STRmix Implementation and Internal Validation-FUSION

Erie County Central Police Services Forensic Laboratory

In-house studies were undertaken to configure, implement and validate STRmix software V2.3. STRmix is a software system that applies a fully continuous approach to DNA mixture profile interpretation. It standardizes the analysis of profiles within a laboratory by using estimates of variance of electropherograms derived from DNA profiling data. The software was developed and is supported by the Institute of Environmental Science and Research Limited, New Zealand (ESR). Selected published peer reviewed articles describing the development and research of STRmix can be found in the References Section. A broader selection of related published articles can be found in the STRmix User’s Manual and in the Developmental Validation Studies\(^4\) paper.

Background

Manufacturer supplied guidelines were followed. Specific formulas and spreadsheets to be used can be found in: STRmix V2.3 Implementation and Validation Guide, December 11, 2014\(^4\), issued by the Institute of Environmental Science and Research Limited. Additional guidance can be found in the STRmix V2.3 User’s Manual, February 2, 2015\(^2\). Additionally, SWGDAM Guidelines for the Validation of Probabilistic Genotyping Systems Draft for Public Comment 03/16/2015\(^3\) were followed.

STRmix has been previously subjected to developmental validation. This involved, in part, the “by hand” confirmation of the calculations behind the software. The results of the developmental validation are included in the STRmix User’s Manual. In addition, a summary of the developmental validation is discussed in Taylor et al.\(^6\)

Laboratory specific parameters to be determined prior to use of STRmix include:

1. Analysis threshold (PAT)
2. Stutter ratios
3. Drop-in parameters
4. Saturation
5. Peak height variance
6. Hyper-parameter for the variance of locus specific amplification effects (LSAE)

These parameters need to be defined for each STR kit, each protocol and CE Platform. The values need to be updated each time there is a significant change to the platform such as a camera or laser change.

The STRmix software suite uses Model Maker V2.3 to determine the optimal allele variance, stutter variance and locus specific amplification efficiency parameters to use. This is based on data generated by the laboratory and is instrument platform specific for each STR multiplex employed.

All samples were extracted using standard in-house extraction protocols. Samples were then quantified using Promega Plexor HY kit. All samples were subsequently amplified using the Promega FUSION STR kit (30 cycles) and run on the 3500 Genetic Analyzer platform.

Analysis Thresholds
An analytical threshold of 145RFU was applied to the data. This value was previously determined to be the maximum Peak Amplitude Threshold (PAT) for the FUSION Kit using the 3500 Genetic Analyzer platform. The value was determined during a previous in-house validation study undertaken prior to the implementation of the FUSION Kit.

**Stutter**

Two stutter parameters within STRmix require optimization. The first is the maximum allowable stutter and the second is a stutter file used to model the expected heights of the stutter alleles assigned by STRmix. 102 profiles were analyzed to determine stutter values. Observed stutter values for each heterozygote and homozygote allele were calculated. The following parameters were determined and applied to the STRmix software.

1. The maximum allowable stutter ratio permitted is set to 0.3.
2. The stutter values to be applied to the software can be found in Table 1, below.
3. Appendix 1 of the Estimation of STRmix Parameters for Erie County Central Police Services contains the graphical representations of the values in Table 1.

**Table 1. Erie County FUSION 30 cycle stutter values for STRmix™**

<table>
<thead>
<tr>
<th>Locus</th>
<th>Intercept</th>
<th>Slope</th>
</tr>
</thead>
<tbody>
<tr>
<td>D3S1358</td>
<td>-0.0559</td>
<td>0.00886</td>
</tr>
<tr>
<td>D1S1656</td>
<td>-0.00143</td>
<td>0.00604</td>
</tr>
<tr>
<td>D2S441</td>
<td>0.0648</td>
<td>-0.00153</td>
</tr>
<tr>
<td>D10S1248</td>
<td>-0.0657</td>
<td>0.0103</td>
</tr>
<tr>
<td>D13S317</td>
<td>-0.0603</td>
<td>0.00971</td>
</tr>
<tr>
<td>Penta E</td>
<td>-0.0186</td>
<td>0.00394</td>
</tr>
<tr>
<td>D16S539</td>
<td>-0.0542</td>
<td>0.0106</td>
</tr>
<tr>
<td>D18S51</td>
<td>-0.0422</td>
<td>0.00809</td>
</tr>
<tr>
<td>D2S1338</td>
<td>-0.0217</td>
<td>0.00504</td>
</tr>
<tr>
<td>CSF1PO</td>
<td>-0.0498</td>
<td>0.0101</td>
</tr>
<tr>
<td>Penta D</td>
<td>-0.0182</td>
<td>0.00317</td>
</tr>
<tr>
<td>Th01</td>
<td>-0.000594</td>
<td>0.00332</td>
</tr>
<tr>
<td>vWA</td>
<td>-0.122</td>
<td>0.0114</td>
</tr>
<tr>
<td>D21S11</td>
<td>-0.134</td>
<td>0.00713</td>
</tr>
<tr>
<td>D7S820</td>
<td>-0.0645</td>
<td>0.0114</td>
</tr>
<tr>
<td>D5S818</td>
<td>-0.0462</td>
<td>0.00889</td>
</tr>
<tr>
<td>TPOX</td>
<td>-0.026</td>
<td>0.00582</td>
</tr>
<tr>
<td>DyS391</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>D8S1179</td>
<td>0.000945</td>
<td>0.00508</td>
</tr>
<tr>
<td>D12S391</td>
<td>-0.0944</td>
<td>0.00979</td>
</tr>
<tr>
<td>D19S433</td>
<td>-0.0801</td>
<td>0.0107</td>
</tr>
<tr>
<td>FGA</td>
<td>-0.0468</td>
<td>0.00521</td>
</tr>
<tr>
<td>D22S1045</td>
<td>-0.145</td>
<td>0.016</td>
</tr>
</tbody>
</table>

**Drop-in Parameters**
Within STRmix, drop-in is modelled using a gamma distribution. The maximum observed drop-in height of 260RFU was applied as previously calculated.

Using the Gamma Model v2.3 Drop-in Calculator described in the Validation Guidelines, the calculated drop-in parameters of (7.98, 21.78) with a Drop-in Frequency of 0.0015, were applied to the data.

**Saturation**

A maximum saturation value of 29,900 RFU was applied to the data. Laboratory testing of samples amplified up to 500 pg of DNA found the instrument to be linear up to this concentration. This amount of amplified DNA yielded a maximum RFU of 29,900. 500 pg of amplified DNA is the target DNA for FUSION in this Laboratory. The saturation value is used to correct for any observed RFU’s that exceed the instrument’s CCD saturation point.

STRmix will adjust any RFU readings above this value qualitatively, not quantitatively.

**Figure 1. 3500 Saturation Experiment.**

**Peak Height Variance, Stutter Variance and LSAE using Model Maker**

This was calculated by the Institute of Environmental Science and Research Limited using data supplied by the Laboratory. 95 single-source samples were amplified with FUSION using established Laboratory protocols and were analyzed in GeneMapper IDx. The profiles ranged from full to low level, partial profiles. For this data, the V2.3 Model Maker function was used. The following values will be used in the software:

- Allelic Variance is: (2.867, 1.164)
- Stutter Variance is: (4.786, 7.641)
LSAE Variance is: 0.033

**Default STRmix Parameters**

Parameters based on the above experimental data were entered as default values to be used in all subsequent internal validation studies and in actual casework once STRmix is implemented. Default settings for STRmix v2.3 are controlled in 2 windows. **Figure 2** shows the default General Settings. **Figure 3** shows the default values for the FUSION DNA Profiling Kit Settings.

**Figure 2. Default STRmix General Settings.**

![STRmix General Settings](image)
Figure 3. Default STRmix FUSION Kit Settings.
**Internal Validation**

The following internal validation studies are written following the format used in the document: SWGDAM Guidelines for the Validation of Probabilistic Genotyping Systems Draft for Public Comment 03/16/2015. Texts written in italics are quotes from the SWGDAM document. The Laboratory’s data and/or response to the specific guideline will follow each SWGDAM standard.

4.1 The laboratory should test the system using representative data generated in-house with the amplification kit, detection instrumentation and analysis software used for casework. Additionally, some studies may be conducted by using artificially created or altered input files to further assess the capabilities and limitations of the software.

All testing was done using the FUSION Amplification Kit. Samples were run on the 3500 Genetic Analyzer platform and subsequently analyzed using GeneMapper IDx v1.4. The GeneMapper stutter filter is turned off prior to generation of the corresponding GeneMapper Table for use by STRmix.

4.1.1 Specimens with known contributors, as well as case-type specimens that may include unknown contributors.

Casework specimens from adjudicated cases were run under standard conditions. This data is shown in 4.2.1.1 on page 9.

Mixtures were prepared in-house using extracted buccal swabs from staff members. Various combinations of 2 persons, 3 persons and 4 person mixtures were prepared at different dilutions, as indicated below. The mixtures were amplified using FUSION and run on the 3500 Genetic Analyzer using standard conditions. All resulting DNA profiles were subsequently analyzed using GeneMapper IDx and the resulting GeneMapper Table was exported for use by STRmix. Following deconvolution by STRmix, various hypotheses were tested. All data related to these sample sets are shown under 4.1.2.

**Experimental Design – Mixture Preparation Set 1**

Six mixtures of known contributors were prepared in-house. The following concentrations, ratios, and dilutions were made and analyzed with STRmix. Set 1 was also used for the Sensitivity and Specificity Experiments.

1) **Mixture Series 1** – A 4:3:2:1 mixture was prepared where the minor contributor drops from 100, 50, 25, 12.5, and 6.25 pg of DNA. The following contributors were used in the following proportions; 4(KMO):3(DC):2(CF):1(TG). The concentrations of all four contributors were first diluted to 100 pg/µl (0.100 ng/µl) and then 4 µl of KMO, 3 µl of DC, 2 µl CF, and 1 µl of TG were combined for a final volume of 10 µl. This sample was then either amplified directly or diluted accordingly.

2) **Mixture Series 2** – A 3:2:1 mixture was prepared where the minor contributor drops from 100, 50, 25, 12.5, and 6.25 pg of DNA. The following contributors were used in the following proportions; 3(DC):2(CF):1(TG). The concentrations of all three contributors were first diluted to 100 pg/µl (0.100 ng/µl) and then 5 µl of DC, 3.3 µl CF, and 1.7 µl of TG were combined for a final volume of 10 µl. This sample was then either amplified directly or diluted accordingly.
Experimental Design – Mixture Preparation Set 2

1) **Mixture Series 1** – A dilution series of two contributors was prepared using DNA extracted from two staff members. The series of dilutions were adjusted so that approximately 0.5 ng of total DNA was amplified: 19:1, 9:1, 3:1, 1:1, 1:3, 1:9, 1:19. The total template quantity was 500 pg per reaction.

2) **Mixture Series 2** – Two sets of three contributors were prepared using DNA extracted from three staff members. The dilutions were adjusted so that approximately 500 pg of total DNA was amplified: 5:1:1, 10:4:1, 1:5:1, 4:1:10, 1:1:5, 1:10:4.

3) **Mixture Series 3** – One set of four contributors were prepared using DNA extracted from four staff members. The dilutions were adjusted so that 500 pg of total DNA was amplified: 17:1:1:1, 14:3:2:1, 1:1:1:1.

4.1.2 **Hypothesis testing with contributors and non-contributors.**

4.1.2.1 The Laboratory should evaluate more than one set of hypotheses for individual evidentiary profiles to aid in the development of policies regarding the formulation of hypotheses.

Various hypotheses were tested using selected samples from the Mixture Set 1. The following hypotheses were tested for the first set of data below seen in Figures 4-7. The results demonstrate that the LR will decrease as the amount of template DNA decreases, where Hp is true (non-contributors). This is intuitively correct based on our training and experience with FUSION.

Hp: For each dilution series, the DNA originated from the person of interest and N-1 unknown individuals.

Hd: For each dilution series, the DNA originated from N unknown contributors.

(where N is the true number of contributors)

Figure 4. Three person mixture effect of DNA concentration on the LR at a 1:1:1 ratio (Concentration).
Effects of DNA Concentration on Likelihood Ratio (3 Person Mixture 1:1:1)

- Caucasian
- Black
- SWH
- Asian

Log (LR) vs DNA Amount (Total pg amplified)
Figure 5. Three person mixture effect of DNA concentration on the LR at a 1:1:1 ratio (# of Alleles).

Figure 6. Four person mixture effect of DNA concentration on the LR at a 4:3:2:1 ratio (Concentration).
4.1.3 Variable DNA typing conditions.

The Laboratory does not use alternative amplification or electrophoresis conditions to increase or decrease the detection of alleles.

4.1.4 Allelic peak height, to include off-scale peaks.

STRmix models peak height. Allelic peak height variations (varying amounts of input DNA) were tested as part of the mixture study. The data above demonstrates that variations in the peak heights of different contributors will change the resulting weights of the genotype combinations that are assigned during the MCMC. Off-scale peaks were not tested since this Laboratory does not amplify extremely excessive amounts of DNA.

4.1.5 Single-source specimens.

The data for this experiment is shown in 4.2.1.2 under the heading “Check of the LR”. It can be seen that based on single source samples, the resulting LR generated by STRmix is consistent with a manually calculated LR.

4.1.6 Mixed specimens.

4.1.6.1 Various contributor ratios.
Mixtures of varying ratios were run as described in 4.1.1. The mixtures were amplified using the above protocols. The resulting GeneMapper Tables were exported for analysis using STRmix.

A test of the weights was conducted by analyzing a two person mixture series. A set of 2 person mixtures was prepared at the following ratios: 19:1, 9:1, 3:1, 1:1 and the reverse. The samples were run through STRmix under the following hypotheses:

- \( H_0 \): The suspect and one unknown are contributors to the DNA mixture
- \( H_1 \): Two unknown individuals are contributors to the DNA mixture

The resulting Log (LR) for the major contributor was plotted against the Mixture Ratio. The results can be seen in Figure 8.

Figure 8. Plot of 2 person mixtures consisting of various ratios versus the Log of the LR.

A second test of the weights for mixtures was conducted by analyzing 3 person mixtures. A set of 3 person mixtures were prepared as follows:

- 3 Person: 5:1:1, 10:4:1, 4:1:10, 1:1:5

The samples were run through STRmix under the following hypotheses:

- \( H_0 \): The suspect and N-1 individuals are contributors to the DNA mixture
- \( H_1 \): N unknown individuals are contributors to the DNA mixture

Where N is the number of contributors to the mixture.

The resulting Log (LR) for the various contributors was plotted against the Mixture Ratio. The results can be seen in Figure 9.
The data above demonstrate that the weights obtained from the STRmix deconvolution are intuitively correct. Where the proportions of DNA from the individual contributors is similar (e.g. 1:1:5) peak heights are no longer informative and therefore individual contributor genotypes are not able to be resolved. It is more pronounced in the 3 person mixtures.

### 4.1.6.2 Various total DNA template quantities.

#### Check of the Weights

A dilution series of a 3 person mixture was run. The resulting LR’s for the known contributor was calculated for each dilution. The data were plotted as the Log(LR) versus DNA Concentration and Log(LR) versus # of Alleles. The results can be seen in Figures 4 and 5, above. The data demonstrates the expected drop in LR that is seen with decreasing drop in the quantity/quality of DNA.
4.1.6.3 Various numbers of contributors.

Mixtures were run at varying ratios and concentrations as indicated in 4.1.1. The data is shown in the graphs below and was also reported above. 2 person and 3 person mixtures were deconvoluted in STRmix using N and N+1 contributors. 4 person mixtures were only run in STRmix at N contributors. It is Laboratory policy to not deconvolute or generate a statistical comparison for mixtures that contain more than 4 contributors.

4.1.6.4 Over and Under estimation of the number of contributors.

Selected 2 person, 3 person and 4 person mixtures that were previously described above were deconvoluted using STRmix. STRmix was configured with the number of contributors set at N, N-1 and N+1, were N equals the true number of contributors. Evaluation of the results found that for 2, 3 and 4 person mixtures, STRmix did not generate data when the number of contributors was set to N-1. That is, the profiles could not be explained by a reduced number of contributors (N-1) and therefore STRmix did not run.

The N+1 data can be seen in Figures 10 and 11. Figure 10 shows the results of running 2 person mixtures from Mixture Set 2 with 3 contributors set in STRmix. Figure 11 shows the results of running 3 person mixtures from Mixture Set 2 with 4 contributors set in STRmix.

It can be seen in Figure 10 that there is a significant decrease in the Log (LR) in the samples in which it is typically more difficult to deconvolute. The sample exhibiting the greatest reduction in LR is the 3:1 ratio mixture with Donor 1.

Figure 11 also shows a slight decrease in the Log (LR) in the samples that exhibit the greatest difficulty in deconvolution. These are the samples that are 1:1:5. The LR reported for Donor #1 was highest in the mixture where Donor #1 was the higher contributor to the mixture (5:1:1) and was lowest in the mixture where Donor #1 was the lowest contributor to the mixture (1:1:5).

The reduction in LR is expected as STRmix has to ‘create’ an additional but unseen contributor. This N+1 contributor must share alleles with the known contributors. This increases the number of genotype combinations that STRmix considers thus diffusing the weights resulting in lower LR values.
Figure 10. 2 person mixtures run as 3 contributors from Mixture Set #2.

![Graph showing Log (LR) vs Mixture Ratio for 2 and 3 Contributors (Donor 1)].

Figure 11. 3 person mixtures run as 4 contributors from Mixture Set #2.

![Graph showing Log (LR) vs Mixture Ratio for 3 and 4 Contributors (Donor 1 Considered)].

4.1.6.5 Sharing of alleles among contributors.

The mixtures that were set-up for the above experiments consisted of samples prepared from staff members. A review of the epg’s shows sharing of alleles at various loci as seen in the FUSION data. This demonstrates that...
the in-house prepared mixtures had loci in which the various contributors shared alleles. This is designed to test
STRmix’s ability to deconvolute a mixture where there is allele sharing. Based on the data presented in the
graphs for Mixture Set 1 and Mixture Set 2, STRmix is capable of correctly deconvoluting when there is allele
sharing between the various contributors.

Kinship testing was done to further test the sharing of alleles. Various 2-4 person mixtures were tested with
different combinations. Results are as follows:

a) **Mother:Father (3:1).** Two daughters were compared and returned an LR of zero, each.
b) **Father:Son (4:1).** Mother, daughter and second son were compared and returned an LR of zero, each.
c) **Son A:Son B (2:1).** Mother, father and sibling were compared and returned an LR of zero, each.
d) **Father:Son B:Unrelated Male (4:2:1).** Mother, daughter and son A were compared and returned an LR
   of zero, each. All known contributors returned LR’s > 1 and were intuitively correct.
e) **Son B:Son A:Unrelated Female (5:2:1).** Mother, father and daughter were compared and returned an LR
   of zero, each. All known contributors returned LR’s > 1 and were intuitively correct.
f) **Father:Son B:Unrelated Male 1:Unrelated Male 2 (4:3:2:1).** Mother, daughter, son A were compared
   and returned an LR of zero or < 1, each. All known contributors returned an LR > 1 and were intuitively
   correct.
g) **Son A:Son B:Unrelated Female:Unrelated Male (6:2:1:1).** Mother was compared and returned an LR of
   zero. The father returned an LR < 500 and the daughter returned an LR < 50. All known contributors
   returned an LR > 1 and were intuitively correct.

### 4.1.7 Partial profiles, to include the following:

#### 4.1.7.1 Allele and locus dropout

#### 4.1.7.2 DNA degradation

#### 4.1.7.3 Inhibition

Selected mixture DNA profiles, from Mixture Set 1 above, were degraded *in silico* to mimic samples with allele
and locus dropout, DNA degradation and partial inhibition.

**Inhibited/Degraded Experiment**

Six mixtures of known contributors previously prepared in house were artificially degraded/inhibited. The
resulting profiles were consistent with those seen in actual casework that exhibit allele and locus dropout, DNA
degradation and or inhibition.

1) **Scenario 1** – A two person mixture was used at a 3:1 ratio. Donor 1 was degraded and Donor 2 not
   changed.
2) **Scenario 2** – The same two person mixture was used as in scenario 1. Both contributors were
   degraded.
3) **Scenario 3** – A second two person mixture was used at a 9:1 ratio. Donor 2 was degraded and Donor 1
   was not changed.
4) **Scenario 4** – The same two person mixture was used as in Scenario 3. Both contributors were
   degraded.
5) **Scenario 5** – A three person mixture was used at a 1:1:1 ratio and a concentration of 200 pg for each contributor. Donor 1 was degraded and the other two contributors were not.

6) **Scenario 6** – The same three person mixture was used as in scenario 5. All three contributors were degraded.

7) **Scenario 7** – A three person mixture was used at a 1:1:1 ratio and a concentration of 100 pg for each contributor. One of the other two donors was degraded and the other two contributors were not.

8) **Scenario 8** – The same three person mixture in scenario 7 was used. All three contributors were degraded.

9) **Scenario 9** – A four person mixture was used at a ratio of 4:3:2:1 and at concentrations of 100 pg and 50 pg for the minor contributor. The minor contributor was degraded for 100 pg set. The major contributor was degraded for the 50 pg set.

10) **Scenario 10** – The same four person was used as in scenario 9. All contributors were degraded.

The samples were run through STRmix and the resulting LR’s plotted.

**Figure 12. Two person 3:1 mixture. Scenarios 1 and 2.**
Figure 13. Two person 9:1 mixtures. Scenarios 3 and 4.

![LR of Minor Contributor after Degradation (9:1 Mix)](image)

- **Caucasian**
- **Black**
- **SWH**
- **Asian**

Degradation Type:
- No Deg.
- General
- POI (minor)

Figure 14. Three person mixtures. Scenarios 5 and 6.

![LR after Degradation (3 Person 1:1:1 200pg DNA)](image)

- **Caucasian**
- **Black**
- **SWH**
- **Asian**

Degradation Type:
- No Deg.
- Gen Deg.
- POI only

Figure 15. Three person mixtures. Scenarios 7 and 8.
Figure 16. Four person mixtures. Scenario 9 and 10 (100 pg).
4.1.8 Allele drop-in, spikes and pull-up.

Drop-in, spike and pull-up peaks were artificially added in silico to a prepared mixture of two and/or three known contributors.

**Spikes:**

Spikes were added in silico to a 3:1 (17 at Penta E and 12 at Penta D) and 9:1 (16 at Penta D) 2 person mixture. The samples were run through STRmix versus one of the known contributors. In both instances, STRmix returned an LR of zero. It was noted that STRmix was unable to resolve the loci in which a spike was present.

**Drop-in:**

A drop-in allele (13 at D7S820) was added to a 9:1 2 person mixture. STRmix was unable to resolve the locus and returned an LR of zero.

**Pull-up:**

A pull-up peak was added to a 3:1 (TPOX 8 pulled into Penta E 14) and 9:1 (FGA 23 pulled into D7S820 13). STRmix was unable to resolve the locus with pull-up and returned an LR of zero.

A pull-up peak was added to a 3 person mixture (FGA 23 pulled into D7S820 13). STRmix was unable to resolve the locus with pull-up and returned an LR of zero. The same result was obtained when the MCMC was increased to 400,000 iterations.

This demonstrates the need for proper review of the electropherogram prior to analysis with STRmix. The STRmix manual states that “labels must be removed from all other artifacts including forward stutter, pull-up, spikes and dye blobs.” A trained analyst is able to recognize these artifacts and unmark them prior to loading.
into STRmix. If STRmix fails to resolve a profile at a particular the locus, the trained analyst will review the electropherograms to determine if any artifacts were missed.

Mixture used: A 3 person mixture ratio of 1:1:1 and concentration of 200 pg for each contributor was used.

- Drop-in: Allele 15 at locus CSF1PO (RFU 160 bp 354.61)
- Spike: Allele 16 at locus Penta D (RFU 176, bp 444.74)
- Pull-up: Allele 23 at locus D7S820 (RFU 180, bp 301.52)

As shown in Figure 18, STRmix was able to resolve one of the contributors. It can be seen that there was no significant effect on the LR. However, one of the other contributors did return an LR of zero for the spike and pull-up samples. Similar data was obtained for a 4 person mixture (data not shown). This demonstrates that drop-in, spikes and/or pull-up have the greatest effect on 2 person mixtures and will variably impact a 3-4 person mixtures dependent upon the contributor that is being tested.

**Figure 18.** Drop-in/Spike/Pull-up effect on the LR. 3 person mixture.

![Effect of Non-Allelic Peaks on LR (3 Person 1:1:1, 200 pg DNA amped)](image)

4.1.9  **Forward and reverse stutter.**

The in-house prepared mixture samples were examined for the presence of forward and reverse stutter. Reverse stutter was present in the various samples that were analyzed with STRmix. No forward stutter was observed with FUSION.

4.1.10 **Intra-locus peak variance.**

The mixtures that were used in this study exhibited varying degrees of Intra-locus peak variance. It can be seen from the numerous charts throughout that STRmix is capable of analyzing electropherograms that exhibit Intra-locus peak variance.
4.1.11 Inter-locus peak variance.

Inter-locus peak variance is routinely encountered in forensic casework samples. Past experience has found that inter-locus peak variance increases in samples that have an amount of DNA present that is less than the target level. The samples tested and reported in the experiments above contained varying amounts of inter-locus peak variance. STRmix was able to properly deconvolute the samples and return a valid LR.

4.1.12 For probabilistic genotyping systems that require in-house parameters to be established, the internal validation tests should be performed using the same parameters.

As indicated in the Background, STRmix requires several parameters to be established in-house. All parameters were calculated and established prior to running the internal validation studies.

4.1.13 Sensitivity, specificity and precision, as described for Developmental Validation.

Sensitivity and Specificity:

Sensitivity is defined as the ability of the software to reliably resolve the DNA profile of known contributors within a mixed DNA profile for a range of starting DNA template. The Log(LR) for known contributors (Hp true) should be high and should trend to 0 as less information is present within the profile. Information includes the amount of DNA from the contributor or interest, conditioning profiles, replicates and decreasing numbers of contributors.

Specificity is defined as the ability of the software to reliably exclude known non-contributors (Hd true) within a mixed DNA profile for a range of starting DNA template. The LR should trend downwards as less information is present within the DNA profile.

Specificity and sensitivity for FUSION was tested by calculating the LR for a number of three (Figure 19) and four person (Figure 20) mixtures for known contributors. The LR distribution for Hp true was well separated at high template. As the concentration decreases the Log (LR) also decreases. At high template, STRmix correctly and reliably gives a high LR for true contributors and a low LR as the contributors drop out. At low template, STRmix correctly and reliably reported that the analysis of the sample tends toward uninformative or inconclusive.

A review of the electropherograms for the mixtures that were used found that the mixtures in which Hp true approaches zero were very poor quality mixtures, many of which would not qualify for running through STRmix if they were actual casework samples. This data therefore supports the practice of having trained analysts interpret the electropherograms and determine if they have a suitable number of loci that exhibit alleles that will yield a reliable LR when analyzed by STRmix.
Figure 19. Effect of DNA Concentration on LR-3 Person 1:1:1 Mixture. Green squares represent Hp True and Brown dots represent Hd True.
**Figure 20. Effect of DNA Concentration on LR-4 Person 4:3:2:1 Mixture.** Green squares represent Hp True and Pink dots represent Hd True

<table>
<thead>
<tr>
<th>Log (LR)</th>
<th>DNA amount minor contributor (pg)</th>
</tr>
</thead>
<tbody>
<tr>
<td>-20</td>
<td>0</td>
</tr>
<tr>
<td>-15</td>
<td>20</td>
</tr>
<tr>
<td>-10</td>
<td>40</td>
</tr>
<tr>
<td>-5</td>
<td>60</td>
</tr>
<tr>
<td>0</td>
<td>80</td>
</tr>
<tr>
<td>5</td>
<td>100</td>
</tr>
<tr>
<td>10</td>
<td>120</td>
</tr>
<tr>
<td>15</td>
<td>140</td>
</tr>
</tbody>
</table>

**Precision (Reliability):**

The MCMC process is used to determine the weights within STRmix for different genotype combinations. This is a sampling procedure and therefore the weights will vary slightly between each run. 2 person, 3 person and 4 person mixtures from above were run through STRmix a total of 10 times each. The resulting LRs were plotted and can be seen in the charts below. The results demonstrate the variability in the LRs due to the MCMC process. However, they also demonstrate that the variability is random and the values obtained for the various runs remain close. This is consistent with the model described by the developer. The results can be seen in **Figures 21-24.** To further account for the variability of the allele frequency database and the MCMC process, STRmix V2.3 uses the highest posterior density (HPD) method, which is a type of confidence interval.

In most instances, the default MCMC settings of 100,000 accepts with a burn-in of 20,000 will yield STRmix results with acceptable diagnostics. STRmix provides diagnostics for each analyzed sample. For cases in which complex mixtures fail to yield acceptable diagnostics, the analyst is to re-run the sample with an increased MCMC. This will reduce the variability of the resulting LR. To demonstrate the effectiveness of this, a complex 3 person mixture was analyzed 10 times in STRmix with the MCMC increased to 400,000 accepts and a burn-in of 80,000 iterations. The data is plotted in **Figure 24.** It can be seen that the standard deviation and variance of the LR is decreased with the higher number of MCMC iterations. The average Log(LR) is also slightly decreased with increased MCMC accepts. However, evaluation of the average plus the standard deviation shows that the LR obtained with 100,000 MCMC accepts tends to overlap and is not significantly different than that obtained with 400,000 MCMC accepts.
Figure 21. 2 person precision. 100,000 MCMC accepts. Standard Deviation bars are indicated for each Log (LR).

Figure 22. 3 person precision. 100,000 MCMC accepts. Standard Deviation bars are indicated for each Log (LR).
Figure 23. 4 person precision, 100,000 MCMC accepts.

Figure 24. 3 person mixture 400,000 MCMC accepts. Standard Deviation bars are indicated for each Log (LR).
4.1.14 Additional challenge testing (e.g., the inclusion of non-allelic peaks such as bleed-through and spikes in the typing results).

Additional mixtures from one of the sets above were prepared in silico to have pull-up. These were run through STRmix.

This data is reported in 4.1.8, Figure 18.

This demonstrates the need for proper review of the electropherogram prior to analysis with STRmix. The STRmix manual states that “labels must be removed from all other artifacts including forward stutter, pull-up, spikes and dye blobs.” A trained analyst is able to recognize these artifacts and unmark them prior to loading into STRmix. If STRmix fails to resolve a profile at a particular the locus, the trained analyst will review the electropherograms to determine if any artifacts were missed.

4.2 Laboratories with existing interpretation procedures should compare the results of probabilistic genotyping and manual interpretation of the same data, notwithstanding the fact that probabilistic genotyping conclusions are inherently different from and not directly comparable to binary conclusions (e.g., exclusion or inclusion). Match statistics that are generated by these two approaches are based on different assumptions, thresholds and formulae. However, such a comparison should be conducted and evaluated for general consistency.

4.2.1 The Laboratory should determine whether the results produced by the probabilistic genotyping software are intuitive and consistent with expectations based on non-probabilistic mixture analysis methods.

4.2.1.1 Generally, known specimens that are included based on non-probabilistic analyses would be expected to also be included based on probabilistic genotyping.

Case Samples

Adjudicated case samples were analyzed using STRmix. The samples consisted of swabs from various types of evidence. The swabs were run during the normal part of casework. Two person, three person or four person DNA mixtures of varying quality were obtained. Buccal specimens from suspected contributors to the mixtures were obtained and analyzed in a similar fashion. Buccal swabs from at least one suspect were submitted for each case. The resulting evidentiary DNA profiles were visually compared by the analyst to determine if the suspect could be a contributor to the DNA mixture. Laboratory protocol for mixture interpretation can yield one of 3 possible conclusions:

1. Cannot be excluded
2. Inconclusive—does not meet Laboratory criteria for reporting
3. Excluded

A review of the casework data by the original analyst found the following reported results:

1. Fourteen (14) suspects were excluded as contributors to the DNA Mixture.
2. Fourteen (14) could not be excluded as contributors to the DNA mixture

Since these were FUSION cases, no CPI's were calculated. Each profile was analyzed using STRmix with subsequent comparison to the corresponding buccal specimen. The following hypotheses were used:

- Hp: The suspect is a contributor to the DNA mixture
- Hd: An unknown individual is a contributor to the DNA mixture

The resulting LRs were examined and compared to the original results obtained by the analyst. All previously excluded individuals except one returned an LR ≤0. Fourteen (14) individuals previously reported as “cannot be excluded” returned an LR >1. A graphical representation of the Log (LR) vs. the “Human Interpretation” can be found in Figure 25, below. It should be noted that there is overlapping of some data points on the chart and therefore, not all data points are visible.

Figure 25. Log (LR) vs. Human Interpretation of case work samples.
4.2.1.2 For single-source specimens with high quality results, genotypes derived from binary manual interpretation of profiles above the stochastic threshold should be in complete concordance with the results of probabilistic methods.

Check of the LR

Using STRmix and the default settings described above, 5 single source profiles were tested. These consisted of staff DNA profiles that were amplified with the Laboratory’s target amount of DNA (0.5ng) with FUSION using established Laboratory protocols and were subsequently analyzed with STRmix using the above settings.

Two files were generated for each sample. The first sample was a mock crime scene sample consisting of the DNA profile from the individual staff member. A second file was generated as the corresponding known sample from the same individual. The DNA profiles were entered into STRmix and the LR was calculated for each sample using the following two hypotheses:

H₀: The DNA originated from the POI
H₁: The DNA originated from an unknown individual

The LR was also calculated by hand using an Excel spreadsheet provided by ESR. The formulas in the Excel spreadsheet were examined by Laboratory staff and verified to be correct. The population databases used were the NIST Caucasian, Black and Hispanic. The results can be found in Table 3, below. It can be seen that the LR generated by STRmix for each of the samples and in each of the population groups the same as those generated using the Excel spreadsheet.
Table 3. Single Source LR Verification.

<table>
<thead>
<tr>
<th>Staff</th>
<th>Caucasian</th>
<th>Black</th>
<th>Hispanic</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>STRmix LR</td>
<td>Manually Calculated LR</td>
<td>STRmix LR</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Staff 2</td>
<td>1.64E+28</td>
<td>1.64E+28</td>
<td>1.64E+30</td>
</tr>
<tr>
<td>Staff 3</td>
<td>4.74E+24</td>
<td>4.75E+24</td>
<td>1.66E+28</td>
</tr>
<tr>
<td>Staff 5</td>
<td>6.51E+26</td>
<td>6.51E+26</td>
<td>2.89E+28</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Staff</th>
<th>Asian</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>STRmix LR</td>
</tr>
<tr>
<td>Staff 1</td>
<td>5.92E+28</td>
</tr>
<tr>
<td>Staff 2</td>
<td>3.52E+30</td>
</tr>
<tr>
<td>Staff 3</td>
<td>1.28E+27</td>
</tr>
<tr>
<td>Staff 4</td>
<td>2.56E+26</td>
</tr>
<tr>
<td>Staff 5</td>
<td>2.79E+30</td>
</tr>
</tbody>
</table>
Check of the Weights

A dilution series of a single source profile was run. The resulting LR for the known contributor was calculated for each dilution. The data were plotted as the DNA Concentration versus the LR. The results can be seen in Figure 26, below. The data demonstrates the expected drop in LR that is seen with decreasing drop in the quantity/quality of DNA.

Figure 26. DNA Concentration vs LR.
The number of alleles versus the Log LR was also plotted and can be found in Figure 27, below. The results demonstrate an increasing LR as the number of alleles increases. The data obtained for the mixture experiments demonstrates that the weights obtained from the STRmix results are intuitively correct.

Figure 27. Plot of The Number of Alleles vs LR.
References:


