Upgrade of STRmix™ version 2.8 to version 2.9.1 for the Idaho State Police Laboratory

(Fusion 5C 3500)

STRmix™ Implementation

This document describes the comparison of laboratory-specific parameters determined within STRmix™ V2.8 and V2.9.1 for the Idaho State Police Crime Laboratory (hereafter ISP). It also describes the performance check of the adopted parameters. The current STRmix™ parameters in use by ISP were determined for Fusion 5C data generated within the ISP Laboratory using 29 PCR cycles and separated using 3500 CE instrumentation with a single injection protocol in STRmix™ V2.8.

STRmix™ Parameters

There are a number of parameters that are not optimised by the MCMC in a STRmix™ analysis. These parameters must be set by the user and are either determined by analysis of empirical data or modelled within STRmix™ using the Model Maker function. The laboratory-specific parameters that are determined prior to use of STRmix™ are:

- Analytical/detection thresholds,
- Stutter ratios,
- Drop-in parameters,
- Saturation threshold,
- Allelic and stutter peak height variance parameters, and
- Locus Specific Amplification Efficiency (LSAE) variance parameter.

Analytical thresholds, stutter ratios, drop-in parameters and saturation thresholds will not change between STRmix™ V2.8 and V2.9.1 (unless these settings have been updated outside of STRmix™, i.e. ATs in the analysis software have been reassessed).

Allelic and stutter peak height variance parameters and LSAE parameters should be recalculated in STRmix™ 2.9.1 to determine if any difference is observed.

Peak height variance and LSAE using Model Maker

The 99 single-source Fusion 5C profiles of varying quality originally submitted by ISP for Model Maker in STRmix™ V2.8 were re-run in STRmix™ V2.9.1. These profiles had also been previously run in STRmix™ V2.8 as part of ISP's internal validation of STRmix™. No significant difference in variance distributions was found. The Model Maker comparisons can be found in Figure 1. ISP's kit settings determined in V2.8 are appropriate for use in STRmix V2.9.1.

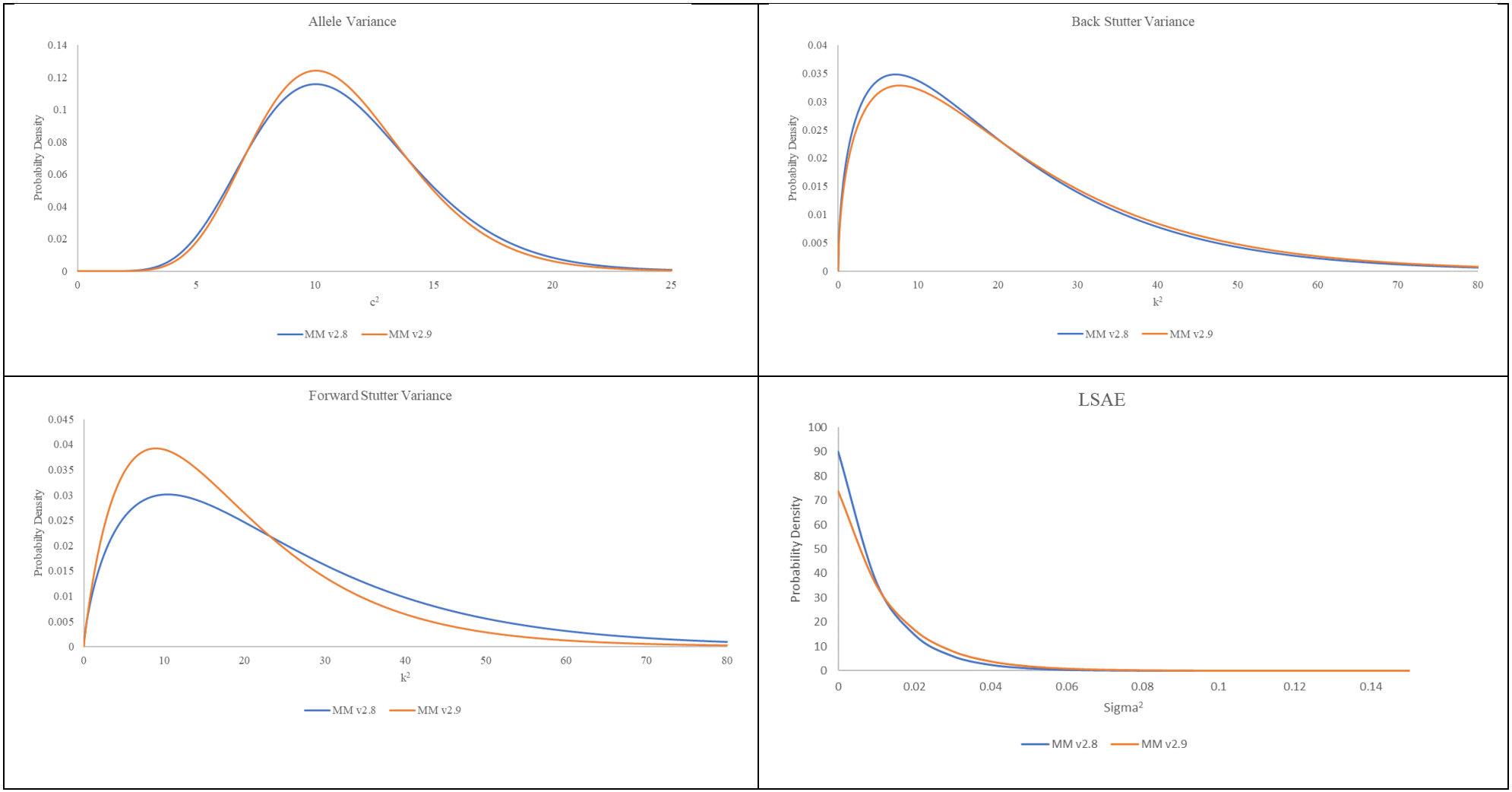


Figure 1: Comparisons of variance distributions for the Fusion 5C data

The diagnostics output by Model Maker were reviewed. In particular, the correlation plots were examined and assessed. These plots are in the Model Maker Reports appended to this document. No obvious correlation was observed (the desired result).

STRmix™ V2.9.1 kit settings

The recommended STRmix™ V2.9.1 default parameters for the interpretation of 29 cycle Fusion 5C profiles analysed on a 3500 CE instrument with a 1.2 kV/15 s injection protocol within the ISP Laboratory are given in Figure 2 to Figure 5. All settings remain unchanged from STRmix V2.8.

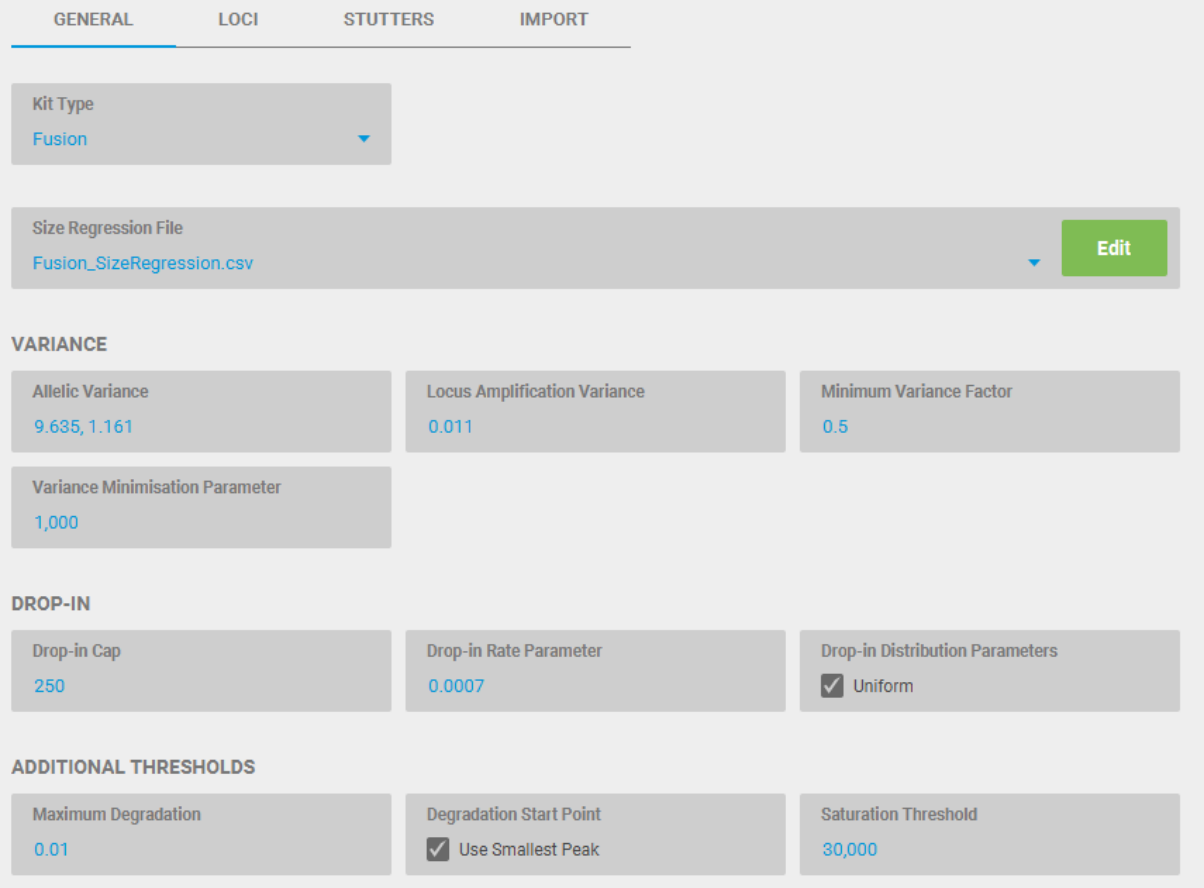


Figure 2: General settings for the Fusion 5C STRmix™ V2.9.1 kit

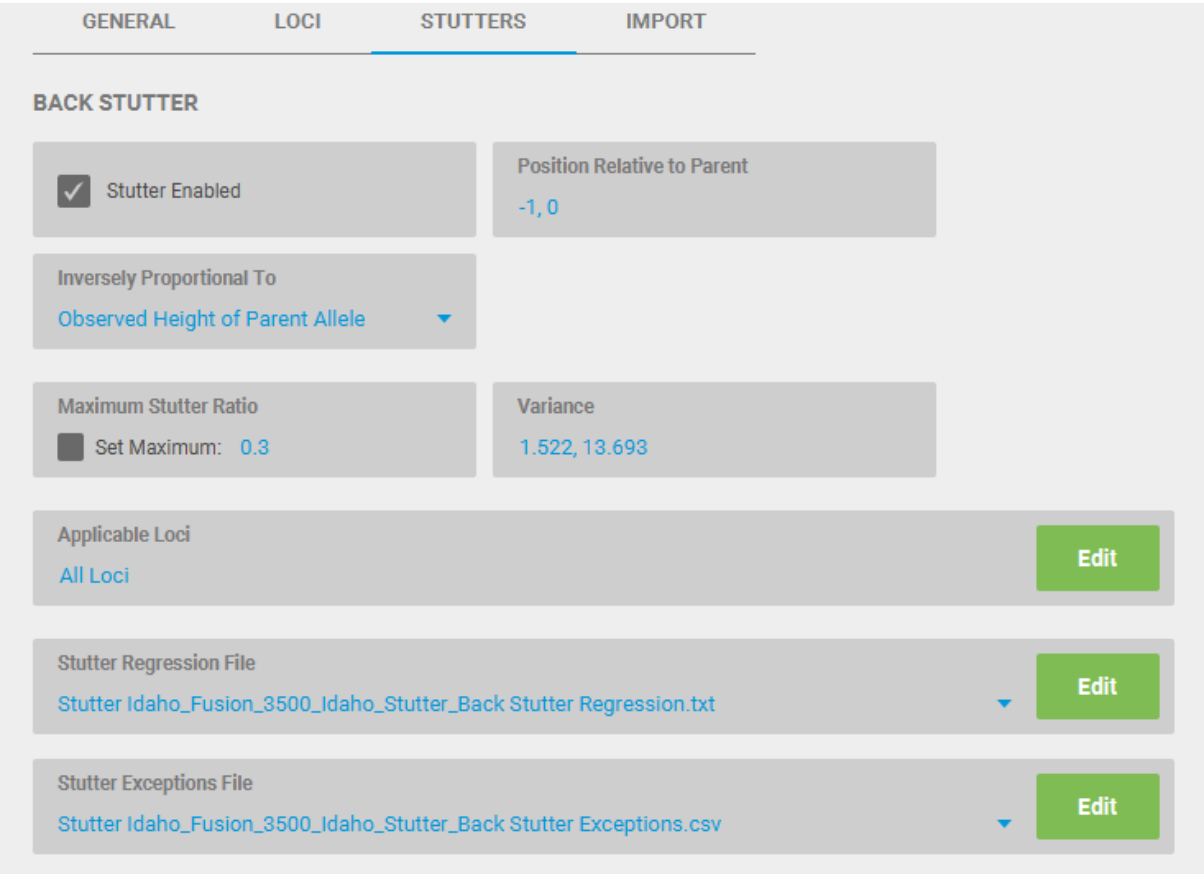


Figure 3: Back stutter settings for the Fusion 5C STRmix™ V2.9.1 kit

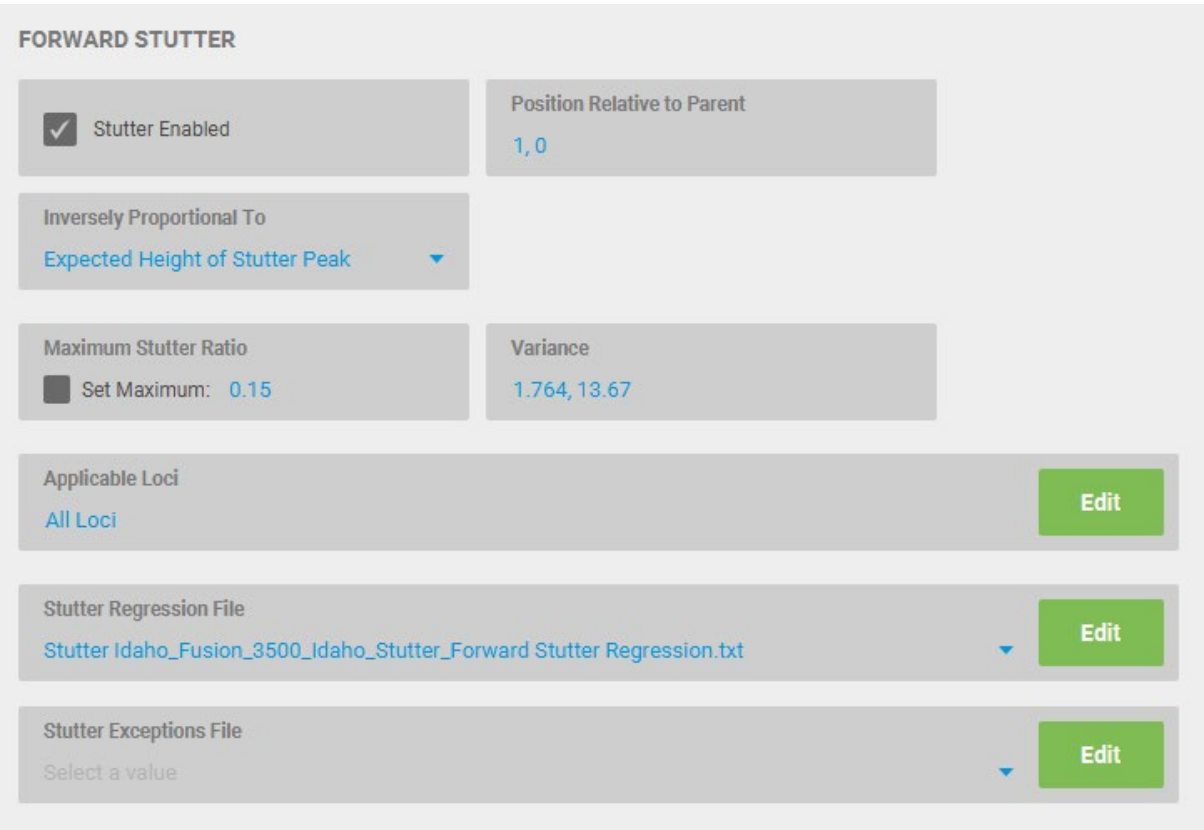


Figure 4:Forward stutter settings for the Fusion 5C STRmix™ V2.9.1 kit

GENERAL | LOCI | STUTTERS | IMPORT

LOCUS NAME	GENDER?	REPEAT LENGTH	IGNORE?	DETECTION THRESHOLD
AMEL	✓			
D3S1358		4		70
D1S1656		4		70
D2S441		4		70
D10S1248		4		70
D13S317		4		70
Penta E		5		70
D16S539		4		70
D18S51		4		70
D2S1338		4		70
CSF1PO		4		70
Penta D		5		70
TH01		4		70
vWA		4		70
D21S11		4		70
D7S820		4		70
D5S818		4		70
TPOX		4		70
DYS391		4	✓	70
D8S1179		4		70
D12S391		4		70
D19S433		4		70
FGA		4		70
D22S1045		3		70

Figure 5: Analytical Threshold information for the Fusion 5C 3500 STRmix™ V2.9.1 kit

Performance check of updated parameters

A suggested performance check for an upgrade from V2.8 to V2.9.1 involves the interpretation of about fifty profiles of varying quality (template) and varying numbers of contributors. A suggested plan is:

a) An unambiguous (high template) single source profile where weights = 1 for a single genotype will result in identical point LRs (including relative propositions) using the same allele frequency database and theta values.
b) Mixed DNA profiles that contain multiple low level (non-assumed) contributors, where multiple genotype sets with dropped alleles for both contributors are being considered. This will result in different but similar LRs due to the expected variability within the MCMC, and changes to the calculations in v2.9.1.
c) Mixed DNA profiles where one contributor is a trace or minor contributor with alleles in stutter positions (back, forward, double back etc.) of the major contributor should be interpreted and the results be intuitive.
d) LR from Previous calculations for deconvolutions carried out in previous versions should result in the same LR.

Section A

An unambiguous (high template) single source profile where weights = 1 for a single genotype will result in identical point LRs (including relative propositions) using the same allele frequency database and theta values.

Fusion 5C sample DNA1_500pg.hid_EV which had been previously interpreted in STRmix™ V2.8 and had an LR assigned using the FBI_Extended_Cauc allele frequencies with a theta value of 1%, was reinterpreted in STRmix™ V2.9.1 with the same allele frequency database and theta value.

The propositions considered were:

H_p: The DNA originates from the person of interest

H_d: The DNA originates from an unknown, unrelated individual

The sub-source LRs were identical (Figure 6). The HPD LRs are slightly different and this expected due to differences in the seed that the two interpretations were run with.

PER LOCUS LIKELIHOOD RATIOS

LOCUS	FBI_EXTENDED_CAUC 0.01b(1.0, 1.0) Pr(E|Hp)	Pr(E|Hd)	LR
D3S1358	1	8.4886E-2	1.1780E1
D1S1656	1	2.4306E-2	4.1142E1
D2S441	1	1.6511E-1	6.0565E0
D10S1248	1	9.9137E-2	1.0087E1
D13S317	1	3.6579E-2	2.7338E1
Penta E	1	1.7504E-2	5.7129E1
D16S539	1	1.1580E-1	8.6353E0
D18S51	1	4.0035E-2	2.4978E1
D2S1338	1	9.9769E-3	1.0023E2
CSF1PO	1	1.9220E-3	5.2029E2
Penta D	1	5.2949E-2	1.8886E1
TH01	1	1.4051E-1	7.1167E0
vWA	1	1.1350E-1	8.8106E0
D21S11	1	2.2967E-2	4.3541E1
D7S820	1	7.0180E-2	1.4249E1
D5S818	1	2.8702E-1	3.4841E0
TPOX	1	2.8027E-1	3.5680E0
DYS391			
D8S1179	1	1.4076E-1	7.1043E0
D12S391	1	3.1127E-2	3.2127E1
D19S433	1	1.9729E-1	5.0686E0
FGA	1	1.8871E-2	5.2990E1
D22S1045	1	2.3188E-1	4.3126E0
SUB-SUB-SOURCE LR			1.2560E27
SUB-SOURCE LR			1.2560E27
99% 1-SIDED LOWER HPD INTERVAL			3.2906E26

STRmix V2.8

PER LOCUS LIKELIHOOD RATIOS

LOCUS	FBI_EXTENDED_CAUC 0.01b(1.0, 1.0) Pr(E|Hp)	Pr(E|Hd)	LR
D3S1358	1	8.4886E-2	1.1780E1
D1S1656	1	2.4306E-2	4.1142E1
D2S441	1	1.6511E-1	6.0565E0
D10S1248	1	9.9137E-2	1.0087E1
D13S317	1	3.6579E-2	2.7338E1
Penta E	1	1.7504E-2	5.7129E1
D16S539	1	1.1580E-1	8.6353E0
D18S51	1	4.0035E-2	2.4978E1
D2S1338	1	9.9769E-3	1.0023E2
CSF1PO	1	1.9220E-3	5.2029E2
Penta D	1	5.2949E-2	1.8886E1
TH01	1	1.4051E-1	7.1167E0
vWA	1	1.1350E-1	8.8106E0
D21S11	1	2.2967E-2	4.3541E1
D7S820	1	7.0180E-2	1.4249E1
D5S818	1	2.8702E-1	3.4841E0
TPOX	1	2.8027E-1	3.5680E0
DYS391			
D8S1179	1	1.4076E-1	7.1043E0
D12S391	1	3.1127E-2	3.2127E1
D19S433	1	1.9729E-1	5.0686E0
FGA	1	1.8871E-2	5.2990E1
D22S1045	1	2.3188E-1	4.3126E0
SUB-SUB-SOURCE LR			1.2560E27
SUB-SOURCE LR			1.2560E27
99% 1-SIDED LOWER HPD INTERVAL			3.0372E26

STRmix V2.9.1

Figure 6: Interpretation of a single source Fusion 5C unambiguous profile in STRmix™ versions 2.8 and 2.9.1

Sections B and C

Mixed DNA profiles that contain multiple low level (non-assumed) contributors, where multiple genotype sets with dropped alleles for both contributors are being considered. This will result in different but similar LRs due to the expected variability within the MCMC, and changes to the calculations in v2.9.1.

Mixed DNA profiles where one contributor is a trace or minor contributor with alleles in stutter positions (back, forward, double back etc.) of the major contributor should be interpreted and the results be intuitive.

All mixed DNA profiles (143 Fusion 5C) that had previously been interpreted by ISP in STRmix™ V2.8 were reinterpreted in STRmix™ V2.9.1.

Profiles contained multiple low level (non-assumed) contributors, where multiple genotype sets with dropped alleles for both contributors were being considered. Profiles were also included that contained at least one trace or minor contributor with alleles in stutter positions of the major contributor.

Likelihood ratios were assigned to true- and non-contributors by searching each deconvolution against a database that contained the DNA profiles of the known donors as well as 200 non-contributor profiles. The non-contributor profiles were simulated from the FBI Caucasian allele frequencies. An LR was assigned for each database individual considering the following propositions:

H_p: The DNA originates from the database individual and *N*-1 unknown, unrelated individuals

H_d: The DNA originates from *N* unknown, unrelated individuals

Where *N* is the experimentally designed NOC. As per the LRs assigned in STRmix™ V2.8, STRmix™ V2.9.1 LRs were assigned using the FBI Caucasian allele frequencies with θ = 0.01 and the sub-source *LR* used as the point of comparison.

Plots of log(*LR*) STRmix V2.8 v log(*LR*) STRmix V2.91 for known contributors to mixtures using the Fusion 5C kit are shown in Figure 7. All results were concordant.

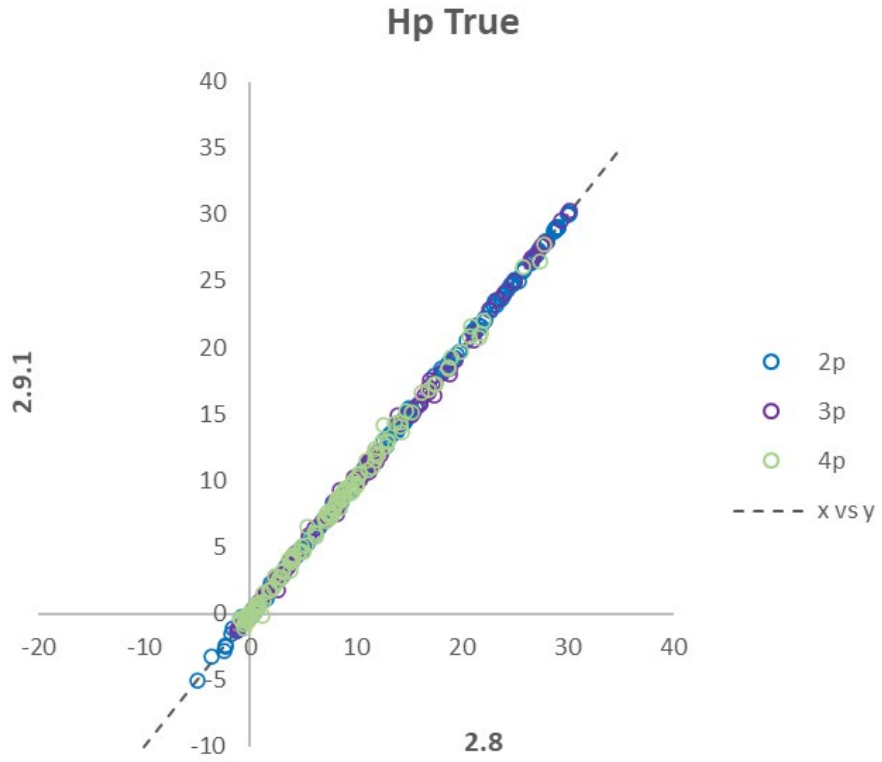


Figure 7: A plot showing the log(LR) of known donors in STRmix V2.8 v V2.9.1 using the same mixtures interpreted in each version of the software for the Fusion 5C kit.

Figure 8 shows the comparison of log(LR)s for non-contributors to the Fusion 5C mixtures interpreted in STRmix™ V2.8 and V2.9.1. As expected we see more variance in the LRs of non-contributors than we saw in the LRs of contributors. This is due to the increase of variation of gentoypes given low weights (i.e. genotypes that get given little weight in the MCMC may not be given any weight if the sample was to be reinterpreted. Conversely, genotypes with large weights should be accepted with a a similar weight if the sample were to be reinterpreted). These results are as expected.

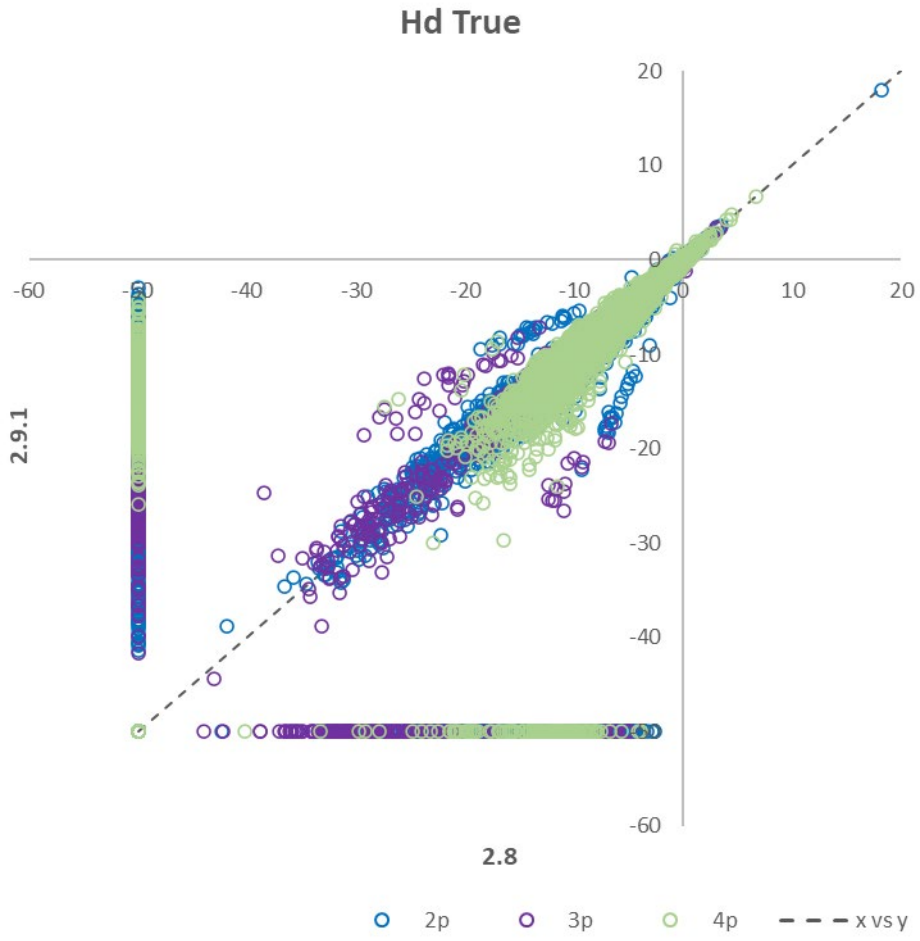
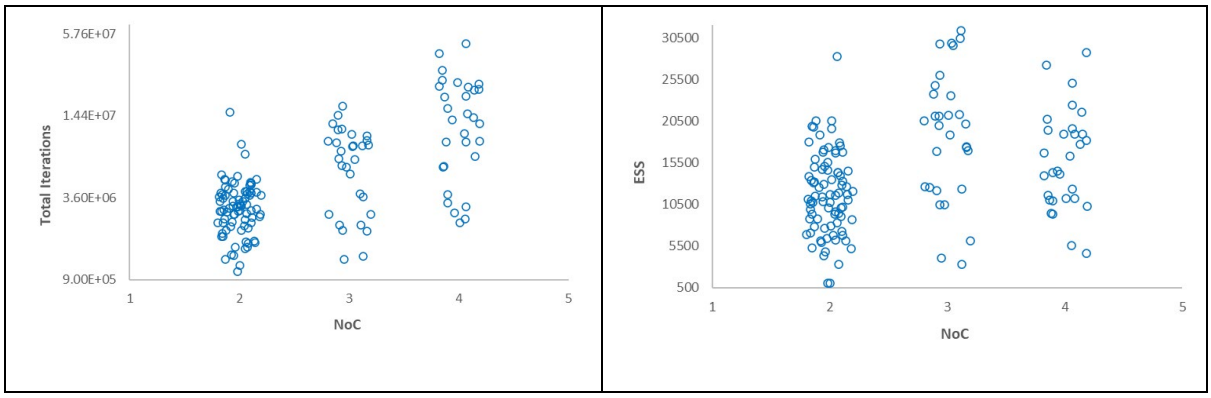


Figure 8: Comparisons of log(LR)s for non-contributors in STRmix V2.8 and V2.9.1 using the Fusion 5C kit

The diagnostics for the Fusion 5C mixtures interpreted in STRmix™ V2.9.1 are in Figure 9. All results are within expectations. Note that there is one outlier in the back stutter posterior mean plot (sample 10.5.1_C1 k^2~69). This outlier was also observed in the STRmix™ V2.8 validation and is due to a missing stutter peak at D22S1045 and is correctly identified on the STRmix™ report under Evidence Peak Issues.



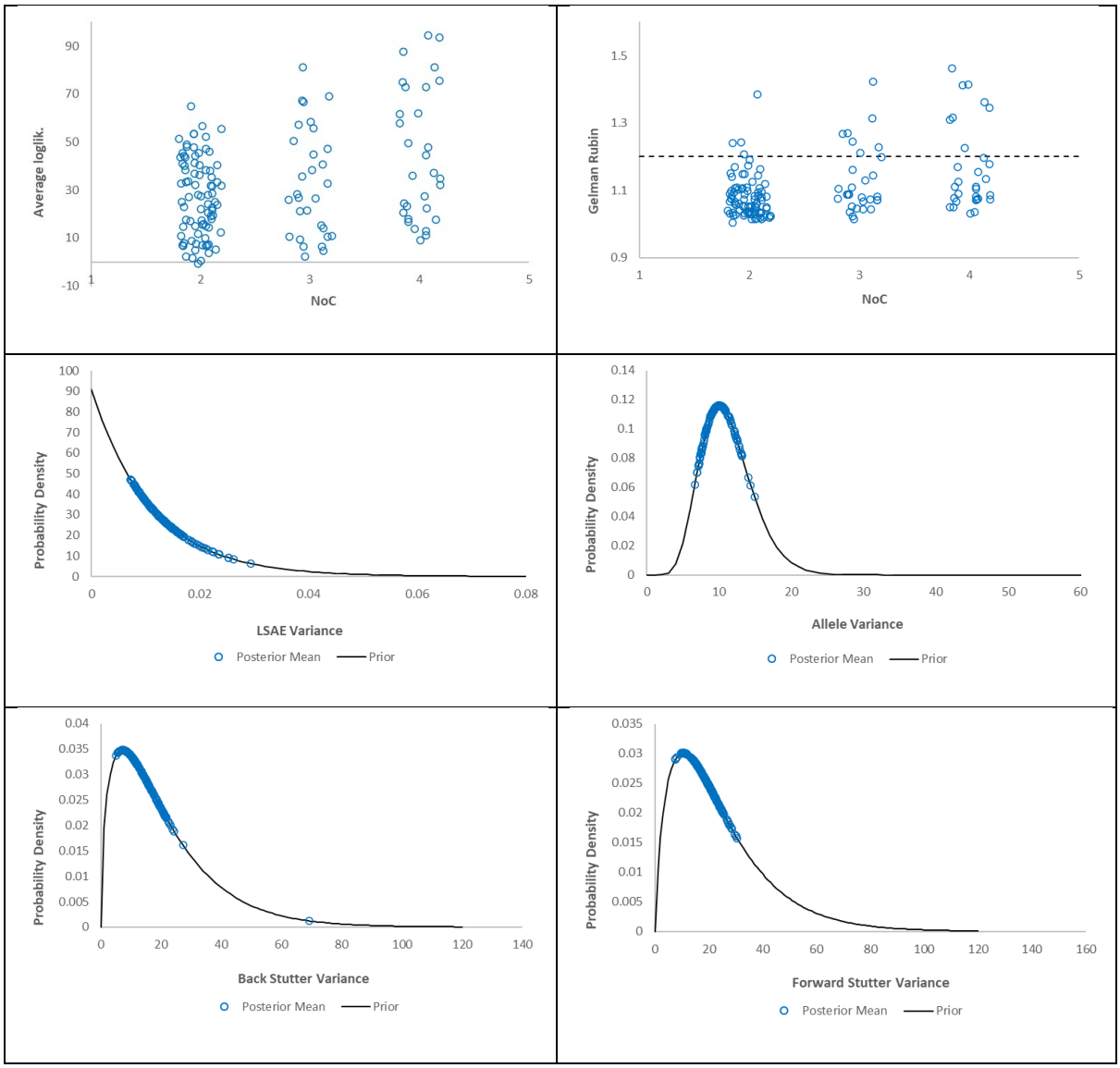


Figure 9:Diagnostics from STRmix V2.9.1 interpretations of Fusion 5C profiles

Section D

LR from Previous interpretations on deconvolutions carried out in previous versions should result in the same LR.

12 mixtures that had previously been interpreted in STRmix™ V2.8, had LRs assigned to each contributor using the LR from Previous function in STRmix™ 2.9.1. Specifically, six two-person, three three-person and four four-person mixtures.

The propositions considered were:

H_p: The DNA originates from the POI and *N*-1 unknown, unrelated individuals

H_d: The DNA originates from *N* unknown, unrelated individuals

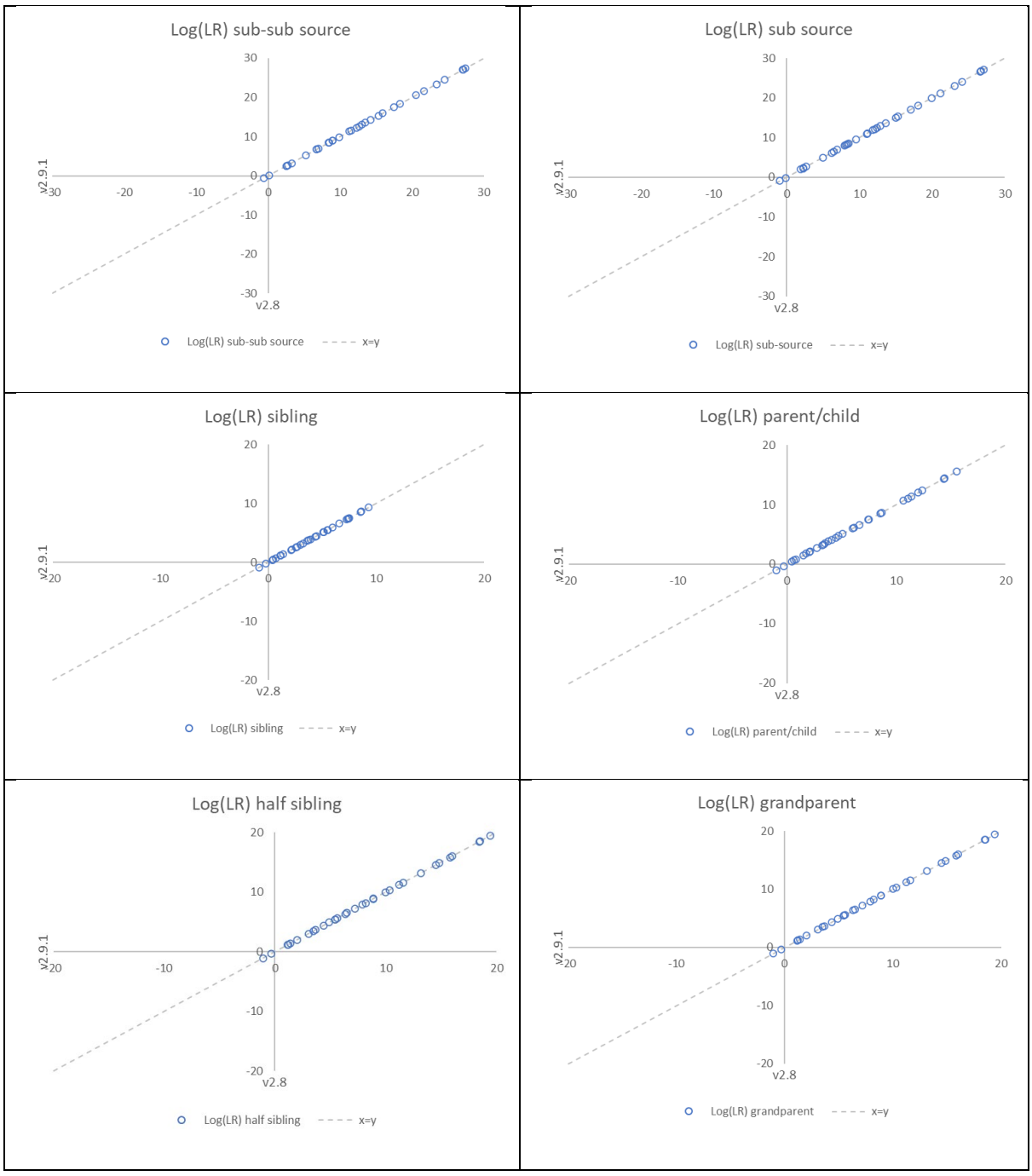
And

H_p: The DNA originates from the POI and *N*-1 unknown, unrelated individual

H_r: The DNA originates from a nominated relative of the POI and *N*-1 unrelated individual

Where the each of the known contributors in each mixture was considered the POI in turn, resulting in 33 *LR*s in total. As per the LRs assigned in STRmix™ V2.8, STRmix™ V2.9.1 LRs were assigned using the FBI Caucasian allele frequencies with θ = 0.01.

The resulting LRs were compared to the LRs obtained in STRmix™ V2.8. The results are in Figure 10 below. All LRs were identical. This is the expected result.



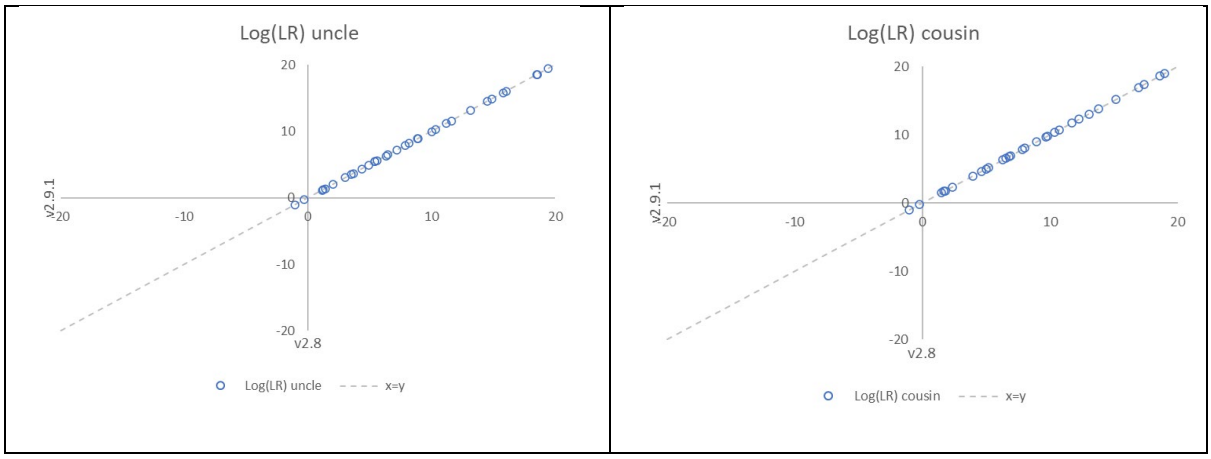


Figure 10: Comparison of STRmix™ 2.8 log(LR)s and STRmix™ 2.9.1 log(LR)s. The top pane is considering an unrelated individual under H_d and the remaining panes are considering a relative of the POI under H_d. The nominated relationship is in the title.

Conclusion

This document describes the upgrade and performance check of STRmix™ V2.9.1 using Fusion 5C PCR kits within the ISP Laboratory. It has been shown that it is suited for its intended use for the interpretation of profiles generated from crime scene samples.

Signatures

ISP Laboratory STRmix™ implementation manager

ISP Laboratory Technical Leader

This work has been reviewed and it has been determined that STRmix™ V2.9.1 is suitable for its intended use for interpretation of crime profiles within the ISP Laboratory. The project work has met the validation requirements as required by A2LA and FBI QAS.