COLORADO
Bureau of Investigation
Department of Public Safety

MEMORANDUM

TO: Clint Thomason, Quality Director

FROM: Sarah Miller, DNA Program Manager/Technical Leader

DATE: September 5, 2018

SUBJECT: Approval of STRmix™ v2.5.11 for Casework use

In August 2018, the Institute of Environmental Science and Research (ESR) submitted the summary and associated data files for the internal validation of STRmix™ v2.5.11 for the CBI-FS laboratories. This validation is in compliance with the current (2011) FBI *Quality Assurance Standards for Forensic Testing Laboratories* as well as the 2015 Scientific Working Group on DNA Analysis Methods (SWGDAM)'s publication *Guidelines for the Validation of Probabilistic Genotypying Systems*. It is comprised of two parts:

- Estimation of STRmix™ Parameters for the CBI Laboratory System (GlobalFiler, 3500xL CE)
- Internal Validation of STRmix™ v2.5 for the CBI Laboratory System (GlobalFiler, 3500xL CE)

The first step of the STRmix™ validation involved gathering DNA profiles created during the original (2015) GlobalFiler validation to establish the software's parameters for the CBI-FS. The second step involved using different DNA profiles from the GlobalFiler validation as well as creating new profiles using the current CBI-FS procedures and instrumentation in all 5 DNA labs. This second validation step incorporated the established CBI-FS STRmix™ parameters.

On this day I have completed my write up and review of this validation and approve of the use of STRmix™ v2.5.11 on CBI-generated GlobalFiler profiles in all CBI-FS laboratories, pending the appropriate competency testing of DNA analysts and publication of applicable DOMs.

Please see attached summaries.

X _____ 09 05 18

Approved **Clint Thomason, Quality Director** **Date**

Estimation of STRmix™ parameters for the Colorado Bureau of Investigation (CBI) Laboratory system (GlobalFiler™, 3500xL CE)

STRmix™ Implementation

This document describes the estimation of the STRmix™ parameters for GlobalFiler™ DNA profiling data (29 cycles, 3500xL CE) from the Colorado Bureau of Investigation Laboratory system (hereafter called CBI) for use in STRmix™ V2.5.

STRmix™ parameters

There are a number of parameters which are not optimized 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 Model Maker. The laboratory specific parameters that are determined prior to use of STRmix™ are:

- Analytical threshold (detection threshold)
- Stutter ratios
- Drop-in parameters
- Saturation
- Allelic and stutter peak height variance
- The hyper-parameter for the variance of locus specific amplification effects (LSAE).

These parameters need to be defined for each STR kit, each protocol (e.g. cycle number variation), and CE platform (e.g. 3130 or 3500), and potentially each time there is a significant change of platform (e.g. a camera or laser change). Stutter settings, stutter ratios, analytical thresholds and drop-in parameters were determined for CBI's 29 cycle GlobalFiler™ data analyzed on 3500xL capillary electrophoresis instruments using empirical data. Peak height variance and locus specific amplification efficiencies were calculated using Model Maker within STRmix™ from analysis of empirical profile data. The results of these analyses are described within this report.

Analytical Thresholds

The assignment of a signal as allelic product as opposed to baseline or noise is important in DNA profile analysis. This differentiation is usually undertaken using a set threshold above which peaks are deemed to be allelic if they also meet certain morphological requirements, and below which they are ignored, regardless of morphology. The issue is to assign a threshold, often termed the limit of detection (LOD) or analytical threshold (AT), to minimize the detection of artifacts whilst maximizing the detection of allelic peaks.

Optimum AT values have previously been determined by CBI for all the GlobalFiler™ loci, see the *2016 GlobalFiler and YFiler Plus AT and ST summary* on Qualtrax. These dye-specific thresholds have been set as follows:

Table 1: Analytical thresholds for use by CBI

Channel color	Analytical threshold (RFU)
Blue	60
Green	90
Yellow	40
Red	70
Purple	70

These values were used for all data analysis within this report.

Stutter

110 profiles from CBI were used to inform STRmix™ stutter parameters. The data was analyzed in GeneMapper® (ThermoFisher) using dye channel specific Analytical Thresholds (AT) of 30 relative fluorescence units (RFU) for blue, 45 RFU for green, 20 RFU for yellow, and 35 RFU for red and purple to capture back and forward stutters. These values are half the intended casework AT. Peak heights from the analyzed data generates the stutter parameters and files for CBI's GlobalFiler™ system.

Back Stutter

There are three parameters within STRmix™ that calculate expected back stutter rates and therefore require optimization. The first is the **maximum allowable stutter ratio**. The maximum allowable stutter ratio reduces run time by only permitting peaks in a stutter position below a certain percentage to be considered stutter. The highest back stutter observed in this data set was 17.8% at SE33. As such this parameter has been set conservatively high at **0.3** (30%) based on inspection of this empirical data.

The second parameter is a file used to model the expected heights of the stutter peaks based on their partner allele designation. The values used to determine expected stutter heights are 'per allele'. Per allele stutter ratios (*SR*) are calculated using a linear equation and regressing stutter ratio against allele.

Within STRmix™, stutter is estimated using the model $SR = m \times Allele + c$ where the intercept (*c*) and slope (*m*) are determined using linear regression.

Values for *m* and *c* were calculated. A plot of *SR* versus Allele (and *SR* versus LUS, which is discussed below) for each locus is provided in **Appendix 1**. A summary of the STRmix™ allelic stutter file for the CBI's data is given in Table 2.

Table 2: Linear equation parameters for estimating CBI's per allele GlobalFiler™ stutter ratios for STRmix™. The associated text file has been titled: **"CBI_GlobalFiler_Stutter.txt"** for use in STRmix™ at the CBI laboratories.

Locus	Intercept	Slope
D3S1358	-0.05756	0.00865
vWA	-0.08711	0.00939
D16S539	-0.04778	0.00924
CSF1PO	-0.05143	0.00978
TPOX	-0.02805	0.00563
Yindel	N/A	N/A
D8S1179	0.01281	0.00383
D21S11	-0.08816	0.00537
D18S51	-0.0378	0.00707
DYS391	N/A	N/A
D2S441	0.0376	0.0008
D19S433	-0.05875	0.00896
TH01	0.01527	0.00054
FGA	-0.07689	0.00675
D22S1045	-0.12612	0.01374
D5S818	-0.04778	0.00956
D13S317	-0.05701	0.00964
D7S820	-0.05228	0.00951
SE33	0.03802	0.00229
D10S1248	-0.03345	0.0078
D1S1656	0.02982	0.00304
D12S391	-0.0964	0.00923
D2S1338	-0.01054	0.00428

A better explanatory variable for stutter ratio for some loci with compound and complex repeat structure has been shown to be the longest uninterrupted stretch of common repeats (LUS) within the allele [1-3] and not the allele designation itself. Values for LUS are determined by sequencing alleles. A number of common alleles for forensic loci have been typed. A summary of these appear on STRBase [4, 5]. A plot of *SR* versus LUS for compound and complex loci within GlobalFiler™ is provided within **Appendix 1**. Some of the plots of *SR* versus LUS are provided for comparison only as for some loci where neither a linear model based on Allele nor LUS well describes the data, a decision was made to adopt an average of the observed *SR* data for each allele at that locus. Please refer to Table 3 for details.

The third parameter within STRmix™ that determines expected stutter peak heights is an exception file based on either LUS or an average observed stutter ratio. LUS is used where it is a good explanatory variable for *SR* otherwise the average of the observed *SR* is used. A stutter exception file based on laboratory data has been created and was used in this analysis. Where a 0 appears in a column for a given allele in this file the expected stutter rates are calculated from the allele file (Table 2). A summary of the source of the predicted *SR* for each locus is given in Table 3.

Table 3: A summary of the explanatory variables for the predicted SR for each of the GlobalFiler™ loci

Locus	Source
D3S1358	Allele
vWA	Average
D16S539	Allele
CSF1PO	Allele
TPOX	Allele
Yindel	N/A
D8S1179	Average
D21S11	Average
D18S51	Allele
DYS391	N/A
D2S441	Average
D19S433	LUS
TH01	LUS
FGA	Average
D22S1045	Allele
D5S818	Allele
D13S317	Allele
D7S820	Allele
SE33	Average
D10S1248	Average
D1S1656	LUS
D12S391	Average
D2S1338	Average

A copy of the Stutter Exceptions file for the GlobalFiler™ multiplex is provided within **Appendix 2**. The file format in STRmix is .csv and this file has been titled: "**CBI_GlobalFiler_Exceptions.csv**" for use at the CBI laboratories.

Forward Stutter

There are two parameters within STRmix™ that calculate expected forward stutter ratios and therefore require optimization. The first is the **maximum allowable forward stutter ratio**. The maximum allowable stutter ratio reduces run time by only permitting peaks in a forward stutter position below a certain percentage to be considered forward stutter. This parameter has been set conservatively high at **0.15** (15%) based on inspection of the laboratory's forward stutter ratio data. Please see comments regarding D22S1045 and SE33 below.

The second parameter is a file used to model the expected heights of the forward stutter peaks based on their parent allele designation. Within STRmix™ V2.5 forward stutter ratios (FSR) are modelled using a per allele model, if applicable. The same 110 single source profiles, as described above, were analyzed (with the AT's described above) and forward stutter peaks were captured. A summary of the numbers of observed forward stutter peaks is given in Table 4. The markers demonstrating the highest rates of forward stutter were D22S1045 and SE33. This is not unexpected for D22S1045 as this is a tri-nucleotide repeat marker which are known to stutter more. The highest observed FSR was observed in D22S1045 ~6% and SE33 ~7%.

It is assumed that all loci are stuttering in a forward position (N+1), however most of these peaks are below the analytical threshold and therefore not visible. Missing forward stutter peaks were inserted at half the analytical threshold for each channel. This is a statistical method to account for data paucity.

D22S1045 is the only locus where there is both sufficient data and it supports a linear model based on allele designation as a good predictor of FSR. For this locus, as for back stutter ratios, the intercept and slope within the parameters was determined by regression. For all other loci the average observed FSR was calculated after missing data was reinserted into the dataset at half the AT used, provided the parent peak height was 1000 RFU or greater. The FSR average is then used as the intercept value, and the slope set to 0 within the Forward Stutter file.

Plots of the individual loci forward stutter data are given in **Appendix 3**.

Table 4: Forward stutter observations and average forward stutter ratios (where intercept = average). This text file has been titled: "**CBI_GlobalFiler_Forward Stutter (AT adjusted).txt**" for use at the CBI laboratories

Marker	Count of FS observations	Intercept	Slope
D3S1358	38	0.0053	0
vWA	31	0.00595	0
D16S539	38	0.00735	0
CSF1PO	40	0.00645	0
TPOX	1	0.00466	0
Yindel	N/A	N/A	N/A
D8S1179	43	0.00654	0
D21S11	39	0.0089	0
D18S51	24	0.00797	0
DYS391	N/A	N/A	N/A
D2S441	88	0.0069	0
D19S433	11	0.00489	0
TH01	3	0.0034	0
FGA	57	0.00576	0
D22S1045[#]	97	-0.04894	0.00548
D5S818	79	0.00774	0
D13S317	52	0.00605	0
D7S820	16	0.00604	0
SE33	38	0.00809	0
D10S1248	33	0.00559	0
D1S1656	74	0.00679	0
D12S391	22	0.00673	0
D2S1338	13	0.00465	0

For information purposes a plot of *FSR* versus allele is provided in **Appendix 3**.

[#]Note: Where sufficient real observations of forward stutter were available and the data supported a linear model (D22S1045), then the regression was taken from only the observed data without reinserting missing data.

Drop-in parameters

Drop-in is non-reproducible, unexplained peaks observed within a profile. There are four parameters used for the modelling of drop-in in STRmix™. These are:

1. Z: the detection threshold or analytical threshold
2. A cap on the maximum allowed drop-in peak height
3. The drop-in frequency
4. α,β: two parameters for the prior distribution model.

Drop-in rates for a laboratory platform (multiplex and instrument combination) should be monitored. This is done by recording counts and corresponding heights of drop-in peaks observed in negative controls and counts of negative controls without drop-in peaks. Within STRmix™ drop-in can be modelled using either a uniform (uninformed) or a gamma distribution.

The CBI laboratories have observed and monitored drop-in peaks within GlobalFiler™ profiles. The drop-in parameters for STRmix™ were determined based on empirical observations from the five laboratories (Boulder, Denver, Grand Junction, Pueblo, and NCRFL) at CBI using GlobalFiler™, 29 cycles & 3500xL CE. A total of 27 drop-in peaks were observed in 1065 negative profiles (23430 loci [21 autosomal loci plus Amelogenin per profile]).

CBI's drop-in parameters were determined using an Excel spreadsheet obtained from the STRmix™ training team [refer to the Excel workbook titled '**CBI Drop-in calculator II**']. Within this worksheet, the parameters are fitted by minimizing the squared differences of the modelled and observed data. The optimized parameters are provided in Table 5. Within STRmix™ drop-in for use in casework at the CBI laboratories will be modelled using a gamma distribution.

Table 5: Colorado Bureau of Investigation drop-in parameters for STRmix™ for the GlobalFiler™ data

Drop-in cap	**250**
Drop-in frequency	**0.0012**
Drop-in parameters	**15.52,5.21**

Saturation

The peaks in a DNA profile are measured using fluorescence. The amount of fluorescence is proportional to the amount of DNA present. This fluorescence is captured by a camera. It is expected that as more DNA is added into a PCR the resulting peak height (measured in relative fluorescent units, RFU) in an electropherogram will increase. The camera can become saturated when there is too much fluorescence detected. This means we can no longer accurately measure the height of the peaks observed or estimate how much DNA is really represented by this result. Following this we can no longer accurately model over-saturated peak heights using STRmix™. Saturation, like the analytical threshold, is mostly instrument related and not kit or method dependent.

Common saturation thresholds for a 3500 CE instrument are in the range of 28,000 – 32,000 RFU. A value of **30,000 RFU** has been selected for use across the CBI laboratories, which is based on CBI's internal validation of GlobalFiler™ and the experience of the STRmix™ scientific support team's empirical laboratory observations.

Peak height variance and LSAE using Model Maker

Empirical observations and experience suggests that profiles differ in variance (hereafter "quality"). Within STRmix™ the variability of peaks within profiles is described using a model containing a variance constant. Allele and stutter peaks have separate variances; c^2 and k^2, respectively. The c^2 and k^2 terms are variables which are determined after sampling from a gamma distribution within the MCMC.

The gamma distribution priors that STRmix™ samples from during an interpretation are optimized in Model Maker, an add-on to the STRmix™ software. Model Maker works by using a component wise MCMC. In component 1 each DNA profile has its mass parameters optimized and uses a stable gamma distribution for allele, stutter and LSAE variance constants. In component 2 the mass parameters for each profile are held constant and the hyperparameters for each gamma distribution are varied. Components are 1000 accepts long and they cycle through a number of times depending on the user input value.

The Model Maker analysis within STRmix™ for the CBI laboratories utilized a 39-42 single source profiles of varying quality for 100 cycles (100,000 accepts total).

CBI makes use of 9 different CE instruments spread across its 5 laboratories. See *DNA DOM 05 Equipment* for general Instrument Protocol parameters and *DNA DOM 10-10 STR Typing via 3500xL Capillary Electrophoresis* for instrument-specific injection times. One Model Maker run was undertaken per CE [n=9], then per laboratory, for those labs with greater than 1 CE [n=3] and then as a whole system with data from all 9 CE instruments in one run.

A known donor (7057) was selected and a 14 step dilution series was created from (A) 4ng input template with intervals down to (N) 0.49pg. These were amplified, using the standard casework procedures, in triplicate. This yielded up to 42 profiles available to be run through the various CE instruments. The resultant CE data was analyzed in GeneMapper™ and then run in Model Maker. For the majority of the runs the 4ng data was omitted prior to running in Model Maker, as it exhibited saturation. A breakdown of the samples used is provided in Table 6.

Table 6: Summary of Model Maker (MM) input files for CBI

Laboratory	**Samples analysed in GeneMapper**	**Samples omitted at GeneMapper stage**	**No. read in to MM run**	**No. used by STRmix in MM run**	**Samples not used by STRmix in MM run**
Boulder B1	39	2 (G2 & M3)	37	32	(5) B1, M2 & N1 to 3
Denver D1	42	1 (J2)	41	31	(10) A1 to 3, B2, M1 to 3 & N 1 to 3
Denver D2	39	0	39	37	(2) N1 & N3
Denver D3	39	0	39	38	(1) N1
Denver D1 to D3			*119*	*106*	*See above*
Grand Junction W1	39	0	39	34	(5) B2, B3 & N1 to 3
Grand Junction W2	39	0	39	34	(5) M2, M3, N1 to 3
Grand Junction W1 to W2			*78*	*68*	
NCRFL N1	39	0	39	36	(3) N 1 to 3
Pueblo P1	39	0	39	33	(6) B1 to 3, M3, N2 & N3
Pueblo P2	39	0	39	36	(3) B1, B3 & N2
Pueblo P1 to P2			*78*	*69*	*See above*
Combined			*351*	*311*	*See above*

Note: The samples not used by STRmix™ in MM either contained peaks above 30,000RFU (A 4ng and B 2ng input template samples) or contained 10 or fewer peaks (the trace M 0.98pg and N 0.49pg input template samples).

A summary of the results for both c^2 and k^2 for the dataset is provided in Table 7 and the resultant gamma distributions are plotted in Figure 1 & Figure 2. The Model Maker results files can be found at P:/Lab/Discipline Files/DNA/STRmix/Validation/Phase 1/Model Maker.

Table 7: Summary of allele and stutter prior variance values for the CBI system

Multiplex	Number profiles analysed	Allele variance parameters	Stutter variance parameters	Mean LSAE variance
		α,β (Mode)	α,β (Mode)	
Boulder B1	32	9.798, 1.031 (9.072)	2.418, 6.434 (9.124)	0.007
Denver D1	31	10.353, 1.033 (9.658)	1.938, 8.714 (8.176)	0.011
Denver D2	37	20.643, 0.551 (10.814)	2.165, 7.085 (8.253)	0.013
Denver D3	38	7.337, 1.539 (9.754)	3.63, 4.051 (10.655)	0.022
Denver D1 to 3	106	6.726, 1.764 (10.098)	2.852, 5.569 (10.316)	0.013
Grand Junction W1	34	19.603, 0.662 (12.323)	3.684, 4.239 (11.376)	0.008
Grand Junction W2	34	9.502, 1.101 (9.363)	1.511, 16.025 (8.195)	0.008
Grand Junction W1 & 2	68	12.261, 0.868 (9.779)	1.818, 10.921 (8.937)	0.009
NCRFL N1	36	7.509, 1.515 (9.862)	5.076, 2.601 (10.603)	0.009
Pueblo P1	33	9.603, 1.2 (10.321)	1.862, 8.388 (7.232)	0.009
Pueblo P2	36	8.41, 1.815 (13.449)	2.179, 6.065 (7.153)	0.021
Pueblo P1 & 2	69	6.024, 2.364 (11.875)	2.498, 6.283 (9.415)	0.015
Combined	311	**8.240, 1.437** (**10.404**)	**2.377, 6.751** (**9.296**)[#]	**0.011**

[#]Note: Due to rounding on the report, the outputs created by the software are rounded to two decimal places and will therefore show a combined value of 9.30.

Figure 1: Summary of all the allele variance prior distributions (scales adjusted to be consistent with the stutter variance)

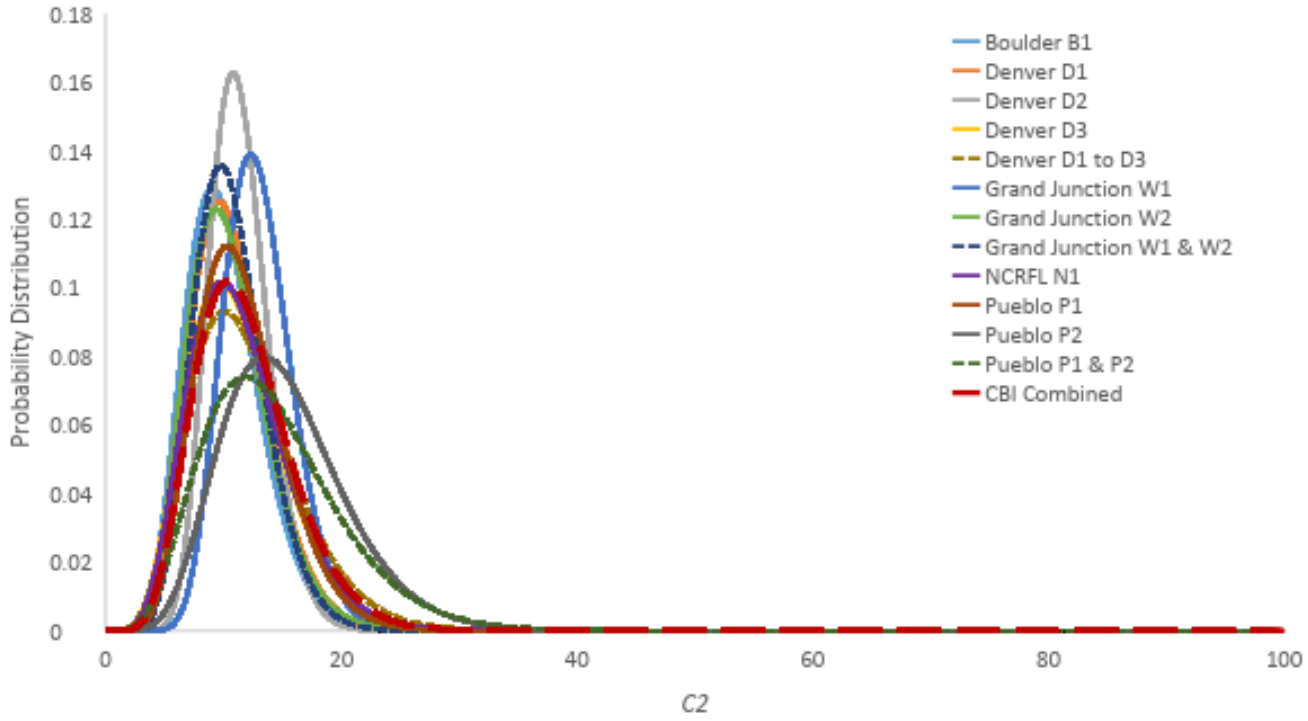
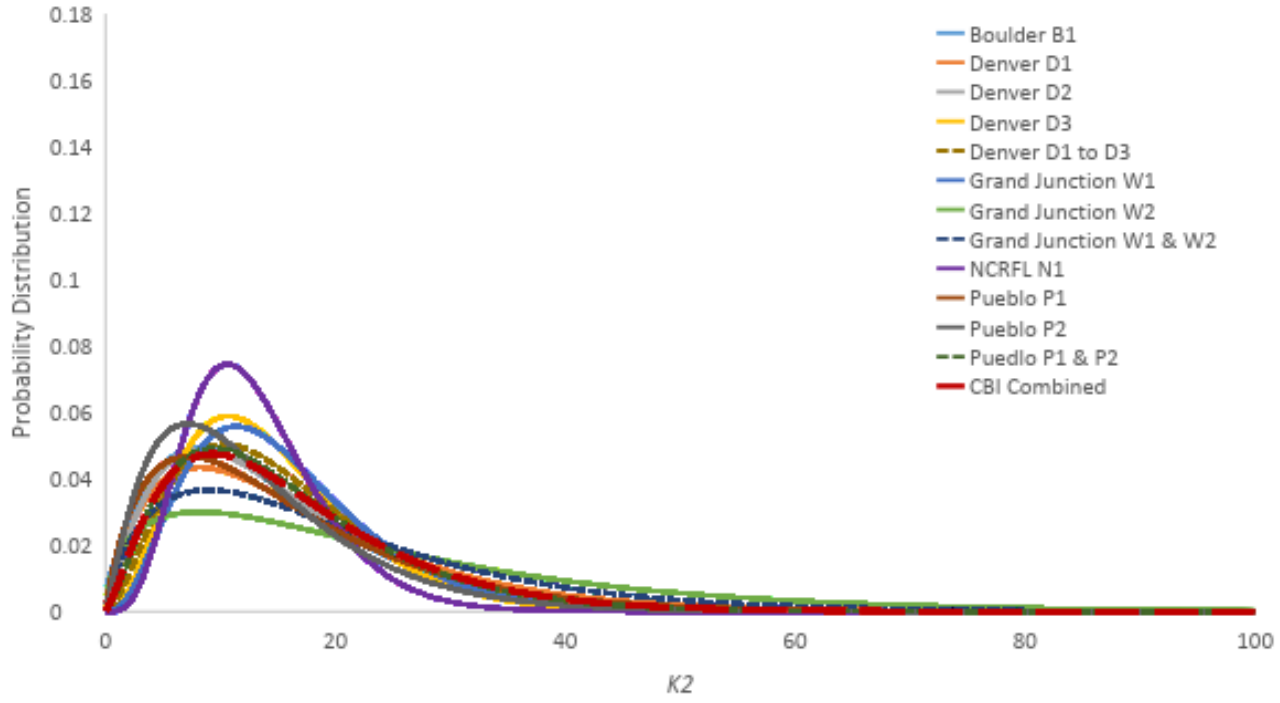


Figure 2: Summary of all the stutter variance prior distributions



Focusing on CBI's 5 different laboratories, below are further plots to show the prior gamma distributions observed for the CE instruments per site, and then the combined data set.

Figure 3: Summary of allele variance prior distributions obtained per laboratory & a combined data set (scales unadjusted)

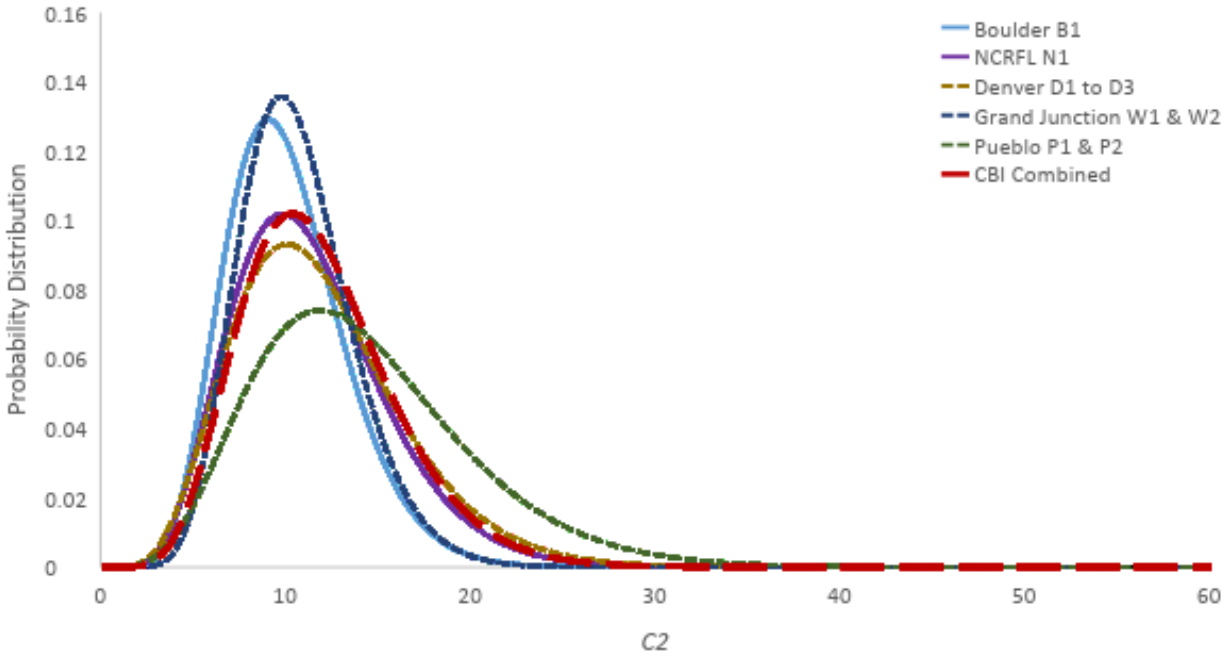
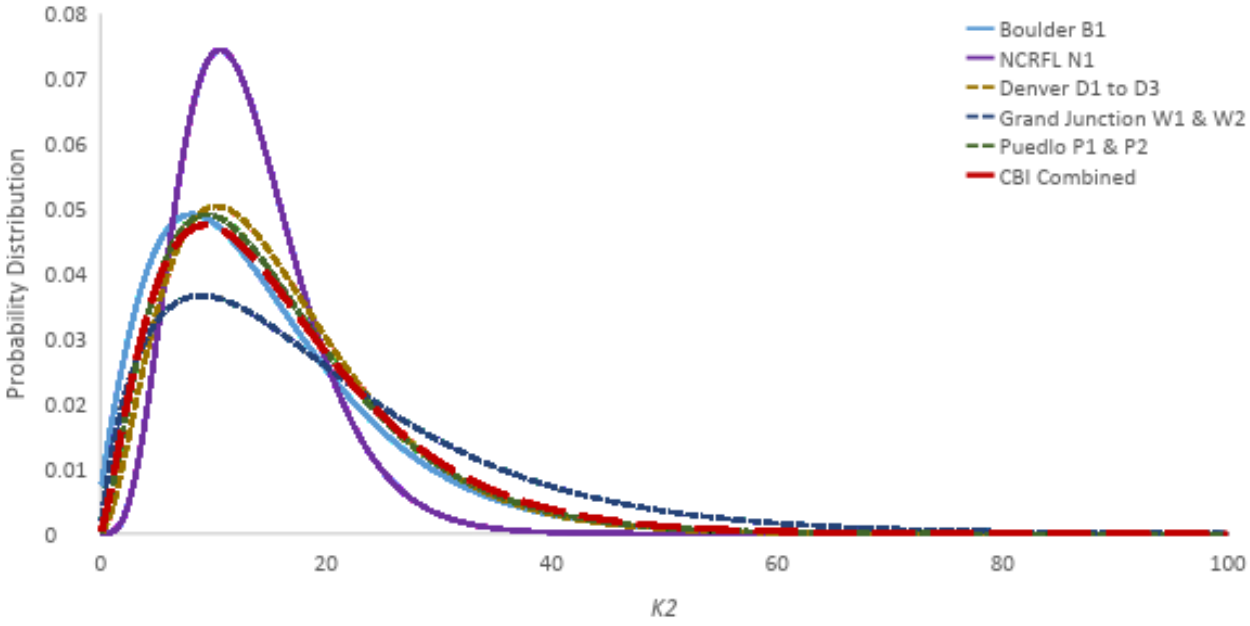


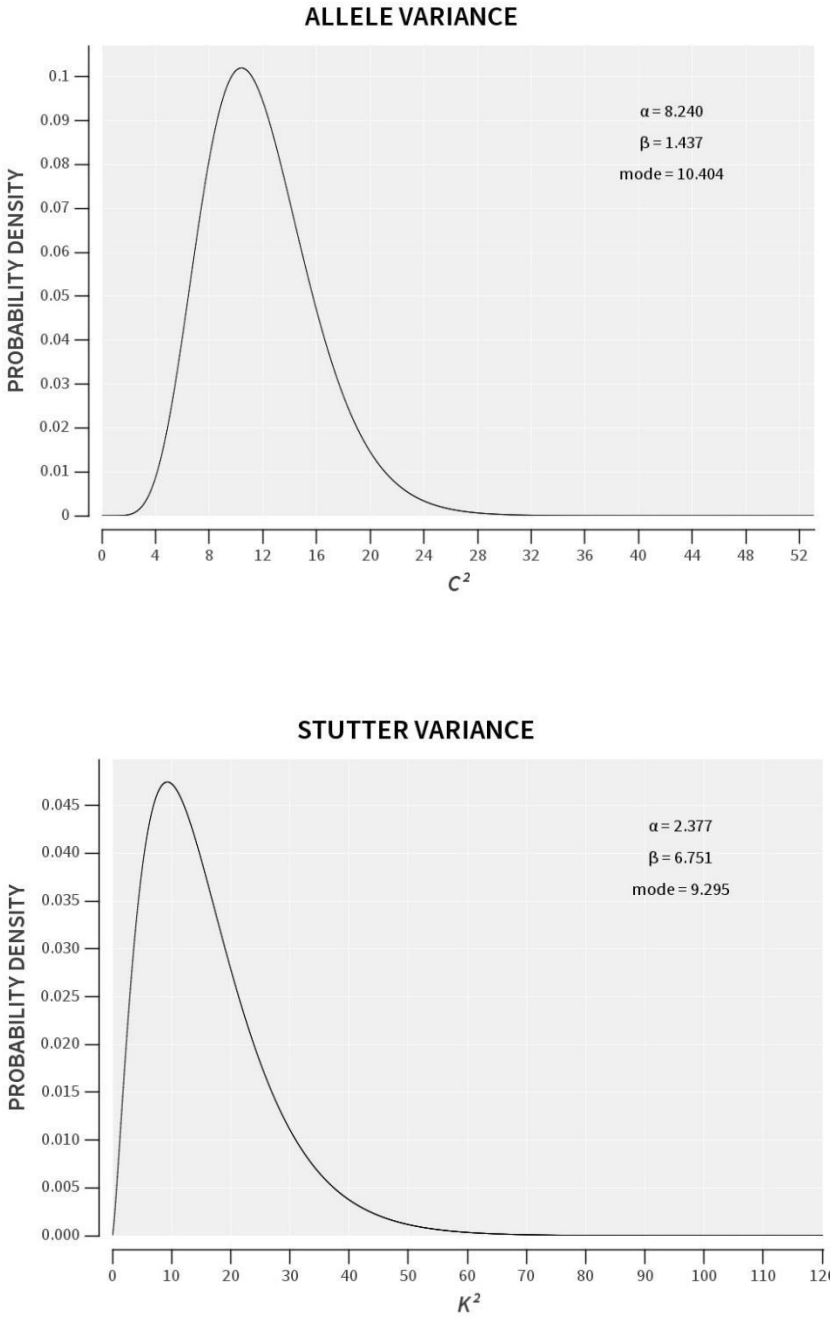
Figure 4: Summary of stutter variance prior distributions obtained per laboratory & a combined data set



The distributions are reasonably consistent between laboratories (and CE's) and as such the CBI laboratory system will implement one set of values to use across all sites/CE instruments. As such the **combined data set** has been chosen as the values to use in casework.

Provided below in Figure 5 are summary plots of the prior gamma distributions for allele and stutter variance to be used in casework across all the CBI sites. Note, the scales differ for allele and stutter variance.

Figure 5: A summary plot of the allele and stutter prior gamma distributions



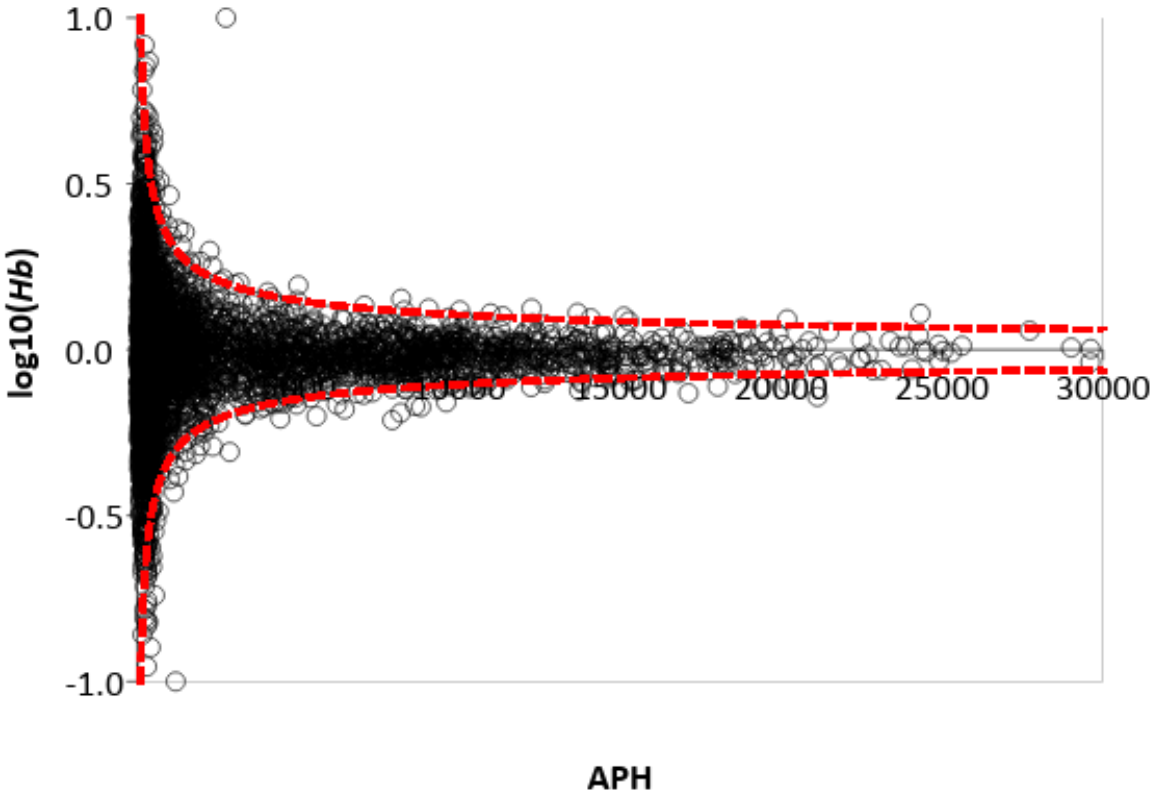
Heterozygote balance was calculated for all heterozygote loci from the combined data set of Model Maker profiles. Heterozygote balance (*Hb*) was calculated as:

$$Hb = \frac{O_{HMW}}{O_{LMW}}$$

Where O_{HMW} refers to the observed height of the high molecular weight allele and O_{LMW} the observed height of the low molecular weight allele. Previous work has suggested that there is a relationship between the variation in peak height and the variation in *Hb* [6, 7]. In single source profiles, variability in *Hb* reduces as the average peak height (APH) at a locus increases. The variance of *Hb* can be used as a proxy for the variance of individual peaks. This allows an approximate comparison between the variance from the STRmix™ MCMC approach and a readily determined variable from empirical data (*Hb)*.

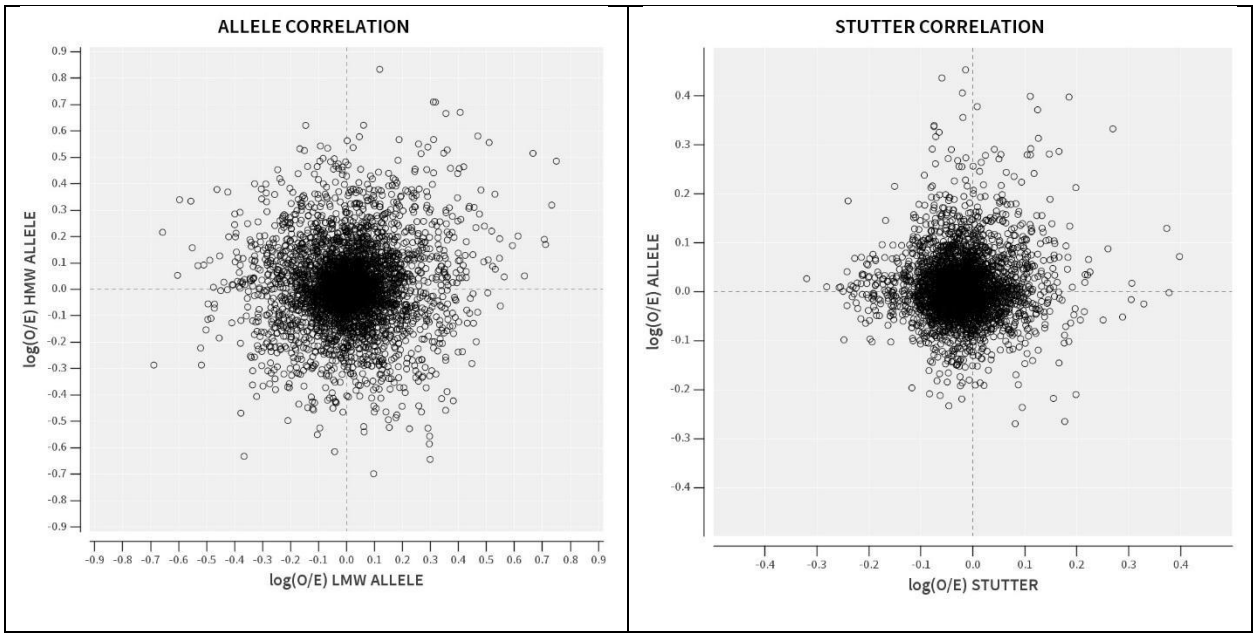
The plot of log(*Hb)* versus APH (the black circles) for the dataset described above and the expected 95% bounds (plotted as dotted red lines) calculated at $\pm\sqrt{2}\times 1.96\times\sqrt{\frac{c^2}{APH}}$ where c^2 = 14.3 (the 75th percentile from the prior gamma distribution from the CBI combined dataset) is provided in Figure 6. Under our assumption of a normal distribution we would expect ~95% of data points to fall within +/- 2 standard deviations (95% bounds) of the mean. The 95% bounds encapsulate sufficient data as demonstrated in the graph (coverage = 96.2%) demonstrating that the values for variance appear sufficiently optimized (Figure 6). This plot is a useful check of the Model Maker output.

Figure 6: Log(*Hb*) versus APH for the single source profiles used in the combined MM at the CBI laboratories



In Figure 7 we plot the correlation plots for Low Molecular Weight (LMW) versus High Molecular Weight (HMW) alleles, and stutter versus allele peaks for the combined Model Maker dataset. The distribution of the points within the figures is as expected, with no observed correlation.

Figure 7: CBI's GlobalFiler™ correlation plots

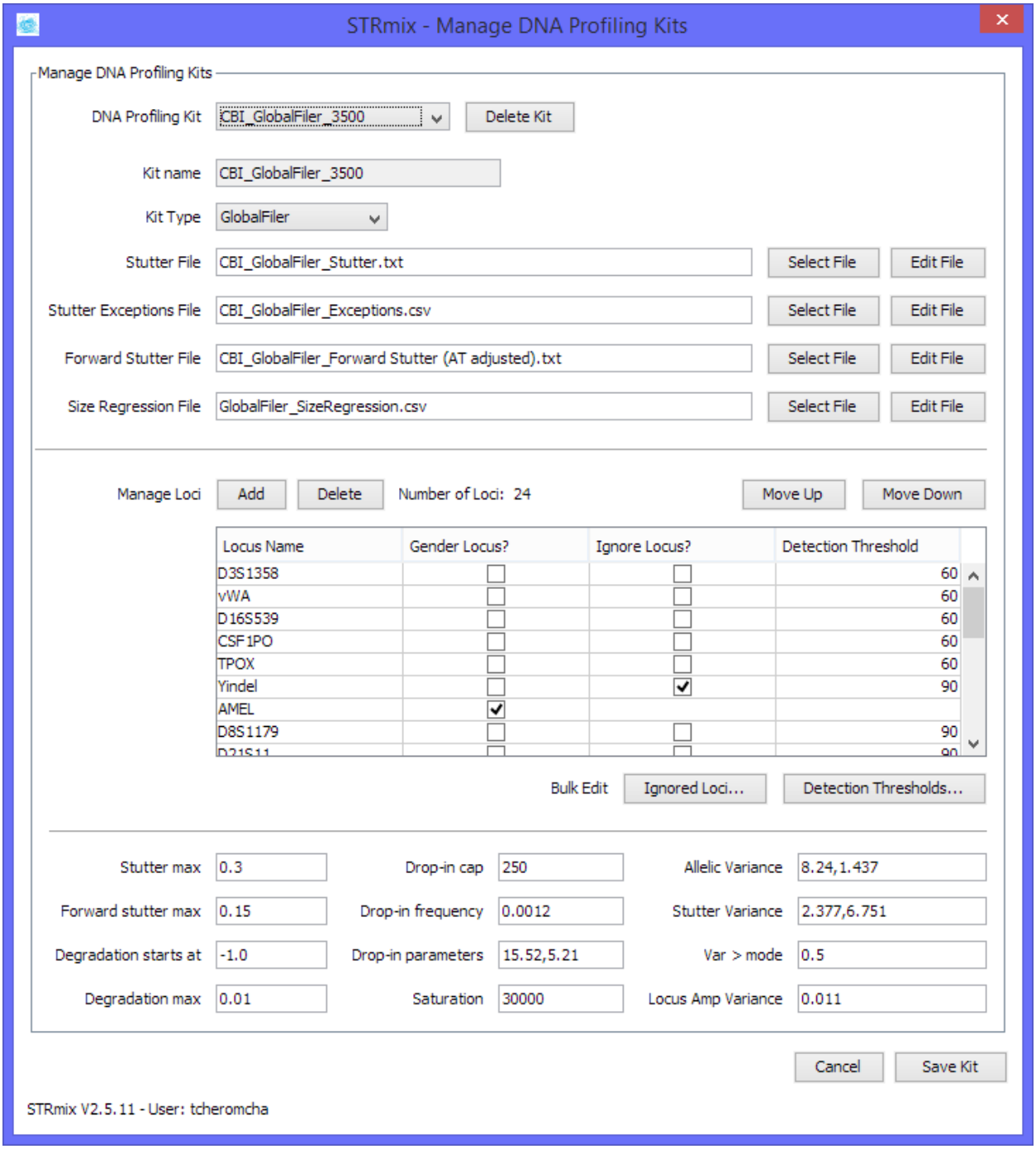


Inspection of the right hand plot of Figure 7 suggests a mild shift of the data to the left for the stutter correlation. This may be a consequence of the use of the same one profile diluted multiple times, rather than numerous individuals and this individual happens to have allele(s) with one or more stutters slightly off trend compared to the stutter files created from multiple individuals. This appears to have little to no effect on the overall Model Maker output.

Conclusions

The recommended STRmix™ V2.5 default parameters for the interpretation of the CBI's 29 cycle GlobalFiler™ profiles run on any of our 3500xL CE instruments are given in Figure 8.

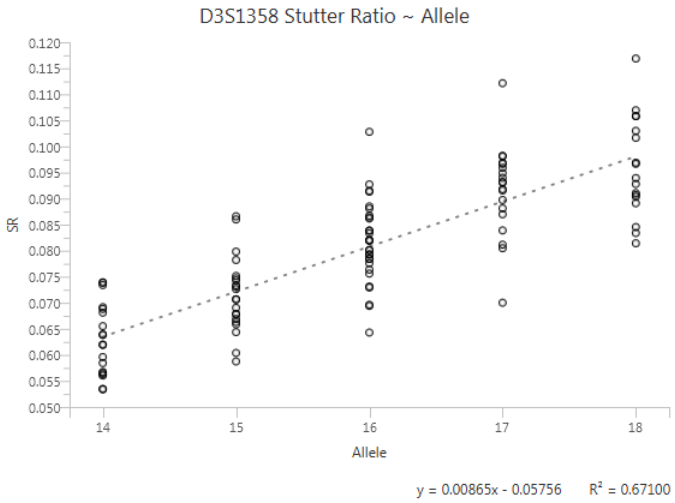
Figure 8: STRmix™ recommended default parameters for CBI GlobalFiler™ profile interpretation

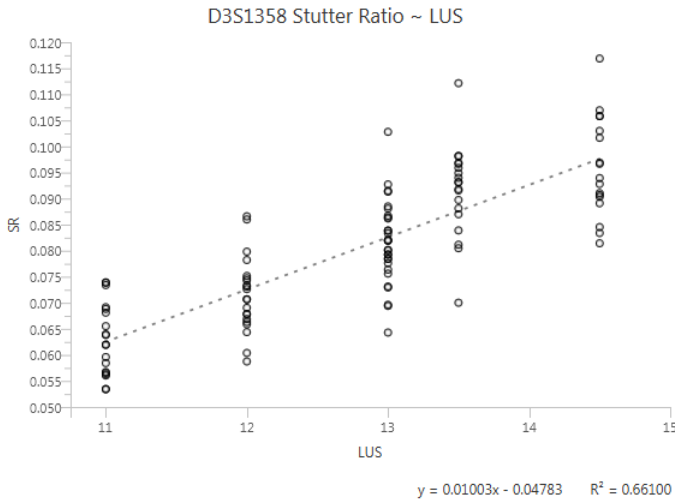


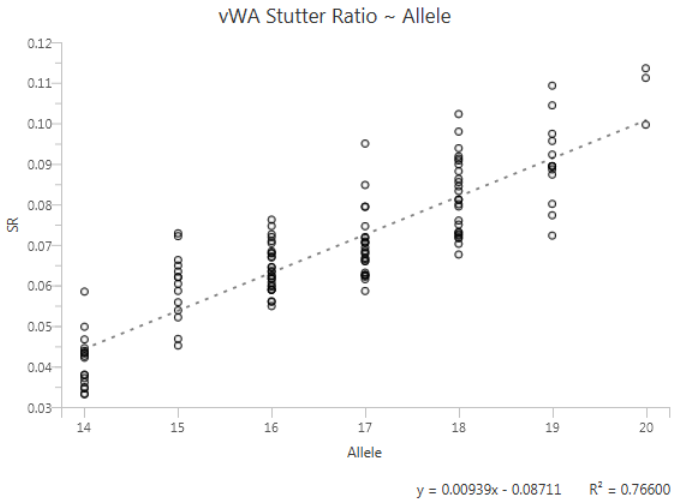
References

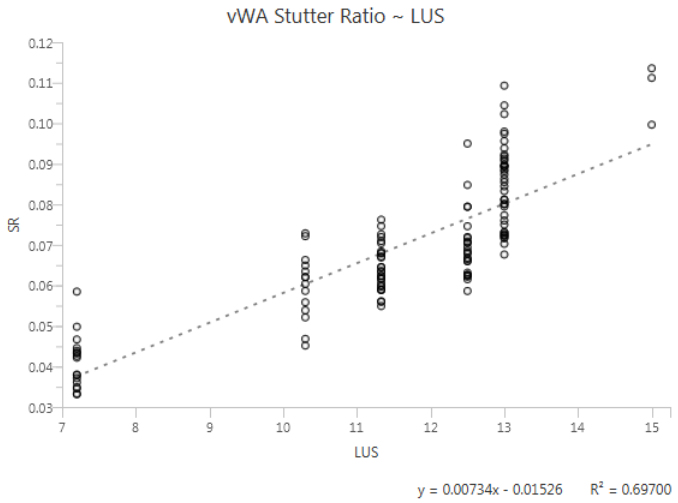
[1] Bright J-A, Taylor D, Curran JM, Buckleton JS. Developing allelic and stutter peak height models for a continuous method of DNA interpretation. Forensic Science International: Genetics. 2013;7:296-304.
[2] Brookes C, Bright J-A, Harbison S, Buckleton J. Characterising stutter in forensic STR multiplexes. Forensic Science International: Genetics. 2012;6:58-63.
[3] Walsh PS, Fildes NJ, Reynolds R. Sequence analysis and characterization of stutter products at the tetranucleotide repeat locus vWA. Nucleic Acids Research. 1996;24:2807-12.
[4] Butler JM, Reeder DJ. Short Tandem Repeat DNA Internet DataBase.
[5] Ruitberg CM, Reeder DJ, Butler JM. STRBase: a short tandem repeat DNA database for the human identity testing community. Nucleic Acids Research. 2001;29:320 - 2.
[6] Bright J-A, Huizing E, Melia L, Buckleton J. Determination of the variables affecting mixed MiniFiler™ DNA profiles. Forensic Science International: Genetics. 2011;5:381-5.
[7] Bright J-A, Turkington J, Buckleton J. Examination of the variability in mixed DNA profile parameters for the Identifiler multiplex. Forensic Science International: Genetics. 2009;4:111-4.

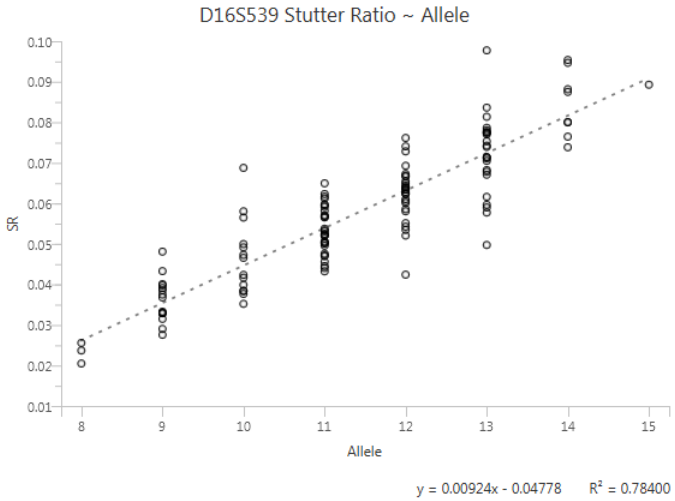
Appendix 1: *SR* versus Allele and *SR* versus LUS

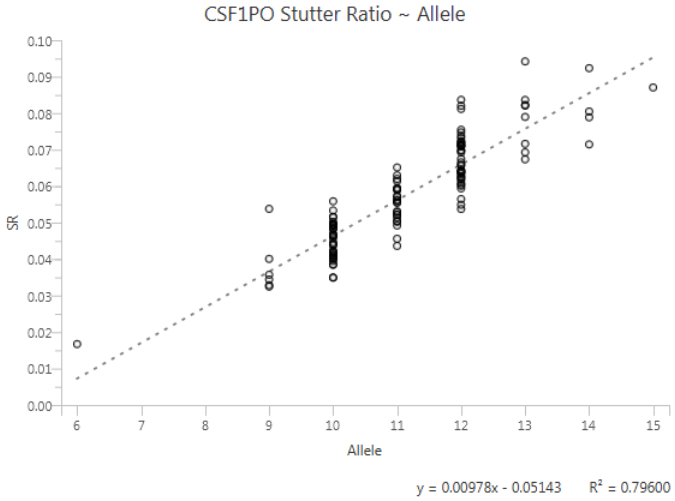


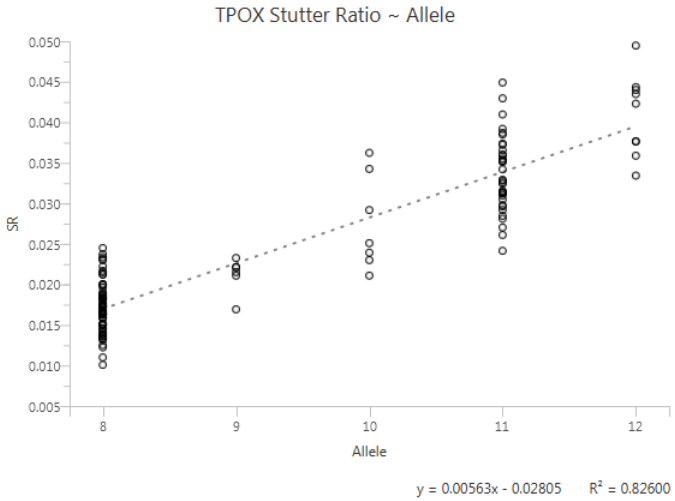


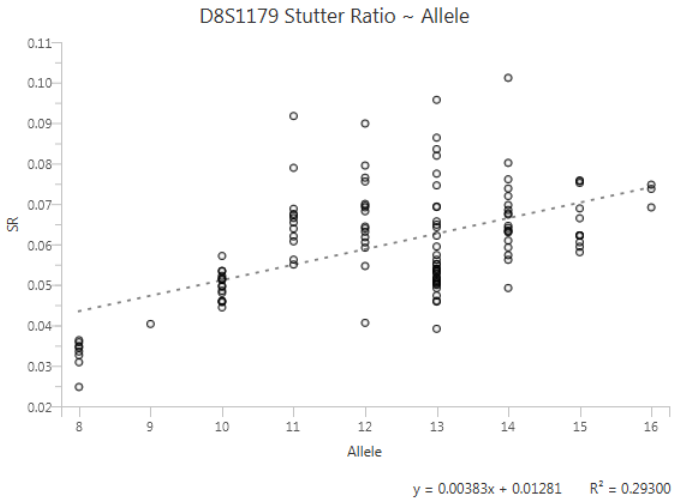


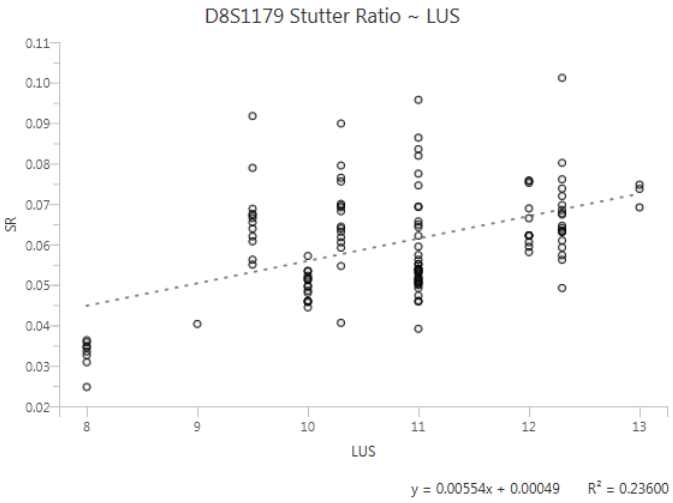


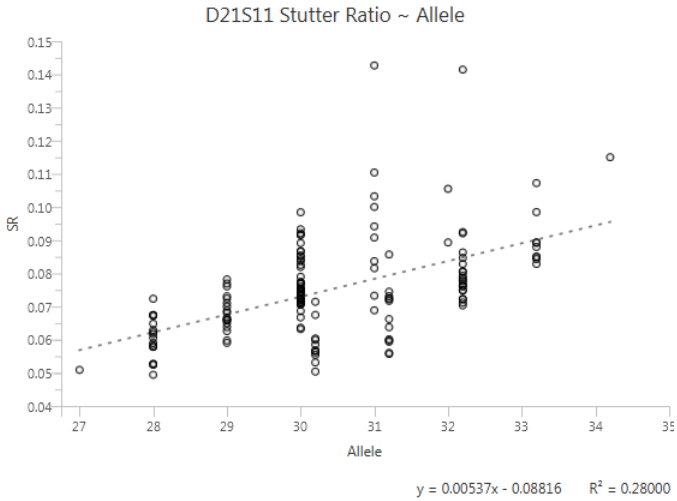

D16S539 Stutter Ratio ~ Allele
SR
Allele
y = 0.00924x - 0.04778
R² = 0.78400

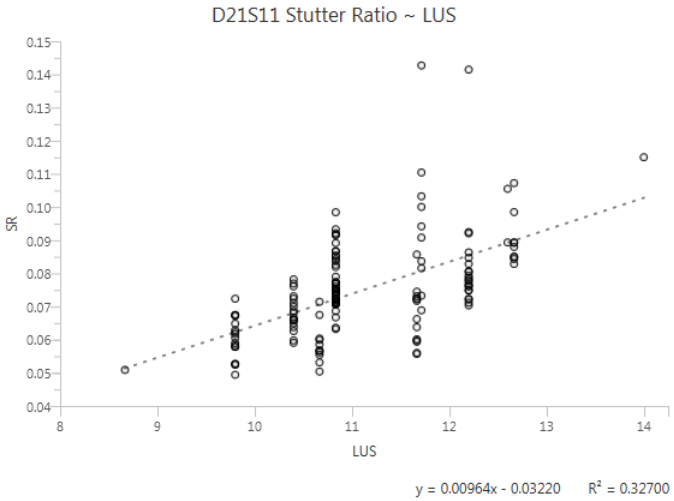

CSF1PO Stutter Ratio ~ Allele
SR
Allele
y = 0.00978x - 0.05143
R² = 0.79600

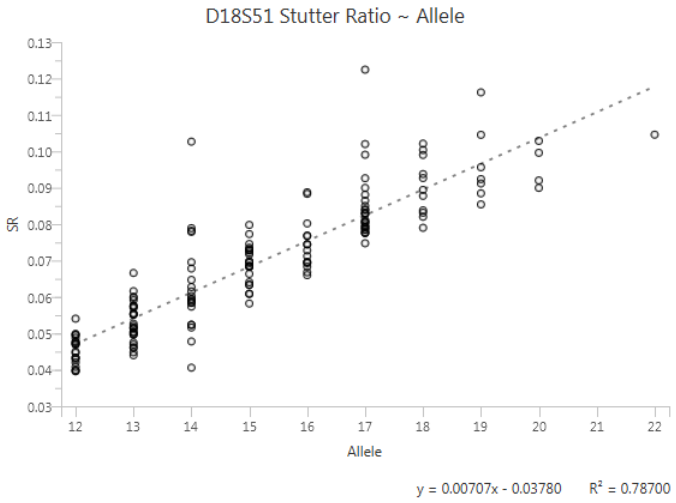

TPOX Stutter Ratio ~ Allele
SR
Allele
y = 0.00563x - 0.02805
R² = 0.82600

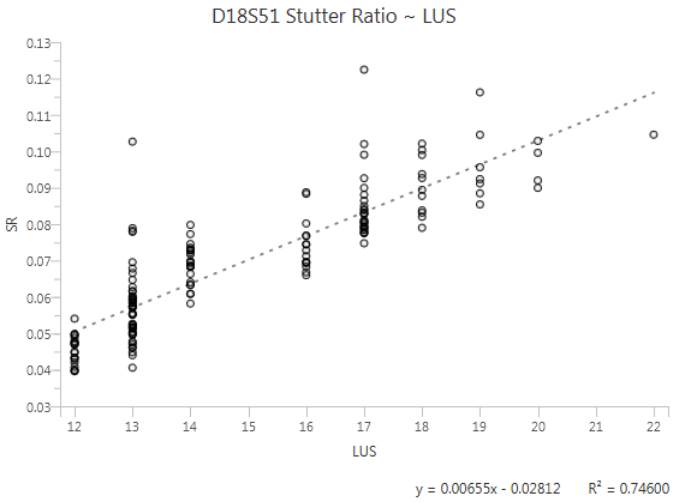

D8S1179 Stutter Ratio ~ Allele
SR
Allele
y = 0.00383x + 0.01281
R² = 0.29300

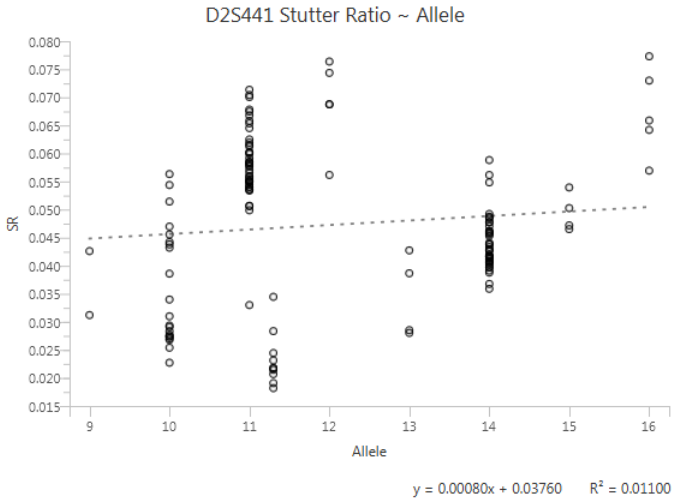

D8S1179 Stutter Ratio ~ LUS
SR
LUS
y = 0.00554x + 0.00049
R² = 0.23600

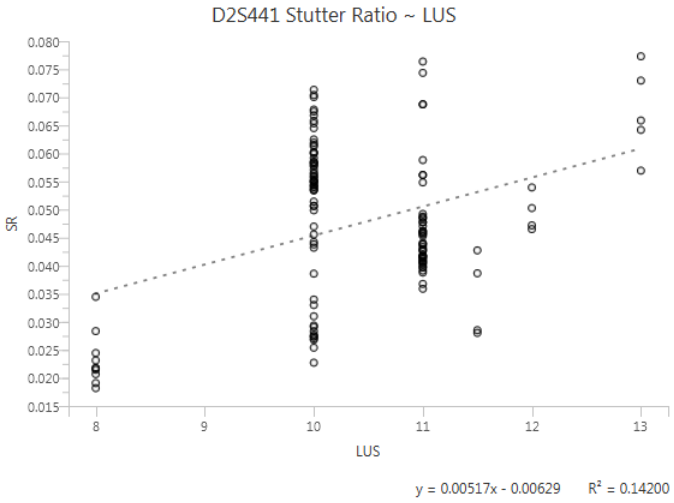

D21S11 Stutter Ratio ~ Allele
SR
Allele
y = 0.00537x - 0.08816
R² = 0.28000

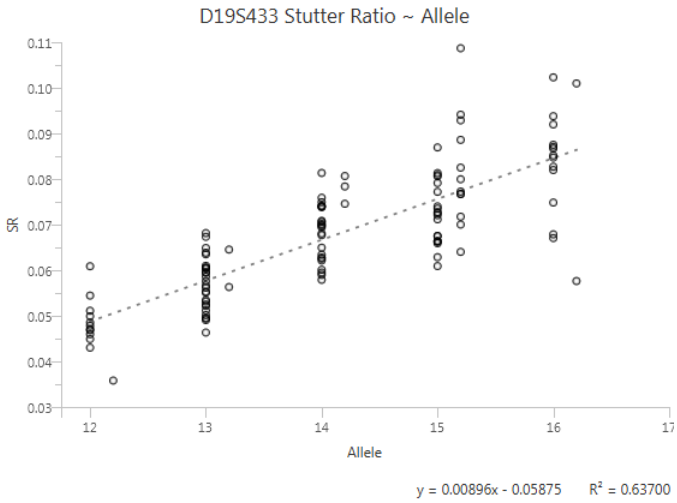

D21S11 Stutter Ratio ~ LUS
SR
LUS
y = 0.00964x - 0.03220
R² = 0.32700

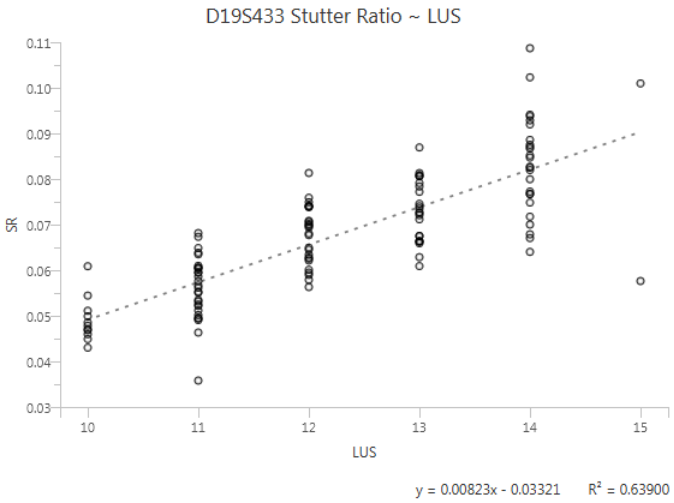

D18S51 Stutter Ratio ~ Allele
SR
Allele
y = 0.00707x - 0.03780
R² = 0.78700

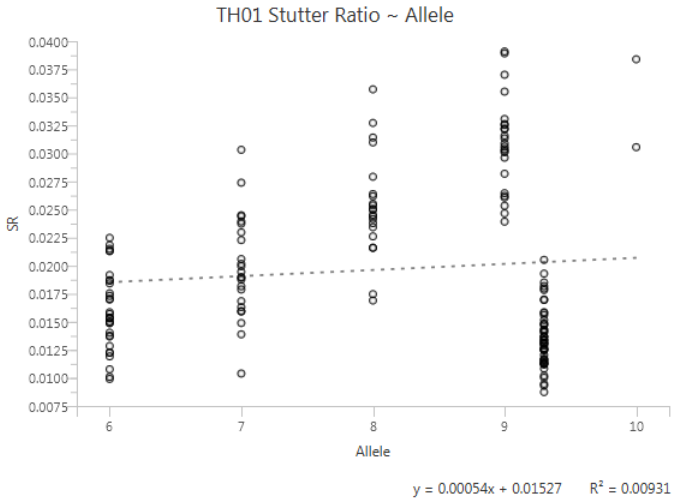

D18S51 Stutter Ratio ~ LUS
SR
LUS
y = 0.00655x - 0.02812
R² = 0.74600

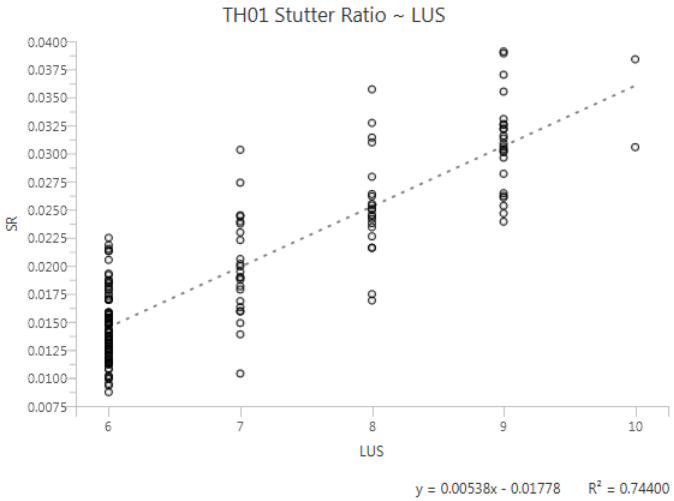

D2S441 Stutter Ratio ~ Allele
SR
Allele
y = 0.00080x + 0.03760
R² = 0.01100

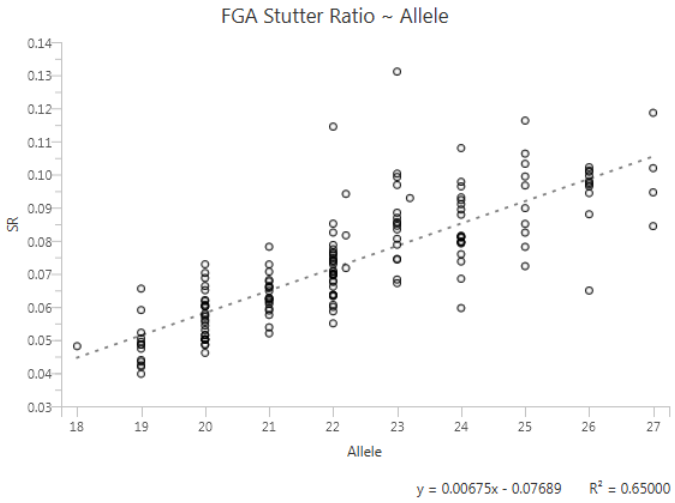

D2S441 Stutter Ratio ~ LUS
SR
LUS
y = 0.00517x - 0.00629
R² = 0.14200

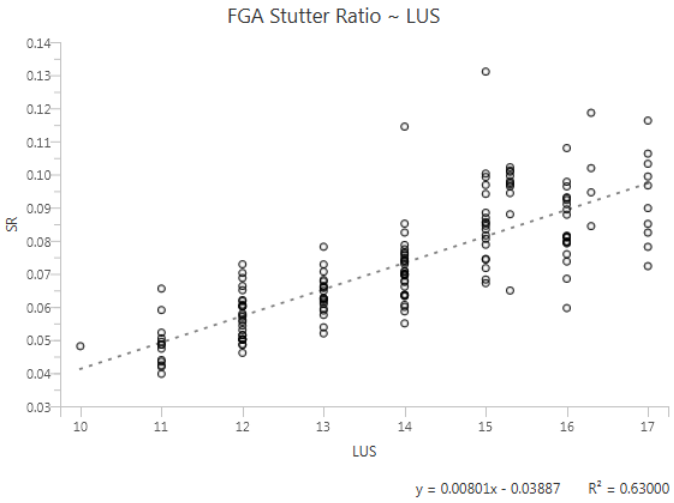

D19S433 Stutter Ratio ~ Allele
SR
Allele
y = 0.00896x - 0.05875
R² = 0.63700

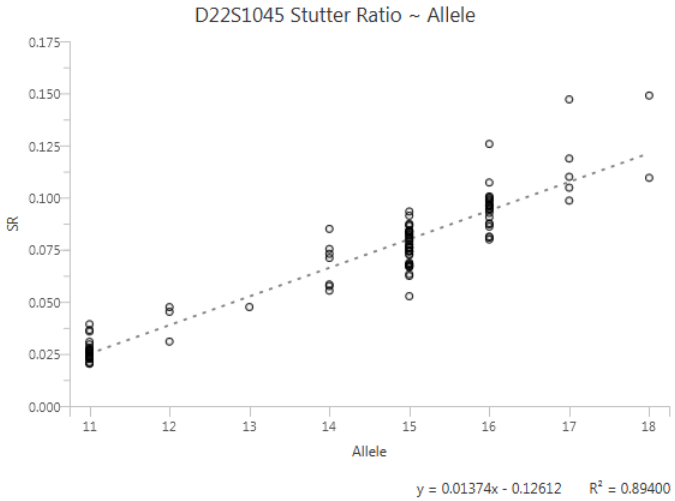

D19S433 Stutter Ratio ~ LUS
SR
LUS
y = 0.00823x - 0.03321
R² = 0.63900

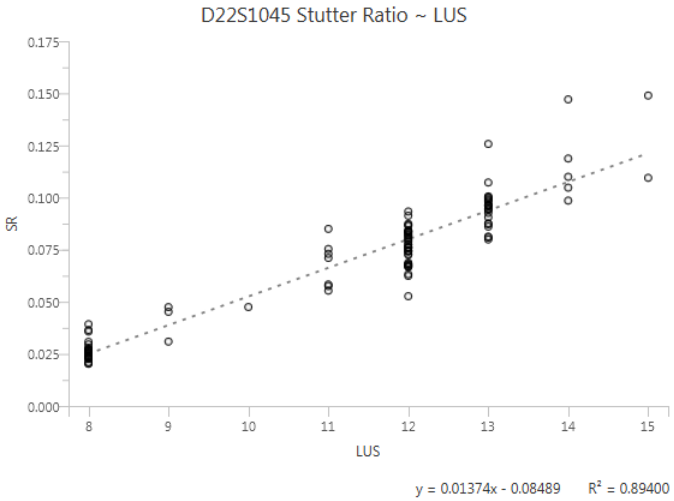

TH01 Stutter Ratio ~ Allele
SR
Allele
y = 0.00054x + 0.01527
R² = 0.00931

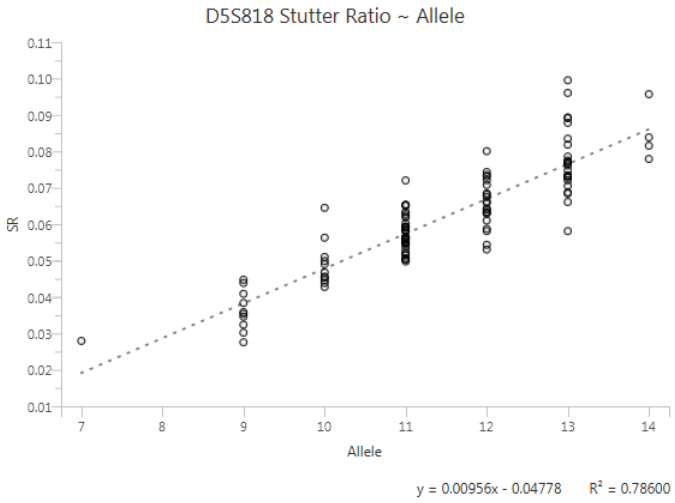

TH01 Stutter Ratio ~ LUS
SR
LUS
y = 0.00538x - 0.01778
R² = 0.74400

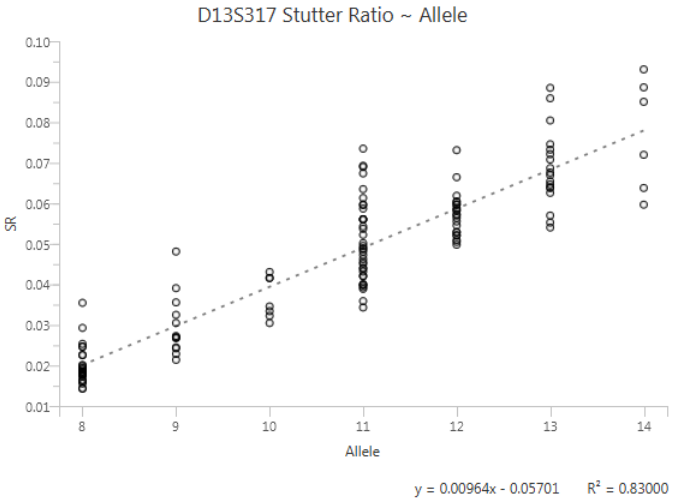

FGA Stutter Ratio ~ Allele
SR
Allele
y = 0.00675x - 0.07689
R² = 0.65000

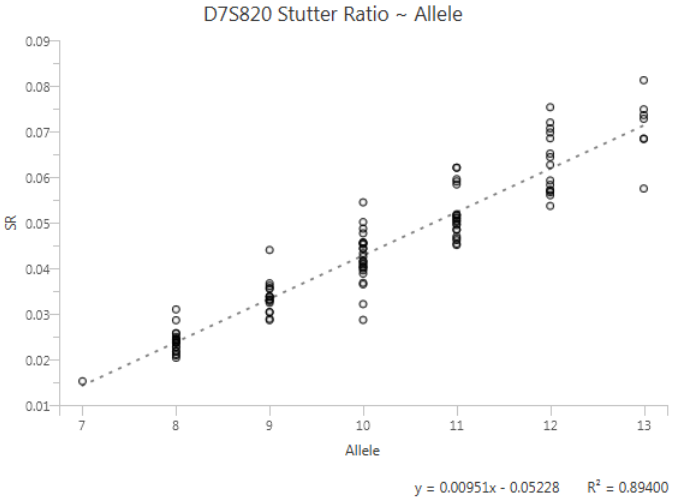

FGA Stutter Ratio ~ LUS
SR
LUS
y = 0.00801x - 0.03887
R² = 0.63000

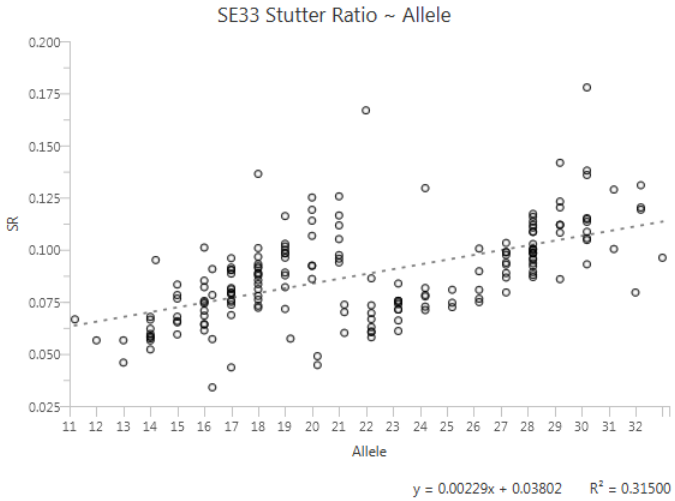

D22S1045 Stutter Ratio ~ Allele
SR
Allele
y = 0.01374x - 0.12612
R² = 0.89400

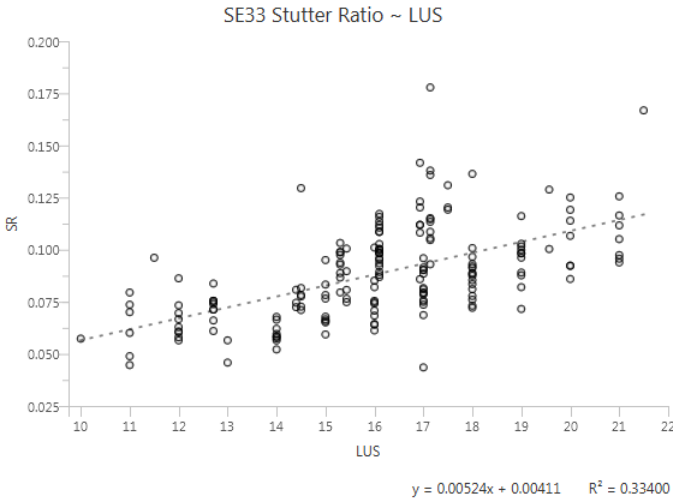

D22S1045 Stutter Ratio ~ LUS
SR
LUS
y = 0.01374x - 0.08489
R² = 0.89400

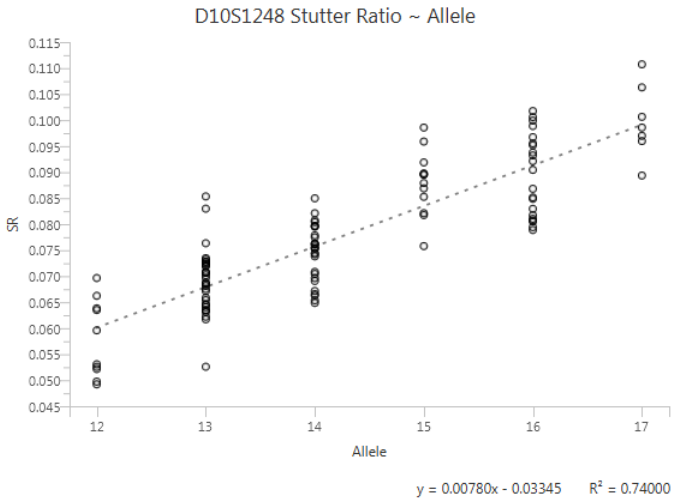

D5S818 Stutter Ratio ~ Allele
SR
Allele
y = 0.00956x - 0.04778
R² = 0.78600

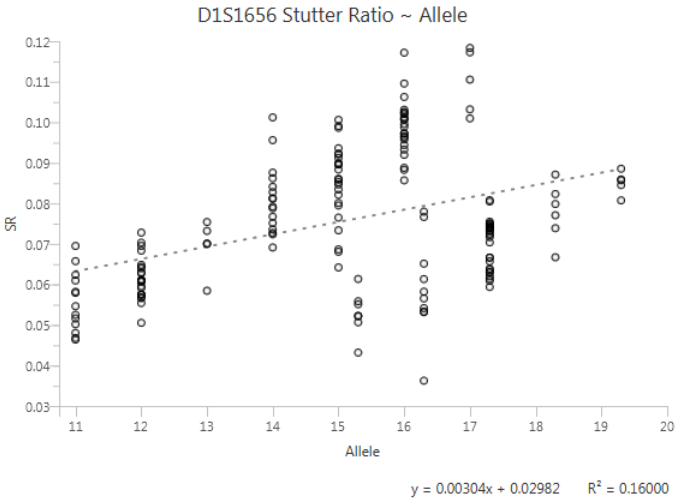

D13S317 Stutter Ratio ~ Allele
SR
Allele
y = 0.00964x - 0.05701
R² = 0.83000

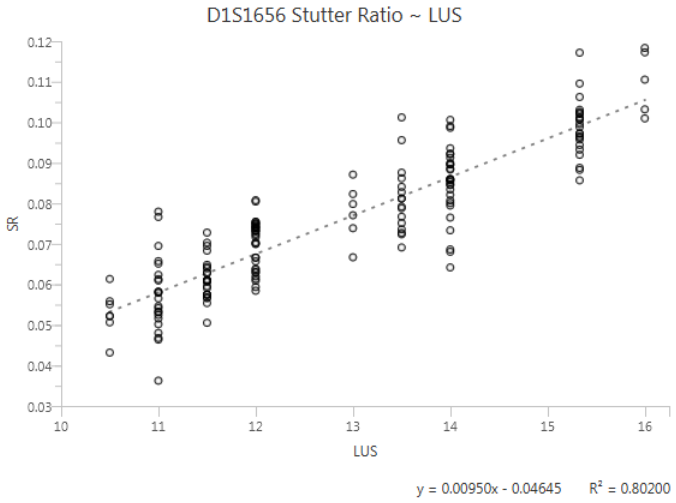

D7S820 Stutter Ratio ~ Allele
SR
Allele
y = 0.00951x - 0.05228
R² = 0.89400

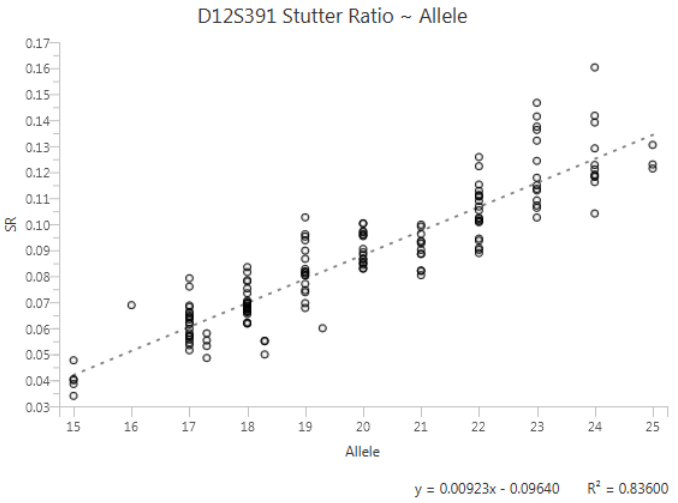

SE33 Stutter Ratio ~ Allele
SR
Allele
y = 0.00229x + 0.03802
R² = 0.31500

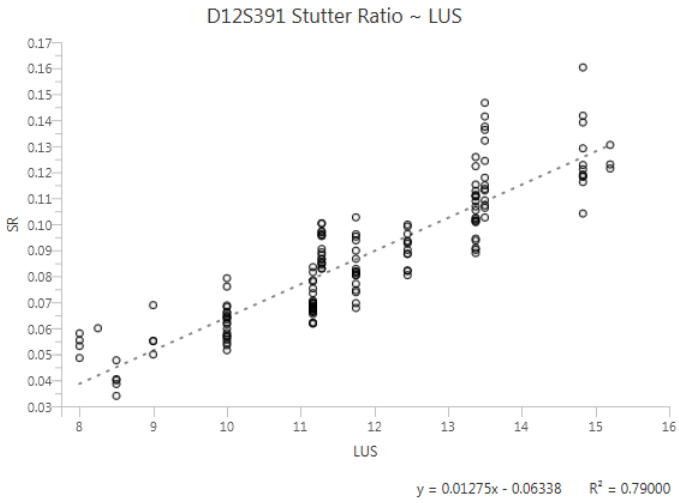

SE33 Stutter Ratio ~ LUS
SR
LUS
y = 0.00524x + 0.00411
R² = 0.33400

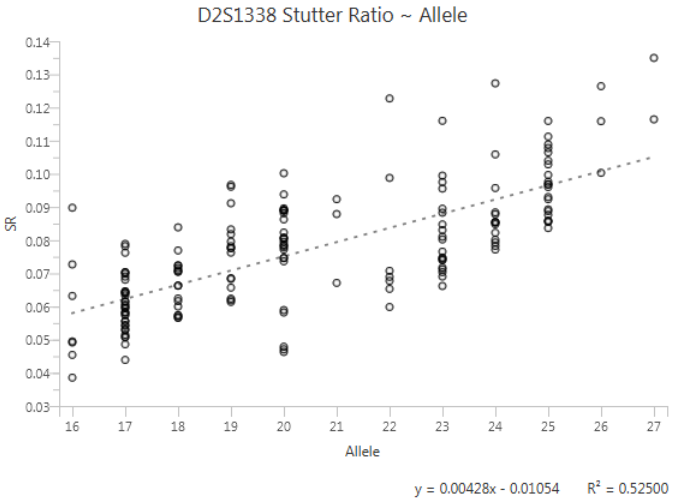

D10S1248 Stutter Ratio ~ Allele
SR
Allele
y = 0.00780x - 0.03345
R² = 0.74000

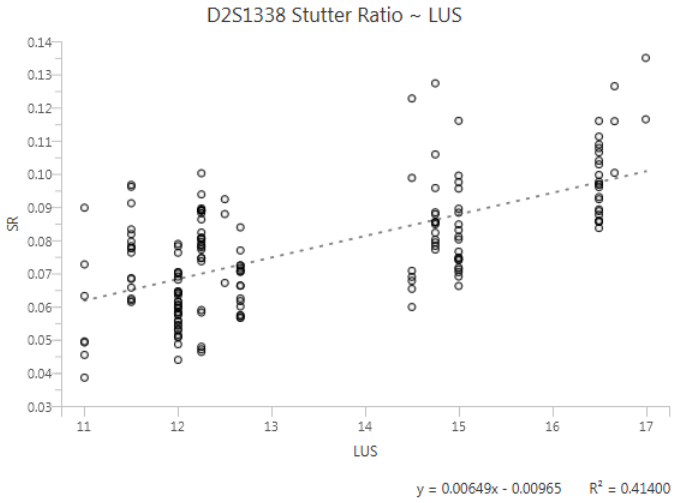

D1S1656 Stutter Ratio ~ Allele
SR
Allele
y = 0.00304x + 0.02982
R² = 0.16000


D1S1656 Stutter Ratio ~ LUS
SR
LUS
y = 0.00950x - 0.04645
R² = 0.80200


D12S391 Stutter Ratio ~ Allele
SR
Allele
y = 0.00923x - 0.09640
R² = 0.83600


D12S391 Stutter Ratio ~ LUS
SR
LUS
y = 0.01275x - 0.06338
R² = 0.79000


D2S1338 Stutter Ratio ~ Allele
SR
Allele
y = 0.00428x - 0.01054
R² = 0.52500


D2S1338 Stutter Ratio ~ LUS
SR
LUS
y = 0.00649x - 0.00965
R² = 0.41400

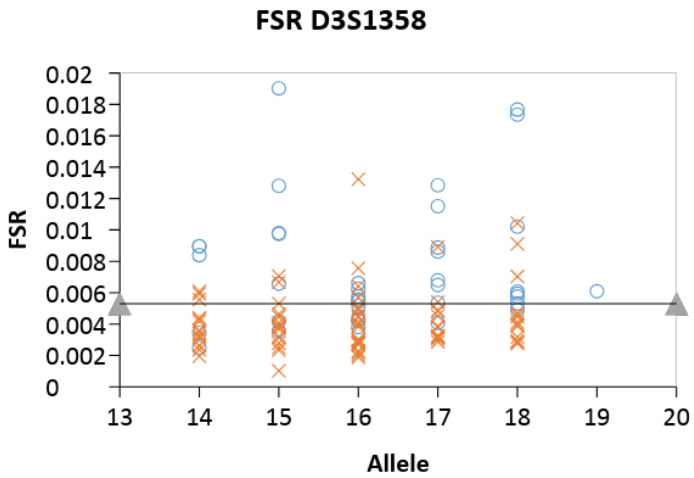
Appendix 2: Stutter Exceptions file for GlobalFiler™ analysis. This .csv file is titled "CBI_GlobalFiler_Exceptions.csv" for use within the CBI laboratories

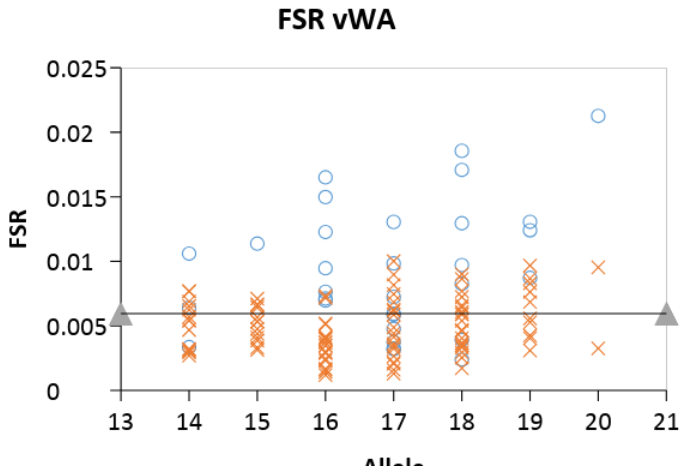
Allele	D3S1358	vWA	D16S539	CSF1PO	TPOX	Yindel	D8S1179	D21S11	D18S51	DYS391	D2S441	D19S433	TH01	FGA	D22S1045	D5S818	D13S317	D7S820	SE33	D10S1248	D1S1656	D12S391	D2S1338
2.2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
4	0	0	0	0	0	0	0	0	0	0	0	0	0.00374	0	0	0	0	0	0	0	0	0	0
5	0	0	0	0	0	0	0	0	0	0	0	0	0.00912	0	0	0	0	0	0	0	0	0	0
5.3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
6	0	0	0	0	0	0	0	0	0	0	0	0	0.0145	0	0	0	0	0	0	0	0	0	0
6.2	0	0	0	0	0	0	0	0	0	0	0	0.00794	0	0	0	0	0	0	0	0	0	0	0
6.3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
7	0	0	0	0	0	0	0	0	0	0	0	0	0.01988	0	0	0	0	0	0	0	0	0	0
7.3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
8	0	0	0	0	0	0	0.032901	0	0	0	0	0.01617	0.02526	0	0	0	0	0	0	0	0.02955	0	0
8.3	0	0	0	0	0	0	0	0	0	0	0	0	0.00912	0	0	0	0	0	0	0	0	0	0
9	0	0	0	0	0	0	0	0	0	0	0	0.01617	0.03064	0	0	0	0	0	0	0	0.03905	0	0
9.3	0	0	0	0	0	0	0	0	0	0	0	0	0.0145	0	0	0	0	0	0	0	0	0	0
10	0	0	0	0	0	0	0.049899	0	0	0	0.03622	0.03263	0.03602	0	0	0	0	0	0	0	0.0438	0	0
10.1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
10.3	0	0	0	0	0	0	0	0	0	0	0	0	0.0145	0	0	0	0	0	0	0	0	0	0
11	0	0	0	0	0	0	0.066992	0	0	0	0.058411	0.04086	0.0414	0	0	0	0	0	0	0	0.05805	0	0
11.1	0	0	0	0	0	0	0	0	0	0	0	0.03263	0	0	0	0	0	0	0	0	0	0	0
11.2	0	0	0	0	0	0	0	0	0	0	0	0.04909	0	0	0	0	0	0	0	0	0	0	0
11.3	0	0	0	0	0	0	0	0	0	0	0.023325	0	0	0	0	0	0	0	0	0	0	0	0
12	0	0	0	0	0	0	0.06666	0	0	0	0.068879	0.04909	0.04678	0	0	0	0	0	0	0.057945	0.0628	0	0
12.1	0	0	0	0	0	0	0	0	0	0	0	0.00794	0	0	0	0	0	0	0	0	0	0	0
12.2	0	0	0	0	0	0	0	0	0	0	0	0.05732	0	0	0	0	0	0	0	0	0	0	0
12.3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
13	0	0	0	0	0	0	0.059279	0	0	0	0	0.05732	0	0	0	0	0	0	0	0.068997	0.06755	0	0
13.2	0	0	0	0	0	0	0	0	0	0	0	0.06555	0	0	0	0	0	0	0	0	0	0	0
13.3	0	0	0	0	0	0	0	0	0	0	0	0	0.02526	0	0	0	0	0	0	0	0.05805	0	0
14	0	0.041158	0	0	0	0	0.06724	0	0	0	0.04413	0.06555	0	0	0	0	0	0	0.059817	0.074437	0.0818	0	0
14.2	0	0	0	0	0	0	0	0	0	0	0	0.07378	0	0	0	0	0	0	0	0	0	0	0
14.3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0.03905	0	0
15	0	0.059749	0	0	0	0	0.066056	0	0	0	0	0.07378	0	0	0	0	0	0	0.070864	0.087901	0.08655	0.040115	0
15.1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
15.2	0	0	0	0	0	0	0	0	0	0	0	0.08201	0	0	0	0	0	0	0	0	0	0	0
15.3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0.0533	0	0
16	0	0.064391	0	0	0	0	0	0	0	0	0.067455	0.08201	0	0	0	0	0	0	0.074648	0.089523	0.099185	0	0.058326
16.1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
16.2	0	0	0	0	0	0	0	0	0	0	0	0.09024	0	0	0	0	0	0	0	0	0	0	0
16.3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0.05805	0	0
17	0	0.06964	0	0	0	0	0	0	0	0	0	0.09024	0	0	0	0	0	0	0.079482	0.099789	0.10555	0.062288	0.060806
17.1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
17.2	0	0	0	0	0	0	0	0	0	0	0	0.094355	0	0	0	0	0	0	0	0	0	0	0
17.3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0.06755	0	0
18	0	0.081955	0	0	0	0	0	0	0	0	0	0.09847	0	0	0	0	0	0	0.08897	0	0.11505	0.069879	0.066684
18.1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
18.2	0	0	0	0	0	0	0	0	0	0	0	0.1067	0	0	0	0	0	0	0	0	0	0	0
18.3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0.07705	0	0
18.4	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
19	0	0.090227	0	0	0	0	0	0	0	0	0	0	0	0.048676	0	0	0	0	0.09484	0	0	0.083246	0.076653
19.1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
19.2	0	0	0	0	0	0	0	0	0	0	0	0.11493	0	0	0	0	0	0	0	0	0	0	0
19.3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0.08655	0	0
19.4	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0

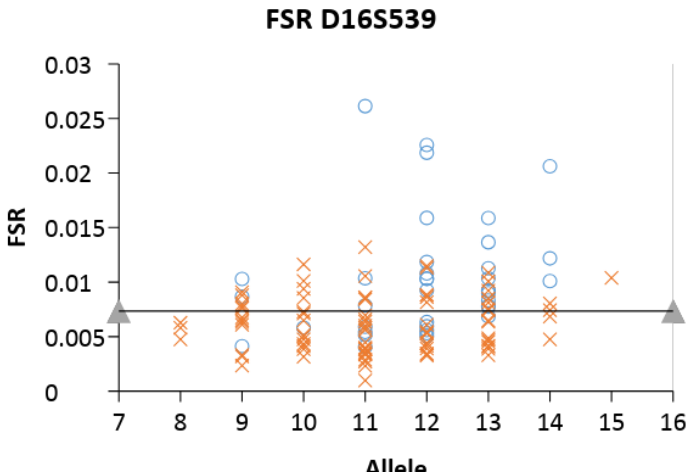
20	0	0	0	0	0	0	0	0	0	0	0	0	0	0.057287	0	0	0	0	0.105028	0	0	0.090797	0.077335
20.2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
20.3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
21	0	0	0	0	0	0	0	0	0	0	0	0	0	0.063707	0	0	0	0	0.106529	0	0	0.090513	0
21.1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
21.2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
21.3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
22	0	0	0	0	0	0	0	0	0	0	0	0	0	0.07148	0	0	0	0	0	0	0	0.104776	0.07916
22.2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0.067205	0	0	0	0
22.3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
23	0	0	0	0	0	0	0	0	0	0	0	0	0	0.086579	0	0	0	0	0	0	0	0.121623	0.081895
23.2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0.072534	0	0	0	0
24	0	0	0	0	0	0	0	0	0	0	0	0	0	0.083825	0	0	0	0	0	0	0	0.126374	0.088802
24.1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
24.2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0.085101	0	0	0	0
25	0	0	0	0	0	0	0	0	0	0	0	0	0	0.092969	0	0	0	0	0	0	0	0	0.096942
25.2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
26	0	0	0	0	0	0	0	0	0	0	0	0	0	0.094546	0	0	0	0	0	0	0	0	0
26.2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0.08443	0	0	0	0
27	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
27.2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0.093306	0	0	0	0
28	0	0	0	0	0	0	0	0.0604	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
28.2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0.101176	0	0	0	0
28.3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
29	0	0	0	0	0	0	0	0.06829	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
29.2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0.114687	0	0	0	0
29.3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
30	0	0	0	0	0	0	0	0.077741	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
30.2	0	0	0	0	0	0	0	0.058744	0	0	0	0	0	0	0	0	0	0	0.120628	0	0	0	0
30.3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
31	0	0	0	0	0	0	0	0.094869	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
31.2	0	0	0	0	0	0	0	0.066963	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
32	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
32.2	0	0	0	0	0	0	0	0.081859	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
33	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
33.1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
33.2	0	0	0	0	0	0	0	0.089922	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
34	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
34.1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
34.2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
35	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
35.1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
35.2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
36	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
36.2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
37	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
37.2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
38	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
38.2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
39	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
39.2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
40	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
40.2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
41	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
41.2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
42	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
42.2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
43	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
43.2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0

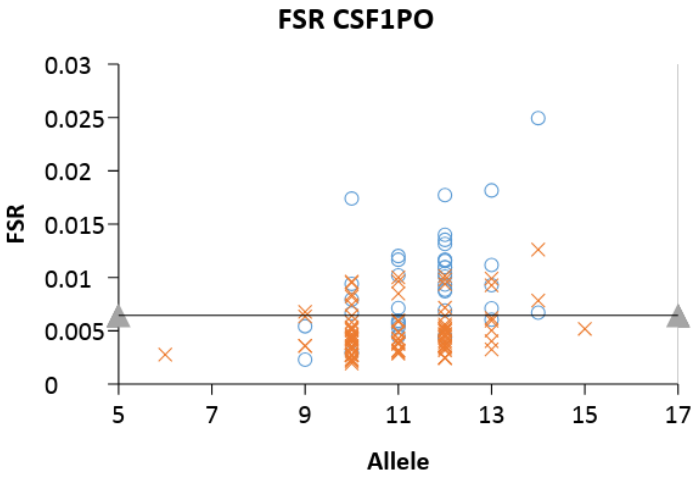
44	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
44.2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
45	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
45.2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
45.3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
46	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
46.2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
47	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
47.2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
48	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
48.2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
49	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
49.2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
50	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
50.2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0

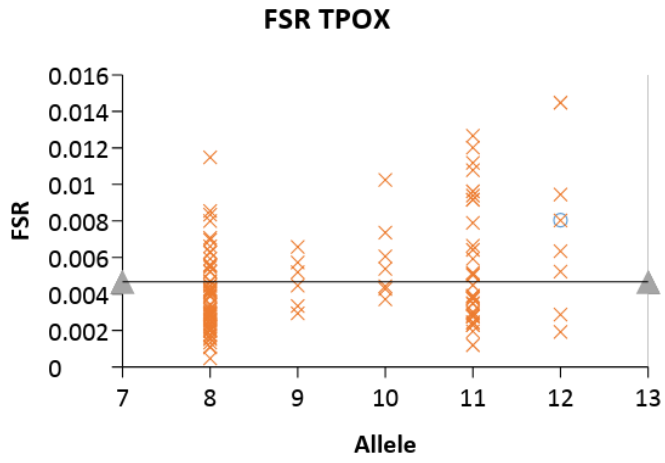
Appendix 3: Forward Stutter Ratio (FSR) versus Allele. The circles represent observed forward stutter and the crosses represent the FSR obtained with data inserted at half the AT used during analysis for each channel. The black solid line is the average (y intercept) of FSR for the locus

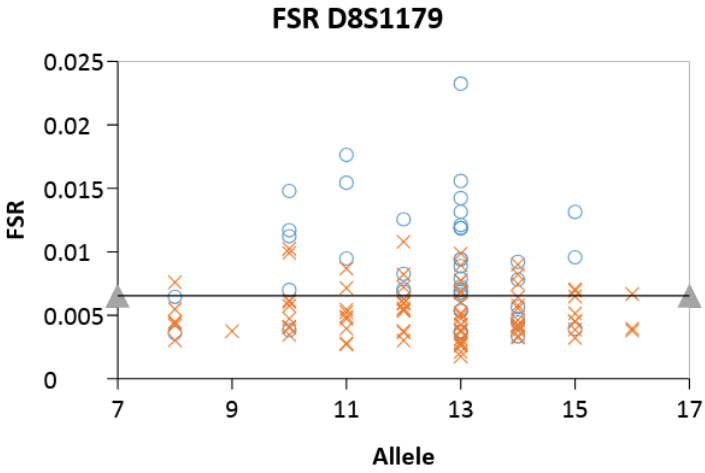


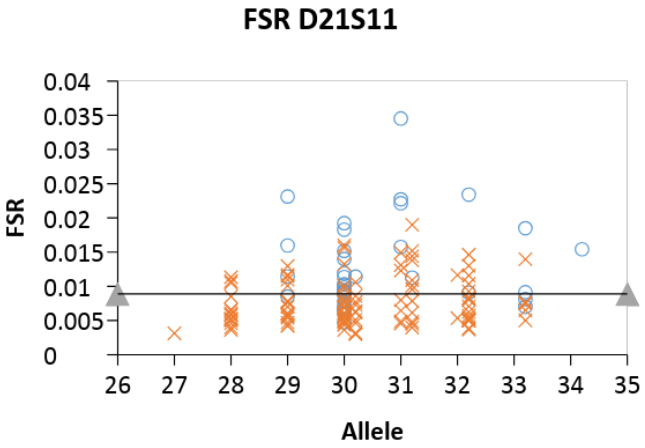


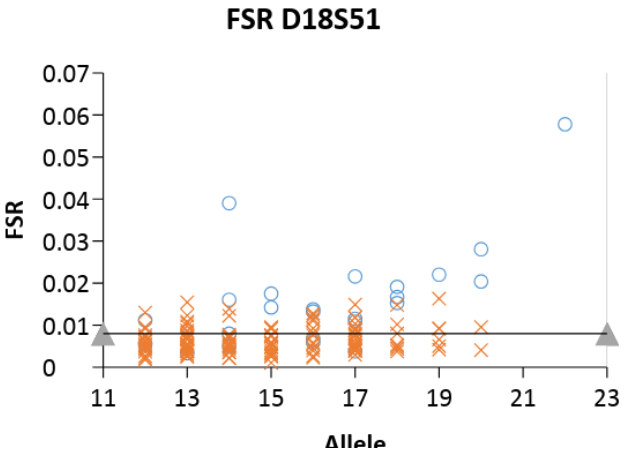


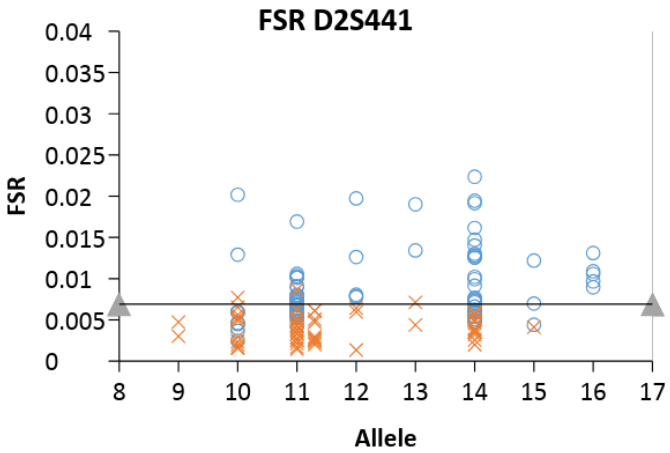


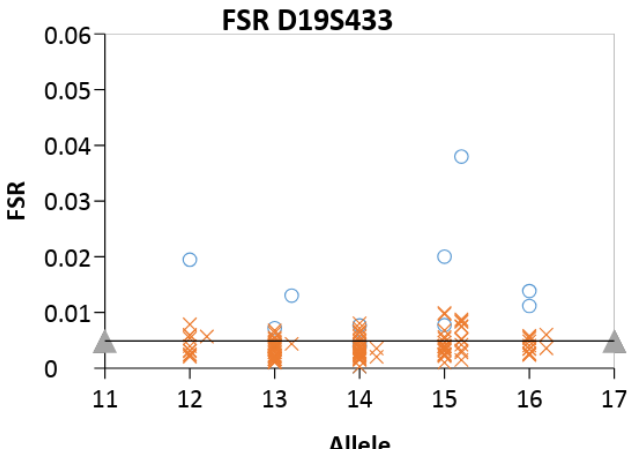

FSR TPOX
FSR
0.016
0.014
0.012
0.01
0.008
0.006
0.004
0.002
0
7
8
9
10
11
12
13
Allele

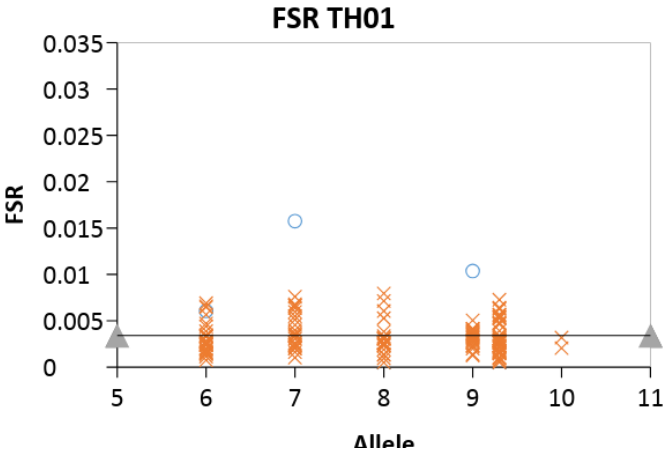

FSR D8S1179
FSR
0.025
0.02
0.015
0.01
0.005
0
7
9
11
13
15
17
Allele

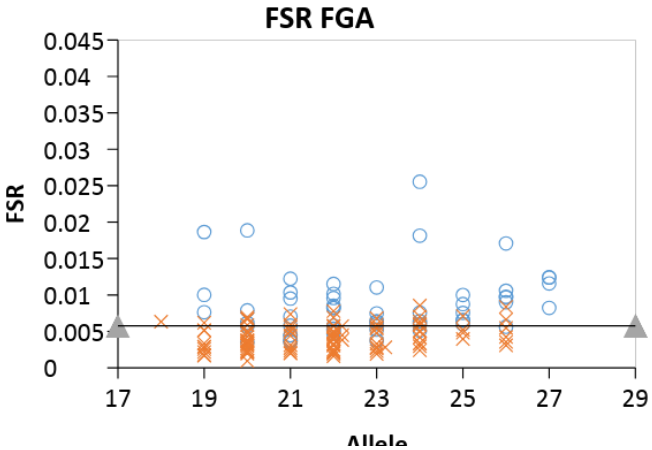

FSR D21S11
FSR
0.04
0.035
0.03
0.025
0.02
0.015
0.01
0.005
0
26
27
28
29
30
31
32
33
34
35
Allele

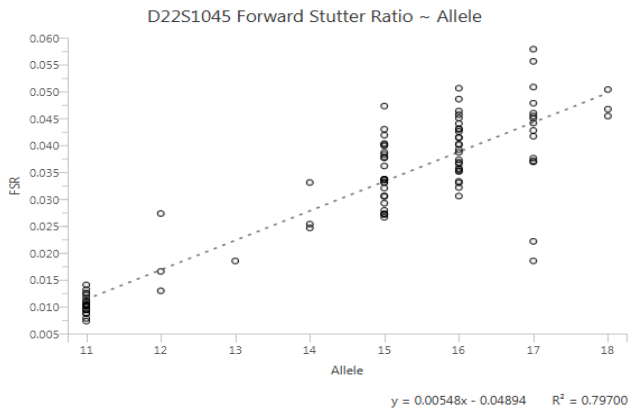

FSR D18S51
FSR
0.07
0.06
0.05
0.04
0.03
0.02
0.01
0
11
13
15
17
19
21
23
Allele

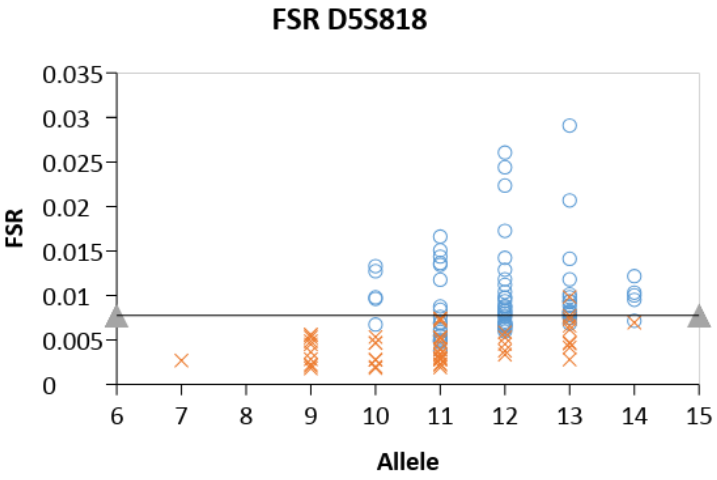

FSR D2S441
FSR
0.04
0.035
0.03
0.025
0.02
0.015
0.01
0.005
0
8
9
10
11
12
13
14
15
16
17
Allele

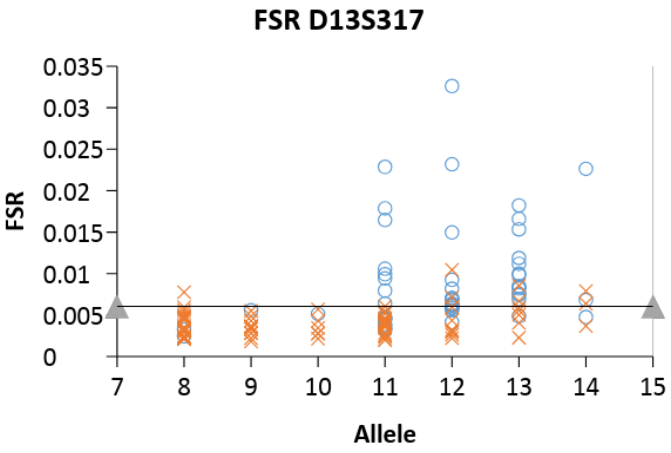

FSR D19S433
FSR
0.06
0.05
0.04
0.03
0.02
0.01
0
11
12
13
14
15
16
17
Allele

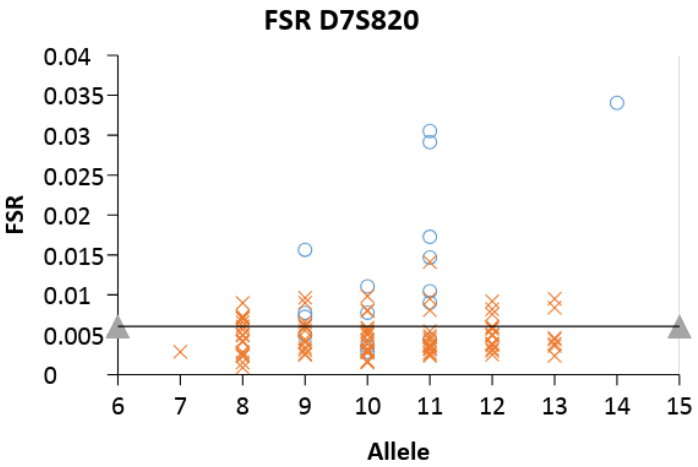

FSR TH01
FSR
0.035
0.03
0.025
0.02
0.015
0.01
0.005
0
5
6
7
8
9
10
11
Allele

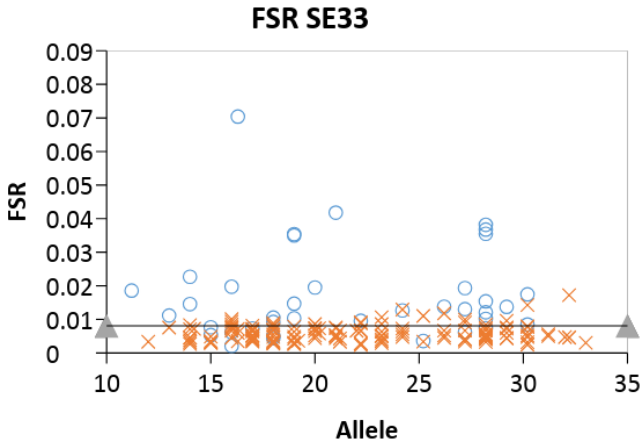

FSR FGA
FSR
0.045
0.04
0.035
0.03
0.025
0.02
0.015
0.01
0.005
0
17
19
21
23
25
27
29
Allele

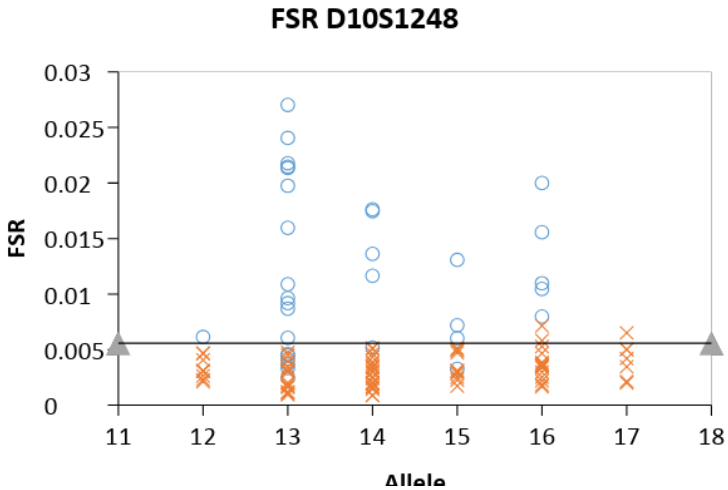

D22S1045 Forward Stutter Ratio ~ Allele
FSR
Allele
y = 0.00548x - 0.04894
R² = 0.79700

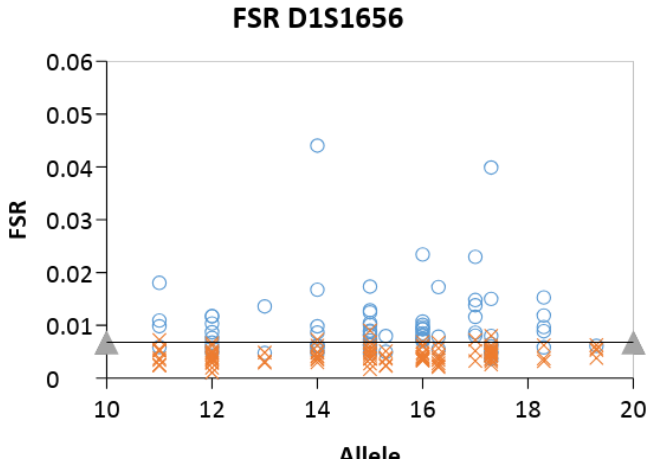

FSR D5S818
FSR
0.035
0.03
0.025
0.02
0.015
0.01
0.005
0
6
7
8
9
10
11
12
13
14
15
Allele

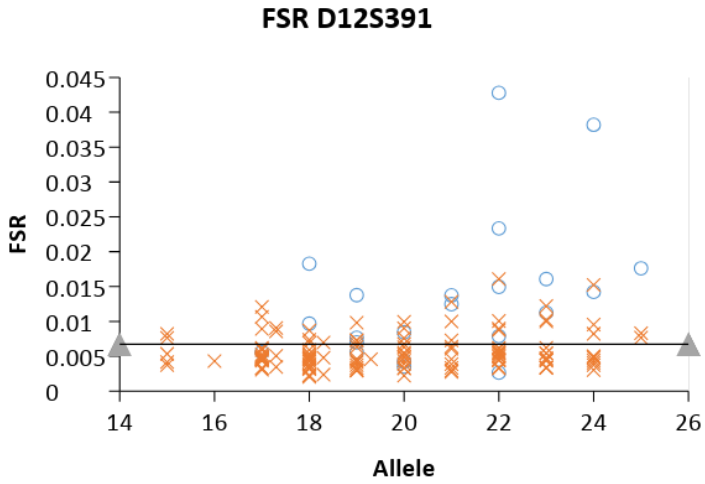

FSR D13S317
FSR
0.035
0.03
0.025
0.02
0.015
0.01
0.005
0
7
8
9
10
11
12
13
14
15
Allele

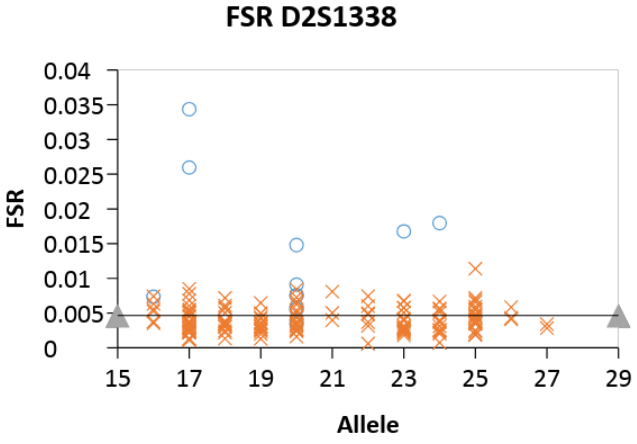

FSR D7S820
FSR
0.04
0.035
0.03
0.025
0.02
0.015
0.01
0.005
0
6
7
8
9
10
11
12
13
14
15
Allele


FSR SE33
FSR
0.09
0.08
0.07
0.06
0.05
0.04
0.03
0.02
0.01
0
10
15
20
25
30
35
Allele


FSR D10S1248
FSR
0.03
0.025
0.02
0.015
0.01
0.005
0
11
12
13
14
15
16
17
18
Allele


FSR D1S1656
FSR
0.06
0.05
0.04
0.03
0.02
0.01
0
10
12
14
16
18
20
Allele


FSR D12S391
FSR
0.045
0.04
0.035
0.03
0.025
0.02
0.015
0.01
0.005
0
14
16
18
20
22
24
26
Allele


FSR D2S1338
FSR
0.04
0.035
0.03
0.025
0.02
0.015
0.01
0.005
0
15
17
19
21
23
25
27
29
Allele

Internal Validation of STRmix™ V2.5 for the Colorado Bureau of Investigation (CBI) Forensic Laboratories (GlobalFiler™, 3500xL CE)

STRmix™ internal validation

This document describes the internal validation of STRmix™ V2.5 at the Colorado Bureau of Investigation Laboratory system (hereafter CBI). This validation was done in conjunction with ESR in New Zealand. This covers the use of STRmix™ across all 5 CBI sites. STRmix™ has previously been subjected to developmental validation. This involved, in part, the complete '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 Bright et al. [1]. A list of all papers describing the theory behind different aspects of STRmix™ is provided in Appendix 1 of this document.

Internal validation describes the activities CBI has undertaken in-house before the implementation of STRmix™ into routine casework. This document follows the internal validation section of the SWGDAM Guidelines for the Validation of Probabilistic Genotyping Systems [2]. This included the examination of known and non-probative evidence samples, investigations into reproducibility and precision, sensitivity and stochastic studies, and mixture studies. The section where specific SWGDAM guidelines are discussed in this document is cross referenced in Appendix 2.

The results of all experiments related to the internal validation of STRmix™ at CBI are retained on the P: drive. The summaries are retained in Qualtrax.

STRmix™ parameters

The parameters described in the document 'CBI STRmix parameters GF_3500_V2.5 FINAL' were used for all internal validation checks presented in this report. All other run parameters have been optimised by the STRmix™ developers.

Section A: Single source profiles

Inspection of weights

This section covers the following standards:

> 4.1.5. Single-source specimens
>
> 4.2.1.2. For single-source specimens with high quality results, genotypes derived from non-probabilistic analyses of profiles above the stochastic threshold should be in complete concordance with the results of probabilistic methods.

Within this section we demonstrate how the weights assigned by STRmix™ to different genotype combinations are appropriate. The weights can be used as a diagnostic of the deconvolution process and should be intuitively correct, where the most supported genotypes have the highest weights.

The addition of information to an analysis can aid in the ability to deconvolute the sample. For example, using replicates and or other profiling kits can reduce ambiguity and increase weightings of individual genotype sets.

A dilution series of a single source profile where the peak heights ranged from above the level where dropout is observed to below was constructed. The sample used was IMR-90 (female). Profiles were amplified using the GlobalFiler™ multiplex following the CBI Laboratory's standard operating procedure for amplification of questioned/crime samples. The template DNA in picograms for the serial dilution was: 1000, 500, 250, 125, 63, 31, and 16 pg. The profiles were analysed following the CBI Laboratory's standard operating procedure for the analysis of crime samples.

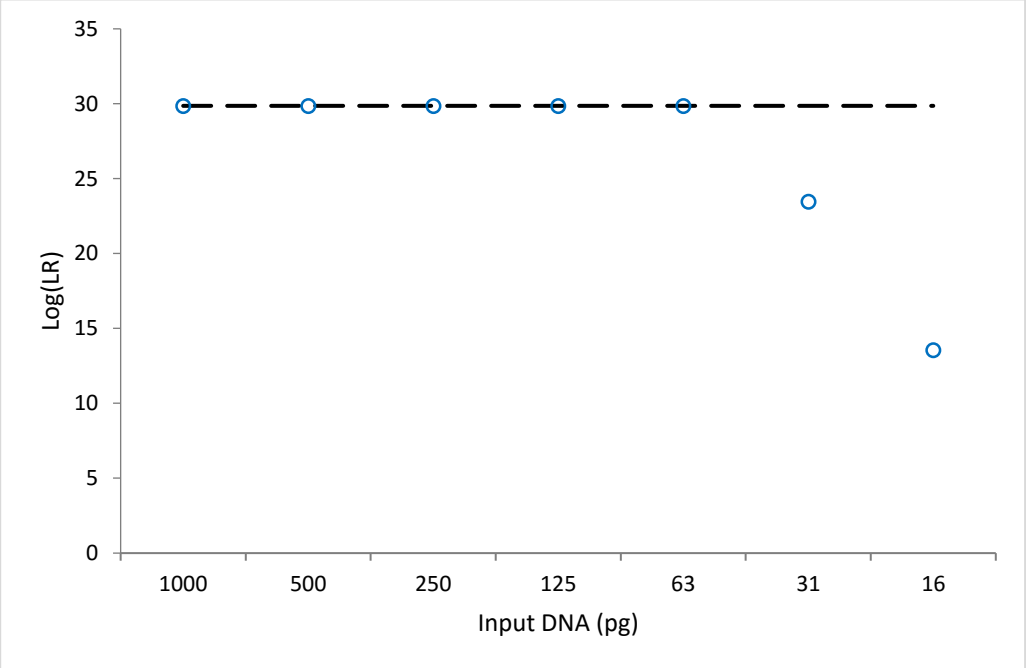
The profiles were interpreted in STRmix™. The propositions were:

H_p: The DNA originated from the person of interest (IMR-90)

H_d: The DNA originated from an unknown individual

The Likelihood Ratio (*LR)* was calculated for the known contributor using the NIST1036 Caucasian July 2017 allele frequencies and an F_{ST} (θ) of 1%. In relation to the STRmix output the point estimate value (*LR* Total) has been selected for direct comparison. A plot of log(*LR*) versus input DNA for GlobalFiler™ is provided in Figure 1. The dashed line represents the log(*LR*) for the full profile, where no dropout is observed.

Figure 1: Plot of log(LR) versus input DNA amount (pg)



Inspection of the plot shows the *LR* progressing from the value for the single source *LR* calculated for a full profile towards *LR* = 1 as the DNA template decreases. As expected, the weights for genotypes considering dropout increased as template drops. In addition, the DNA amounts from the STRmix™ output (*t* or template mass parameter) declined steadily in line with peak heights.

The genotype probability for the 16 pg profile using GlobalFiler™ replicate 1, and GlobalFiler™ replicates 1 and 2, for D8S1179 are displayed in Table 1 below demonstrating that the addition of information to an analysis can reduce the ambiguity and increase the weighting of genotype combinations.

Table 1: Genotype combinations and weightings for D8S1179

	Genotype Combination	Weighting
16 pg rep 1	13,13	5.69E-01
	Q, 13	4.31E-01
	Q, Q	3.78E-04
16 pg rep 1&2	13,14	1

Reproduction of single source LR

There is a small subset of profiles where the 'answer' is known or can be estimated easily [3]. These include single source profiles where the weight is one (or 100%) for one genotype at each locus. The *LR* was calculated 'by hand' at each locus for one single source profile (one of the full IMR-90 samples described above) and the individual locus *LR*s compared with the STRmix™ results. This was undertaken twice; once using an F_{ST} (or θ) value of 0 and once with F_{ST}=0.01. Setting θ to zero returns the product rule where:

$2p_ip_j$ for heterozygote loci

p_i^2 for homozygote loci

Where p_i is the allele frequency for allele *i*, p_j the allele frequency for allele *j*. When θ > 0, the Balding and Nichols [4] formulae (or equations 4.10 from NRCII [5]) are applied. For single source profiles:

$$\frac{2\left[\theta+(1-\theta)\,p_i\right]\left[\theta+(1-\theta)\,p_j\right]}{(1+\theta)(1+2\theta)} \quad \text{for heterozygote loci} \qquad (1)$$

$$\frac{\left[3\theta+(1-\theta)\,p_i\right]\left[2\theta+(1-\theta)\,p_i\right]}{(1+\theta)(1+2\theta)} \quad \text{for homozygote loci} \qquad (2)$$

Where p_i is the allele frequency for allele *i*, p_j the allele frequency for allele *j* and θ is the F_{ST} value. The allele frequencies used within equations 1 and 2 are posterior mean frequencies. These are calculated using the following equation:

$$\frac{x_i+\frac{1}{k}}{N_a+1} \qquad (3)$$

Where for the given locus, x_i is the number of observations of allele *i* in a database, N_a is the number of alleles in that database and *k* is the number of allele designations with non-zero observations in the database at that locus.

The 'by hand' calculated and STRmix™ results for a single source profile for θ=0 and θ=0.01 are given in Table 2.

Table 2: 'By hand' (Excel) calculation of LR versus STRmix™ results for a single source profile with varying F_{ST} values

Locus	Excel θ=0	STRmix™ θ=0	Excel θ=0.01	STRmix™ θ=0.01
D3S1358	17.19565	17.19565	15.92105	15.92105
vWA	23.98497	23.98497	21.87565	21.87565
D16S539	53.77058	53.77058	45.20479	45.20479
CSF1PO	19.80438	19.80438	17.94158	17.94158
TPOX	7.48319	7.48319	7.15023	7.15023
D8S1179	9.13996	9.13996	8.78695	8.78695
D21S11	238.34808	238.34808	163.14452	163.1445
D18S51	52.20327	52.20327	39.28204	39.28204
D2S441	8.49288	8.49288	7.74656	7.74656
D19S433	1331.19822	1331.19822	171.52003	171.52
TH01	15.17729	15.17729	14.01779	14.01779
FGA	240.12901	240.12901	161.80823	161.8082
D22S1045	103.73086	103.73086	59.00237	59.00237
D5S818	9.04921	9.04921	8.65695	8.65695
D13S317	13.21365	13.21365	12.39469	12.39469
D7S820	18.74759	18.74759	17.47645	17.47645
SE33	167.31577	167.31577	124.04824	124.0482
D10S1248	5.47033	5.47033	5.38413	5.38413
D1S1656	43.13702	43.13702	37.57720	37.5772
D12S391	22.63765	22.63765	20.83683	20.83683
D2S1338	40.51269	40.51269	35.76324	35.76324
Total	**1.29673E+32**	**1.29673E+32**	**7.093E+29**	**7.093E+29**

The results in Table 2 show that STRmix™ is giving the expected answer based on the population genetic model being used.

Section B: Use of peak heights/Off-scale peaks, saturation

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

Two single source profiles were amplified with above-optimal DNA input (4ng) and analyzed in STRmix™ in order to review the impact of saturated data on profile interpretation. Each profile had peaks present above the STRmix™ saturation threshold (30,000 RFU) set by CBI. Likelihood Ratios (*LR*s) were calculated using the NIST Caucasian July 2017 allele frequencies and F_{ST} = 0.01. Both interpretations resulted in intuitive genotypes where the weight = 1.0 for the known contributor genotype at each locus. A summary of log(*LR*) and STRmix™ outputs of average proposed template (RFU) and stutter variance (k^2) for each profile is given in Table 3. The LR selected is the point estimate/LR total on the STRmix output. As expected, the observed stutter variance is significantly above the mode of the prior gamma distribution (9.30) for high template profiles. This is because as alleles are more likely to be above the camera saturation limit (30,000 RFU) and their corresponding stutter peaks appear larger than expected. In these cases, within STRmix™ the expected height of the stutter peak is

calculated from the expected height of the allele and not the observed height which leads to slightly higher than expected variance between the observed and expected stutter peaks.

Table 3: Log(LR) for two profiles amplified after the addition of above-optimal DNA

Sample	Log(LR)	Template (RFU)	k^2
4_F-A1_EV.csv	29.85	29613	47.25
4_F-A3_EV.csv	29.85	17880	61.77

The impact of interpreting samples containing some peaks at/above saturation could be more significant if you were dealing with mixed DNA profiles, particularly if there are minor(s) present of similar heights as potential stutters to the saturated peak. More weight could be assigned to a peak in such a stutter position as being allelic.

Overall, it is not recommended that saturated profiles are interpreted within STRmix™ as a profile that exceeds the CE saturation threshold is unlikely to have the true peak heights recorded. Thus the models used within STRmix™ are no longer optimal and we will likely observe higher stutter peak heights than expected, resulting in an elevated k^2 value.

At the CBI laboratories caution will be taken if any profile contains peak(s) at/above the saturation setting in STRmix (30,000 RFU). A thorough review of the weights across the profile will be undertaken for their intuitiveness.

Section C: Weights

This section covers the following standard:

> 4.2.1.3. Generally, as the analyst's ability to deconvolute a complex mixture decreases, so do the weightings of individual genotypes within a set determined by the software.

The weights are described as the primary output from STRmix™. They can be used as a diagnostic of the deconvolution process and should be intuitively correct, where the most supported genotypes have the highest weights.

A two person mixture series (7057 & IMR-90) was constructed in the following ratios 10:1, 4:1, 2:1 and 1:1. The total amount of DNA in the profiles was approximately 1ng.

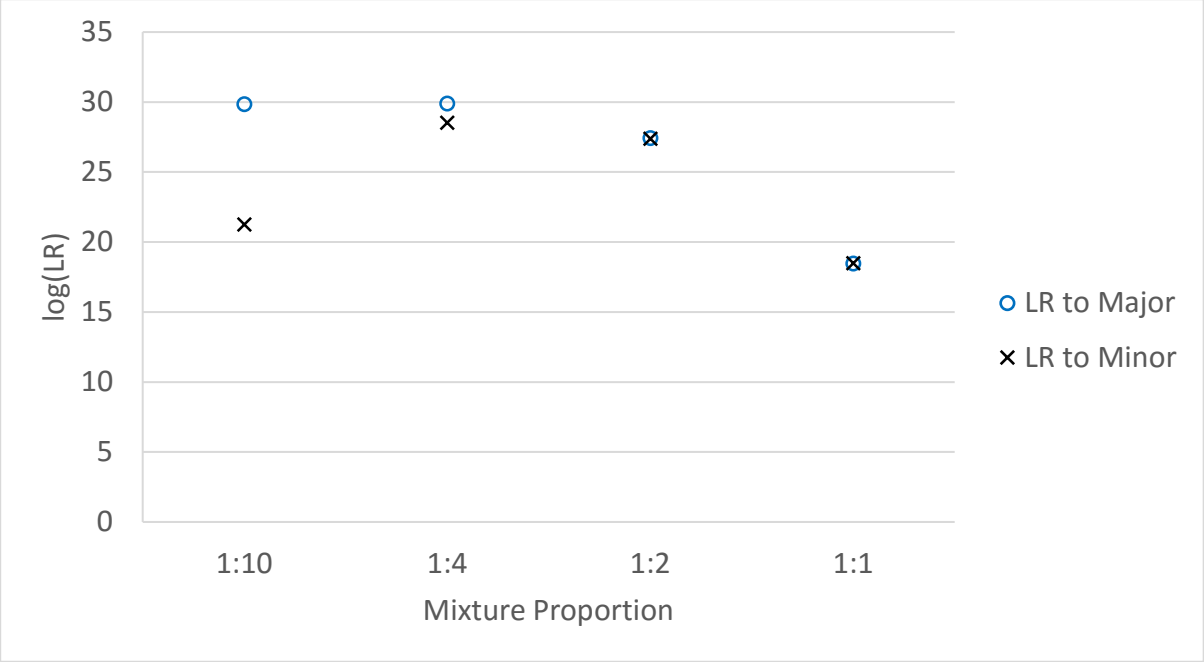
The profiles were interpreted in STRmix™ under the following propositions and a *LR* calculated:

> H_p: The DNA originated from the person of interest (known major or minor) and an unknown individual
> H_d: The DNA originated from two unknown individuals

*LR*s were calculated using NIST Caucasian July 2017 allele frequencies and F_{ST} = 0.01. The point estimate/LR Total from the STRmix output was selected for comparison.

A plot of log(*LR*) for each mixture type considering both the major (blue circle data points) and minor (black cross data points) is provided in Figure 2.

Figure 2: Log(*LR*) for each mixture type considering both the major (blue circles) and minor (black crosses)



Inspection of Figure 2 shows that the *LR* decreases by approximately one third (~11 orders of magnitude) for the 1:1 mixture when compared to the resolved major contributor (see 1:10 & 1:4). The decrease starts where the resolution of major and minor become less clear (approximately 1:2). The *LR* for the minor contributor reduces as the amount of DNA template from them also reduces. This is most evident for the 1:10 mixture. In addition, the mixture proportions in the STRmix™ output changed appropriately as the mixture ratios varied.

Section D: Sensitivity and specificity and mixtures

This section covers the following standards:

4.1.2. Hypothesis testing with contributors and non-contributors

4.1.6. Mixed specimens

4.1.6.1. Various contributor ratios (e.g., 1:1 through 1:20, 2:2:1, 4:2:1, 3:1:1, etc)

4.1.6.2. Various total DNA template quantities

4.1.6.3. Various numbers of contributors. The number of contributors evaluated should be based on the laboratory's intended use of the software. A range of contributor numbers should be evaluated in order to define the limitations of the software.

4.1.6.5. Sharing of alleles among contributors

4.1.7. Partial profiles, to include the following:

4.1.7.1. Allele and locus drop-out

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

A demonstration of sensitivity and specificity for a range of CBI GlobalFiler™ mixtures was undertaken as per Taylor [6], however making use of apparent number of contributors, rather than experimental design and average peak heights (in RFU), rather than template. With respect to interpretation methods, 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 templates. The log(*LR*) for known contributors (H_p true) should be high and should trend to 0 as less information is present within the profile. Inversely the log(*LR*) should trend away from 0 as we add more relevant information in to an analysis. Information includes amount of DNA from the contributor of interest, conditioning profiles (for example the victim's profile on intimate samples), and replicates. Typically dealing with mixtures of fewer contributors also leads to greater sensitivity. Specificity is defined as ability of the software to reliably exclude known non contributors (H_d true) within a mixed DNA profile for a range of starting DNA templates. The log(*LR*) should trend upwards to 0 as less information is present within the profile.

Specificity and sensitivity were tested by calculating the *LR* for a number of two-, three-, four-, and five person profiles for both known contributors and known non-contributors. The plots in [6] have been reproduced for CBI's GlobalFiler™ data. A total of 188 mixtures with varying number of contributors (two-, three-, four-, and five-person mixtures) and mixture proportions were generated in the CBI laboratory using GlobalFiler™. Within this dataset, the contributors to the mixtures are unrelated. A summary of the different types of mixtures is shown in Table 4 and the proposed contributors are shown in Table 5.

Table 4: Summary of experimental design for specificity and sensitivity tests. Values in brackets are the intended dilution series.

				Target DNA Amount (ng) in a dilution series					
Mixture	**Amp.**	**Donor Ratio**	**N**	**Set 1**	**Set 2**	**Set 3**	**Set 4**	**Set 5**	**Set 6**
M2-1	1	20:1	2	2.00	1.50	1.00	0.50	0.25	
	2								
M2-2	1	10:1							
	2								
M2-3	1	5:1							
	2								
M2-4	1	3:1							
	2								
M3-1	1	10:5:1	3						
	2								
M3-2	1	10:10:1							
	2								
M3-3	1	3:2:1							
	2								
M4-1	1	4:3:2:1	4						
	2								
M4-2	1	10:5:2:1							
	2								
M4-3	1	10:10:5:1							
	2								
M4-4	1	10:5:5:1							
	2								
M5-1	1	5:4:3:2:1	5						
	2								
M5-2	1	10:5:5:2:1							
	2								
M5-3	1	10:10:5:2:1							
	2								
M2-5	1	1:1	2	0.800	0.400	0.158 (0.2)	0.094 (0.1)	0.042 (0.5)	0.025
	2								
M3-4	1	1:1:1	3	1.200	0.561 (0.6)	0.300	0.150	0.0615 (0.075)	0.0195 (0.0375)
	2								
M4-5	1	1:1:1:1	4	1.600	0.800	0.400	0.189 (0.2)	0.060 (0.1)	0.036 (0.05)
	2								
M5-4	1	1:1:1:1:1	5	2.000	0.780 (1)	0.464 (0.5)	0.236 (0.25)	0.087 (0.125)	0.047 (0.0625)
	2								

Note: This constitutes 94 samples amplified in duplicate, hence 188 mixtures. The value for N described above is the experimental design number of contributors.

Table 5: Summary of contributors from experimental design

Mixture	Contributors					
Person (P)	Pre-fix	P1	P2	P3	P4	P5
2 Person (M2)	M2	R3	R1			
3 Person (M3)	M3	R3	R1	R4		
4 Person (M4)	M4	R3	R5	R1	R4	
5 Person (M5)	M5	R3	R2	R6	R1	R4

Some of these profiles represent some of the 'worst' types of profiles likely to be encountered by the laboratory. The profiles are of varying DNA quantity and mixture proportions. The contributors include homozygote and heterozygote alleles and there is varying amounts of allele sharing across the different loci (standard 4.1.6.5). Given the template amounts allele and/or locus dropout was expected to occur within the profiles containing the lower DNA amounts (standard 4.1.7.1).

Prior to interpretation in STRmix™ the profiles were assessed by an experienced analyst with regards to the 'apparent' number of contributors (N) visible in the electropherogram. This was without knowledge of the experimental design. The apparent N was assigned to each profile. The true number of contributors to a questioned/crime scene profile is *always* unknown and to ensure these results are applicable to casework, this approach was used. On the whole the apparent N was the same as experimental design with the exception of a small number (18) being under-assigned (usually due to the trace individual dropping out/simply not being visible to the analyst) and a small number (3) being over-assigned N (usually due to some stochastic imbalance event or high stutter suggesting a possible additional individual), in relation to experimental design.

A summary of the samples where apparent N and experimental design differed is provided in Table 6.

Table 6: Summary of samples where apparent N and experimental designed differed

Δ in apparent N to experimental design N	Sample File name	Comments
Under-assignment		
-3	M5_B02_STRmix_M5-4_0.0625ng_C_0.0625ng_P2_GF_CBI_20s.hid	5 person run as a 2 person
-3	M5_D04_STRmix_M5-4_0.0625ng_C_CBI_24s.hid	5 person run as a 2 person
-2	M5_A02_STRmix_M5-4_0.125ng_C_0.125ng_P2_GF_CBI_20s.hid	5 person run as a 3 person
-1	M2_D02_STRmix_M2-5_0.025ng_C_B1_CBI_24sec.hid	2 person run as single source
-1	M3_B02_STRmix_M3-4_0.0375ng_C_N1_CBI GF.hid	3 person run as a 2 person
-1	M4_B02_STRmix_M4-5_C_0.1ng_D3_CBI_J6.hid	4 person run as a 3 person
-1	M4_C02_STRmix_M4-5_C_0.05ng_D3_CBI_J6.hid	4 person run as a 3 person
-1	M4_E03_STRmix_M4-5_0.05ng_C_CBI_24s.hid	4 person run as a 3 person
-1	M5_A03_M5-3_0.25ng_amp1_CBI_24s.hid	5 person run as a 4 person
-1	M5_A04_STRmix_M5-4_0.25ng_C_CBI_24s.hid	5 person run as a 4 person
-1	M5_A06_M5-3_0.25ng_amp2_CBI_24s.hid	5 person run as a 4 person
-1	M5_A08_M5-3_0.5ng_amp2_CBI_24s.hid	5 person run as a 4 person
-1	M5_C04_STRmix_M5-4_0.125ng_C_CBI_24s.hid	5 person run as a 4 person
-1	M5_H01_STRmix_M5-4_0.25ng_C_0.25ng_P2_GF_CBI_20s.hid	5 person run as a 4 person
-1	M5_H02_M5-2_0.25ng_amp1_CBI_24s.hid	5 person run as a 4 person
-1	M5_H02_M5-2_0.5ng_amp1_CBI_24s.hid	5 person run as a 4 person
-1	M5_H05_M5-2_0.25ng_amp2_CBI_24s.hid	5 person run as a 4 person
-1	M5_H07_M5-2_0.5ng_amp2_CBI_24s.hid	5 person run as a 4 person
Over-assignment		
+1	M3_B02_STRmix_M3-4_1.2ng_C_CBI_24s.hid	3 person run as a 4 person
+1	M3_G01_STRmix_M3-4_0.3ng_C_N1_CBI GF.hid	3 person run as a 4 person
+1	M3_H04_M3-1_0.25ng_amp2_CBI_24s.hid	3 person run as a 4 person

Each profile was interpreted in STRmix™ using the apparent N values and compared to the known contributors and 500+ known non-contributors using the Database Search function within STRmix™ to calculate a likelihood ratio (*LR*). The non-contributors were artificially generated using the NIST Caucasian Allele Frequencies made available in July of 2017. For any given mixture there were over 500 (517 – true N) non-contributors compared to each mixture output.

Using the NIST Caucasian Allele Frequencies (July 2017) and an F_{ST} of 0.01 (1%), a *sub-source* (*Factor-of-N!*) LR was calculated where the following propositions considered were:

H_p: The DNA originated from the database individual and *N*-1 unknown individuals

H_d: The DNA originated from *N* unknown individuals

Where *N* is the apparent number of contributors to the profile.

Plots of log(*LR*) versus the average peak height (*APH*) per known contributor (blue circles) for the apparent one-, two-, three-, four-, and five-person mixtures are given in Figure 3. *APH* was calculated using unmasked,

unshared, and non-stutter affected alleles for each contributor in the mixed profiles. Where the contributor had completely dropped out of the mixture, an *APH* of half the lowest analytical threshold (AT) used by CBI is applied (half of 40 RFU = 20 RFU). One of the two-person mixtures (M2-5 [1:1] at 0.025ng) appeared as single source with information from the contributors dropping out. Hence a small set of *LR*s were created for an apparent single source profile in addition to the two- to five-contributor mixtures. The per contributor amount of DNA for known non-contributors (red crosses) is taken as the lowest APH of the known contributors per mixture, which may also be the half AT value in some instances.

Exclusions (*LR* = 0) are plotted as log(*LR*) = -40.

Upon review of all the initial results (data not provided but retained with the MS Excel™ workbook [Section D plots VI.xlsx) a small number (15) of false exclusions were observed; that is true contributors giving a *LR* of 0 were noted in the two- to five-person mixtures. The majority (8) of these corresponded with saturated samples (i.e. with peaks in excess of 30,000 RFU, the saturation threshold set for CBI data).

As detailed in Section B, it is not recommended that saturated profiles are interpreted within STRmix™. As a DNA profile exceeds saturations, the height of a stutter peak is no longer linearly-proportional to the parent peak height and a different model is invoked. This alternate model proposes that stutter peaks are proportional to the expected peak height of the parent allele. As a result, elevated stutter variance values are proposed and STRmix™ will also propose some of the stutter peak information as allelic. This may cause STRmix™ to propose genotypes that do not align with the minor contributor resulting in potential false exclusions.

In the Federal Bureau of Investigation's (FBI) published internal validation of STRmix™ [7], they also observed false exclusions when saturated profiles were interpreted through STRmix™. Within their internal validation it has stated that *'while some saturated peaks may have a nominal effect on LRs and weights in some STRmix™ analyses, it is advisable to reprocess the sample'*. The STRmix™ developers do not advocate running saturated samples though STRmix™. At the CBI-FS, analysts are advised to reprocess the sample (e.g. amplify at lower template input amounts). If this is not possible due to sample availability, then the primary diagnostics will be reviewed to ensure data is intuitive, and caution will be used when interpreting saturated data.

With this in mind plots were created omitting the most saturated samples. This constituted 17 of the 188 mixtures developed. Please refer to Figure 3. The 17 samples which were omitted were:

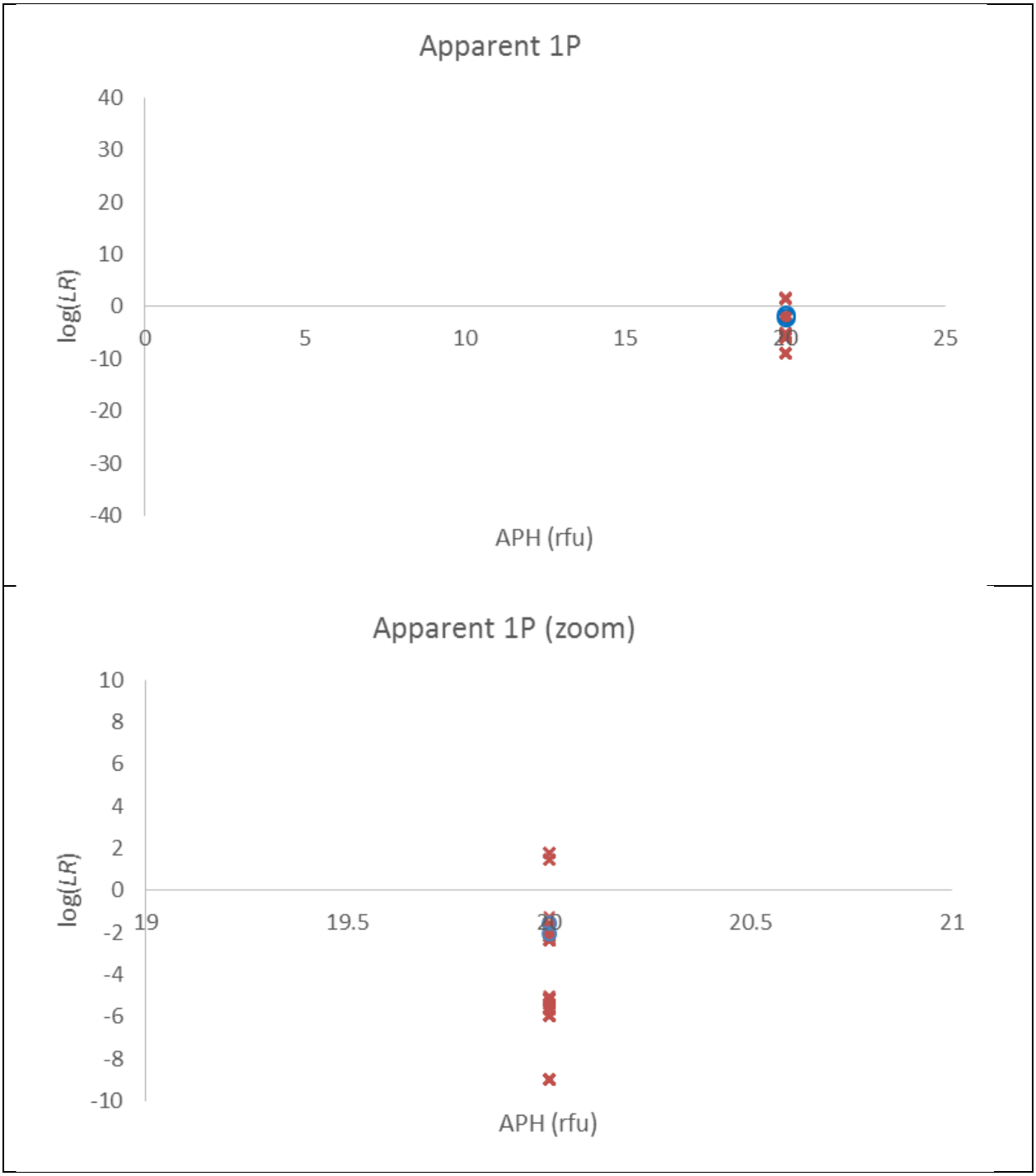
- M2_B01_M2-1_Amp1_CBI_24s.hid
- M2_H01_M2-1_Amp2_CBI_24s.hid
- M3_E05_STRmix_M3-1_1.5ng_CBI_24s.hid
- M3_E06_STRmix_M3-3_C_2ng_CBI_24s.hid
- M3_F02_M3-1_CBI_24s.hid
- M3_F02_M3-1_CBI_24s_2.hid
- M3_F03_M3-2_1.5ng_P2_GF_CBI_20s.hid
- M3_F05_STRmix_3-2_1.5ng_CBI_24s.hid
- M3_G02_M3-2_CBI_24s.hid
- M3_G02_M3-2_CBI_24s_2.hid
- M4_C04_M4-4_CBI_24s.hid

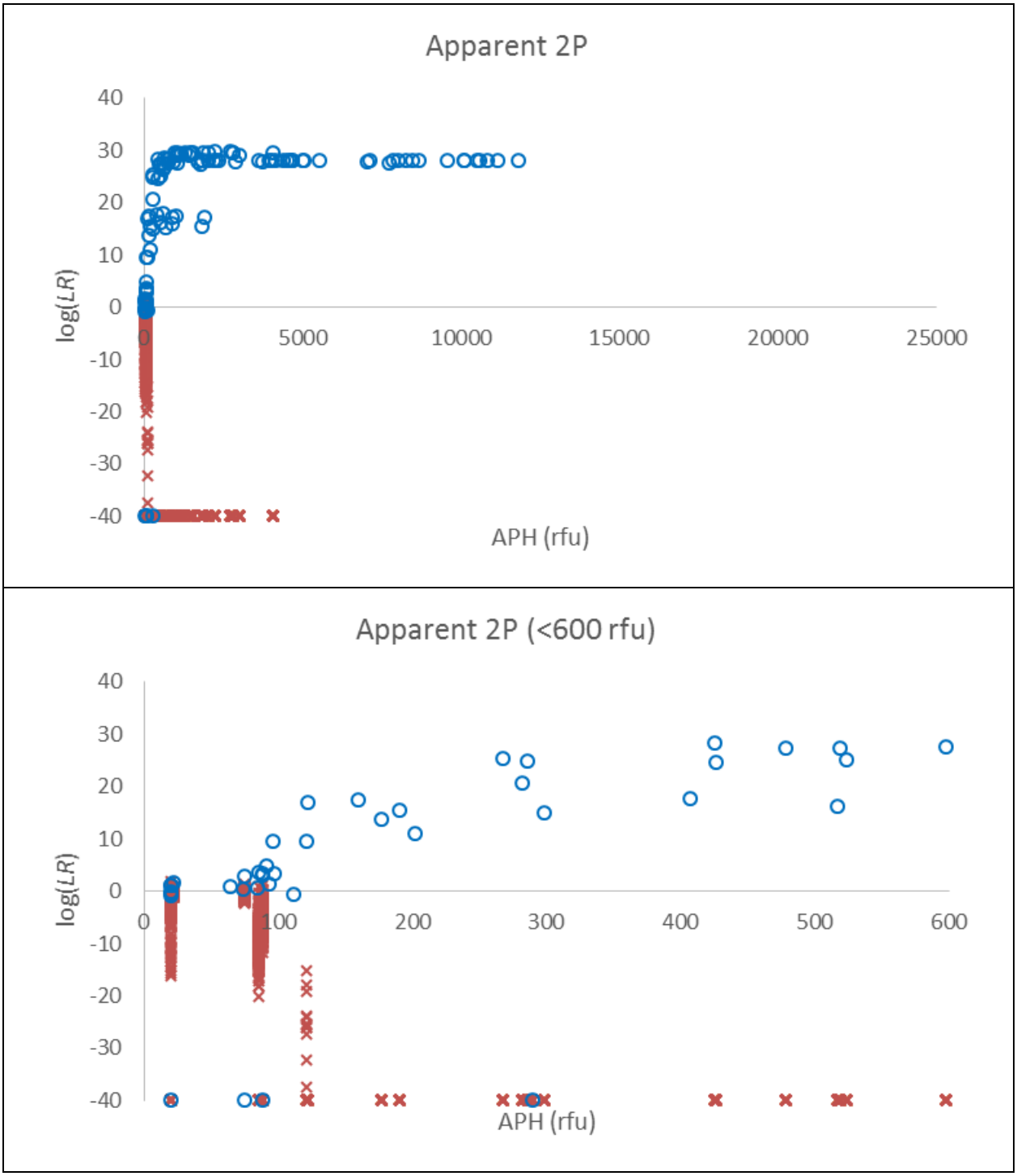
- M4_G01_M4-4_Amp2_CBI_24s.hid
- M5_A02_M5-3_Amp2_CBI_24s.hid
- M5_C01_M5-3_Amp1_CBI_24s.hid
- M5_C05_M5-2_CBI_24s.hid
- M5_C05_M5-2_CBI_24s_2.hid
- M5_D01_STRmix_M5-1_C_1.5ng_P2_GF_CBI_20s.hid

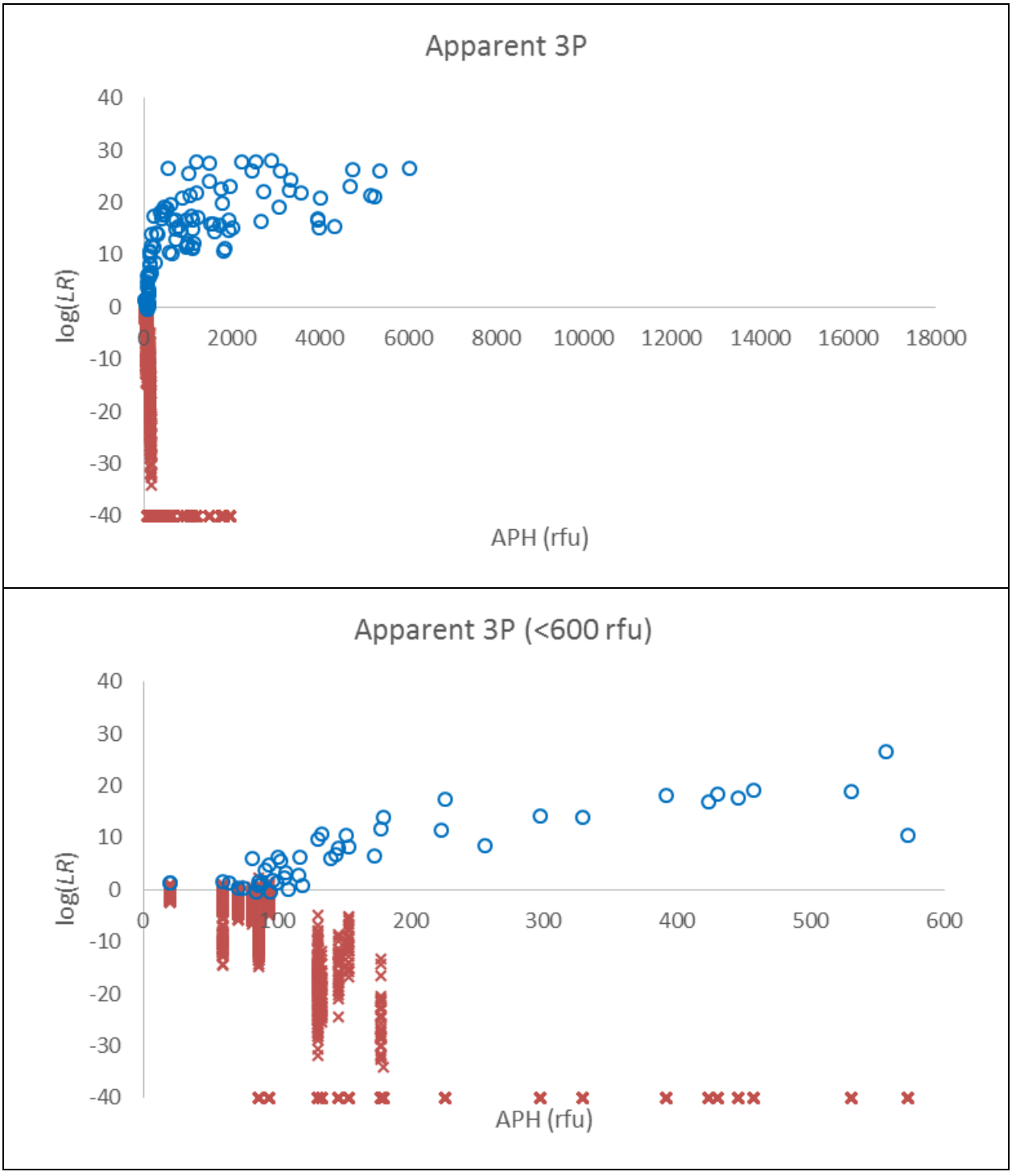
The results of the 171 (188 total minus 17 saturated) comparisons are provided in Figure 3, where the blue circles indicate a known contributor to the mixture (H_p true), and red crosses indicate a known non-contributor (H_d true).

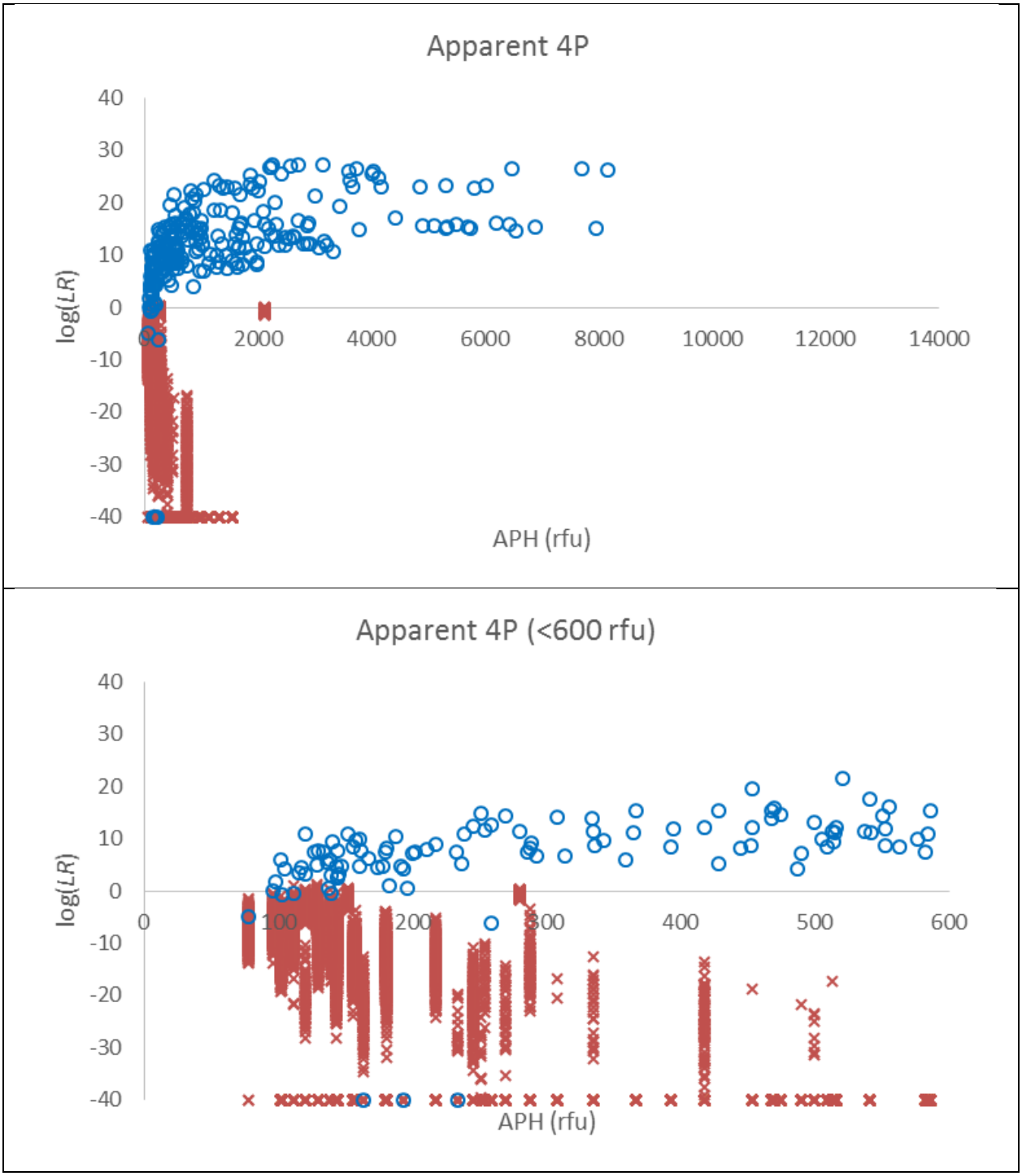
Inspection of the plots in Figure 3 shows that as template (and hence *APH*) increases the *LR* distributions for H_p true and H_d true are very well separated for two-, three-, four-, and five-person mixtures. As the number of contributors increased and the template (and hence *APH*) lowered the two distributions converged on a *LR* = 1 or log(*LR*) = 0. At high template STRmix™ correctly and reliably gave a high *LR* for true contributors and a low *LR* for false contributors.

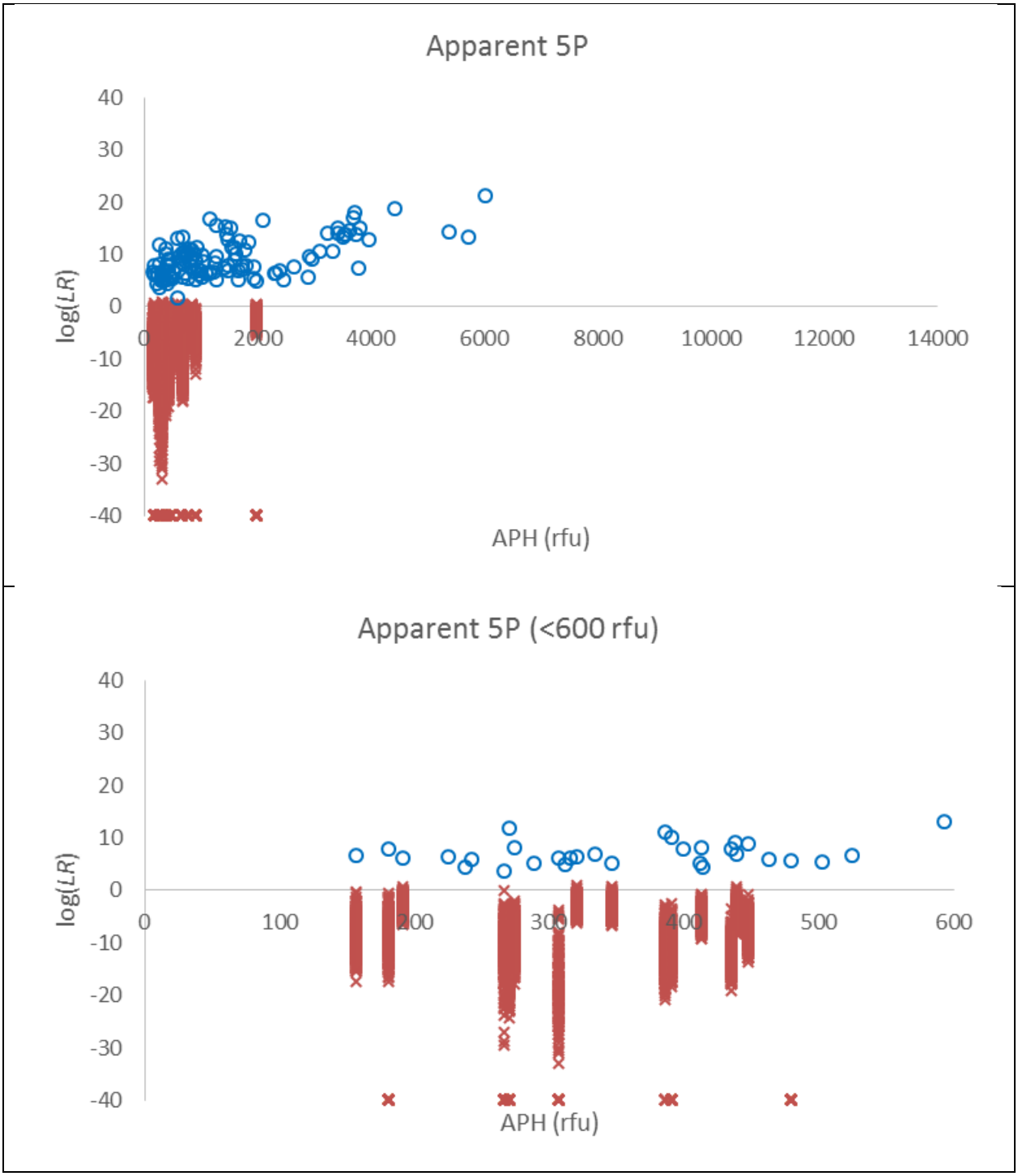
Figure 3: Log(*LR*) versus APH (RFU) for one-, two-, three-, four-, and five-person GlobalFiler™ mixtures. Every second plot is a close-up to better illustrate the data. Blue circles indicate a known contributor to the mixture (H_p true), and red crosses indicate a known non-contributor (H_d true)




Apparent 2P
log(LR)
APH (rfu)
40
30
20
10
0
-10
-20
-30
-40
0
5000
10000
15000
20000
25000
Apparent 2P (<600 rfu)
log(LR)
APH (rfu)
0
100
200
300
400
500
600


Apparent 3P
log(LR)
APH (rfu)
40
30
20
10
0
-10
-20
-30
-40
0
2000
4000
6000
8000
10000
12000
14000
16000
18000
Apparent 3P (<600 rfu)
log(LR)
APH (rfu)
0
100
200
300
400
500
600


Apparent 4P
log(LR)
APH (rfu)
40
30
20
10
0
-10
-20
-30
-40
2000
4000
6000
8000
10000
12000
14000
Apparent 4P (<600 rfu)
log(LR)
APH (rfu)
0
100
200
300
400
500
600


Apparent 5P
log(LR)
APH (rfu)
40
30
20
10
0
-10
-20
-30
-40
0
2000
4000
6000
8000
10000
12000
14000
Apparent 5P (<600 rfu)
log(LR)
APH (rfu)
0
100
200
300
400
500
600

At low template or high contributor number STRmix™ correctly and reliably reported that the analysis of the sample tends towards uninformative or inconclusive (*LR* = 1 or log(*LR*) = 0).

At low template amounts, it can also be difficult to deduce the number of contributors to a mixture. This is most notably shown in the 1p plots in Figure 3, where the number of contributors assigned to the mixture was less than the experimental design (sample [M2-5 [1:1] at 0.025ng]). This is because at low template, allelic information is more likely to drop below the analytical threshold, therefore low template samples are less informative. As a result *LR*s calculated for contributors and known non-contributors fluctuates around the uninformative range. Upon re-assessment of this mixture that was interpreted through STRmix™ as a single-source profile, sub-threshold peaks indicated the presence of two contributors. Therefore the number of contributors was under-assigned and the impact of this is further explored in Section E.

In relation to Figure 3 excluding the most saturated mixtures 7 false exclusions remain. In addition there are known contributors who produce *LR*s lower than expected in samples M5_A06_M5-3_0.25ng_amp2_CBI_24s.hid and M5_A08_M5-3_0.5ng_amp2_CBI_24s.hid. Furthermore, there are groups of H_d true data which are close to an *LR* of 1 in the apparent 4 person mixtures (around an APH of 2131 RFU) and in the 5 person mixtures (around an APH of 1978 RFU). Each of these will be discussed further here.

Of the 7 instances of false exclusions observed (spread across 4 mixtures), 6 false exclusions arose from mixtures that were designed as five-person mixtures. However, due to the complexity of these five-person mixtures, the experienced analyst under-assigned the number of contributors, with the most extreme being assigned as a 2 person mixture, instead of a 5, which led to 3 of the exclusions. The remaining 3 mixtures were assigned as 4 person mixtures resulting in one known contributor being excluded each time. Therefore, when each five-person mixture was interpreted with lower number of contributors, less genotype combinations are proposed. As a result, the genotypes that corresponded to true contributors were not proposed for minor components of the mixture. The effect of under-assigning the number of contributors is explored further in Section F.

The remaining false exclusion that was identified originated from a two-person mixture; D01_M2-1_0.5ng_amp1_CBI_24s.hid. Using the *LR From Previous* function within STRmix™, it was possible to determine that the false exclusion was a result of the D1S1656 and D2S1338 loci. Screenshot images of these two loci for this mixture are shown in Figure 4. A review of the primary diagnostics; namely the individual locus *LR*s, highlighted that inclusionary *LR*s were obtained at all loci except the two described. This allows us to focus on these loci and explore further.

The genotype for the true minor contributor at D1S1656 is [15,15.3] and at D2S1338 [17,25]. Looking at the screenshots of these two loci in Figure 4, we see that the 15 allele at D1S1656 and 25 allele at D2S1338 have dropped below the analytical threshold (70 RFU for this dye channel). This drop-out is likely due to increased stochastic effects. STRmix™ would need to propose drop-out alleles (Q), to consider appropriate genotypes for the true minor contributor to be included. However, after reviewing another primary diagnostic; the weights, in the initial deconvolution of the profile, STRmix™ did not accept any proposed genotype combinations with drop-out at D1S1656 and D2S1338. This caused the false exclusion observed.

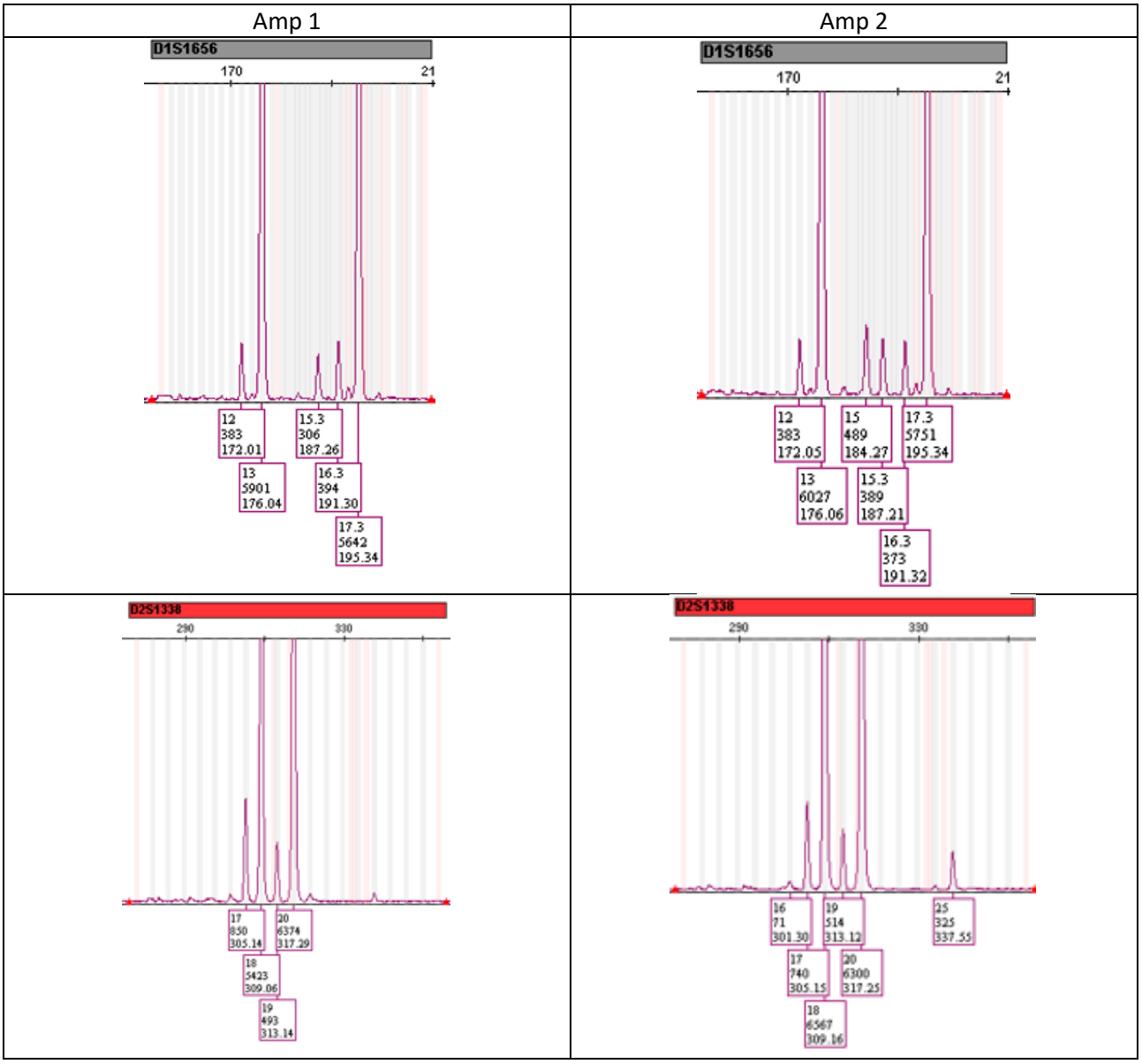
If we compare the electropherogram to that obtained from the second amplification (see Figure 4) of the same extract we see the presence of these alleles.

For investigative purposes, the same profile was deconvoluted again in STRmix™ with an increased number of accepts. This was a ten-fold increase for burnin and post burnin (1,000,000 burn-in accepts per chain and 500,000 post burn-in accepts per chain). This is to allow STRmix™ to more thoroughly explore the sample space and possibly sample more genotype combinations. As expected the genotypes [Q, 15.3] and [Q, 17] were proposed for D1S1656 and D2S1339; albeit, with a low weight of around 10^{-6} and 10^{-5}, respectively. With the possibility of drop-out taken into account, the *LR* for the minor contributor increased from 0 to ~6×10^{14}. Due to MCMC variability, further explored in Section M, the *LR* for the major contributor slightly changed from 1.269×10^{28} to 1.266×10^{28}. It is also important to note that despite increasing the number of accepts the *LR* calculated for all 515 known non-contributors remained at 0.

A review of the primary diagnostics (weights, mixture proportions and, where available, the individual locus *LR*s) can be informative on when re-work options such as extending the number of accepts could be explored.

Furthermore, the impact of using replicate analyses was investigated for this sample. A further STRmix™ deconvolution was under taken using the default number of accepts, but this time using both the input files from amp 1 and amp 2 for this sample. The results support inclusion of a [15,15.3] genotype combination at D1S1656 and [17,25] at D2S1338. The use of replicates can assist with complex or low level DNA profile interpretation; see *DNA DOM 13.1 STRmix Interpretation* for use of replicates.

Figure 4: Screenshot of the loci D1S1656 and D2S1338 in the mixture D01_M2-1_0.5ng_amp1_CBI_24s.hid causing a false exclusion to the minor contributor, R1. The true minor contributor's genotype for D1S1656 is [15,15.3] and for D2S1338 the genotype is [17,25].



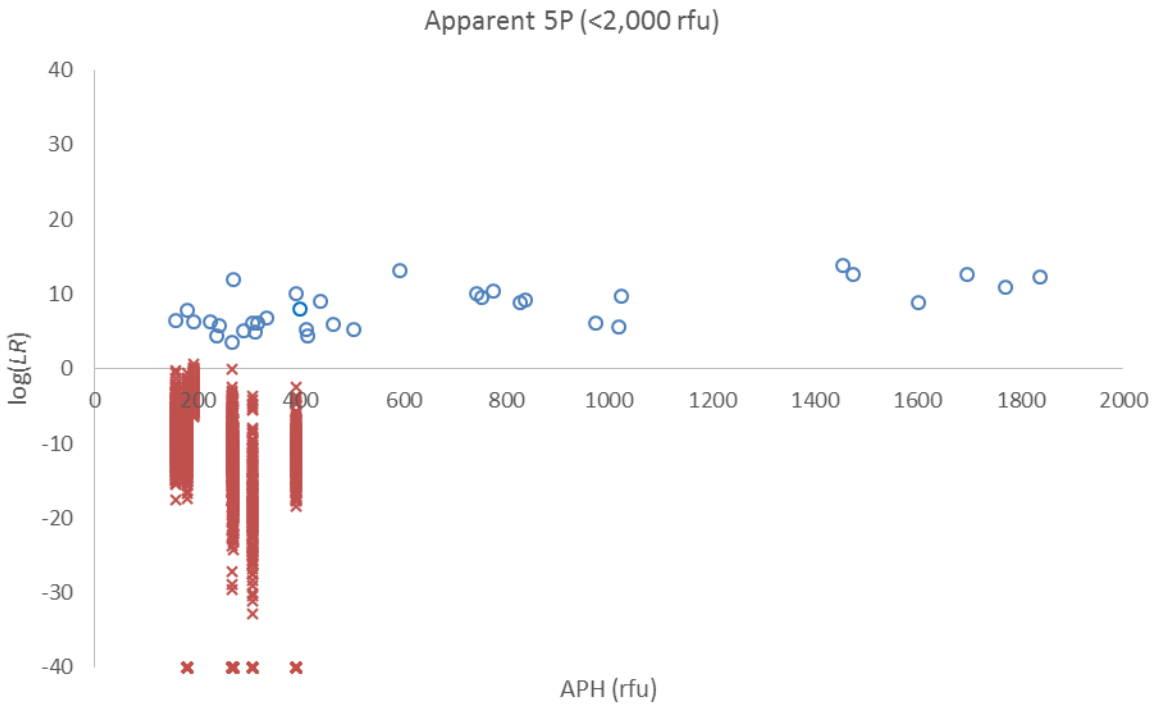
There were a small number of LRs generated to known contributors which were below 1 and so supported exclusion (without being an out-right exclusion). In particular these were in relation to contributor R4 in M5_A06_M5-3_0.25ng_amp2_CBI_24s.hid (*LR* = ~1x10^{-5}, log(LR) = -4.95), which can be seen as the blue circle data point, below the x-axis line at 78 RFU, in the apparent 4 person plot and contributor R1 within sample M5_A08_M5-3_0.5ng_amp2_CBI_24s.hid (*LR* = ~5x10^{-7}, log(LR) = -6.29), which can be seen as the blue circle data point below the x-axis line at 259 RFU in the same 4 person plot. These are to individuals contributing only small amounts of DNA within five-contributor mixtures that have been under-assigned as apparent 4 person mixtures. This type of behaviour is observed with the *LR*s when we under-assign N and is further explored in section F.

The final talking point from Figure 3 is in relation to small groups of known non-contributor (H_d true) *LR*s which are close to an *LR* of 1 (0 on a Log scale) and hence uninformative within the four- and five- contributor mixtures.

For the four-person mixtures at 2131 RFU there is a cluster of known non-contributors with an uninformative *LR* and this relates to mixture B02_STRmix_M3-4_1.2ng_C_CBI_24s.hid. This was designed as a three-person mixture, however due to the number of peaks and imbalances, the experienced analyst over-assigned the number of contributors. Therefore, when the three-person mixture was interpreted as a four-person mixture in STRmix™, more genotype combinations than truly explain the profile are proposed allowing the adventitious links to some non-contributors to occur. The effect of over-assigning the number of contributors is also further explored in Section F.

In the five-person mixtures at 601 RFU up to 1978 RFU there are also clusters of known non-contributors with uninformative *LR*s calculated. These mixtures happened to contain more than 1 ng of template, which may suggest some further saturated peaks within these profiles and arguably could have been omitted from the plots. As discussed previously, saturated profiles are not suitable for STRmix™ interpretation. The log(*LR*) was plotted against the APH again in Figure 5 below for five-person mixtures that did not exceed 1 ng of template. When compared with Figure 3 above it further demonstrates how saturated profiles can generate unintuitive results.

Figure 5: Log(*LR*) versus APH (rfu) for five-person GlobalFiler™ mixtures with less than 1 ng of template.



Overall, the plots in Figure 3 can help inform the limits of STRmix™, particularly the lower limit of DNA where an H_p true hypothesis results in a *LR* greater than 1 and the limit where false positives may arise (a *LR* greater than 1 where H_d is true). These plots also highlight the limitation STRmix™ to interpret saturated data.

Within STRmix™ the primary diagnostics used to assess the appropriateness are the genotype weights, mixture proportions (Mx) and where undertaken, the individual locus *LR*s. These values should be intuitive and resemble the DNA profiles. The secondary diagnostics include the number of iterations, log(likelihoods), Gelman-Rubin convergence diagnostic, and posterior mean of the allele and stutter variances. A summary of these secondary diagnostics for Section D can be found in Appendix 3. These secondary diagnostics will also show the impact of saturated data, reinforcing the premise that saturated profiles should not be interpreted in STRmix™, where possible.

Mixtures of related individuals

In a further extension of this validation a series of mixtures of were created from known relatives. As above specificity and sensitivity was tested by calculating the *LR* for a number of two-, three-, and four-person profiles for both known contributors and known non-contributors. To test the system yet further, some of the non-contributors were additional relatives of the known donors to the mixtures. The plots in [6] have been reproduced for CBI's GlobalFiler™ data as above. A total of 216 mixtures of relatives with varying number of contributors (two-, three-, four-person) and mixture proportions were generated in the CBI laboratory using GlobalFiler™. A summary of the different types of mixtures is shown in Table 7 and the intended contributors is shown in Table 8.

Table 7: Summary of experimental design for specificity and sensitivity tests of related individuals. The values provided are the intended dilution series (note: some actual targets differ slightly)

				Target DNA Amount (ng) in a dilution series					
Mixture	**Amp.**	**Donor Ratio**	**N**	**Set 1**	**Set 2**	**Set 3**	**Set 4**	**Set 5**	**Set 6**
• Sib • PC • GPGC • AUNN • 1_2Sib • COCO	1	5:1	2	1.5	1	0.5			
	2								
	1	3:1							
	2								
• GPGC • AUNN • 1_2Sib • COCO	1	1:1							
	2								
• GPGC • AUNN	1	1:3							
	2								
• GPGC • AUNN	1	1:5							
	2								
• Sib • PC	1	1:1	2	0.8	0.4	0.2	0.1	0.05	0.025
	2								
• PCC	1	7:2:1	3	1.5	1	0.5			
	2								
	1	1:3:1							
	2								
	1	4.5:4.5:1							
	2								
	1	1:8:1							
	2								
• PPC	1	7:2:1	3						
	2								
	1	1:1:3							
	2								
	1	1:2.5:1							
	2								
	1	1:4.5:4.5							
	2								
• PPCC	1	4:4:1:1	4						
	2								
	1	1:1:4:2							
	2								
	1	1:6:2:1							
	2								
	1	14:1:4:1							
	2								

Note: This constitutes 108 samples amplified in duplicate, hence 216 mixtures. The value for N described above is the experimental design number of contributors.

Table 8: Summary of contributors from experimental design

Mixture	Contributors				
Person (P)	Prefix	P1	P2	P3	P4
Siblings (Sib)	Sib	C3	C2		
Parent-child (PC)	PC	P2	C3		
Parent-child-child (PCC)	PCC	P1	C1	C2	
Parent-parent-child (PPC)	PPC	P1	P2	C2	
Parent-parent-child-child (PPCC)	PPCC	P1	P2	C1	C2
Grandparent-grandchild (GPGC)	GPGC	GM1	O1		
Aunt/uncle-niece/nephew (AUNN)	AUNN	M4	O5		
Half siblings (1_2Sib)	1_2Sib	OF2M3	O3		
Cousins (COCO)	COCO	OF2M3	O1		

These profiles will likely represent the most challenging types of profiles to be encountered by the laboratory. The profiles are of varying DNA quantity and mixture proportions. The contributors include significant amounts of allele sharing across the different loci (standard 4.1.6.5). Given the template amounts allele and/or locus dropout was expected to occur within the profiles containing the lower DNA amounts (standard 4.1.7.1).

As above, prior to interpretation in STRmix™ the profiles were assessed, by an experienced analyst, with regards to the 'apparent' number of contributors (N) rather than the true number of contributors (or the number of contributors in the experimental set up). The apparent N was assigned to each profile. The true number of contributors to a questioned/crime scene profile is *always* unknown and to ensure these results are applicable to casework, this approach was used. However, a review of the *LR*s created by apparent N and experimental design is provided.

Even with these mixtures of relatives the majority of the time the assigned apparent N was the same as experimental design, however a number (46, 21%) were under-assigned in relation to experimental design due to the high levels of allelic overlap. All under assigned samples involved parent(s)/child mixtures, which could be anticipated. Five mixtures were under-assigned as 2 contributors. These were all 4 person, parent:parent:child:child (PPCC) mixtures, which arguably are the hardest we could encounter and we could expect a mixing of these 4 people could appear as a 2 person mixture, depending on the mixture proportions. In fact none of the 4 person PPCC (24 mixtures in total) were assigned as an apparent 4 person. Most were assigned as a 3 person with the five, as mentioned, assigned as 2 person.

Each profile was interpreted in STRmix™ using the apparent N values and compared to the known contributors and 500+ known non-contributors using the Database Search function within STRmix™, as above. Again *LR*s were calculated using the NIST Caucasian Allele Frequencies (July 2017) and an F_{ST} of 0.01 (1%). The *sub-source* (*Factor-of-N!*) *LR* was calculated where the following propositions considered were:

H_p: The DNA originated from the database individual and *N*-1 unknown individuals

H_d: The DNA originated from *N* unknown individuals

Where *N* is the apparent number of contributors to the profile.

As above plots of log(*LR*) versus the average peak height (*APH*) per contributor were created. Again *APH* was calculated using unmasked, unshared, and non-stutter affected alleles for each contributor in the mixed profiles. Where the contributor had completely dropped out of the mixture, an *APH* of half the lowest analytical threshold (AT) used by CBI is applied (20 RFU). The per contributor amount of DNA for known non-contributors (red crosses) is taken as the lowest APH of the known contributors per mixture, which may also be the half AT value in some instances.

Exclusions (*LR* = 0) are plotted as log(*LR*) = -40.

Similar to the mixtures of unrelated individuals covered above, during the GeneMapper™ analysis of the 216 profiles, the analyst(s) noted that some mixtures appeared saturated, with peaks in excess of 30,000 RFU; the saturation threshold set for CBI data. For example one mixture (E01_PPC1_Amp1_CBI_24s.hid) was extremely saturated when analysed and hence was not taken further. Upon review of all the initial STRmix™ results of the remaining 215 mixtures (data not provided here but retained with the MS Excel™ workbook [Section D Plots Relatives IV]) a number of false exclusions; that is true contributors giving a LR of 0 were noted in the mixtures. Many of these corresponded with saturated samples. Therefore, the plots in Figure 6 have the results from the saturated mixtures removed.

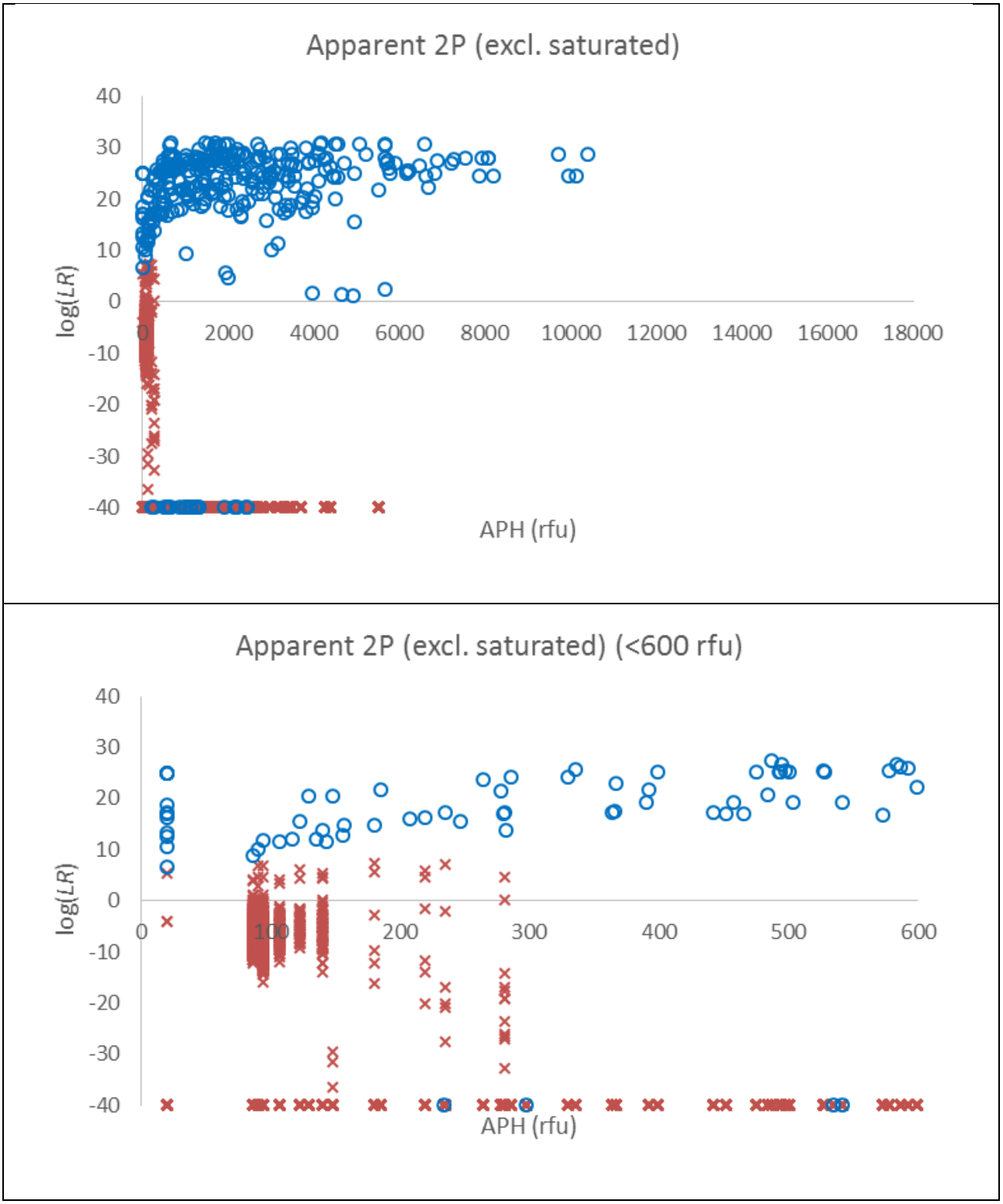
This constituted 30 of the 215 mixtures progressed. The 30 samples which were omitted were:

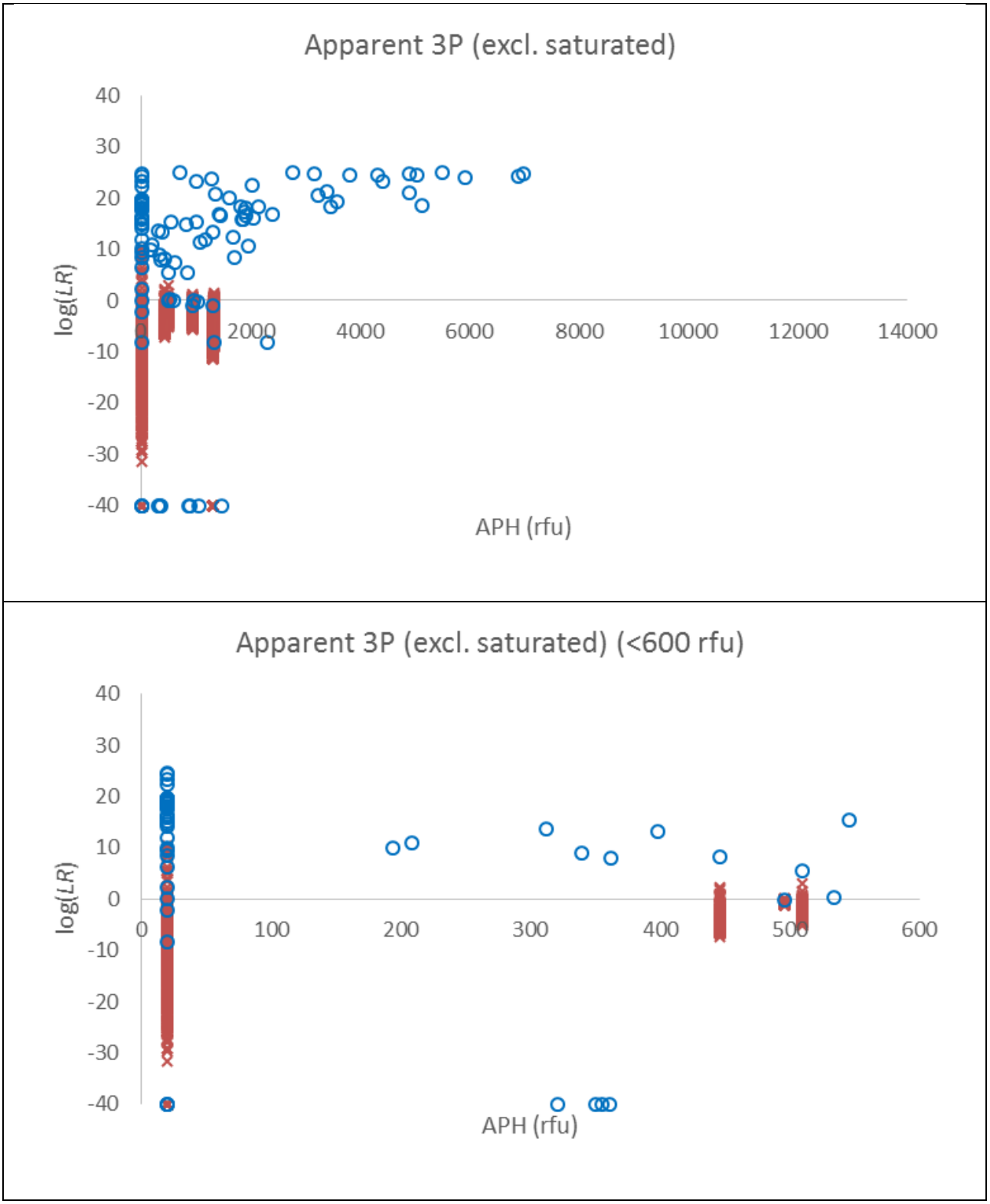
- PC_B08_PC1_1ng_amp2_CBI_24s.hid
- PC_D07_PC1_CBI_24s.hid
- PC_D07_PC1_CBI_24s_2.hid
- PC_E07_PC2_CBI_24s.hid
- PC_E07_PC2_CBI_24s_2.hid
- PCC_D01_PCC1_Amp1_CBI_24s.hid
- PCC_D08_PCC1_CBI_24s.hid
- PCC_E08_PCC2_CBI_24s.hid
- PCC_E08_PCC2_CBI_24s_2.hid
- PCC_F08_PCC3_1ng_amp2_CBI_24s.hid
- PCC_F08_PCC3_CBI_24s.hid
- PCC_F08_PCC3_CBI_24s_2.hid
- PCC_G08_PCC4_CBI_24s.hid
- PCC_G08_PCC4_CBI_24s_2.hid
- PPC_B09_PPC3_CBI_24s.hid
- PPC_B09_PPC3_CBI_24s_2.hid
- PPC_C09_PPC4_CBI_24s.hid
- PPC_C09_PPC4_CBI_24s_2.hid
- PPC_D04_PPC2_0.5ng_amp1_CBI_24s.hid
- PPC_H08_PPC1_CBI_24s.hid
- PPCC_D09_PPCC1_CBI_24s.hid

- PPCC_E09_PPCC2_CBI_24s.hid
- PPCC_E09_PPCC2_CBI_24s_2.hid
- PPCC_F09_PPCC3_CBI_24s.hid
- PPCC_F09_PPCC3_CBI_24s_2.hid
- PPCC_G09_PPCC4_CBI_24s.hid
- PPCC_G09_PPCC4_CBI_24s_2.hid
- Sib_C06_STRmix_Sib2_reprep_1ng_amp2_CBI_24s.hid
- Sib_E01_STRmix_Sib2_reprep_1.5ng_amp1_CBI_24s.hid
- Sib_E04_STRmix_Sib2_reprep_1.5ng_amp2_CBI_24s.hid

The results of the 185 (215 total minus 30 saturated) comparisons are provided in Figure 6, where the blue circles indicate a known contributor to the mixture (H_p true), and red crosses indicate a known non-contributor (H_d true).

Figure 6: Log(*LR*) versus APH (RFU) for the two- and three-person GlobalFiler™ mixtures of related individuals. The saturated samples have been omitted. Every second plot is a close-up to better illustrate the data. Blue circles indicate a known contributor to the mixture (H_p true), and red crosses indicate a known non-contributor (H_d true).




Apparent 3P (excl. saturated)
log(LR)
APH (rfu)
40
30
20
10
0
-10
-20
-30
-40
2000
4000
6000
8000
10000
12000
14000
Apparent 3P (excl. saturated) (<600 rfu)
log(LR)
APH (rfu)
0
100
200
300
400
500
600

Inspection of Figure 6 shows that on the whole inclusionary *LR*s were observed for most of the known contributors to these mixtures and *LR*s supporting exclusion were obtained for most non-contributors.

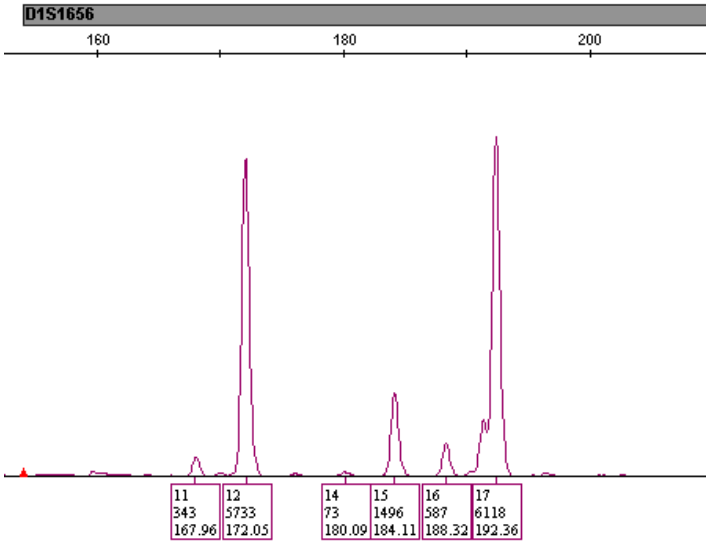
The calculation of APH is more problematic with mixtures of related individuals as there are fewer unshared alleles to consider. Hence there are groups of data and many data points plotted at the lowest value (half AT) of 20 RFU. This may be somewhat misleading for some mixtures. If you had a mixture of closely related individuals there may be few unshared peaks to consider and hence some contributors who may have a reasonable template could still be plotted at 20 RFU.

There were a number of H_p true contributors giving exclusionary *LR*s (*LR* = 0, log(*LR*) = -40). Specifically there were 25 of these observations in the apparent 2 person mixtures and 12 in the apparent 3 person mixtures.

The vast majority of these (24) were from mixtures where the number of contributors was under-assigned. As discussed previously and in Section F, we are aware that if we under-assign we can observe false exclusions to true contributors. This also likely accounts for the group of LRs close to 1 around 4000-6000 RFU in the apparent 2 person mixtures.

The remaining false exclusion of a true contributor (O1) was seen in sample COCO_G06_STRmix_COCO1_1ng_CBI_24s.hid. This was sourced back to the D1S1656 locus. There appears to be a single base-pair resolution issue in GeneMapper™. Below is a screenshot of the locus.

Figure 7: Screen shot of D1S1656 locus from COCO_G06_STRmix_COCO1_1ng_CBI_24s.hid



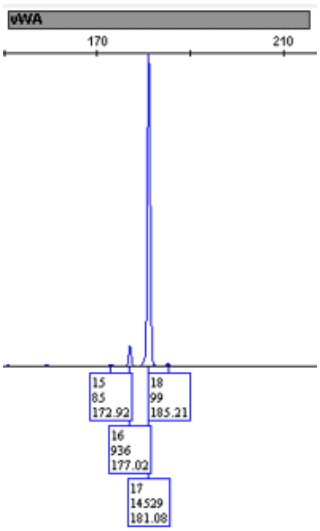
This is a known mixture of OF2M3 & O1. OF2M3 is a [12, 17] and O1 is [15, 16.3]. There appears to be a 16.3 peak however, it is sitting in the shoulder of the strong 17 peak and is not resolved. Therefore, this is not on the input file and STRmix™ has not considered a drop peak pairing with the 15 allele, hence O1 is excluded at this locus. A review of the individual locus *LR*s on the STRmix™ output can be a useful guide, in situations like this. If all the other loci provide inclusionary LRs and one exclusion that locus can be investigated further.

This is a becoming a more common occurrence especially with the larger multiplexes containing more loci with point variant alleles. This is particularly prevalent in major:minor mixtures. It is recommended that work is

undertaken to resolve this issue, whether that is a re-run in the CE instrument or re-amplification of the sample (note this issue is not seen in the other amp of the same sample, where the single allele variant is resolved) or adjusting the analysis settings. If none of these enable the peak to be labelled then the remaining option is to ignore the locus, either during the deconvolution or during 'LR from Previous'. We advocate ensuring concordance of the proposed allelic information to any given person of interest prior to ignoring a locus in any analysis.

In addition to the false exclusions of some known contributors, a couple of blue circle data points are visible around 1300 RFU and 2300 RFU, below the x-axis (LR = 1) line, in the apparent 3 person mixtures. These were sourced back to mixture PCC_E08_PCC2_1ng_amp2_CBI_24s.hid and are the *LR*s calculated to P1 and C2. The input file created contains peaks at vWA and D2S1338 which are likely minus two full repeat (i.e. N-8) back stutters. A screenshot of vWA is provided below.

Figure 8: Screenshot of vWA locus from PCC_E08_PCC2_1ng_amp2_CBI_24s.hid



Double back stutter is not currently modelled in this version of the STRmix™ software. This mixture has a strong major component and hence these types of stutters may be expected. STRmix™ does consider the 15 allele as being potential stutter of the 16 and also drop-in, but favours the 15 as being allelic in this deconvolution for the minor. All the known contributors are [17, 17] and these have much lower weights for the minor contributors in this mixture deconvolution. Exhibit caution when undertaking GeneMapper™ analysis for situations such as this.

These results differ to the mixtures of the unrelated individuals in that there are some H_d true contributors giving elevated *LR*s, for example over 10^3. A review of these shows that all LRs over 10^3 for known non-contributors were obtained to individuals who are related to the known contributors but were not actual contributors themselves. The highest of these being ~$2x10^{10}$ seen in PPC_E09_PPC3_0.5ng_amp2_CBI_24s.hid. This was obtained in a mixture of parent:parent:child (P1:P2:C2) and the false inclusion LR was generated to another child, C1. Adding template and/or changes in mixture proportions allows better resolution of even these related individuals.

However, overall it is clear STRmix™ performs well when comparing mixture of related individuals to non-contributors who are unrelated to the true donors. Mixtures that originate from circumstances in which close relatives could be involved require careful consideration.

Section E: Alternate propositions

This section covers the following standard:

> 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. For example, if there are two persons of interest, they may be evaluated as co-contributors and, alternatively, as each contributing with an unknown individual. The hypotheses used for evaluation of casework profiles can have a significant impact on the results obtained.

A sub-set of the same profiles as used in Section D above were reinterpreted in STRmix™ with alternate propositions.

A total of thirteen samples of varying template, mixture proportions, and complexity were deconvoluted conditioning on the major contributor. These mixtures are listed in Table 9.

Table 9: Samples selected for re-interpretation assuming the major contributor

Samples selected	Comments
M3_E02_STRmix_M3-4_0.15ng_C_CBI_24s.hid	3p 1:1:1 0.15ng
M3_F01_STRmix_M3-4_0.6ng_C_N1_CBI GF.hid	3p 1:1:1 0.6ng
M3_A05_M3-2_0.25ng_amp2_CBI_24s.hid	3p 10:10:1 0.25ng
M3_H01_M3-1_0.25ng_amp1_CBI_24s.hid	3p 10:5:1 0.25ng
M3_H01_M3-1_0.5ng_amp1_CBI_24s.hid	3p 10:5:1 0.5ng
M3_B02_M3-3_0.25ng_amp1_CBI_24s.hid	3p 3:2:1 0.25ng
M4_E03_STRmix_M4-5_0.05ng_C_CBI_24s.hid	4p 1:1:1:1 0.05ng
M4_C03_STRmix_M4-5_0.2ng_C_CBI_24s.hid	4p 1:1:1:1 0.2ng
M4_H01_STRmix_M4-5_C_0.4ng_D3_CBI_J6.hid	4p 1:1:1:1 0.4ng
M4_E02_M4-3_0.5ng_amp1_CBI_24s.hid	4p 10:10:5:1 0.5ng
M4_D07_M4-2_1ng_amp2_CBI_24s.hid	4p 10:5:2:1 1ng
M4_F05_M4-4_0.25ng_amp2_CBI_24s.hid	4p 10:5:5:1 0.25ng
M4_C07_M4-1_0.5ng_amp2_CBI_24s.hid	4p 4:3:2:1 0.5ng

The same profiles as in Section D above were reinterpreted in STRmix™ with alternate propositions. In these interpretations one of the contributors is assumed as a known under both H_p and H_d. The different propositions being considered are:

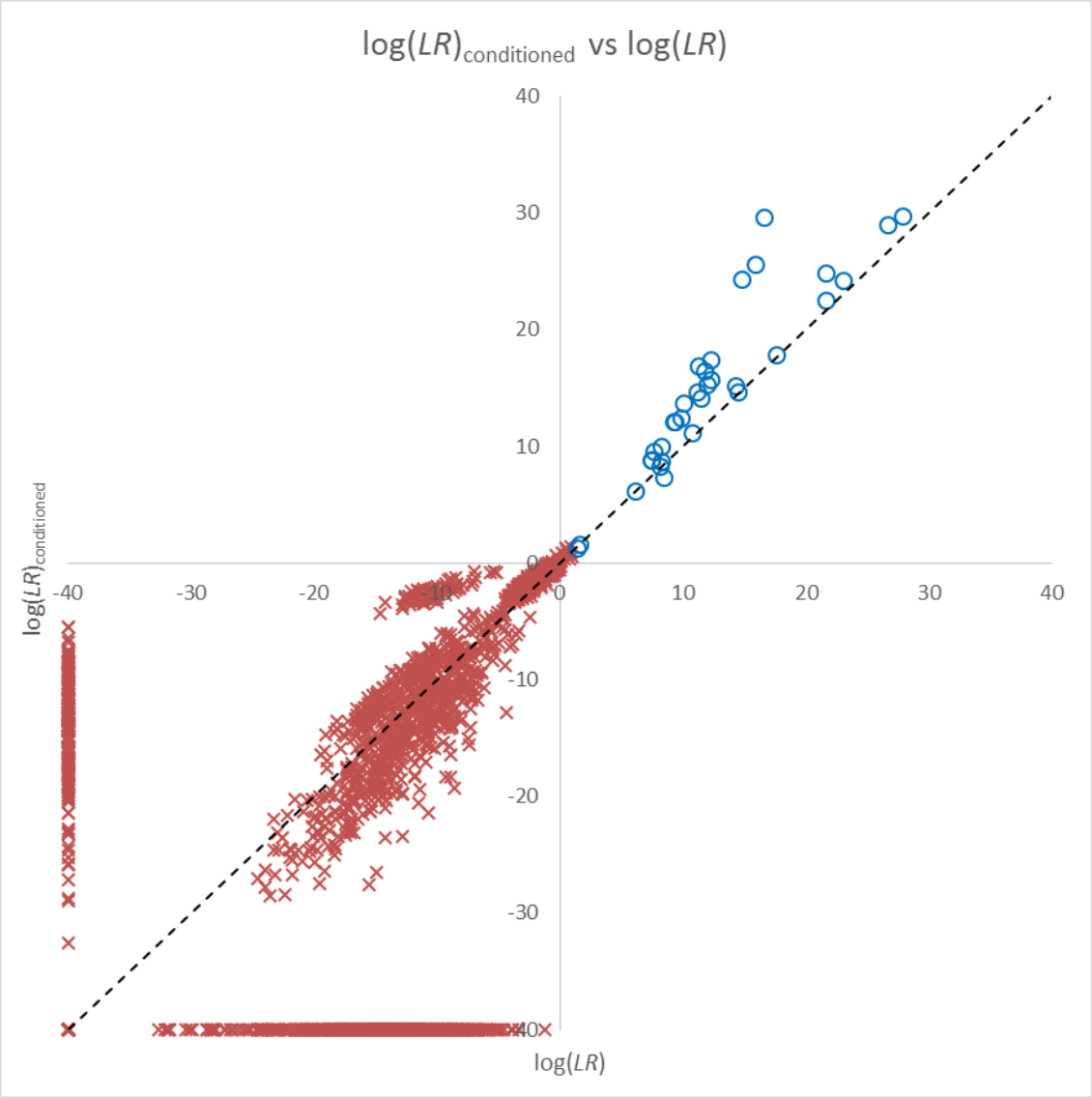
> H_p: The DNA originated from the known individual, the database individual and N-2 an unknown individuals

H_d: The DNA originated from the known individual and *N*-1 unknown individuals

Where, as per section D, N relates to the apparent number of contributors.

To mimic typical casework the major contributor was assumed each time and then likelihood ratios (*LR*s) calculated to the remaining potential contributors. *LR*s were calculated as per section D. An example plot of the log(*LR*) calculated under the two different propositions is provided in Figure 9.

Figure 9: Comparison of the log(*LRs*) obtained from mixtures when assuming a contributor to when there is no person assumed (i.e. section D results). *LR*s obtained to the known contributors are shown as the blue cirles and the non-contributors as red crosses



Values above the line at $x=y$ for the H_p true *LR*s (blue circles) indicate that the *LR* generally increases when conditioning on, or assuming, a true contributor. Also numerous H_d true *LR*s (red crosses) have also gone from supporting exclusion to an outright exclusion, when a true contributor is assumed. There are some results where fixing one contributor has taken some of the H_d true *LR*s from outright exclusion to supporting exclusion, which may be a consequence of the MCMC process. Overall, this shows that the addition of more relevant information (such as the addition of assumed contributors) is shown to improve the performance of STRmix™.

Section F: Assigning number of contributors

This section covers the following standard:

> 4.1.6.4. If the number of contributors is input by the analyst, both correct and incorrect values (i.e., over- and under-estimating) should be tested.

The effect of the uncertainty in the number of contributors within STRmix™ has previously been reported for a number of profiles with *N* and *N*+1 assumed contributors, where *N* is the number of contributors [8, 9]. Increasing the number of contributors to one greater than truly present in the profile had the effect of lowering the *LR* for trace contributors within the profile. STRmix™ adds the additional (unseen) profile at trace levels which interacts with the known trace contribution, diffusing the genotype weights and lowering the *LR*. There was no significant effect on the *LR* of the major or minor contributor within the profiles.

The effect was tested using CBI data by comparing STRmix™ interpretations where the apparent number of contributors to a mixture (assigned in Section D above) differs from the experimental design number of contributors to the mixture.

In casework, the true number of contributors to a questioned/crime profile is *always* unknown. Analysts are likely to add contributors in the presence of an artifact, high stutter, or forward stutter peaks. The assumption of one fewer contributor than that actually present may be made when contributors are at very low levels and dropping out (or visible below the analytical threshold), in profiles where DNA is from individuals with similar profiles at the same concentrations, or family scenarios, such as DNA from a father, mother and their child where the child was the minor contributor.

For the purposes of Section F, N is defined as the experimental design number of contributors. N+1 would indicate the addition of one contributor to the experimental design N, and N-1 would indicate the subtraction of one contributor to N.

Addition of one contributor

During the analysis of the profiles in Table 6 (see Section D), there were three mixtures where the experienced analyst assigned one more contributor than the experimental design number of contributors (hence N+1).

The three samples in question were (also listed in Table 6):

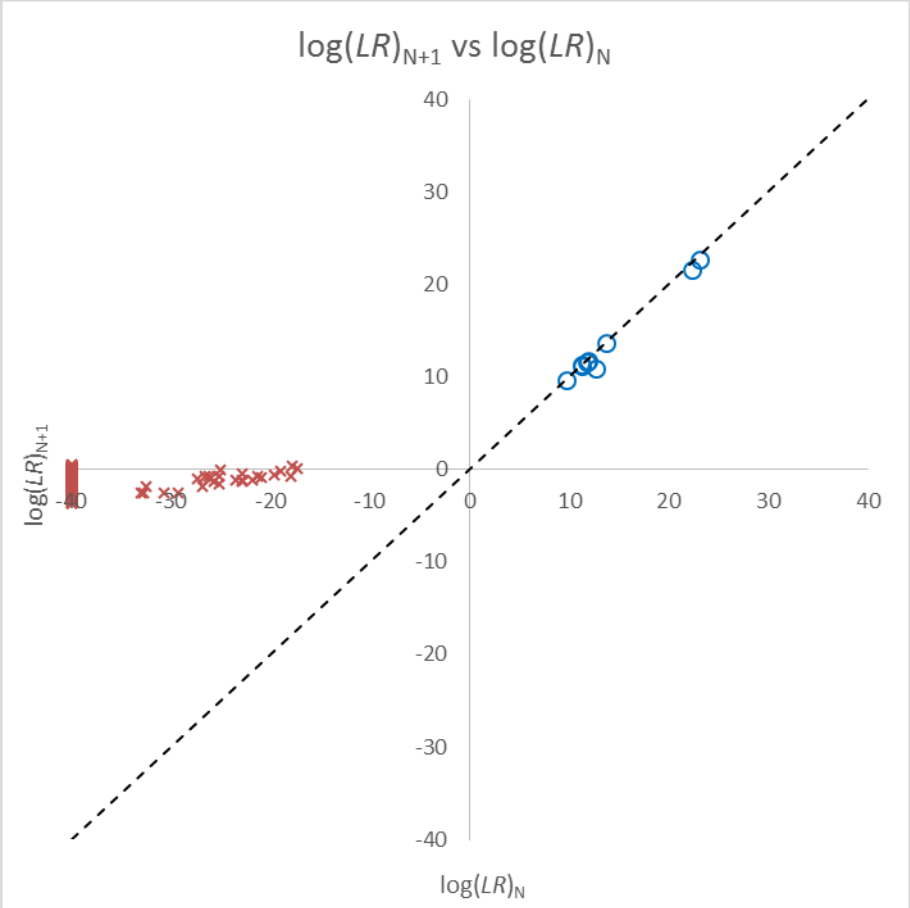
- B02_STRmix_M3-4_1.2ng_C_CBI_24s.hid
- G01_STRmix_M3-4_0.3ng_C_N1_CBI GF.hid
- H04_M3-1_0.25ng_amp2_CBI_24s.hid.

These were all created as 3 person mixtures during experimental design, but were assigned as 4 person mixtures by an experienced analyst, mainly due to stochastic effects causing imbalance at some loci.

The *LR* for the N known contributors and the 517-N known non-contributors (as used for the specificity and sensitivity studies, Section D) were calculated. The same allele frequency database and F_{ST} were used and the sub-source *LR* (Factor-of-N! *LR*) was used as the point of comparison. The log(*LR*) was compared for the known

contributors and known non-contributors under the assumption of *N* (experimental design) and *N*+1 (apparent N, section D) contributors. A plot of log(*LR*) of apparent *N* (Section D, which would be N+1 here) and N is provided in Figure 10.

Figure 10: log(*LR*) when apparent N is an over-assigned (N+1) versus log(*LR*) for experimental design N for the known and non-contributors (where H_p true are in blue circles and H_d true in red crosses)



The plot above demonstrates that there is little effect on the *LR* of the H_p true major or clear minor contributors (blue circles) when the number of contributors is over-estimated. There is a small reduction in a couple of comparisons, where the weights would be slightly diffused due to the additional contributor. The inclusion of an additional contributor beyond that present in the profile has the effect of decreasing the discriminatory power for the H_d true individuals (red crosses). This is because when we 'over-assign' STRmix™ adds the additional (unseen) profile at low DNA amount (template) levels, diffusing the genotype probabilities. This allows more genotype combinations at the loci, albeit with very low genotypic weight. This is visible with the red crosses grouped around the left hand side of the *x*-axis. This essentially means the *LRs* strongly support or out-right exclude (*LR* = 0), when we use the experimental design N, with *LR*s ranging from approx. 10^{-18} to exclusion (*LR* = 0). However, this reduces to be around uninformative (*LR* ~1) to supporting exclusion when we over-assign N, with *LR*s ranging between approx. 10^{-4} to 2.9 (Log(*LR*) = -3.66 to 0.46).

Broadly speaking, over-estimating the number of contributors can result in false inclusions of non-contributors, albeit often these are low *LR*s.

Subtraction of one contributor

The assumption of one fewer contributor than is actually present may be made when contributors are at very low levels and dropping out or there is allele sharing, perhaps due to relatedness.

During the analysis of the profiles in Table 4 (see Section D), there were eighteen mixtures where the experienced analyst assigned at least one less contributor than the experimental design number of contributors.

The eighteen samples in question were (also listed in Table 6):

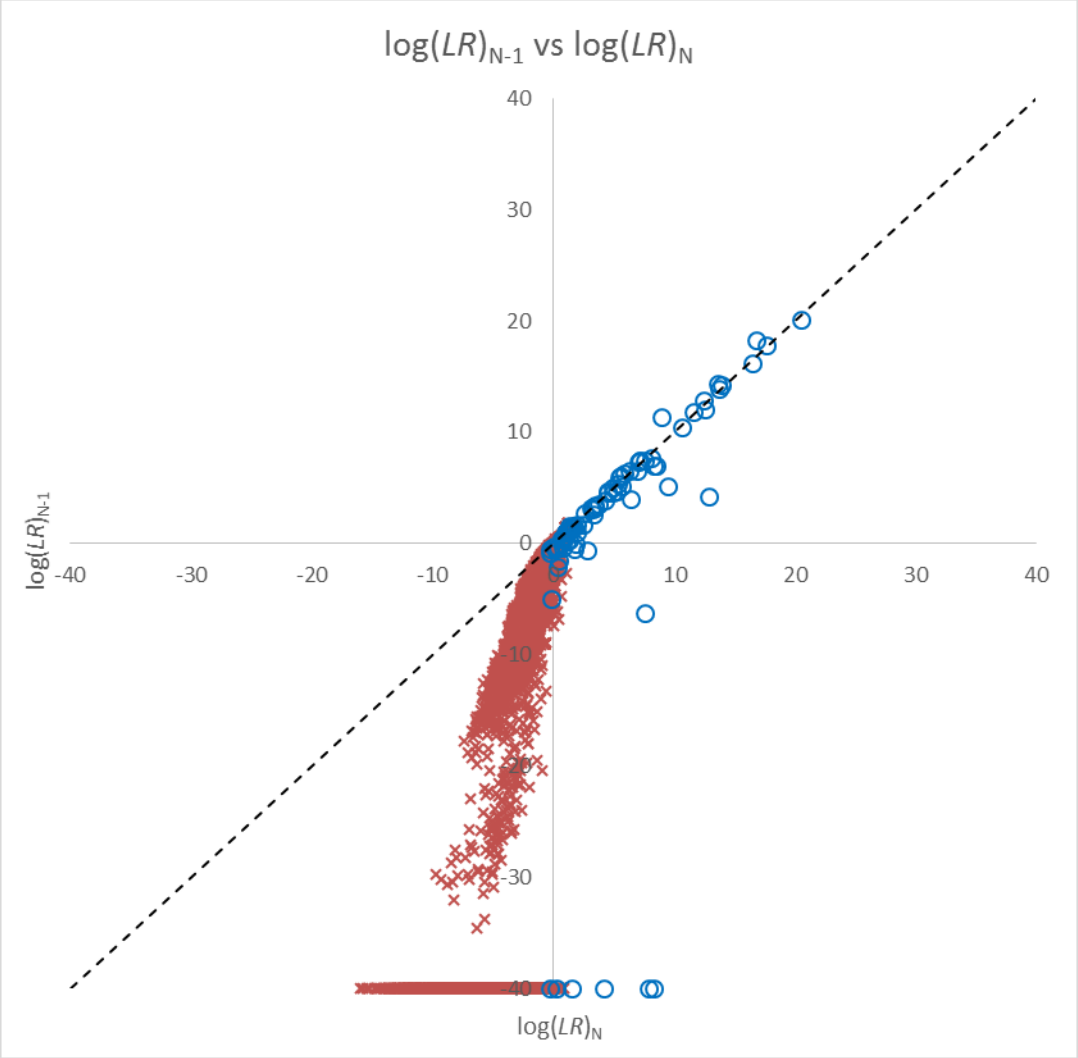
- B02_STRmix_M5-4_0.0625ng_C_0.0625ng_P2_GF_CBI_20s.hid *
- D04_STRmix_M5-4_0.0625ng_C_CBI_24s.hid *
- A02_STRmix_M5-4_0.125ng_C_0.125ng_P2_GF_CBI_20s.hid #
- D02_STRmix_M2-5_0.025ng_C_B1_CBI_24sec.hid
- B02_STRmix_M3-4_0.0375ng_C_N1_CBI GF.hid
- B02_STRmix_M4-5_C_0.1ng_D3_CBI_J6.hid
- C02_STRmix_M4-5_C_0.05ng_D3_CBI_J6.hid
- E03_STRmix_M4-5_0.05ng_C_CBI_24s.hid
- A03_M5-3_0.25ng_amp1_CBI_24s.hid
- A04_STRmix_M5-4_0.25ng_C_CBI_24s.hid
- A06_M5-3_0.25ng_amp2_CBI_24s.hid
- A08_M5-3_0.5ng_amp2_CBI_24s.hid
- C04_STRmix_M5-4_0.125ng_C_CBI_24s.hid
- H01_STRmix_M5-4_0.25ng_C_0.25ng_P2_GF_CBI_20s.hid
- H02_M5-2_0.25ng_amp1_CBI_24s.hid
- H02_M5-2_0.5ng_amp1_CBI_24s.hid
- H05_M5-2_0.25ng_amp2_CBI_24s.hid
- H07_M5-2_0.5ng_amp2_CBI_24s.hid

Note: *= 5 person mixture assigned as a 2 person mixture. # = 5 person mixture assigned as a 3 person mixture. For the purposes of this exercise these mixtures were re-run as 3 person and 4 person, respectively rather than experimental design 5 person mixtures. The effects shown are likely to be magnified if ran according to the true (5) number of contributors.

These were all under-assigned mainly due to the smaller/trace contributors dropping out.

Similar to above, the *LR* for the N known contributors and the 517-N known non-contributors (as used for the specificity and sensitivity studies, Section D) were calculated. The log(*LR*) was compared for the known contributors and known non-contributors under the assumption of *N* and *N*-1 contributors (the section D runs). A plot of log(*LR*) of apparent N (which is N-1 here) and experimental design N (with exception of those described above) is provided in Figure 11.

Figure 11: log(*LR*) when apparent N is under-assigned (N-1) versus log(*LR*) for N for the known and non-contributors (where H_p true are in blue circles and H_d true in red crosses)



The plot above shows that when an analyst under-estimates the number of contributors you can get false exclusions of true contributors. Here we see six *LR*s to known contributors sitting at an *LR*=0 (log(*LR*)=-40). By under-assigning the number of contributors to a mixture, STRmix™ may not propose genotype combinations aligning to the true minor or trace contributor. As an example the greatest difference is shown on the *y*-axis, where the known contributor (R4) to a five-person mixture [H02_M5-2_0.5ng_amp1_CBI_24s.hid] shifted from an *LR* of approx. 10^8 (log(*LR)* of 8.35205), when ran using N to a *LR* of 0 (log(*LR*) of -40) with N-1.

Under-estimating the number of contributors results in lower *LR*s (or greater support for exclusion) for H_d true comparisons as STRmix™ is not having to explain any additional trace components to the mixture as potentially allelic. Note that STRmix™ will not run if there are too many peaks present that cannot be explained using stutter modelling, or drop-in and can only be explained via an extra contributor being present in the mixture.

To summarize, broadly speaking, under-estimating the number of contributors can result in false exclusions of true contributors.

Section G: Drop-in

This section covers the following standard:

4.1.8. Allele drop-in

Observed drop-in rates at the CBI Laboratories have been modelled and the appropriate parameters based on this modelling are in place within STRmix™. To test these settings three experiments were undertaken.

In the first experiment, a realistically sized drop-in peak was artificially added to a *high template* single source STRmix™ input file (sample 4_F-C1 [IMR-90]).

The input file was edited at D18 to contain the following 14 peak:

Locus	Allele	Height	Size
D18S51	14	200	289.7
D18S51	16	1490	297.74
D18S51	17	18945	301.68
D18S51	18	105	305.72

Note: The drop-in cap at the CBI laboratories is 250 RFU.

The edited profile was then interpreted in STRmix™ as a single source profile. As expected STRmix™ completely modelled the additional peak as drop-in because it could not realistically pair with the high template 17 allele (>10,000 RFU) and the peak was below the drop-in cap. The resolved genotype was 17,17 with 100% weight.

The resulting *LR* was the same as that obtained from the original input (without the drop-in added).

This 14 drop-in peak is written to the genotype probability txt output as drop-in. This can been seen below, where the first two numbers are the accepted genotype [17,17], the third is the allele determined to be drop-in [14] and then the weight (1.0 = 100%):

```
Locus 9 (D18S51)
17.0  17.0  14    1.0
```

In the second experiment, a realistically sized (smaller than the drop-in cap) drop-in peak was artificially added to a *low template* single source STRmix™ input file (sample 4_F-H1 [IMR-90]).

The input file was edited at D18 to contain the following 14 peak:

Locus	Allele	Height	Size
D18S51	14	200	289.7
D18S51	17	414	301.75

The edited profile was interpreted in STRmix™ as a single source profile. As expected STRmix™ modelled the additional peak as drop-in and as part of the allelic component of the profile as it was of a more comparable

height to the other low template allele at that locus (<1000 RFU). STRmix™ favours the allelic component due in part to the rarity of drop-in events.

Below is a copy of the STRmix™ results output for the D18S51 locus (taken from the GenotypePdf1T text file):

Locus 9 (D18S51)		Dropin	Weight
14.0	17.0		0.997
17.0	17.0	14	0.003
-1.0	17.0	14	<<0.001

The resulting *LR* is less than obtained from running the original input as the weight is now spread over other possible genotype combinations.

In the third experiment, a 'drop-in' allele was added to the input used in the second experiment above, again at D18S51, outside CBI's parameters (i.e. greater than the 250 RFU drop in cap).

The input file was as follows at D18S51:

Locus	Allele	Height	Size
D18S51	14	300	289.7
D18S51	17	414	301.75

As expected, assuming a single source STRmix™ put 100% weight on a 14,17 genotype. This resulted in an exclusionary LR at this locus and hence and overall LR of 0.

Had a peak of this nature been added to a heterozygote locus in a single source sample then STRmix would not progress the interpretation as the profile could no longer be explained by one contributor.

Finally, a drop-in allele, 10 at TH01, (within CBI's drop-in parameters) was added to a 4:1 ratio two person mixture at a locus with style AB:C (where the AB= 8, 9.3, is the major).

The input file was as follows at TH01:

Locus	Allele	Height	Size
TH01	6	1073	187.3
TH01	7	98	191.32
TH01	8	3529	195.43
TH01	8.3	54	198.48
TH01	9.3	4792	202.47
TH01	10	249	203.5

This did not change the deconvolution of the major, which remained [8, 9.3] and the genotype combinations considered for the minor now included [6, 9.3], [6, 8], [6, 10] & [6, 6] as expected, with the appropriate weights.

Section H: Forward and reverse stutter

This section covers the following standard:

4.1.9. Forward and reverse stutter

STRmix™ implements a 'per allele' back stutter model. This is alternatively based on the longest uninterrupted stretch (LUS) of common repeats in the allele or the allele designation itself. STRmix™ implements a 'per allele' forward stutter (*N*+1) model which can be set to a uniform locus for those loci where allele was not a strong explanatory variable for FSR. Stutter peak labels are retained at analysis and within the STRmix™ input file. The modelling of stutter peaks can be seen in the interpretation of single source profiles where stutter peaks are retained at interpretation. As part of the MCMC process they are considered as alleles in the genotype but those combinations are not accepted and therefore receive no weight. In mixed DNA profiles, where the minor contributor is of a similar height as the stutter peaks they start to be considered as minor alleles. This is as expected.

STRmix™ implements a 'per allele' forward stutter (*N*+1) model which can be set to a uniform locus model by setting the slope to 0 for those loci where allele was not a strong explanatory variable for FSR. The modelling of forward stutter peaks can be seen in the interpretation of profiles where stutter peaks are retained at interpretation. As part of the MCMC process they are considered as alleles in the genotype but those combinations are not accepted and therefore receive no weight. In mixed DNA profiles, where the minor contributor is of a similar height as the stutter peaks they start to be considered as minor alleles. This is as expected.

Section I: Intra-locus peak height

This section covers the following standard:

4.1.10. Intra-locus peak height variance

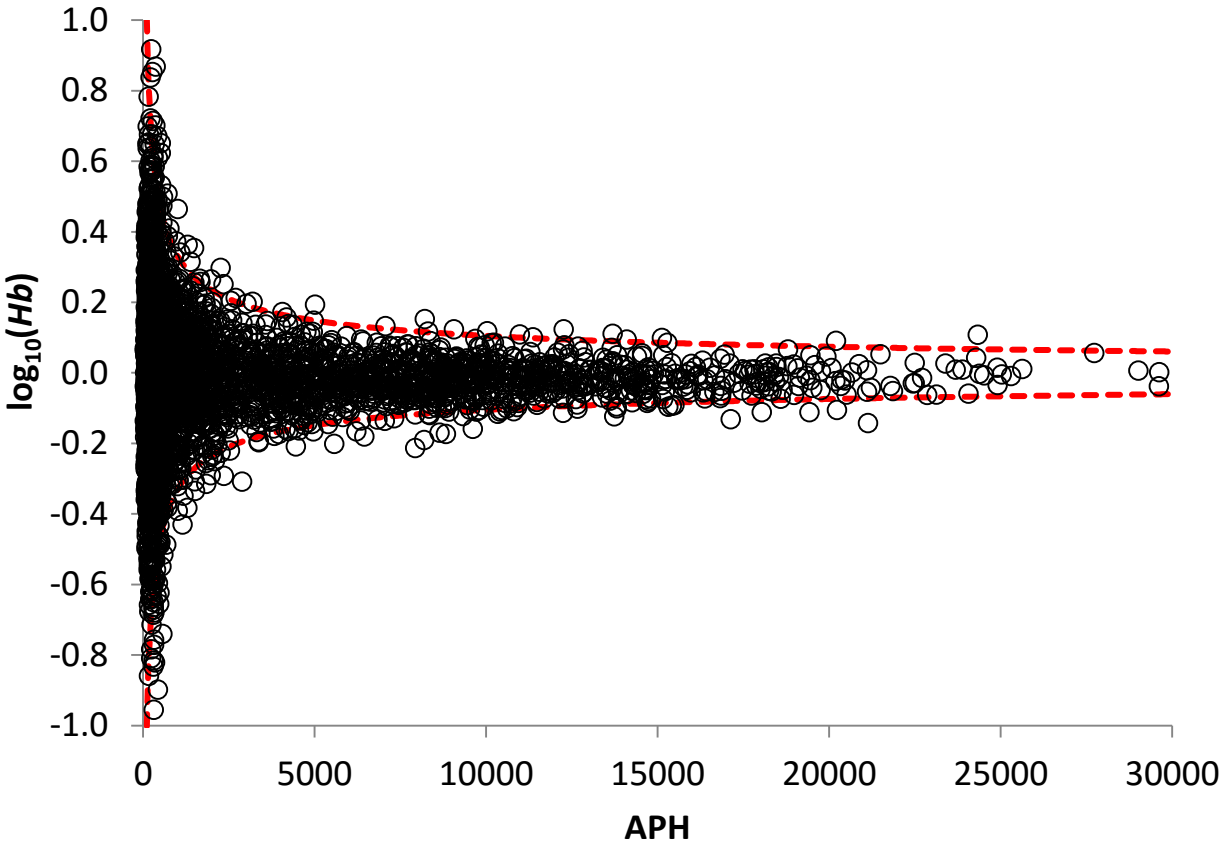
STRmix™ models the variability of single peaks. The variance of this model is determined by directly modelling laboratory data. This is undertaken within STRmix™ using the Model Maker (MM) function. Traditionally we investigate heterozygote balance (*Hb*), which can be thought of as the variability of two alleles at a heterozygous locus. The performance of Model Maker is checked by plotting the bounds informed by the Model Maker results (refer to CBI STRmix™ parameters GF_3500 V2.5 FINAL report for further details).

In single source profiles, variability in *Hb* reduces as the average peak height (APH) at a locus increases. The variance of *Hb* can be used as a proxy for the variance of individual peaks. This allows an approximate comparison between the variance from the STRmix™ MCMC approach and a readily determined variable from empirical data (that being *Hb*).

The plot of log(*Hb*) versus APH (the black circles) for the single source dataset used in Model Maker and the expected 95% bounds (plotted as dashed red lines) calculated at $\pm\sqrt{2}\times 1.96\times\sqrt{\frac{c^2}{APH}}$ where c^2 = 14.303 (the 75th percentile from the prior gamma distribution from the CBI combined dataset) is provided in Figure 12.

Under our assumption of a normal distribution we would expect ~95% of data points to fall within +/- 2 standard deviations (95% bounds) of the mean. The 95% bounds encapsulate sufficient data as demonstrated in the graph (coverage = 96.2%) demonstrating that the values for variance appear sufficiently optimised (Figure 12). This plot is a useful check of the Model Maker output and demonstrates the model performs as we'd expect for intra-locus peak heights.

Figure 12: Log(*Hb*) versus APH for the single source profiles used in MM at the CBI laboratories



Section J: Inter-Locus peak heights

This section covers the following standard:

4.1.11. Inter-locus peak height variance

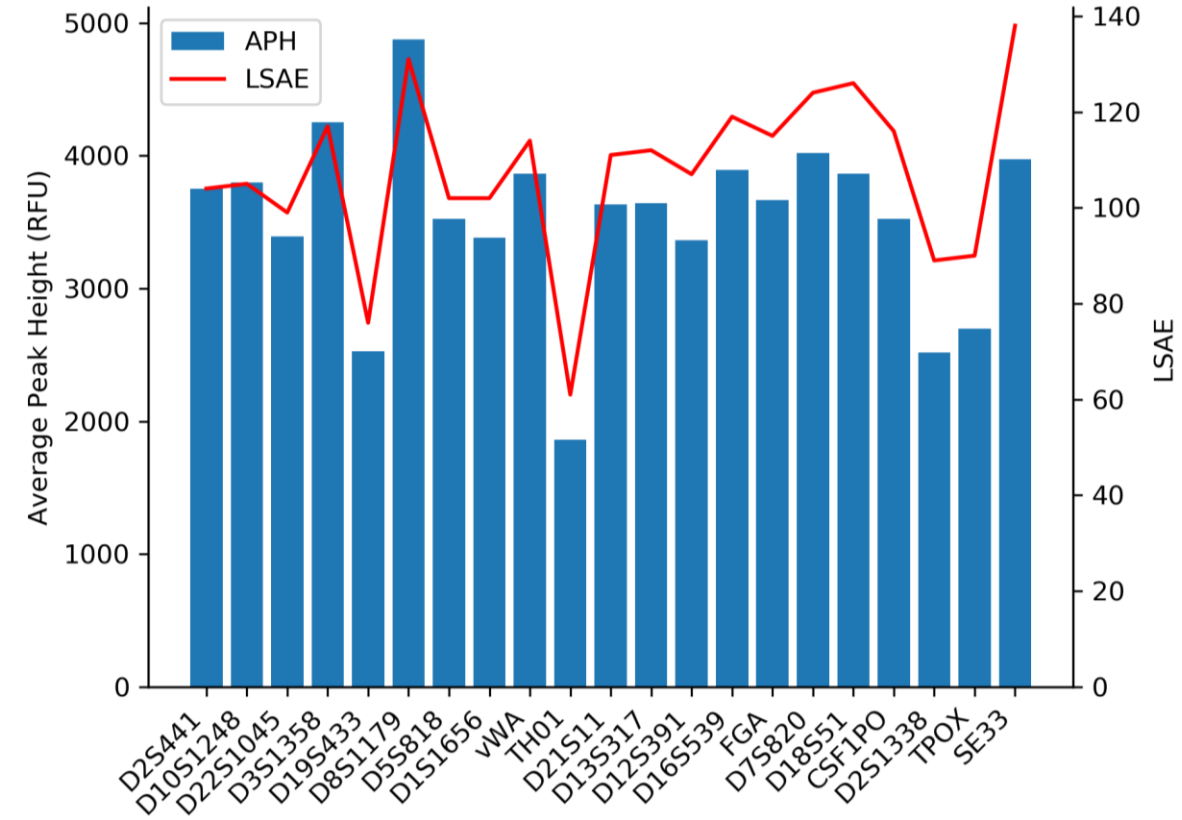
4.1.7.3 Inhibition

4.1.7.2. DNA degradation

Inter-locus peak variance is modelled in STRmix™ using locus specific amplification efficiencies (LSAE). The LSAE model reflects the observation that even after template DNA amount, degradation and variation in peak height within loci are modelled, the peak heights between loci are still more variable than predicted. The variance of this model is determined by directly modelling laboratory data and is one of the outputs from

Model Maker. However, individual LSAE values for each STRmix™ interpretation appear within the results. We can demonstrate the relationship of LSAE values to average peak heights (APH) via a simple plot. The LSAE values should mimic the average peak height of the locus. This is demonstrated by running one single source GlobalFiler™ profile (profile 4_F-D1 [IMR-90]) through STRmix™ and plotting the resultant LSAE values and the APH. This can be seen in Figure 13 where the x axis represents the different loci by increasing molecular weight across the profile.

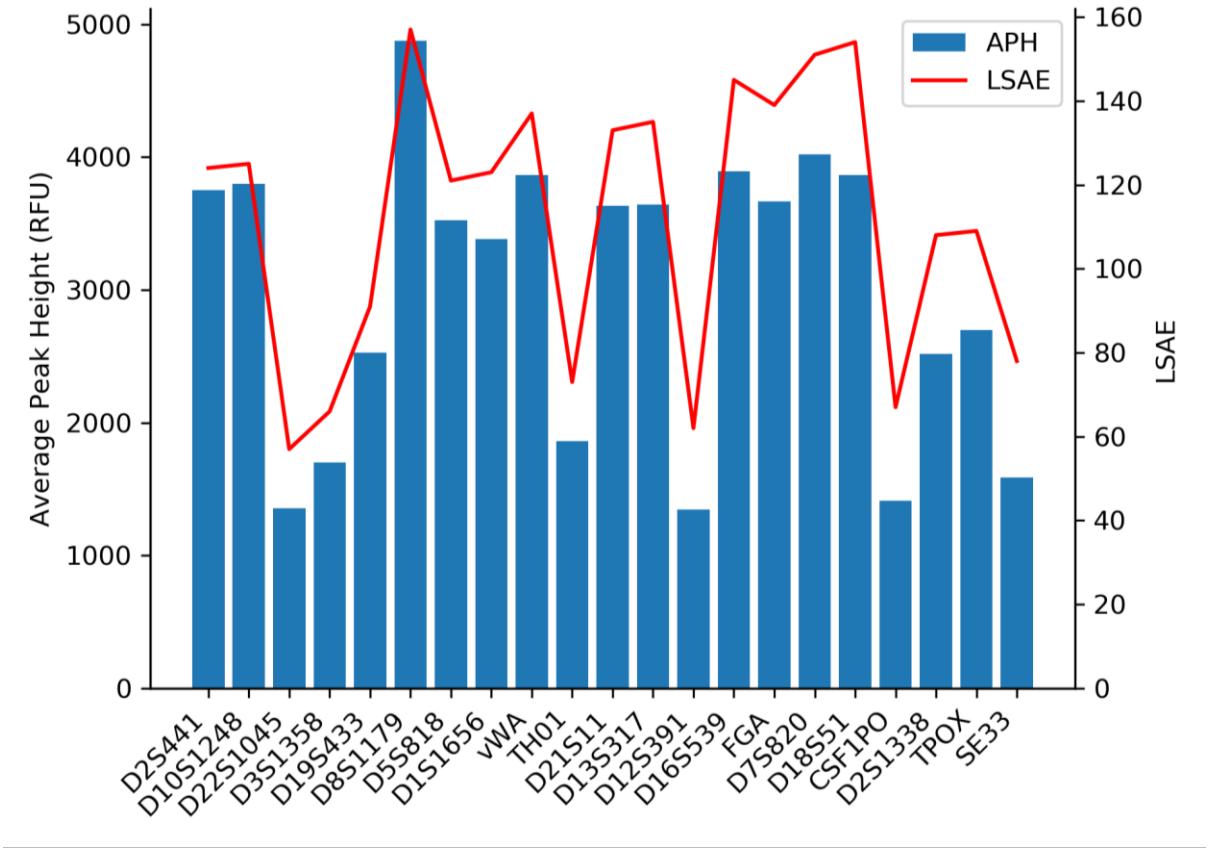
Figure 13: Plot of APH and LSAE value for each locus for a single source GlobalFiler™ profile (4_F-D1 [IMR-90])



Inspection of Figure 13 demonstrates the expected relationship between APH and LSAE.

The same single source GlobalFiler™ input file was then artificially inhibited, by amending the peak heights across 5 loci. Listed here by increasing molecular weight, these 5 loci were D22S1045, D3S1358, D12S391, CSF1PO & SE33. All the heights at these loci were reduced by 60% on the input file and the remainder were unedited. The resultant impact on LSAE and APH can be seen in Figure 14.

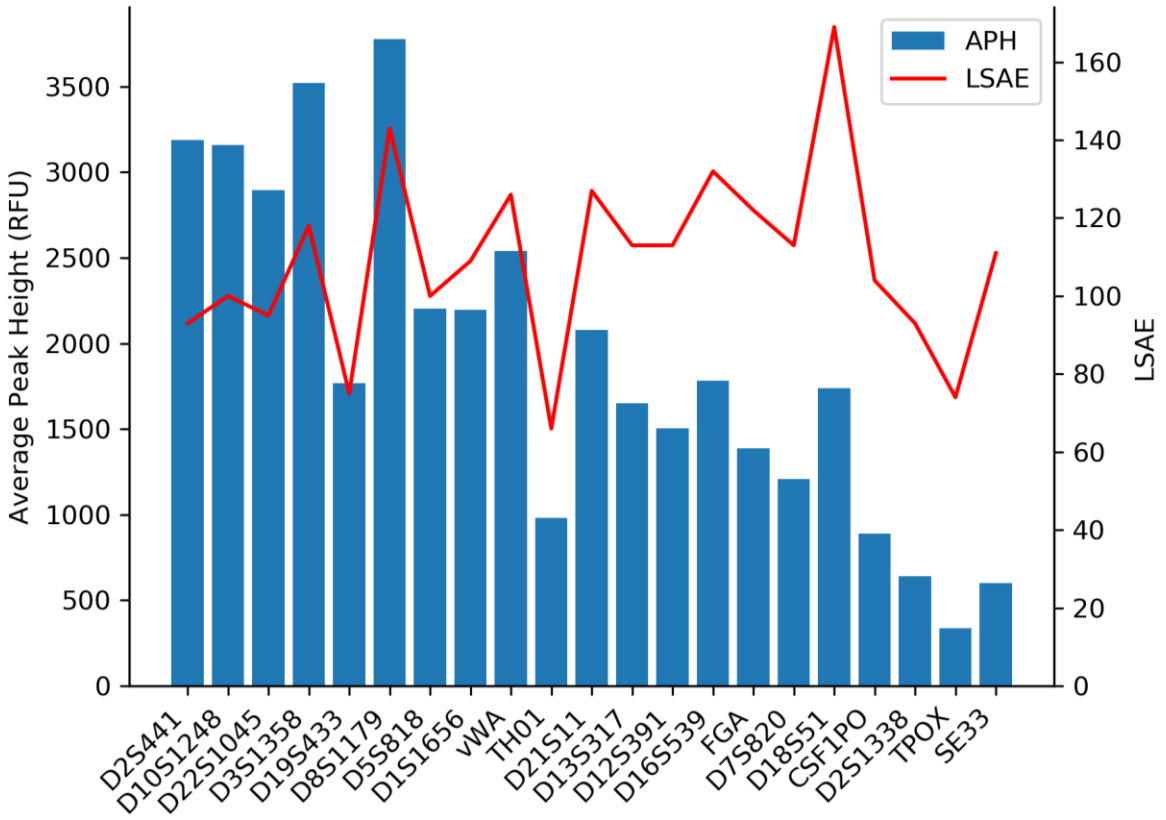
Figure 14: Plot of APH and LSAE value for each locus for a single source GlobalFiler™ profile (4_F-D1 [IMR-90]) with inhibition applied



Inspection of Figure 14 demonstrates the expected relationship between APH and LSAE holds with a certain level of inhibition.

The same profile was artificially degraded in Microsoft™ Excel by amending the peak heights in a stepwise fashion from low to high molecular weight. The highest molecular weight loci were degraded by up to 90%, the lower molecular weight loci by 5%. A plot of APH and LSAE for the artificially degraded profile is given in Figure 15 below. The general degradation within the profile can be seen by the downwards slope of the APH. LSAE values are independent of this degradation and the individual locus efficiencies can be seen by the LSAE values. This is most obvious at TPOX and SE33.

Figure 15: APH and LSAE for artificially degraded profile



Inspection of Figure 15 demonstrates the expected relationship between APH and LSAE holds even with degraded samples. The fact that degradation is also taken in to account counters the differences seen above between APH and the LSAE.

The average of the post burn-in degradation values for the original (un-degraded) profile was 2.760 RFU/bp. The same parameter was 14.112 rfu/bp for the artificially degraded profile.

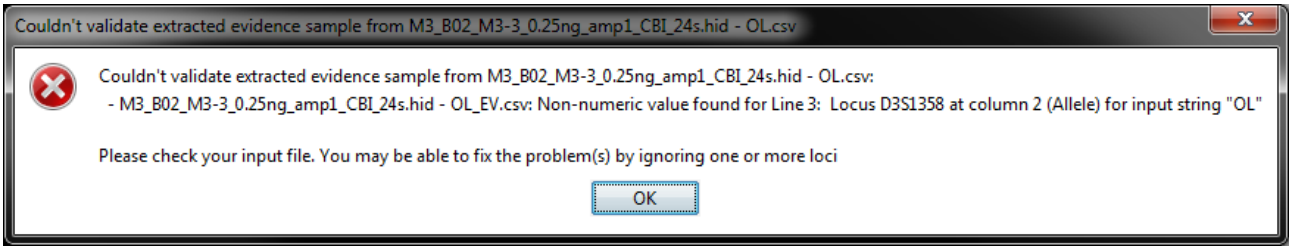
Section K: Challenge testing

This section covers the following standard:

> 4.1.14. Additional challenge testing (e.g., the inclusion of non-allelic peaks such as bleedthrough and spikes in the typing results)

STRmix™ requires that only numeric values are retained within the input file. Any values that are not numeric (such as OL alleles not removed at analysis) will cause STRmix™ to halt the interpretation. A sample (B02_M3-3_0.25_amp1_CBI_24s.hid) was manually edited to include an OL peak on an input file and an attempt was made to run this in STRmix™. The following error message was obtained when the file was used as the input evidence file:

Figure 16: Error message from STRmix™ when an OL peak is present in the input file.



STRmix™ will not allow you to progress past the input file stage until the issue is resolved.

The presence of a non-allelic peak (or peaks) that has sized within an allelic bin position and is retained within the input file can cause a number of results depending on the scenario. These include:

- An exclusionary *LR*. If the artifact is modelled as having originated from the person of interest (for example if the peak is of a similar height to the alleles corresponding to the person of interest in a single source or mixed DNA profile) this may result in an exclusion.
    - A single source sample (4_F-I1) was manually edited to include an additional 11 allele peak at D16S539 of 86 RFU (see screenshot below). The person of interest reference at D16S539 is [10,13], however the 13 allele had dropped out of the original profile. This resulted in STRmix™ modelling the genotypes at this locus as [10,11] with a weight of 1.00, instead of a [10,Q] with a weight of 0.2085 (seen in the original unedited analysis), and hence produced an exclusionary *LR* to the person of interest at D16S539.

Locus	Allele	Height	Size
D3S1358	14	268	117.47
vWA	16	144	177.03
vWA	19	167	189.22
D16S539	10	113	248.04
D16S539	11	86	252.2

- No effect. If drop-in is observed within a laboratory, the artifact may be modelled as a drop-in peak if it less than the drop-in height threshold.
    - A strong single source sample (4_F-E1) was manually edited to include a drop-in peak at D3S1358 allele 17 with a peak height of 100 RFU (see screenshot below). The point/sub source *LR* calculated to the person of interest was the same as the *LR* calculated to the single source profile without the artificial drop-in.

Locus	Allele	Height	Size
D3S1358	13	65	113.43
D3S1358	14	1257	117.47
D3S1358	15	1780	121.28
D3S1358	17	100	129.3

    - Within the GenotypePdf1T text file [4_F-E1_EV – Dropin_GenotypePdf1T.txt] within the results folder you can see the 14, 15 alleles designated as allelic with the 17 put in a separate column as drop-in (screenshot below).

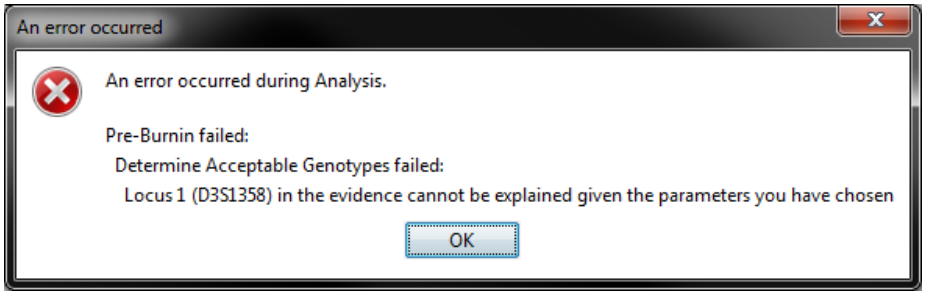
```
Locus 1 (D3S1358)
14.0  15.0  17    1.0
```

- Failure to interpret. If an artifact within an allelic bin is retained in a profile it may artificially increase the minimum number of contributors within the profile. For example an artifact at a heterozygous locus in a single source profile (not modelled as stutter or drop-in) will increase the minimum number of contributors by one. STRmix™ will not proceed assuming only one contributor.
    - A single source sample (4_F-E1) was manually edited (at D3S1358) to include 3 (allelic) peaks at a locus, all above the drop-in cap (see screen shot below).

Locus	Allele	Height	Size
D3S1358	13	65	113.43
D3S1358	14	1257	117.47
D3S1358	15	1780	121.28
D3S1358	17	1000	129.3

    - Attempts were made to run this in STRmix™ as a single source and the following error message was obtained in Figure 17.

Figure 17: Error message from STRmix™ when there are more alleles present than can be described by the number of contributors set.



Each of these expected outcomes was demonstrated by editing a three-person mixture or a single source input file and calculating an *LR* within STRmix™.

Section L: Casework profiles

This section covers the following standards:

4.2. Laboratories with existing interpretation procedures should compare the results of probabilistic genotyping and of manual interpretation of the same data, notwithstanding the fact that probabilistic genotyping is inherently different from and not directly comparable to binary interpretation. The weights of evidence 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.

4.1.7. Partial profiles, to include the following:

4.1.7.2. DNA degradation

4.1.7.3. Inhibition

Previously interpreted GlobalFiler™ profiles from validation studies and adjudicated cases were re-examined in STRmix™. Approximately 21 profiles covering a range of profile types, including a sample exhibiting degradation and inhibition, were interpreted in STRmix™ where the person of interest (POI) was previously considered to be included, excluded, or inconclusive using our traditional interpretation methods (i.e. RMP/CPI statistical analysis). Table 10 summarizes the results from the initial interpretation and the STRmix™ interpretation to include multiple LR propositions and values (99% 1-sided Lower HPD, Unified, and Stratified LRs).

Table 10: Chart comparing profiles calculated in STRmix™ versus prior qualitative assessment of the profile

Item Description	Initial Interpretation	Statistical Value Reported	STRmix™ Results - 99% 1-Sided Lower HPD Interval (lowest)	STRmix™ Results - Unified LR US population	STRmix™ Results - Stratified/Unified LR US population proportions
Apple iPhone swabs	mixture of 4 with suspect 1 and suspect 2 references; inconclusive due to the number of contributors	N/A	suspect 1: LR: 1.32E6; suspect 2 excluded	suspect 1 LR: 1.16E6; suspect 2 excluded	suspect 1 Stratified ULR: 6.33E7 suspect 1 Stratified TLR: 3.14E8 suspect 2 excluded
swabs from leash end	mixture of 3 with victim, suspect 1, and suspect 2 references; partial major matches suspect 2; minor inconclusive due to limited genetic information	RMP: 1 in 23 nonillion (10^30)	victim LR: 7.99E1 suspect 2 LR: 3.59E28 combined victim and suspect 2 LR: 1.16E31; suspect 1 excluded	victim LR: 8.45E1 suspect 2 LR: 1.49E17 suspect 1 excluded	victim Stratified ULR: 2.46E2 victim Stratified TLR: 2.46E2 suspect 2 Stratified ULR: 2.06E17 suspect 2 Stratified TLR: 9.60E28 combined victim and suspect 2 Stratified TLR: 1.02E32 suspect 1 excluded

external genitalia swabs	unresolved mixture of 2 (intimate) with victim, suspect 1, and suspect 2 references; victim and suspect 2 not excluded	CPI: 1 in 10 trillion	victim LR: N/A suspect 2 LR: 7.61E27; suspect 1 excluded	victim LR: N/A suspect 2 LR: 7.11E16 suspect 1 excluded	victim LR: N/A suspect 2 Stratified ULR: 9.4E16 suspect 2 Stratified TLR: 3.10E28 suspect 1 excluded
cutting from underwear	mixture of 3 with victim reference; inconclusive due to limited genetic information	N/A	victim LR: 9.28E17	victim LR: 1.01E15	victim Stratified ULR: 1.05E15 victim Stratified TLR: 1.14E18
penile swabs	mixture of 2 (intimate) with victim and suspect references; deduced foreign matches victim	RMP: 1 in 1.4 sextillion (10^21)	victim LR: 3.63E24	victim LR: 5.47E15	victim Stratified ULR: 6.23E15 victim Stratified TLR: 4.94E24
gas can handles	mixture of 3 with victim, suspect, and elimination references; major matches elimination; minor inconclusive due to limited genetic information	RMP: 1 in 5.7 trillion	victim LR: 9.85E9 elimination LR: 4.99E24 combined victim and elimination LR: 1.86E35 suspect excluded	victim LR: 1.03E10 elimination LR: 5.40E16 suspect excluded	victim Stratified ULR: 1.54E10 victim Stratified TLR: 1.56E10 elimination Strat ULR: 6.08E16 elimination Strat TLR: 6.55E24 comb. victim and elimination Strat TLR: 5.24E35 suspect excluded
swabs of scissor blades	partial mixture of 3 with victim and suspect references; inconclusive due to limited genetic information	N/A	suspect LR: 3.06E17; victim excluded	suspect LR: 4.19E14 victim excluded	suspect stratified ULR: 4.54E14 suspect stratified TLR: 3.16E17 victim excluded
swabs of neck of tank top	mixture of 4 with victim, suspect, and elimination references; inconclusive due to the number of contributors	N/A	victim LR: 4.99E6 suspect LR: 2.28E13 combined victim and suspect LR: 2.68E22; elimination excluded	victim LR: 5.28E6 suspect LR: 7.42E12 elimination excluded	victim Stratified ULR: 6.21E6 victim Stratified TLR: 6.21E6 suspect Stratified ULR: 1.17E13 suspect Stratified TLR: 5.33E13 comb. victim and suspect Stratified TLR: 6.34E22 elimination excluded
cutting of M-vac filter of driver's seat	partial mixture of 4 with victim (expected contributor) and suspect references; inconclusive due to the number of contributors	N/A	suspect excluded	suspect excluded	suspect excluded

degraded and inhibited reference	partial single source	RMP: 1 in 6.0 septillion (10^24)	reference LR: 9.21E22	reference LR: 1.38E15	reference Stratified ULR: 3.53E15 reference Stratified TLR: 3.47E23
swabs of magazine of firearm	mixture of 4, inconclusive due to number of contributors	N/A	suspect 1 LR: 3.73E9; suspect 2 LR: 3.59E28; combined LR: 4.75E38	suspect 1 LR: 3.39E9 suspect 2 LR: 1.76E17	suspect 1 Stratified TLR: 3.71E9 suspect 1 Stratified ULR: 3.61E9 suspect 2 Stratified TLR: 3.72E28 suspect 2 Stratified ULR: 1.74E17 combined Stratified TLR 7.90E38
swabs of a knife handle	mixture of 3, inconclusive to victim and suspect 1, exclude elimination	N/A	suspect LR: 4.35E14 victim LR: 6.86E15 elimination excluded	suspect LR: 1.24E13 victim LR: 9.06E13	suspect Stratified TLR: 4.55E14 suspect Stratified ULR: 1.20E13 victim Stratified TLR: 7.41E15 victim Stratified ULR: 9.22E13 suspect victim combined Stratified TLR 5.49E41
swabs of shoe	mixture of 3, CPI not excluded victim and elimination, suspect excluded	CPI: 1 in 90	victim LR: 1.74E24 elimination LR: 1.95E22 suspect excluded conditioned on elimination, victim LR: 1.43E27	victim LR: 4.05E16 elimination LR: 1.55E16 conditioned on elimination, victim LR: 1.44E17	victim Stratified TLR: 1.71E24 victim Stratified ULR: 3.99 E16 elim Stratified TLR: 2.16E22 elim Stratified ULR: 1.49E16 conditioned on elim victim Stratified TLR: 1.97E27 Stratified ULR: 1.53E17
swab of stain on towel	mixture of 2, partial major match suspect, partial minor matches victim	RMP major: 1 in 4.6 octillion (10^27) RMP minor: 1 in 970 billion	suspect LR: 2.31E28 victim LR: 3.299E30 combined LR: 5.08E59 conditioned on victim, suspect LR: 5.11E28	suspect LR: 3.13E17 victim LR: 1.44E17 conditioned on victim, suspect LR: 3.17E17	suspect Stratified TLR: 2.61E28 suspect Stratified ULR: 3.11E17 victim Stratified TLR: 3.88E30 victim Stratified ULR: 1.45E17 combined Stratified TLR: 6.16E59 conditioned on victim, suspect Stratified TLR: 7.94E28 Stratified ULR: 3.41E17

swab of screwdriver handle	mixture of 2, victim and suspect included	CPI: 1 in 39 million	suspect LR: 2.64E14 victim LR: 1.03E16 combined LR: 1.05E45 conditioned on victim, suspect LR: 5.23E28	suspect LR: 2.37E11 victim LR: 8.32E10 conditioned on victim, suspect LR: 3.12E17	suspect Stratified TLR: 5.71E14 suspect Stratified ULR: 3.03E11 victim Stratified TLR: 1.06E16 victim Stratified ULR: 8.34E10 combined Stratified TLR: 1.55E45 conditioned on victim, suspect Stratified TLR: 6.60E28 Stratified ULR: 3.34E17
hair recovered from hand towel	partial inhibited/degraded single source matched suspect, victim excluded	RMP: 1 in 56 quintillion (10^18)	suspect LR: 4.92E20 victim excluded	suspect LR: 6.02E14	suspect Stratified TLR: 6.37E20 suspect Stratified ULR: 6.51E14
hair recovered from bath towel	degraded single source, matched suspect, victim excluded	RMP: 1 in 6.2 nonillion (10^30)	suspect LR: 5.09E28 victim excluded	suspect LR: 3.13E17	suspect Stratified TLR: 6.71E28 suspect Stratified ULR: 3.41E17
swabs of left hand palm	mixture of 3, victim expected, no conclusions for minor	N/A	conditioned on victim, suspect LR: 1.10E16 elimination excluded	conditioned on victim, suspect LR: 5.57E12	conditioned on victim, suspect Stratified TLR: 2.30E17 Stratified ULR: 7.49E12
anal swab	mixture of 2, victim expected, partial deduced matched suspect, excluded elimination	RMP: 1 in 28 septillion (10^24)	conditioned on victim, suspect LR: 5.21E26 elimination excluded	conditioned on victim, suspect LR: 1.08E17	conditioned on victim, suspect Stratified TLR: 3.55E27 Stratified ULR: 2.75E17
swabs of handgun grip/trigger/slide	mixture of 3, not suitable for statistical comparisons	N/A	suspect LR: 1.55E10	suspect LR: 1.58E10	suspect Stratified TLR: 1.81E10 suspect Stratified ULR: 1.78E10
swabs of handgun magazine	mixture of 3, major matched suspect, minor not suitable	RMP: 1 in octillion (10^27)	suspect LR: 3.28E25	suspect LR: 7.76E16	suspect Stratified TLR: 3.06E25 suspect Stratified ULR: 8.08E16

Based on these results, previous binary interpretation methods are generally consistent with STRmix™ analysis for inclusions and exclusions of references. Additionally, STRmix™ provides lower LR values for mixtures that were previously determined inconclusive (due to the number of contributors and/or potential allelic drop out throughout the profile) which reflects the limited amount of information available in the evidence profile(s). Note that as the LR values decreased, the differences between Unified and Stratified LR values for the same propositions were minimized. If multiple POIs were included in Hp, the LR value increased significantly.

These tests helped further explore the limits of the software for known and non-probative casework samples. The Stratified Total LR (Factor of N! enabled, 99% 1-sided lower HPD interval, MCMC uncertainty on) will be reported for the majority of CBI casework profiles as applicable.

Section M: Precision

This section covers the following standard:

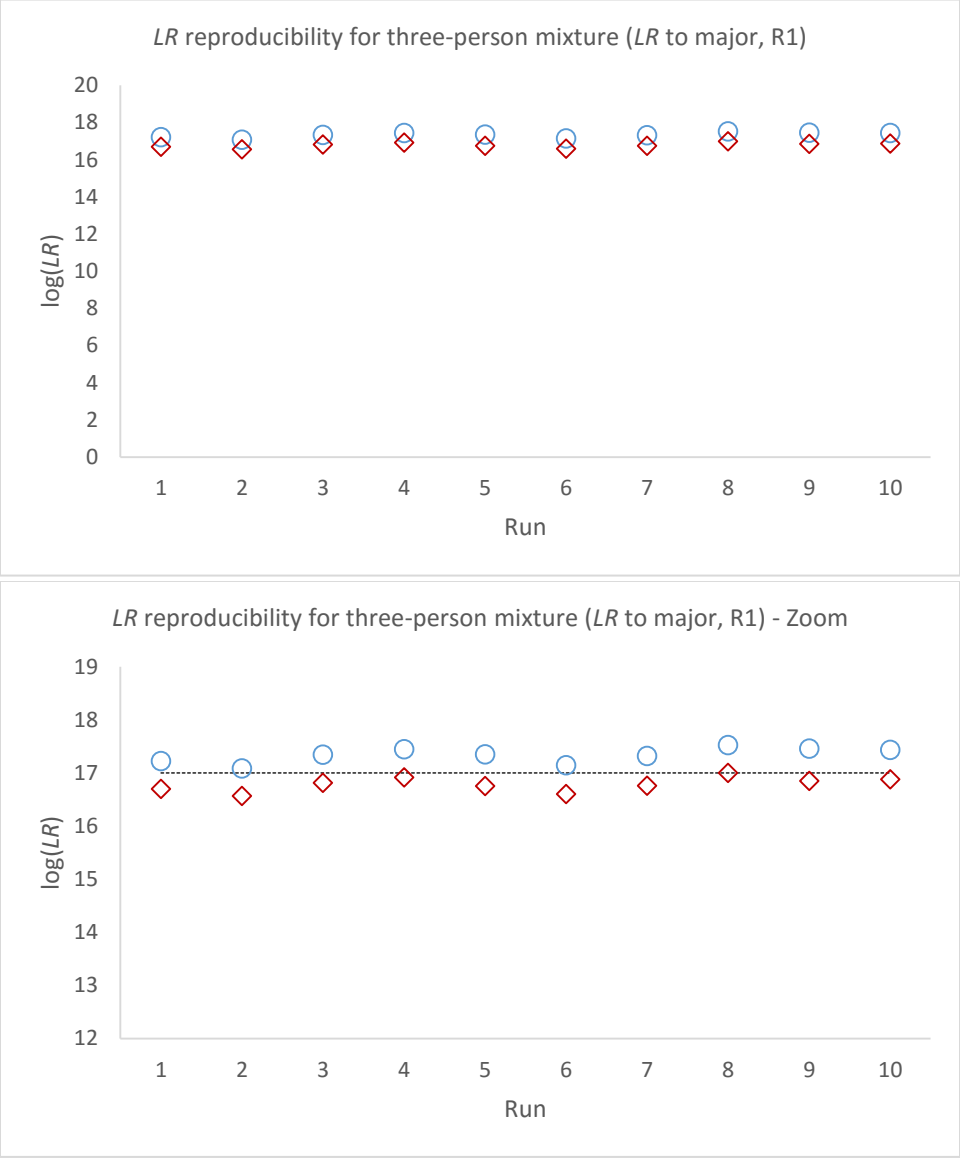
4.1.13. Sensitivity, specificity and *precision*, as described for Developmental Validation

Refer to section D above for details of sensitivity and specificity tests.

The MCMC process is used to generate the weights within STRmix™ for different genotype combinations. This is a sampling procedure and therefore the weights will vary slightly between each run (with the exception of high template, full single source profiles). The variability in *LR*s between replicate interpretations has previously been explored [10]. The MCMC process was shown to be a small source of variability compared with other lab variables including the PCR and CE process. The variability due to the size of the allele frequency database and the MCMC process is taken into account within STRmix™ V2.5 using the highest posterior density (HPD) method [9, 11, 12] (a type of confidence interval).

The extent of STRmix™ run variability was investigated by CBI laboratories by interpreting one of the mixed DNA profiles from Section D (B02_M3-3_0.25ng_amp1_CBI_24s.hid) [3:2:1 mixture], where there was ambiguity in the genotype combinations, ten times. A plot of log(*sub-source LR*) obtained from the major contributor (R3) for each replicate is given in Figure 18. The blue circles indicate the *sub-source LR* values and the red diamonds are the 99.0% 1-sided lower bound of the HPD.

Figure 18: Plot of replicate log(*LR*) demonstrating reproducibility of STRmix™ (upper pane) and zoom of *y*-axis (lower pain). The blue circles indicate the *sub-source LR* values and the red diamonds are the 99.0% 1-sided lower bound of the HPD. The dash line is the highest HPD value.



Inspection of Figure 18 shows that the *LR*s are very reproducible and that the 99.0% 1-sided lower bound of the HPD is below the point *LR* values.

Parameters within STRmix™ that affect run variability include the number of iterations and the RWSD (random walk standard deviation). The default number of iterations is set to 100,000 burn-in and 50,000 post burn-in accepts per chain. These will be suitable for many different types of profiles. Decreasing the number of iterations may mean that the MCMC chains within STRmix™ will not converge and more variability is expected. Increasing the number of iterations may mean convergence is achieved (if it hasn't already) with the trade-off of higher run times. A two-, three-, and four-person mixture was interpreted using three different sets of accepts, five times each, with *sub-source LRs* calculated to the known contributor of the second component of

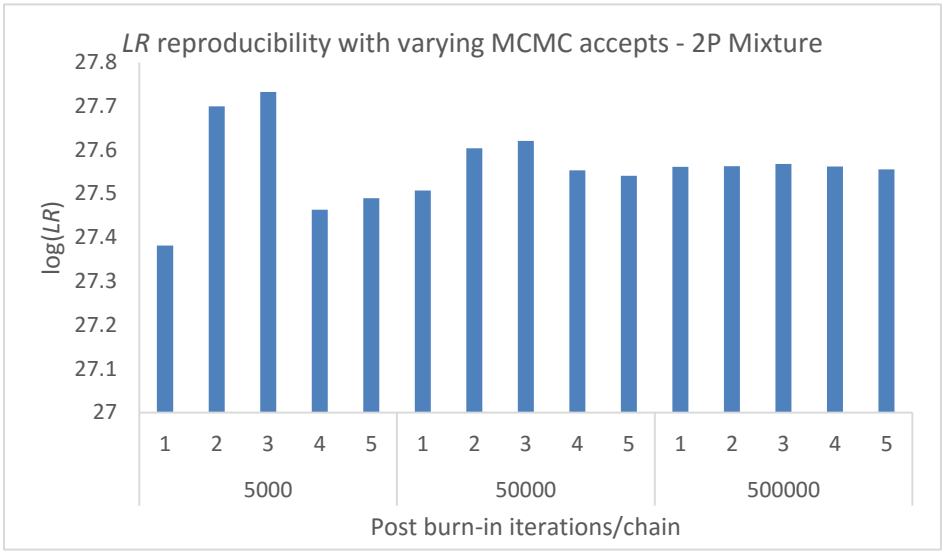
the DNA mixtures. A summary of the accepts settings are displayed in Table 11. A plot of log(*LR*) for each replicate is given in Figure 19.

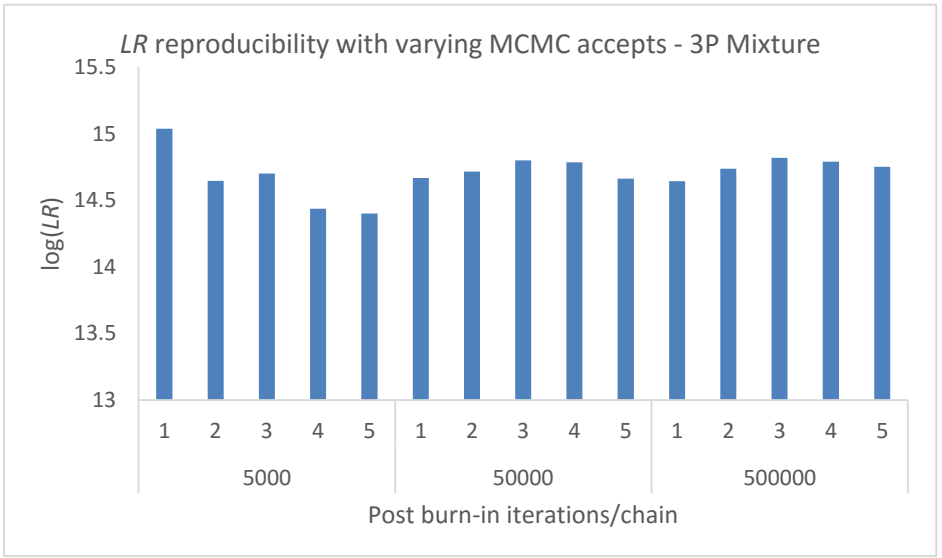
Table 11: Summary of accepts settings for each sample ran five times

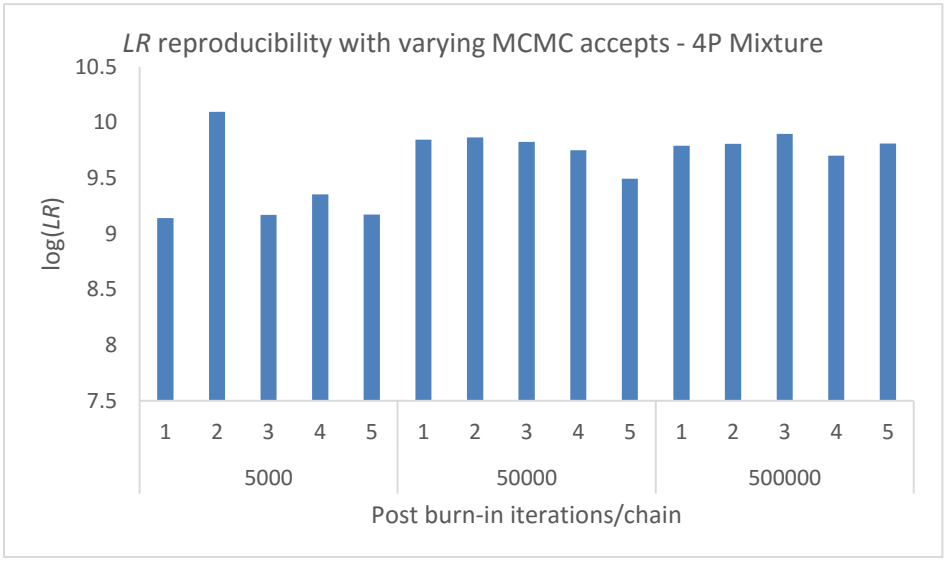
Sample	N	Chains	Burn-in Accepts/Chain	Post Burn-in Accepts/Chain
E02_M4-3_0.5ng_amp1_CBI_24s.hid (10:10:5:1)	4	8	10,000	5,000
			100,000	50,000
			1,000,000	500,000
B02_M3-3_0.25ng_amp1_CBI_24s.hid (3:2:1)	3	8	10,000	5,000
			100,000	50,000
			1,000,000	500,000
F06_M2-3_0.5ng_amp2_CBI_24s.hid (5:1)	2	8	10,000	5,000
			100,000	50,000
			1,000,000	500,000

Inspection of Figure 19 shows a reduction in run to run variability as we increase accepts. Even so there is typically less than one order of magnitude difference in the *LR*s. The STRmix™ developers propose that default settings of 100,000 burn-in and 50,000 burn-in per chain (using 8 chains) should be sufficient for most samples. However, the option remains to up this value as and when required.

Figure 19: Log(*LR*) to the contributor of the second component of a two-, three-, and four-person mixture interpreted five times in STRmix™ using different numbers of accepts







Conclusion

This document describes the CBI's internal validation activities for STRmix™ V2.5. It has been shown that it is suited for its intended use for the interpretation of GlobalFiler™ profiles generated from evidence samples at the CBI.

9/5/18

Sarah Miller, DNA Program Manager / Technical Leader

References

[1] J.-A. Bright, D. Taylor, C. McGovern, S. Cooper, L. Russell, D. Abarno, J. Buckleton, Developmental validation of STRmix™, expert software for the interpretation of forensic DNA profiles, Forensic Science International: Genetics 23 (2016) 226-239.

[2] Scientific Working Group on DNA Analysis Methods (SWGDAM), Guidelines for the Validation of Probabilistic Genotyping Systems, (2015).

[3] J.-A. Bright, I.W. Evett, D. Taylor, J.M. Curran, J. Buckleton, A series of recommended tests when validating probabilistic DNA profile interpretation software, Forensic Science International: Genetics 14 (2015) 125-131.

[4] D.J. Balding, R.A. Nichols, DNA profile match probability calculation: how to allow for population stratification, relatedness, database selection and single bands, Forensic Science International 64 (1994) 125-140.

[5] National Research Council II, National Research Council Committee on DNA Forensic Science, The Evaluation of Forensic DNA Evidence, National Academy Press, Washington, D.C., 1996.

[6] D. Taylor, Using continuous DNA interpretation methods to revisit likelihood ratio behaviour, Forensic Science International: Genetics 11 (2014) 144-153.

[7] T.R. Moretti, R.S. Just, S.C. Kehl, L.E. Willis, J.S. Buckleton, J.-A. Bright, D.A. Taylor, A.J. Onorato, Internal validation of STRmix™ for the interpretation of single source and mixed DNA profiles, Forensic Science International: Genetics 29 (2017) 126-144.

[8] J.-A. Bright, K.E. Stevenson, J.M. Curran, J.S. Buckleton, The variability in likelihood ratios due to different mechanisms, Forensic Science International: Genetics 14(0) (2015) 187-190.

[9] D. Taylor, J.A. Bright, J. Buckleton, J. Curran, An illustration of the effect of various sources of uncertainty on DNA likelihood ratio calculations, Forensic Science International: Genetics 11 (2014) 56-63.

[10] J.-A. Bright, K.E. Stevenson, J.M. Curran, J.S. Buckleton, The variability in likelihood ratios due to different mechanisms, Forensic Science International: Genetics 14 (2015) 187-190.

[11] C.M. Triggs, J.M. Curran, The sensitivity of the Bayesian HPD method to the choice of prior, Science & Justice 46(3) (2006) 169-178.

[12] J.M. Curran, J.S. Buckleton, An investigation into the performance of methods for adjusting for sampling uncertainty in DNA likelihood ratio calculations, Forensic Science International: Genetics 5(5) (2011) 512-516.

Appendix 1: List of papers that support STRmix™

The following is a list of papers that directly support STRmix™.

1. D. Taylor, J.-A. Bright and J.S. Buckleton, The interpretation of single source and mixed DNA profiles. Forensic Science International: Genetics, 2013 7(5): 516-528 **(Core maths paper)**
2. J.-A. Bright, D. Taylor, J.M. Curran and J.S. Buckleton, Developing allelic and stutter peak height models for a continuous method of DNA interpretation. Forensic Science International: Genetics, 2013. 7(2): 296-304 **(Core models paper)**
3. J.-A. Bright, D. Taylor, J.M. Curran and J.S. Buckleton, Degradation of forensic DNA profiles, Australian Journal of Forensic Sciences, 2013. 45(4): 445-449
4. D. Taylor. Using continuous DNA interpretation methods to revisit likelihood ratio behaviour. Forensic Science International: Genetics, 2014. 11: 144-153
5. J.-A. Bright, D. Taylor, J.M. Curran and J.S. Buckleton, Searching mixed DNA profiles directly against profile databases. Forensic Science International: Genetics, 2014. 9: 102-110
6. D. Taylor, J.-A. Bright, J.S. Buckleton, J. Curran, An illustration of the effect of various sources of uncertainty on DNA likelihood ratio calculations. Forensic Science International: Genetics, 2014. 11: 56–63
7. J.-A. Bright, J.M. Curran and J.S. Buckleton, The effect of the uncertainty in the number of contributors to mixed DNA profiles on profile interpretation. Forensic Science International: Genetics, 2014. 12: 208-214
8. J.-A. Bright, K.E. Stevenson, J.M. Curran and J.S. Buckleton, The variability in likelihood ratios due to different mechanisms. Forensic Science International: Genetics, 2015. 14:187-190
9. D .Taylor, J.-A. Bright and J.S. Buckleton, Considering relatives when assessing the evidential strength of mixed DNA profiles. Forensic Science International: Genetics, 2014. 13: 259-263
10. D. Taylor, J-A. Bright and J.S. Buckleton. Interpreting forensic DNA profiling evidence without specifying the number of contributors. Forensic Science International: Genetics, 2014. 13: 269-280

The following is a subset of other papers that support the theory within STRmix™:

1. J.-A. Bright, J.M. Curran. Investigation into stutter ratio variability between different laboratories. Forensic Science International: Genetics, 2014. 13: 79-81
2. C. Brookes, J.-A. Bright, S.A. Harbison, and J.S. Buckleton, Characterising stutter in forensic STR multiplexes. Forensic Science International: Genetics, 2012. 6(1): 58-63
3. H. Kelly, J.-A. Bright, J.M. Curran, and J.S. Buckleton Identifying and modelling the drivers of stutter in forensic DNA profiles. Australian Journal of Forensic Sciences, 2014. 46(2): 194-203
4. J.-A. Bright, S. Neville, J.M. Curran, and J.S. Buckleton. Variability of mixed DNA profiles separated on a 3130 and 3500 capillary electrophoresis instrument. Australian Journal of Forensic Sciences, 2014. 46(3): 304-312

5. J.-A. Bright, K.E. Stevenson, M.D. Coble, C.R. Hill, J.M. Curran, and J.S. Buckleton Bright, Characterising the STR locus D6S1043 and examination of its effect on stutter rates. Forensic Science International: Genetics, 2014. 8(1): p. 20-23.
6. D. Taylor, J.S. Buckleton. Do low template DNA profiles have useful quantitative data? Forensic Science International: Genetics, 2015. 16: 13-16.
7. Taylor D, Buckleton J. Do low template DNA profiles have useful quantitative data? Forensic Science International: Genetics. 2015;16:13-6.
8. Taylor D, Buckleton J, Bright J-A. Factors affecting peak height variability for short tandem repeat data. Forensic Science International: Genetics. 2016;21:126-33.

The following is a subset of other papers that support the validation and use of STRmix™:

1. J.-A. Bright, I.W. Evett, D. Taylor, J.M. Curran and J.S. Buckleton, A series of recommended tests when validating probabilistic DNA profile interpretation software. Forensic Science International: Genetics, 2015. 14: 125-131
2. T.W. Bille, S.M. Weitz, M.D. Coble, J.S. Buckleton, J.-A. Bright. Comparison of the performance of different models for the interpretation of low level mixed DNA profiles. ELECTROPHORESIS. 2014;35:3125-33.
3. S.J. Cooper, C.E. McGovern, J.-A. Bright, D. Taylor, J.S. Buckleton. Investigating a common approach to DNA profile interpretation using probabilistic software. Forensic Science International: Genetics, 2014. 16: 121-131.
4. Moretti TR, Just RS, Kehl SC, Willis LE, Buckleton JS, Bright J-A, et al. Internal validation of STRmix for the interpretation of single source and mixed DNA profiles. Forensic Science International: Genetics. 2017;29:126-44.
5. Bright J-A, Taylor D, McGovern CE, Cooper S, Russell L, Abarno D, et al. Developmental validation of STRmix™, expert software for the interpretation of forensic DNA profiles. Forensic Science International: Genetics. 2016;23:226-39.
6. Taylor D, Bright J-A, McGoven C, Hefford C, Kalafut T, Buckleton J. Validating multiplexes for use in conjunction with modern interpretation strategies. Forensic Science International: Genetics. 2016;20:6-19.

Appendix 2: Cross reference for document sections and SWGDAM recommendations

Standard	Text	Refer section
4.1	Test the system using representative data	Preamble
4.1.1	Specimens with known contributors	Preamble
4.1.2	Hypothesis testing with contributors and non-contributors	D
4.1.2.1	More than one set of hypotheses	E
4.1.3	Variable DNA typing conditions	Preamble
4.1.4	Allelic peak height, to include off-scale peaks	B
4.1.5	Single-source specimens	A
4.1.6	Mixed specimens	D
4.1.6.1	Various contributor ratios	D
4.1.6.2	Various total DNA template quantities	D
4.1.6.3	Various numbers of contributors	D
4.1.6.4	Both correct and incorrect number of contributors (i.e., over- and under-estimating)	F
4.1.6.5	Sharing of alleles among contributors	D
4.1.7	Partial profiles	D
4.1.7.1	Allele and locus drop-out	D
4.1.7.2	DNA degradation	L
4.1.7.3	Inhibition	L
4.1.8	Allele drop-in	G
4.1.9	Forward and reverse stutter	H
4.1.10	Intra-locus peak height variance	I
4.1.11	Inter-locus peak height variance	J
4.1.12	In-house parameters	Preamble
4.1.13	Sensitivity, specificity and precision	D and M
4.1.14	Additional challenge testing	K
4.2	Compare the results of probabilistic genotyping and of manual interpretation	L
4.2.1	Intuitive and consistent with expectations	L
4.2.1.1	Known specimens that are included based on non-probabilistic analyses would be expected to also be included based on probabilistic genotyping	L
4.2.1.2	Concordance of single-source specimens with high quality results	A
4.2.1.3	Generally, as the analyst's ability to deconvolute a complex mixture decreases, so does the weighting of a genotype set determined by the software	C

Appendix 3: Review of Secondary Diagnostics of Section D

This section reviews the secondary diagnostics for each of the 188 mixtures of unrelated individuals in Section D. Please note this includes the saturated profiles.

These diagnostics include the total number of iterations, log(likelihood), Gelman-Rubin convergence diagnostic, and the posterior mean allele and stutter variances. Secondary diagnostics are a useful guide to provide confidence the interpretation has progressed as expected. These are termed secondary, coming after the primary diagnostics which are the weights, mixture proportions and, if available, the *LR*s.

Individual secondary diagnostics may indicate whether a review is warranted, however analysts should not rely on these diagnostic alone. Further review of the other diagnostics, the primary diagnostics and the profile itself could indicate that STRmix™ is performing as expected.

Posterior Mean Allele Variance

This is calculated from the average of the accepted allele variance values from the 400,000 post burn-in accepts. Figure 20 shows the spread of the allele variance values from the STRmix™ outputs for each sample. The red dotted line represents the 50th percentile of the allele variance prior distribution, however the mode is also a useful point of reference. The green dotted line represents the 75th percentile of the allele variance prior distribution. It can also be helpful to visualise where the values sit compared to the prior distribution, determined during parameter setting and hence this is displayed in the lower pane. This is also displayed on the individual STRmix™ reports.

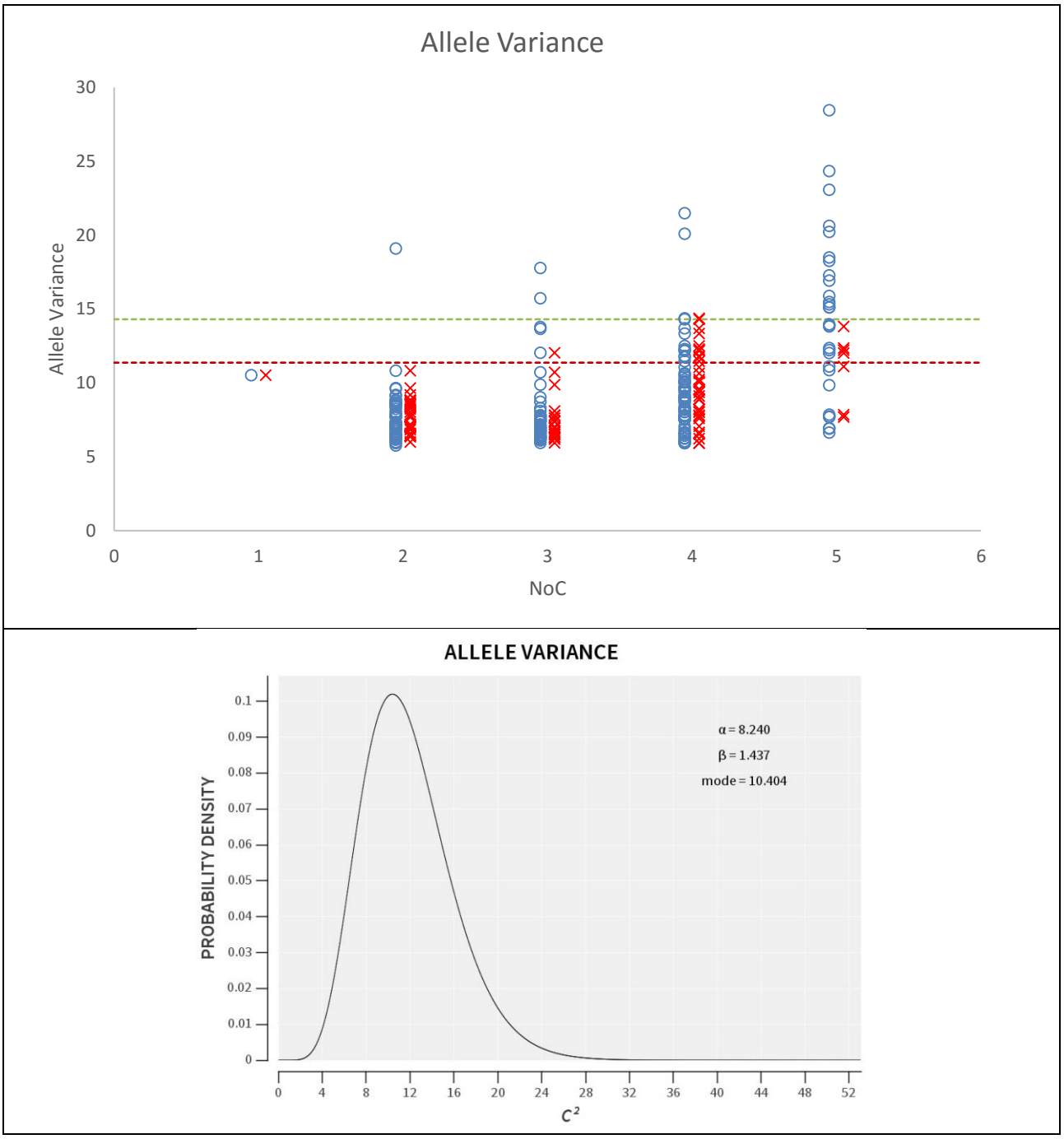
The prior distribution is modelled by a gamma distribution, Γ(8.24,1.437) and was determine by Model Maker. Ideally we want to see the posterior mean variance value on the STRmix™ output sitting in the main body of the prior distribution.

The values from all 188 mixtures are provided in Figure 20. However, as discussed some of the mixtures were saturated. Therefore, each plot has two groups of data, the blue circles represent all the runs and the red crosses limit those to samples with less than 1ng of template DNA.

Inspection of Figure 20 shows the majority of the runs gave values within the main body of the prior distribution. Some values for the 5 person mixtures have elevated values. Removing potentially saturated data improves the overall spread of these values.

Many of the two- and three- person profiles gave values below the mode or 50th percentile, suggesting little variance is required to describe the data. This may well be as expected. However, going forward it may be worthwhile monitoring these values and the primary diagnostics to ensure the interpretations progress as expected. STRmix™ has a setting (Var > mode) that will automatically reject iterations that propose an allele variance smaller than half the mode of the prior distribution (c^2-mode = 0.5 × 10.4 = 5.2). This means a lower limit is imposed based on the prior distribution.

Figure 20: Plot of the average posterior mean allele variance values for each DNA profile (upper pane). Blue circles show allele variance value for all 188 profiles, red crosses show allele values for data less than 1 ng of template. The red dashed line represents the 50th percentile values and the green dashed line represents the 75th percentile of the prior distribution modelled by a gamma distribution. CBI's prior gamma distribution is also provided (lower pane)



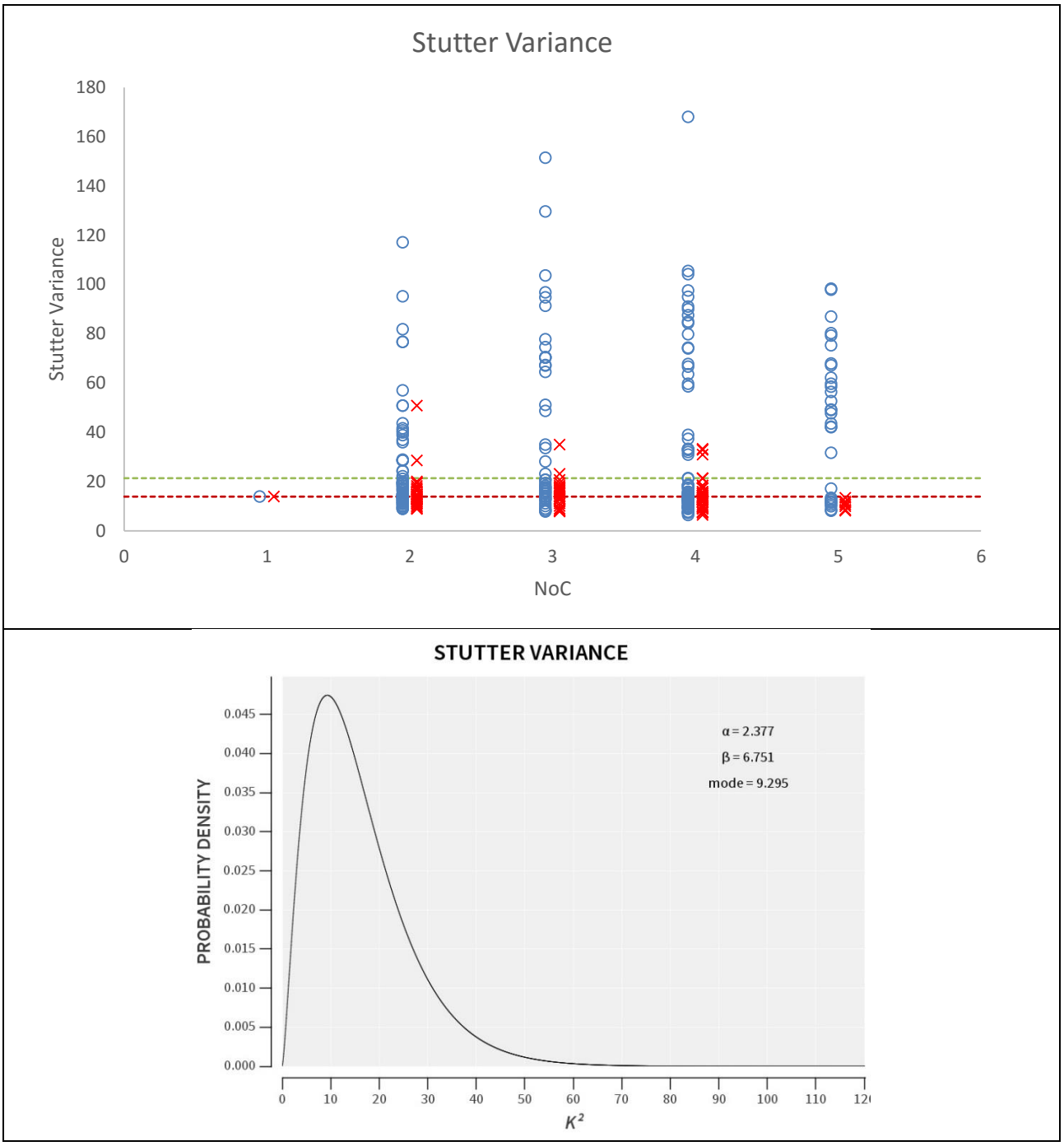
Posterior Mean Stutter Variance

In Figure 21, we show the stutter variance values from the STRmix™ outputs for each sample run. This is the average of the accepted stutter variance values from the 400,000 post burn-in accepts. The red dotted line represents the 50th percentile of the allele variance prior distribution, however the mode is also a useful point of reference. The green dotted line represents the 75th percentile of the allele variance prior distribution. It can also be helpful to consider where the values sit compared to the prior distribution, determined during parameter setting.

The prior distribution is modelled by a gamma distribution, Γ(2.377,6.751). Comparing these stutter variances to the prior distribution which is provided in the lower pane of Figure 21, it can be seen that the spread of all the stutter variance values (blue circles) include inflated values in comparison to the prior distribution. This is largely a result of the saturated profiles that were initially included; as stutters are no longer proportional to the parent peak height. This reinforces the recommendation, stated in Section D, that saturated profiles should be reprocessed where possible. Comparing this to the data set omitting samples with more than 1ng of template (red crosses) significantly improves the spread of these values.

In Section D we have demonstrated the effect of saturation on the primary diagnostics, *LR* and weights. However, it also affects the other secondary diagnostics.

Figure 21: Plot of the average posterior mean stutter variance values for each DNA profile (upper pane). Blue circles show stutter variance value for all 188 profiles, red crosses show stutter values for data less than 1 ng of template. The red dashed line represents the 50th percentile values and the green dashed line represents the 75th percentile of the prior distribution modelled by a gamma distribution. CBI's prior gamma distribution is also provided (lower pane)

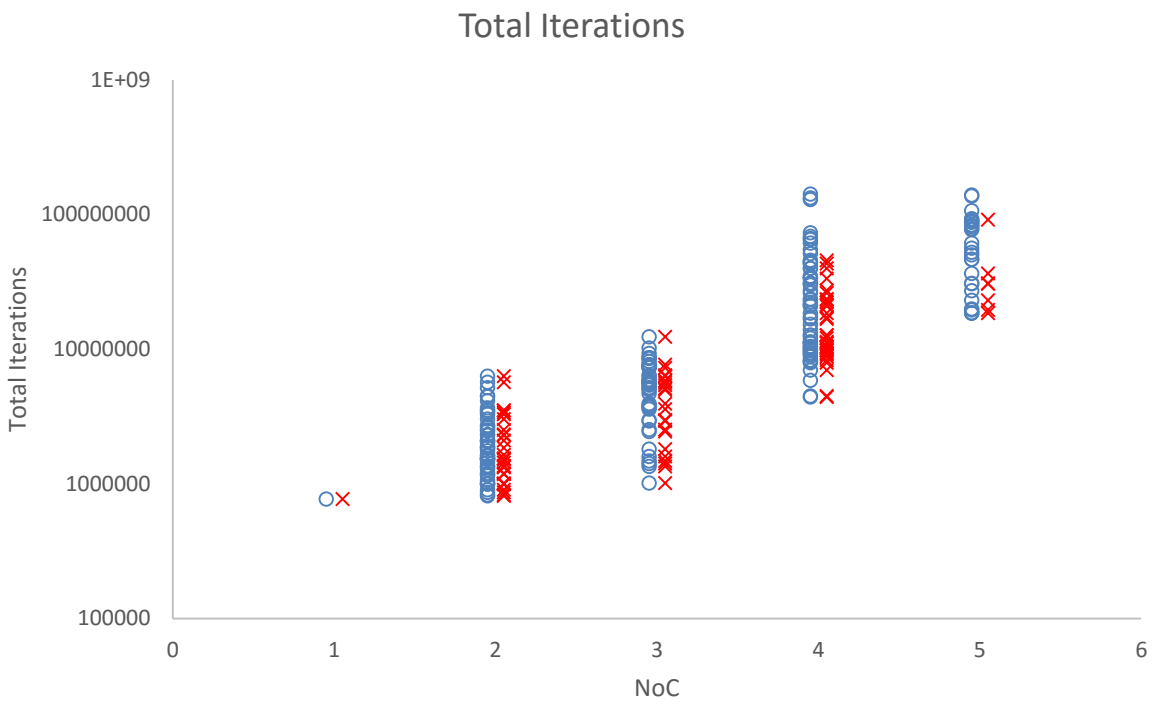


Total Number of Iterations

The total number of iterations simply shows the number of iterations required for 400,000 accepts to be reached (50,000 accepts across eight MCMC chains). As shown in Figure 22, we would expect the number of iterations to increase as DNA profiles become more complex. Excessive number of iterations, or low acceptance rate could indicate that STRmix™ could not converge on a good probability space during MCMC, speaking to the complexity of the profile.

As expected as we increase the complexity of the profiles we observe a general increase in the total number of iterations.

Figure 22: The total number of iterations required for each mixture to yield 400,000 accepts. The blue circles are data from all the 188 deconvolutions in the mixtures of unrelated people from Section D. The red crosses is the same set of data from samples with less than 1 ng of template.

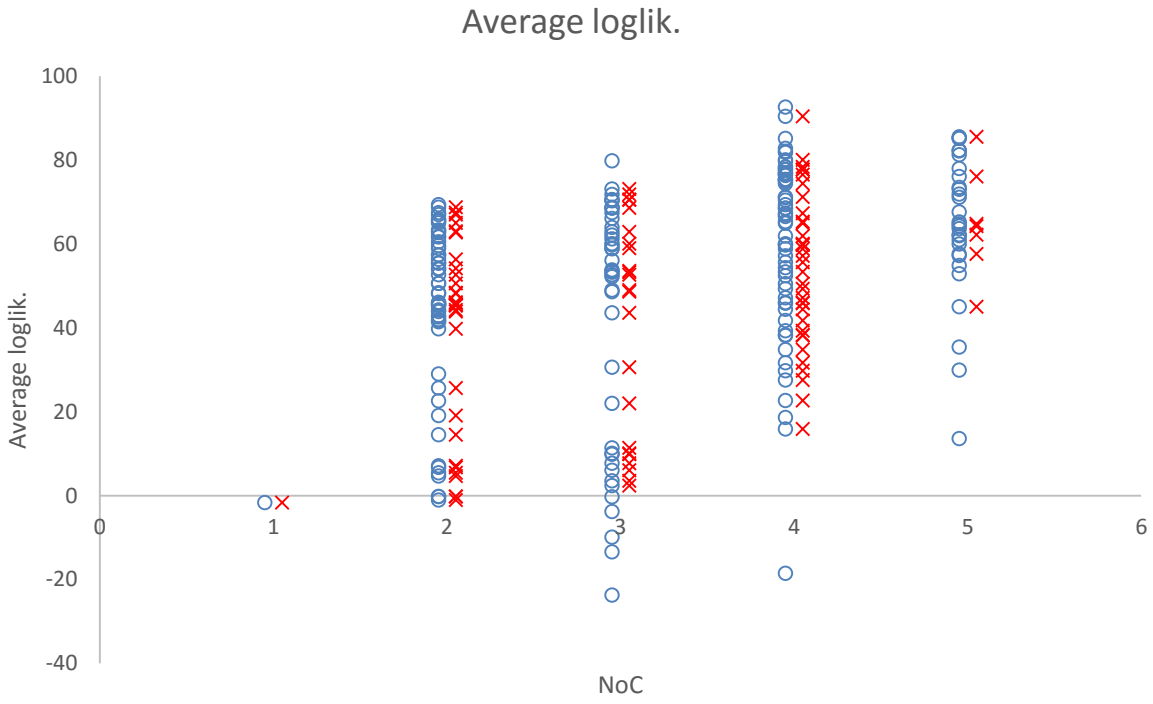


Average log(likelihood)

The average log(likelihood) is the average post burn-in probability density (or likelihood) values leading to an accept across the chains used in a deconvolution. The values shown in Figure 23 shows a good spread of these values, in to the 90's. Generally, a high average log(likelihood) is better, however, low or negative values, such as those seen below, do not necessarily indicate an issue.

As with previous secondary diagnostics removing saturated data improves the overall trend of the values.

Figure 23: The average log(likelihood) output for each mixture. The blue circles are data from all 188 samples. The red crosses are data with less than 1 ng of template.

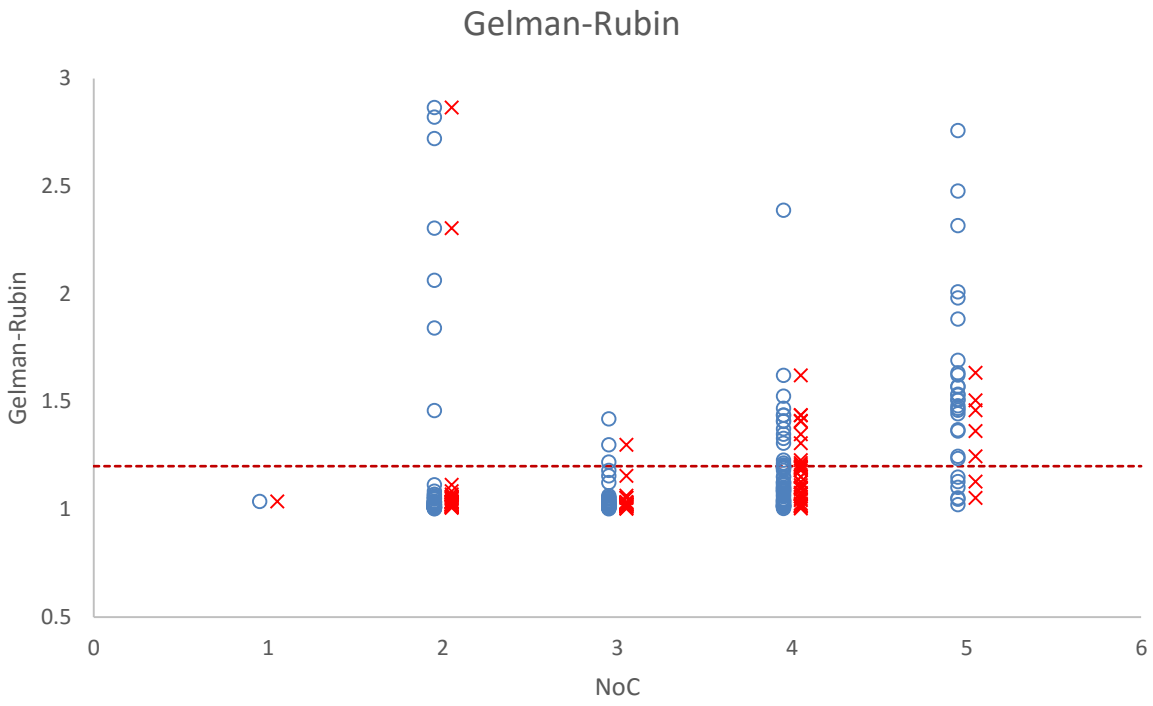


Gelman-Rubin (GR) convergence diagnostic

The Gelman-Rubin (GR) is a diagnostic value that indicates whether there is likely convergence of the MCMC probabilities in each MCMC chain. This value is an average within and across all the chains. Figure 24**Error! Reference source not found.** shows a spread of GR with the majority of the points (84.4%) below 1.2 (shown as the dashed line). A value of 1.2 or less typically indicates a likely convergence of the MCMC chains.

However, GR values greater than 1.2 do not necessarily indicate that the deconvolution is unsuitable. This value could simply be indicating the complexity of the given mixture. As seen in Figure 24, as the number of contributors increases, more GR values are greater than 1.2. When an excessive GR value is observed it is best to investigate the other primary and secondary diagnostics. If these other diagnostics are not within expectations, then the analyst may choose to re-run the sample with the same or an extended number of accepts. If this does not decrease the GR value, it may indicate a further review of the input file or a reassessment of N could be warranted.

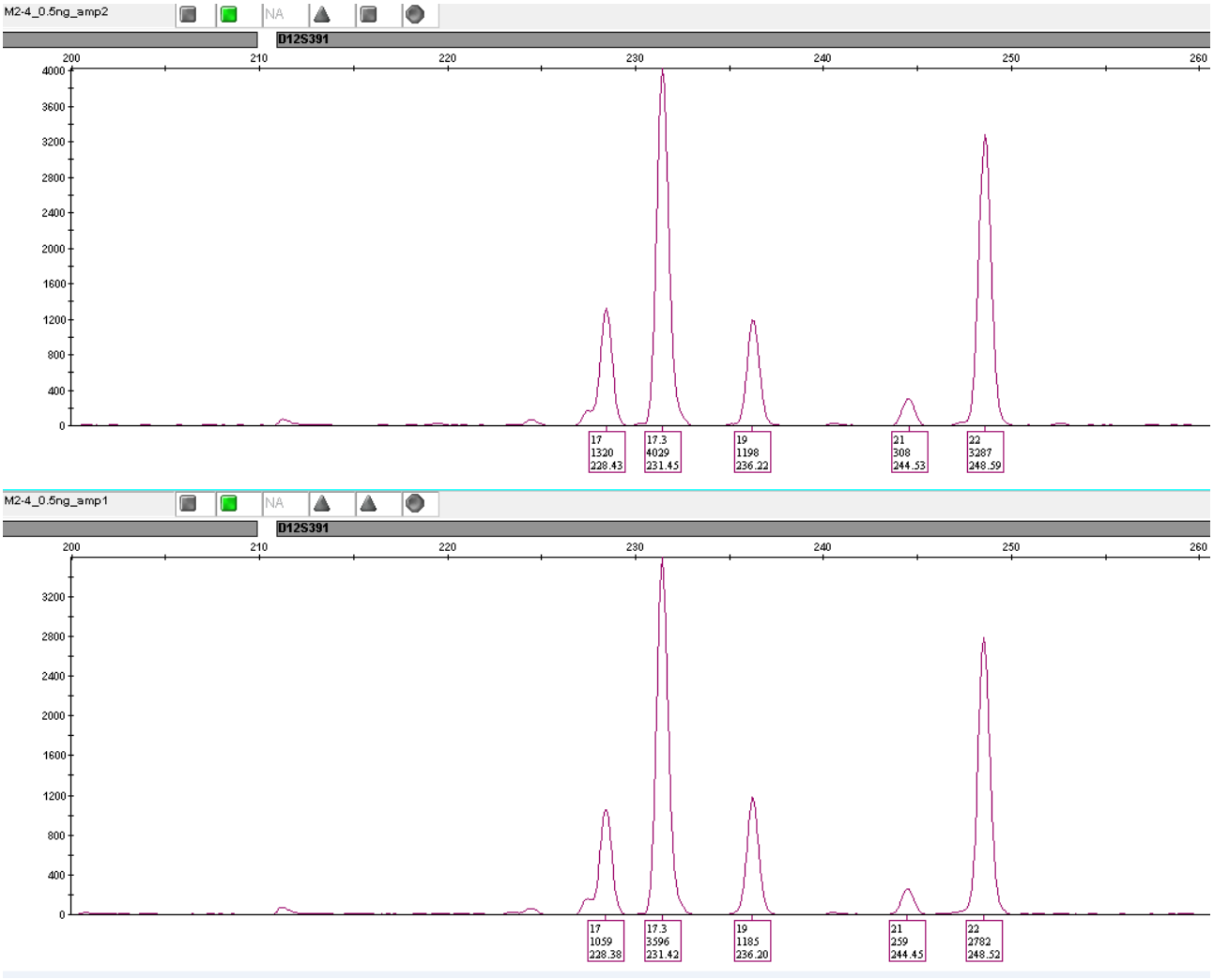
Figure 24: The GR value for each of the 188 mixtures are shown in blue circles. The red crosses are GR values for mixtures with less than 1 ng of template.



Inspection of Figure 24 shows an overall improvement of GR when potentially saturated data is omitted, with the exception of 2 elevated values for the 2 person mixtures. These two values (~2.3 and ~2.8) relate to runs M2_G01_M2-4_0.5ng_amp1_CBI_24s.hid and M2_G06_M2-4_0.5ng_amp2_CBI_24s.hid, respectively. These are amps 1 and 2 of the same sample. Inspection of the primary diagnostics suggest these deconvolution may have progressed as expected. Nevertheless, further inspection of the electropherogram indicates some imbalances in an otherwise strong mixture and also strong 17.3 peaks at the D12S391, with no apparent stutter

labelled. Please refer to Figure 25 below. This again appears to be a single base pair resolution issue of the 16.3 stutter peak.

Figure 25: Screen shot of D12S391 from M2_G01_M2-4_0.5ng_amp1_CBI_24s.hid and M2_G06_M2-4_0.5ng_amp2_CBI_24s.hid



Overall, this appears to have had little impact on the deconvolution but does appear to be driving the GR values up. For investigation purposes only, a re-run was undertaken inserting a data point for a 16.3 peak at 200 RFU in the amp 2 input file. This reduced the GR value to 1.03, raised the log(likelihood) and lowered the stutter variance value in comparison to the initial run. However, the mixture proportions and weights remained similar. As discussed a slightly elevated GR alone may not invalidate a run.

The STRmix™ team advocate a review of the primary and secondary diagnostics for a given run to ensure the findings appear intuitive.