmation, the copy of the STRmix Run folder located in the **OCME_STRmix_Fileshare** folder should be **deleted**.

# 5 LRs for Convicted Offender Match Testimony

5.1 For testimony in relation to a convicted offender match, where a statistic is needed, an LR must be calculated through STRmix™.

5.1.1 Obtain the Convicted Offender profile through the CODIS software:

5.1.1.1 Go to the Specimen Manager window → Click the filter button at the top of the Specimen ID column → Click "custom" → Enter your specimen ID from your evidence sample in the top line in the dialog box → Hit "search"

5.1.1.2 Once your specimen ID pops up, right click on the line and click "view matches" → Right-click and choose "print reports", choose "match details short report"

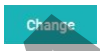
5.1.2 Add this to your case file and use the convicted offender profile from this report in order to create a comparison sample profile for STRmix™ LR calculation using the form: Reference Profile for STRmix Input Creation.

5.1.3 Once the STRmix™ LR report is generated, it will need to undergo technical review (documented using case contacts) and recertification prior to testimony.

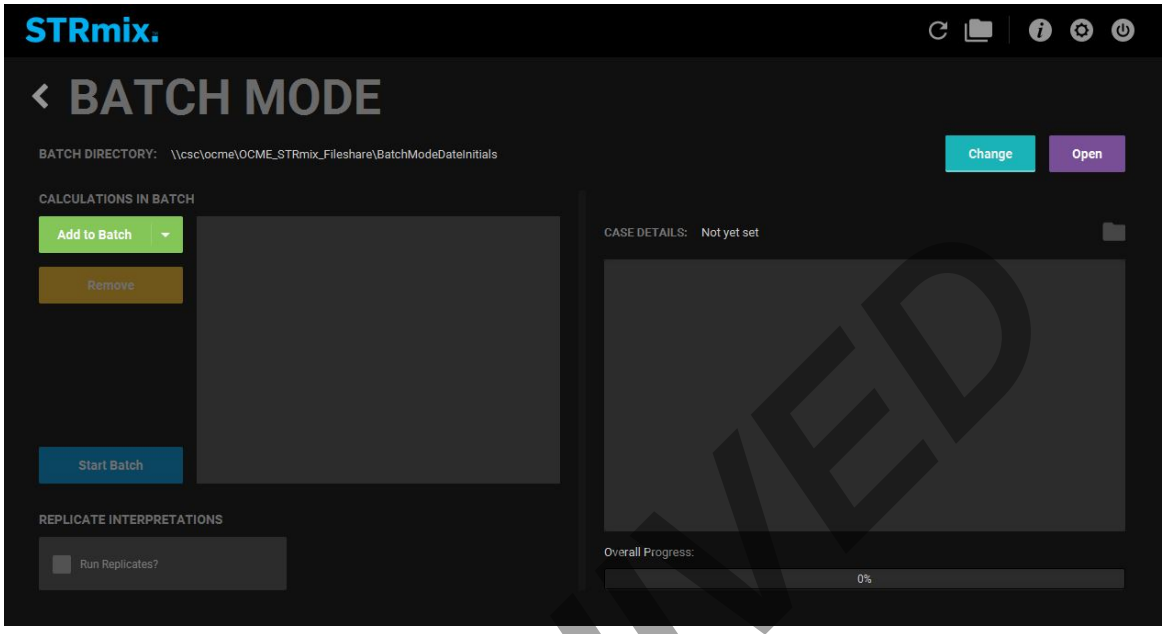
# 6 Running STRmix™ using Batch Mode

6.1 Several STRmix™ runs can be set up and queued to run sequentially. To set up a queued analysis for multiple runs, select **Batch Mode** from the STRmix™ main window. Batch mode should only be used for samples from one individual case within a single batch.

6.2 Before setting up a batch, navigate to the OCME_STRmix_Fileshare folder (\\csc\ocme\OCME_STRmix_Fileshare). Create a new folder within the STRmix Fileshare folder with the name of "Batch Mode [date] [your initials]."

6.3 Select 

to change the Batch Directory. Navigate to and select the created folder inside the **OCME_STRmix_Fileshare**. This new folder will now appear at the top of the screen:

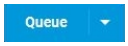



6.4 Use the down arrow next to **Add to Batch** 

and **Add Interpretation** or **Add LR from Previous** to set up each individual run. If you do not click the arrow and click **Add to Batch**, it will automatically open whichever run type was used last.

6.4.1 These batch details will be saved even if you close STRmix™ until you set up a new run; the batch can be stopped and restarted at a later date.

6.5 Complete the analysis set up for the first sample following the corresponding setup instructions: Section 3 for Deconvolutions (Investigation) or Section 4 for LRs (Investigation: LR from Previous).

6.6 After setting up your run, click 

to return to the Batch Mode setup window.

6.7 Repeat steps **6.4**-**6.6** to add additional runs.

6.7.1 If you need to edit a run, you will need to remove it and redo the setup. To remove a sample from the batch mode, highlight the case/sample in the **Calculations in Batch** part of the window and select **Remove**.

6.8 Select 

to start the batch run.

6.9 After completion of analyses, use the left arrow at the top left next to 

to return to the STRmix™ Main Menu. You can click 

to be directed to the Batch Mode folder you created.

6.9.1 Alternatively, if you have additional runs to perform within the same batch, you can add additional runs and continue the batch.

6.9.2 The **Batch Log** will show if any of the individual runs failed and why. If an individual run failed, the run folder will not contain a Results folder, and a .txt file called BATCH CALCULATION FAILED will be made instead.

6.10 Results folders from Batch Mode will be saved in the folder you created within the **OCME_STRmix_Fileshare** folder (\\csc\ocme\OCME_STRmix_Fileshare).

6.11 **COPY your run folders** into the previously created FB sample folder(s) within the **STRmix Data** folder. Place a **copy of the Batch Log** within the FB or FBS folder.

6.12 Once you have copied the folders, **CONFIRM that all files have transferred over correctly** to the **STRmix Data** folders. After confirmation, the copy of the STRmix Run folder and the Batch folder with your initials and date located in the **OCME_STRmix_Fileshare** folder should be **deleted**.

# 7 Evaluation of the STRmix™ Analysis Setup

7.1 Verify the evidence and reference input sections of the STRmix™ printout against the associated electropherograms. Ensure that:

7.1.1 All appropriate edits were made; no artifact peaks were left labeled and no allelic peak labels were removed.

7.1.1.1 If an Evidence profile issue error was generated at the beginning of your run and checked for any issues against the electropherogram prior to running, this will also show up in the **EVIDENCE PEAK ISSUES** section at the end of the report.

7.1.2 Correct input file(s) have been selected.

7.1.3 The correct file was imported into an LR from previous analysis, if applicable.

7.1.4 All suitable replicates have been utilized.

7.1.4.1 Check the **Inter replicate efficiency** that is generated by STRmix™. If there are drastic inconsistencies between the two efficiencies that are consistent with the amount of data present in the replicates, the STRmix analysis may be marked as ‘not reported’ and a new analysis may be performed with the amplification containing the most information. A third amplification may be warranted.

7.2 The number of contributors that best describes the data was chosen, as applicable.

7.3 The correct assumptions (i.e. conditioning) have been made, if applicable.

7.4 The appropriate proposition has been selected (i.e. LR calculation), if applicable.

7.5 The **SEED** value listed at the beginning of the report, with the CASE NUMBER, SAMPLE NAME, and COMMENTS, is the starting number used within the random number generator.

7.6 Check the **SETTINGS** at the end of the report to verify that the STRmix™ run was set up properly; note that any edited settings values are bolded by the program.

7.6.1 The **CASE SETTINGS** will be specific to your case and should be used to check for correct setup with the run name.

7.6.2 The **MCMC SETTINGS** are only for runs that include a deconvolution.

7.6.2.1 Check the remaining settings against the following screenshot keeping in mind the numbers of accepts may be different (and bolded) if this setting has been approved for use.

MCMC SETTINGS

| | |
|---|---|
| Number of contributors | 4 |
| Use Mx priors | N |
| Number of chains | 8 |
| Burn-in accepts per chain | 10,000 |
| Post burn-in accepts per chain | 50,000 |
| Random walk SD | 0.005 |
| Post burn-in shortlist | 9.0 |
| Auto-continue on Gelman-Rubin | N |

7.6.3 The **KIT SETTINGS** will be included in deconvolutions and LRs. If any loci were ignored, they will be listed here.

## OCME_fusion_3500 Kit

**KIT SETTINGS**

| | | |
|---|---|---|
| Ignored loci | DYS391 | |
| Detection thresholds | D3S1358 | 85 |
| | D1S1656 | 85 |
| | D2S441 | 85 |
| | D10S1248 | 85 |
| | D13S317 | 85 |
| | Penta E | 85 |
| | D16S539 | 120 |
| | D18S51 | 120 |
| | D2S1338 | 120 |
| | CSF1PO | 120 |
| | Penta D | 120 |
| | TH01 | 130 |
| | vWA | 130 |
| | D21S11 | 130 |
| | D7S820 | 130 |
| | D5S818 | 130 |
| | TPOX | 130 |
| | D8S1179 | 160 |
| | D12S391 | 160 |
| | D19S433 | 160 |
| | FGA | 160 |
| | D22S1045 | 160 |
| Saturation | 30,000 | |
| Degradation starts at | -1.0 | |
| Degradation max | 0.01 | |
| Drop-in cap | 300 | |
| Drop-in rate parameter | 0.0087 | |
| Drop-in parameters (α, β) | 22.31, 2.65 | |
| Min variance factor | 0.5 | |
| Variance minimization parameter | 1,000 | |
| LSAE variance parameter (1/λ) | 0.03 | |
| Allelic variance parameters (α, β) | 7.427, 3.479 | |

| | |
|---|---|
| ***Back Stutter*** | |
| Maximum stutter ratio | 0.3 |
| Stutter variance parameters (α, β) | 1.799, 19.052 |
| ***Forward Stutter*** | |
| Maximum stutter ratio | 0.2 |
| Stutter variance parameters (α, β) | 1.999, 11.703 |
| ***Half Back (-2bp) Stutter*** | |
| Maximum stutter ratio | 0.1 |
| Stutter variance parameters (α, β) | 2.597, 6.632 |
| ***Double Back (-8bp) Stutter*** | |
| Maximum stutter ratio | 0.1 |
| Stutter variance parameters (α, β) | 2.816, 6.792 |

## OCME_fusion_3130 Kit

**KIT SETTINGS**

| | | |
|---|---|---|
| Ignored loci | DYS391 | |
| Detection thresholds | D3S1358 | 50 |
| | D1S1656 | 50 |
| | D2S441 | 50 |
| | D10S1248 | 50 |
| | D13S317 | 50 |
| | Penta E | 50 |
| | D16S539 | 50 |
| | D18S51 | 50 |
| | D2S1338 | 50 |
| | CSF1PO | 50 |
| | Penta D | 50 |
| | TH01 | 50 |
| | vWA | 50 |
| | D21S11 | 50 |
| | D7S820 | 50 |
| | D5S818 | 50 |
| | TPOX | 50 |
| | D8S1179 | 50 |
| | D12S391 | 50 |
| | D19S433 | 50 |
| | FGA | 50 |
| | D22S1045 | 50 |
| Saturation | 8,000 | |
| Degradation starts at | -1.0 | |
| Degradation max | 0.01 | |
| Drop-in cap | 100 | |
| Drop-in rate parameter | 0.0024 | |
| Drop-in parameters (α, β) | 0.0, 0.0 | |
| Min variance factor | 0.5 | |
| Variance minimization parameter | 1,000 | |
| LSAE variance parameter (1/λ) | 0.007 | |
| Allelic variance parameters (α, β) | 7.092, 0.933 | |
| ***Back Stutter*** | | |
| Maximum stutter ratio | 0.3 | |
| Stutter variance parameters (α, β) | 2.108, 3.509 | |
| ***Forward Stutter*** | | |
| Maximum stutter ratio | 0.1 | |
| Stutter variance parameters (α, β) | 2.312, 3.259 | |



7.6.4 The **PROFILE SETTINGS** can be used to check input file names and that the run was set up properly.

7.6.5 The **LR SETTINGS** will only be present when an LR was run (whether separately or along with the deconvolution).

**OCME_fusion_3500 Kit/OCME_fusion_3130 Kit**

| LR SETTINGS | |
|---|---|
| Number of populations | 4 |
| Assign sub-source LR | Y |
| Calculate HPD | Y |
| HPD Iterations | 1,000 |
| Use MCMC uncertainty | Y |
| Use Allele Frequency uncertainty | Y |
| HPD quantile | 99% |
| HPD sides | 1 |
| ***NIST_AFAM_2.7*** | |
| Proportion | 0.25 |
| FST | 0.03b(1.0, 1.0) |
| Allele frequency file | NIST Fusion AfAm_Amended2017.csv |
| ***NIST_ASIAN_2.7*** | |
| Proportion | 0.25 |
| FST | 0.03b(1.0, 1.0) |
| Allele frequency file | NIST Fusion Asian_Amended2017.csv |
| ***NIST_CAUC_2.7*** | |
| Proportion | 0.25 |
| FST | 0.03b(1.0, 1.0) |
| Allele frequency file | NIST Fusion Cauc_Amended2017.csv |
| ***NIST_HISP_2.7*** | |
| Proportion | 0.25 |
| FST | 0.03b(1.0, 1.0) |
| Allele frequency file | NIST Fusion Hisp_Amended2017.csv |



7.6.6 The **PERFORMANCE SETTINGS** shows the number of threads (specific to the computer used to run) and whether Low Memory Mode was used.

7.6.7 The **ADDITIONAL KIT DETAILS** lists the stutter files used in the analysis.

## OCME_fusion_3500 Kit

### ADDITIONAL KIT DETAILS

| | |
|---|---|
| Size Regression File | fusion_sizeregression.csv |
| ***Back Stutter*** | |
| Position Relative to Parent | -1, 0 |
| Inversely Proportional To | Observed Height of Parent Allele |
| Stutter Regression File | OCME_BackStutterFile3500_022321.txt |
| Stutter Exceptions File | OCME_BSExceptionsFile_020321.csv |
| ***Forward Stutter*** | |
| Position Relative to Parent | 1, 0 |
| Inversely Proportional To | Expected Height of Stutter Peak |
| Stutter Regression File | OCME_ForwardStutterFile3500_022321.txt |
| Stutter Exceptions File | OCME_FSExceptions File_092020.csv |
| ***Half Back (-2bp) Stutter*** | |
| Position Relative to Parent | 0, -2 |
| Inversely Proportional To | Expected Height of Stutter Peak |
| Stutter Regression File | OCME_HalfBackStutterFile3500_022321.txt |
| Stutter Exceptions File | OCME_HBSExceptionsFile_092020.csv |
| ***Double Back (-8bp) Stutter*** | |
| Position Relative to Parent | -2, 0 |
| Inversely Proportional To | Expected Height of Stutter Peak |
| Stutter Regression File | OCME_DoubleBackStutterFile3500_022321.txt |
| Stutter Exceptions File | OCME_DBSExceptionsFile_092020.csv |

**OCME_fusion_3130 Kit**

**ADDITIONAL KIT DETAILS**

| | |
|---|---|
| Size Regression File | fusion_sizeregression.csv |
| *Back Stutter* | |
| Position Relative to Parent | -1, 0 |
| Inversely Proportional To | Observed Height of Parent Allele |
| Stutter Regression File | ocme_fusion3130_stutter.txt |
| Stutter Exceptions File | ocme_fusion3130_exceptions.csv |
| *Forward Stutter* | |
| Position Relative to Parent | 1, 0 |
| Inversely Proportional To | Expected Height of Stutter Peak |
| Stutter Regression File | ocme_fusion3130_forward stutter.txt |
| Stutter Exceptions File | |



# 8 Evaluation of the STRmix™ Analysis Diagnostics

8.1 The presence of a single sub-optimal diagnostic is not always an indication that rework is required. In some instances, a sub-optimal diagnostic(s) may be due to the nature of the sample (ex. low amounts of input DNA and/or stochastic effects) and not due to an issue with the STRmix™ run. Refer to the Troubleshooting Guide (Section 8) for further steps that may be taken to improve a sub-optimal diagnostic result.

8.2 For deconvolutions, verify that the following (**primary**) diagnostics conform to your qualitative expectations when compared to the electropherogram(s):

8.2.1 The **mixture proportions** and **template amounts** assigned to the contributor(s).

8.2.2 The **weights** assigned to the genotypes for each contributor listed in the **COMPONENT INTERPRETATION**.

8.2.3 The **degradation** values and degradation plots (at the beginning of the report). The per contributor plot can be helpful for recognizing extreme degradation its potential influence over a single contributor's genotype weights across a profile.

8.2.4 The **Locus Efficiencies (LSAE)**, following the variance charts.

8.3 For deconvolutions, evaluate the following (**secondary**) diagnostics for the run information listed in the **POST BURN-IN SUMMARY** at the beginning of the report.

8.3.1 **Total iterations:** If the total iterations exceeds 2.14 billion ($2.14 \times 10^9$), this may lead to incorrect genotype weightings being assigned.

8.3.2 **Acceptance rate:** A very low acceptance rate (e.g. 1 in thousands to millions) may, in combination with the other diagnostics, indicate that the analysis needs to be run with additional accepts.

8.3.3 **Effective sample size (ESS):** 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 ESS value (tens or hundreds) means that there is potential for a large difference in weights if the analysis is run again.

8.3.4 **(Log)likelihood:** The larger this value, the better STRmix™ has been able to describe the observed data. A negative value suggests that STRmix™ has not been able to describe the data very well given the information it has been provided. A low or negative value for the log(likelihood) may indicate to users that the analysis requires additional scrutiny.

8.3.5 **Gelman-Rubin diagnostic:** If this value is above 1.2 then it is possible that the analysis has not converged, and the analysis requires additional scrutiny.

8.3.6 **LSAE variance**: The LSAE variance probability density chart in the report may be used in conjunction with the LSAE efficiency plot to identify unusual amplification within a profile (ex. extreme inhibition).

8.3.7 **Allele variance and stutter variance:** These variances should be compared to the modes and variance charts included in the reports. If the numbers are significantly elevated, the analysis may require additional scrutiny.

8.4 For LR comparisons, the overall **category of support** (inclusion, uninformative, exclusion) should conform to your qualitative expectations in comparison to the data.

8.4.1 Evaluate the **Per Locus Likelihood Ratio** table per locus and per sample, as well as the range of LR's **between population subgroups**; pay special attention to outliers and/or zero values.

8.5 For LR comparisons that result in support for an inclusion, check to ensure that the comparison sample falls in the appropriate contributor order. See report section titled **CONTRIBUTOR ORDER GIVING HIGHEST LR** at the beginning of the report.

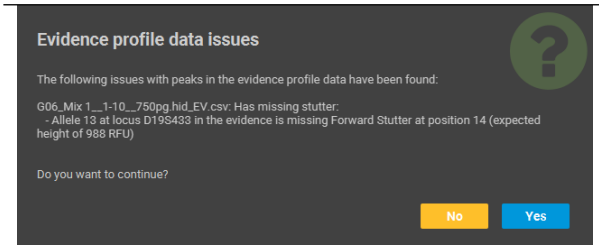
8.5.1 If multiple comparison samples are positively associated with the same sample, results should be evaluated to ensure that they do not align with the same contributor. **See Section 8 for troubleshooting when more than one comparison aligns with the same contributor**.

# 9 Troubleshooting Guide

9.1 The purpose of this guide is to address commonly seen scenarios which arise in casework. These guidelines are based on validation studies, literature references, and casework experience. However, not every situation can be covered by the Troubleshooting Guide. If a diagnostic issue arises that is not covered here, please discuss the issue with your supervisor, manager or the Technical Lead Team as needed.

| **Observations** | **Actions** |
|---|---|
| ***EVALUATION OF DECONVOLUTION DIAGNOSTICS*** | |
| The mixture proportions or template amount do not reflect what is observed<br>AND/OR<br>the degradation does not reflect what is observed<br>AND/OR<br>the interpreted contributor genotypes do not conform to your qualitative expectations<br>AND/OR<br>the category of support for an LR comparison does not conform to your **qualitative expectations** | Re-evaluate the number of contributors; consider another STRmix™ analysis with one higher or one lower number of contributors. |
| | Consider amplifying a replicate if one has not already been done, with increased input amount when available and appropriate.<br>Inhibition has occurred—microcon to clean and reamplify sample. |
| | Consider another STRmix™ analysis at a greater number of accepts (typically, 100,000 burn-in accepts and 500,000 total accepts per chain). **Note:** this requires approval by the Technical Leader (or his/her designee). |
| For a mixture with an **extreme ratio**: ex. 98:2 or 99:1; the mixture proportions and genotype weights do not conform to your qualitative expectations based on the electropherograms, and other trouble-shooting options (including additional iterations) have been exhausted. | Consider another STRmix analysis utilizing user informed Mx priors. This requires approval by the Technical Leader (or his/her designee). |
| The **total iterations** is exceeds 2.14 billion ($2.14 \times 10^9$). | This could indicate the genotype weightings have been incorrectly assigned. Please contact the Technical Lead Team. An additional STRmix™ analysis may be required. |
| A low or negative**(log) likelihood** | Re-evaluate number of contributors; consider another STRmix™ analysis with one higher or one lower number of contributors. |
| | Data has been removed that is allelic and/or stutter; data must be re-imported. |
| | Artifact peaks have been left labeled and must be removed. |
| **Gelman-Rubin** value is greater than 1.2 | Consider another STRmix™ analysis at greater number of accepts (typically, 100,000 burn-in |

| | |
|---|---|
| | accepts and 500,000 total accepts per chain). This may sometimes reduce the GR to below 1.2. **Note:** this requires approval by the Technical Leader (or his/her designee). |
| **Stutter and/or allele variance** significantly elevated usually at or beyond the horizontal asymptote (may be in conjunction with low (log) likelihood) | Check to make sure no data has been omitted. |
| | Re-evaluate number of contributors; consider another STRmix™ analysis with one higher or one lower number of contributors. |
| | Consider amplifying a replicate if one has not already been done, with increased input amount when available and appropriate. |
| ***EVALUATION OF LR DIAGNOSTICS*** | |
| Differences in **Per Locus LRs**: large LR's (>1) obtained for each locus where the comparison profile is qualitatively included in the evidentiary profile, with one locus where the LR = 0 | Data entry problem—check input files. |
| | Allele call not fully resolved at a given locus – ignore locus and perform analysis again. **Note:** discuss with supervisor as needed. |
| | Inhibition has occurred—microcon to clean and reamplify sample. |
| | Consider amplifying a replicate if one has not already been done, with increased input amount when available and appropriate. |
| | Re-evaluate number of contributors; consider another STRmix™ analysis with one higher or one lower number of contributors. |
| | Consider another STRmix™ analysis at greater number of accepts (typically, 100,000 burn-in accepts and 500,000 total accepts per chain). **Note:** this requires approval by the Technical Leader (or his/her designee). |
| Multiple elimination and/or comparison samples gave LRs supporting inclusion to an evidence sample and are **aligning with the same contributor** and/or there is an indication of **relatedness**. | Consult a supervisor and the Tech Lead Team to go over case specifics. Additional deconvolution and/or LR scenarios may need to be considered. |
| The 99% 1-sided HPD and Unified LR result(s) for one or more **population subgroup(s)** is several orders of magnitude lower than the other population subgroups in comparison to the point estimate LR, and the lowest LR of the four population subgroups does not conform to your qualitative expectations of the comparison. | Consult your supervisor or manager and the Technical Lead Team. A stratified likelihood ratio using census data may be calculated and reported; this must be approved by the Technical Leader (or his/her designee). |
| ***ERROR MESSAGES*** | |
| | This is an example of a **missing stutter** flag. You will be notified if there's an expected stutter peak |

| | |
|---|---|
|  | that is not included based on the heights of the allelic peaks in the input file. See Section 3.16.1. Elevated stutter variance may occur in conjunction with this flag. The data should be evaluated to determine if the elevated variance makes intuitive sense with the peaks present in the profile. Consult the Tech Lead Team if further assistance is needed. |
| Evidence profile data issues after pre-checks | Another profile data issue that you may see is if there are peaks within the input file that are below the set analytical threshold. The original data should be re-evaluated in GeneMarker to determine if incorrect settings were applied. |
| Pre-Burnin failed:<br>Determine Acceptable Genotypes failed: Locus # (locus name) in the evidence cannot be explained given the parameters you have chosen. | The number of contributors selected was too few and must be increased. |
| | Consider amplifying a replicate if one has not already been done, with increased input amount when available and appropriate. |
| (OutofMemoryError) GC overhead limit exceeded | The run could not be completed with the computer power supplied. The run should be set up again using the Low Memory Mode setting and/or on a CPU with more computing power. |

## 10 References:

10.1 STRmix™ v.2.7 Operation Manual and previous versions

10.2 STRmix™ v. 2.7 User's Manual and previous versions

10.3 NYC OCME Internal Validation of STRmix™ v2.7 for Fusion 5C/3500xL (September 2021)

10.4 NYC OCME Internal Validation of STRmix™ v2.7 for Fusion 5C/3500xL Data - STRmix™ Parameters (August 2021)

10.5 NYC OCME Stutter Study for GeneMarker® HID 3.0.0 and STRmix™ Version 2.7-PowerPlex® Fusion Data run on 3500xl Genetic Analyzers (September 2021)

10.6 NYC OCME Internal Validation of STRmix™ V2.4 for Fusion (January 2017)

10.7 NYC OCME STRmix™ V2.4.08 Performance Check. (July 2018)