Internal Validation of STRmix™ V2.6 for the analysis of GlobalFiler™ profiles at the Jefferson County Regional Crime Laboratory

STRmix™ internal validation

This document describes the internal validation of STRmix™ V2.6 at the Jefferson County Regional Crime Laboratory (hereafter: JCRCL). 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 Taylor et al. [2]. 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 JCRCL has undertaken in-house, in collaboration with the STRmix™ team at ESR, 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. This included the examination of known and non-probative evidence samples, investigations into reproducibility and precision, sensitivity and stochastic studies, and mixture studies. The sections where specific SWGDAM guidelines are discussed in this document are cross referenced in Appendix 2.

The results of all experiments related to the internal validation of STRmix™ at JCRCL are retained within the laboratory's quality system.

STRmix™ parameters

The parameters described in the document 'Estimation of STRmix™ parameters for JCRCL' 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 are one of the primary diagnostics 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 conditioning profiles can reduce ambiguity and increase weightings of individual genotype sets.

A dilution series of a single-source profile (GF-M3) was prepared, with template amounts of 4ng, 2ng, 0.5ng, 0.25ng, 0.125ng, 0.063ng, 0.031ng, and 0.016ng used. Low template samples where allelic dropout would reasonably be expected were included. Profiles were amplified using the GlobalFiler™ multiplex following JCRCL's standard operating procedure for amplification of crime samples (see CRIME-OPS 4-410). All samples were amplified in triplicate.

All samples with template of 0.125ng and above produced full profiles, with one of the 0.063ng replicates also producing a full profile. The profiles were interpreted in STRmix™ using the JCRCL kit, and a likelihood ratio (*LR*) was assigned for the known donor using the following propositions:

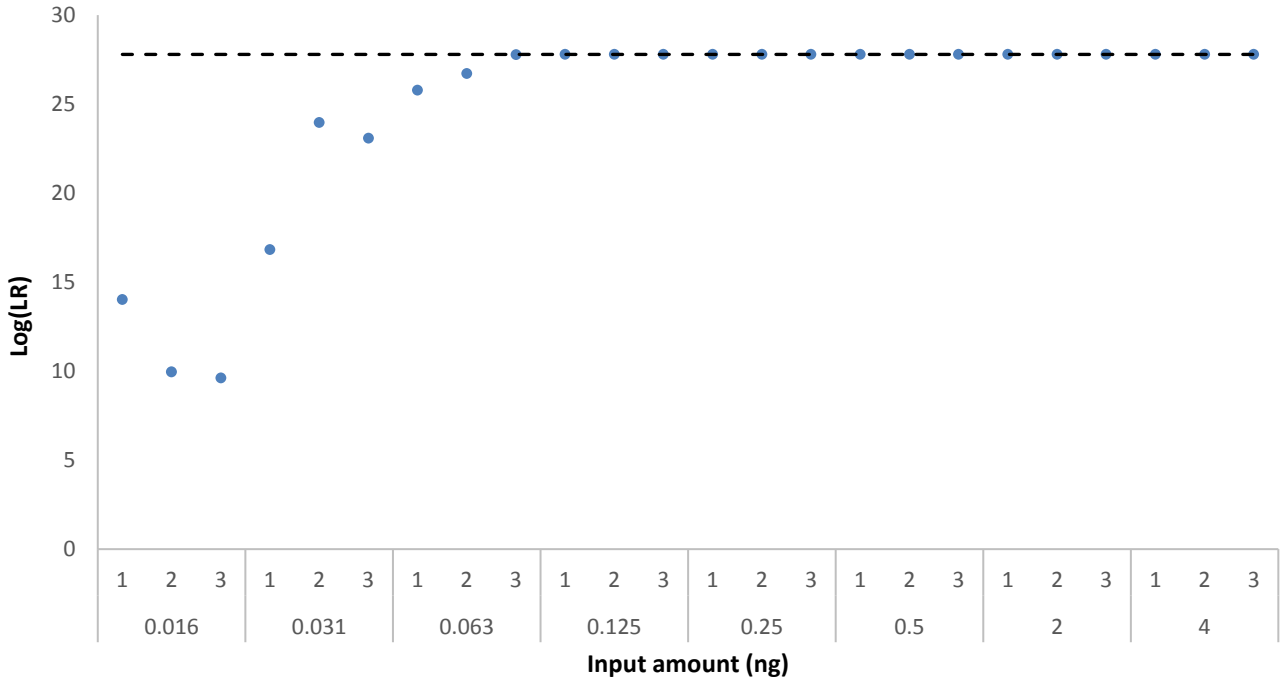
H_p: The DNA originated from the person of interest

H_d: The DNA originated from an unknown individual

The *LR* was calculated for the known contributor (M3) using the NIST1036 Caucasian allele frequencies and an F_{ST} (θ) of 1%. In relation to the STRmix™ output, the sub-source *LR* has been selected for direct comparison between runs. A plot of log(*LR*) versus input DNA is provided in Figure 1. The dashed line represents the log(*LR*) expected for a full, unambiguous single-source profile from donor M3.

When a single-source, high template profile is analysed there is usually only one possible genotype that can describe the observed profile. The genotype at each locus can be considered unambiguous and all of the weight, 1 or 100%, is applied to it. Inspection of the plot in Figure 1 shows, in general, that the *LR* progresses from the value expected for a full, unambiguous profile towards *LR* = 1 (log(*LR*)=0) as DNA template decreases. This is due to STRmix™ considering multiple genotypes including options that consider dropout. As peak heights decrease, the amount of weight applied to genotypes that include dropout increases. There is some variation of the log(*LR*) between the replicates of the low input template intervals which is not unexpected. This is caused by stochastic sampling effects in the laboratory leading to differing amounts of information within each DNA profile. When analysed in STRmix™, this difference in information between replicates leads to differences in the weights applied to the genotypes considered. The weight is used in the calculation of the *LR*, hence the difference observed in *LR*s between replicates.

Figure 1: Plot of log(LR) versus input DNA amount (ng). The dashed line at log(LR) ≈ 27.7 represents the LR expected for a full, unambiguous single-source profile originating from donor M3.



The use of more information during interpretation is expected to result in increased support for inclusion for known donors. This was explored in the present study by re-interpreting each template amount using all three replicates. Following interpretation, an *LR* was calculated for donor M3. The *LR*s produced have been plotted against template amount in Figure 2. Again, the *LR* expected for a full, unambiguous single-source profile from donor M3 is represented as a dashed line. As expected, the use of additional information during interpretation (in this case, the use of replicates) has resulted in increased *LR*s compared with those produced following interpretation of each replicate in isolation.

Figure 2: Plot of combined replicates log(LR) versus input DNA amount (ng). The dashed line at log(LR) ≈ 27.7 represents the LR expected for a full, unambiguous single-source profile originating from donor M3.

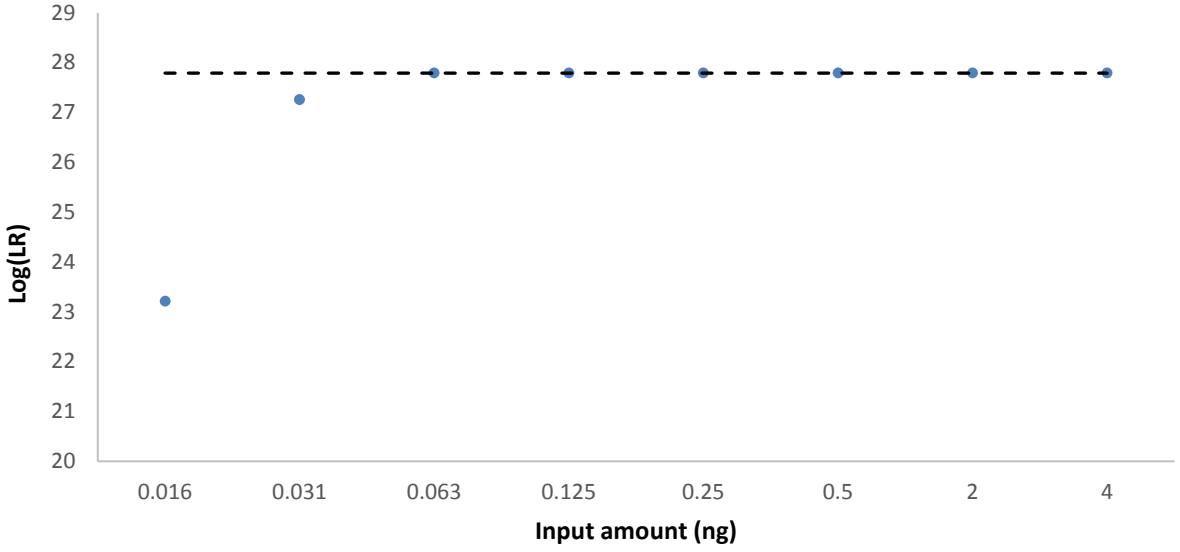


Table 1 summarizes the genotypes accepted at locus D3S1358 and their corresponding weights following STRmix™ interpretation of the 0.031ng profile using replicate 1 only as well as a second interpretation using replicates 1, 2, and 3 combined. The known donor has a genotype of 15,18 at this locus. It may be seen that with the low template, single replicate analysis, STRmix™ considered multiple options: a homozygote 15,15, a 15 paired with an allele that has dropped out (represented as a Q allele in the STRmix™ results), and also dropout of both donor alleles (in this combination the 15 is accounted for as drop-in). The weight is spread across these combinations, with most of the weight assigned to a 15,15 genotype. However, supplying STRmix™ with additional information in the form of replicate amplifications has resulted in a refinement of the list of accepted genotypes with only a single genotype accepted. In this case, STRmix™ has assigned all of its weight to a genotype of 15,18, which corresponds with the genotype of the known donor. In summary, adding relevant information such as multiple PCR replicates can assist STRmix™ by reducing ambiguity within a DNA profile. This can lead to a reduction in the number of accepted genotype combinations, a focussing of the weight, and better ability to distinguish between true donors and non-contributors.

Table 1: Genotypes and weights accepted by STRmix™ at locus D3S1358 for an interpretation using a single replicate amplified using 0.031ng of DNA, versus an interpretation using three 0.031ng PCR replicates. Q represents any allele other than those observed in the profile; in the case of 0.031ng Rep 1, Q represents any allele other than 15

	Genotype	Weight
0.031ng Rep 1	15,15	0.92457
	Q,15	0.075348
	Q,Q	8.3761E-5
0.031ng Rep 1, Rep 2 & Rep 3 combined	15,18	1.0

Reproduction of single-source LR via 'by hand' calculation

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 a single genotype at each locus. The *LR* was calculated 'by hand' at each locus for one single-source profile, and the individual locus *LR*s were 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 a given locus, x_i is the number of observations of allele *i* in a database, N_a is the total 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™ calculated results for a full single-source profile (GF-M3-D1_0.5ng_28c_15s_05_ E10_10.hid) with θ=0 and θ=0.01 are provided in Table 2.

Table 2: 'By hand' (MS Excel) calculation of LR versus STRmix™ results for a full single-source profile (GF-M3-D1_0.5ng_28c_15s_05_E10_10.hid profile). Calculations were carried out using the product rule (ϑ=0) and using the Balding & Nichols formulae with ϑ=0.01

Locus	Excel θ=0	STRmix™ θ=0	Excel θ=0.01	STRmix™ θ=0.01
D3S1358	12.153	12.153	11.545	11.545
vWA	23.508	23.508	21.472	21.472
D16S539	10.132	10.132	9.1258	9.1258
CSF1PO	10.498	10.498	9.4309	9.4309
TPOX	19.083	19.083	16.370	16.370
D8S1179	29.369	29.369	26.490	26.490
D21S11	18.024	18.024	16.586	16.586
D18S51	55.480	55.480	41.354	41.354
D2S441	34.050	34.050	29.469	29.469
D19S433	38.407	38.407	30.677	30.677
TH01	70.472	70.472	50.490	50.490
FGA	35.481	35.481	31.293	31.293
D22S1045	4.0777	4.0777	4.0484	4.0484
D5S818	9.8587	9.8587	9.4100	9.4100
D13S317	5.7262	5.7262	5.6259	5.6259
D7S820	9.5329	9.5329	9.1869	9.1869
SE33	132.37	132.37	100.39	100.39
D10S1248	59.589	56.589	42.049	42.049
D1S1656	269.19	269.19	141.94	141.94
D12S391	31.184	31.184	28.113	28.113
D2S1338	22.384	22.384	20.586	20.586
Total	1.7199E29	1.7199E29	6.1885E27	6.1885E27

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

This section covers the following standard:

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

STRmix™ is a fully continuous model that uses peak heights to inform the genotype combinations of contributors to profiles. As template decreases dropout starts to be considered. In general, as the weights for genotypes considering dropout increase, the weights for genotype combinations for the *true* contributors decrease and consequently the *LR* decreases. This can be observed in Figure 1 and Figure 2; this is the expected result.

It is recommended that saturated profiles are not interpreted using STRmix™. Peak heights observed in such profiles are unlikely to be representative of template amount causing the models used within STRmix™ to become sub-optimal. This may result in non-intuitive genotypes being accepted by STRmix™, particularly when interpreting mixtures where one or more contributors are present at similar levels to the stutter peaks of the major contributor(s). A number of diagnostics are included within STRmix™ which may indicate to the user that further review of the profile should be carried out (for example, the stutter variance diagnostics (k^2), which will likely be elevated in saturated profiles).

To examine the effect of saturated data in STRmix™, one single-source sample (GF_1227.4) was prepared by JCRCL with a deliberately high DNA template amount (8ng). The profile was interpreted in STRmix™ and the weights were reviewed. The profile was interpreted correctly, with all of the weight at each locus being assigned to the known donor's genotype. As expected, this run did result in elevated k^2 values for all stutter variants being modelled when compared to the modes of the respective prior gamma distributions. The results obtained are summarized in Table 3.

Table 3: Template and variance values obtained following STRmix™ interpretation of a high template sample (template amount: 8ng)

Sample	Max peak height in input file (rfu)	Template (*t*) estimated by STRmix™(rfu)	Posterior k^2 variance values (versus mode in brackets)			
			Back stutter	Forward stutter	N-2bp stutter	Double back stutter
GF_1227.4_8ng	31,471	23,483	29.60 (5.52)	14.85 (4.17)	8.42 (1.44)	13.36 (7.01)

The presence of larger than expected stutter peaks can be more problematic in mixed DNA profile interpretation, especially where stutter peak heights are comparable to those of allelic peaks from one or more contributors to the mixture. It is recommended that profiles with off-scale data are corrected biologically, perhaps through re-amplification at lower input template amounts, before interpreting in STRmix™.

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.

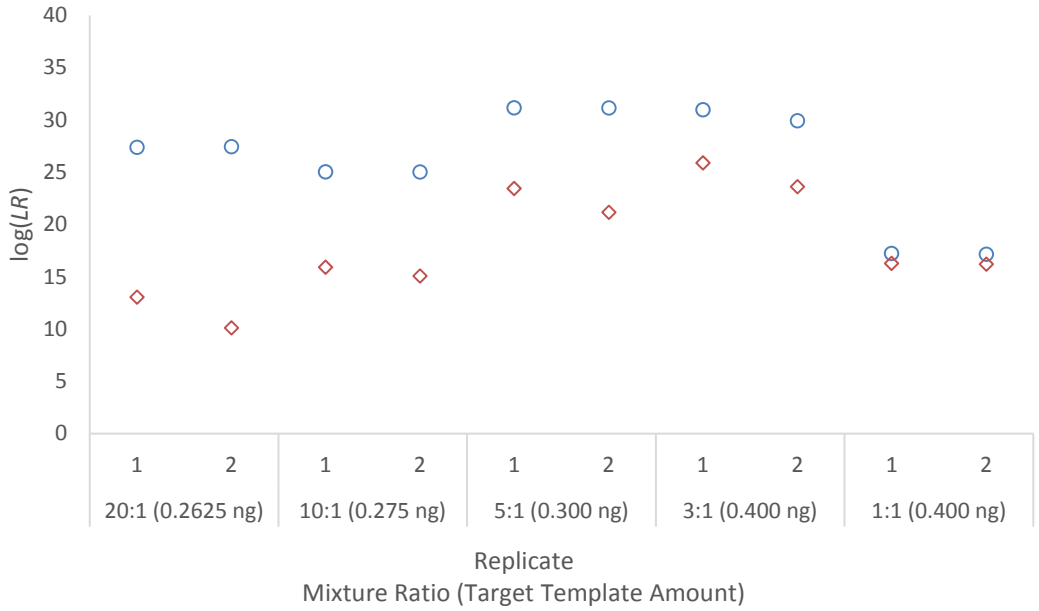
The weights are described as the primary output from STRmix™. They can be used as a diagnostic of the deconvolution process and should be intuitively correct, where the most supported genotypes have the highest weights.

A two-person mixture series was constructed in the following ratios by JCRCL: 20:1, 10:1, 5:1, 3:1, 2:1 and 1:1. A subset of these was selected where the total amount of DNA in the samples was similar (approximately 262.5pg – 400pg). A summary of the profiles used is provided in Table 4. The profiles were interpreted in STRmix™ and an *LR* calculated for each of the known contributors under the following propositions:

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

The *LR* was calculated using the NIST Caucasian allele frequencies and an F_{ST} (θ) of 1%. In relation to the STRmix™ output the sub-source point estimate value (sub-source *LR*) has been selected for direct comparison between runs. A plot of log(*LR*) for each mixture type considering both the major contributor (blue circle) and minor contributor (red rhombus) is provided in Figure 3. The first and second amplifications of each sample were interpreted separately and are plotted below.

Figure 3: Log(LR) for each contributor to each of the mixtures considered. The contributor that corresponds to the 'major' contributor is displayed as blue circle data points and the contributor that corresponds to the 'minor' contributor is shown as red rhombus data points. The lines represent the biological replicates interpreted in STRmix™



Different donors were used to create the mixtures examined, hence some differences in the *LR*s could be driven by differences in the allele frequencies used within the *LR* calculations. Nevertheless, inspection of Figure 3 shows that the *LR* for the major contributor decreased by approximately 14 orders of magnitude as the mixture ratio progressed from fully resolvable (i.e. 5:1) to unresolvable (i.e. 1:1). The *LR* is expected to decrease as the ability to tease apart individual contributors based on peak height diminishes. In addition, we expect the *LR* assigned for the minor contributor to decrease as their contribution to the mixture decreases. From Figure 3, this decrease in *LR* is most evident for the 10:1 and 20:1 mixtures, where the peak heights of the minor donor's alleles are expected to be similar to those of stutter peaks from the major contributor. In Table 4 we provide the mixture proportions determined by STRmix™ along with the proportions targeted when the mixtures were prepared. It can be seen that the mixture proportions output by STRmix™ are largely in agreement with the experimental design mixture proportions.

Table 4: Summary of interpreted profiles and mixture proportions proposed by STRmix™

		STRmix™ M_x		**Experimental Design M_x**	
Sample	**Rep**	**Contributor 1**	**Contributor 2**	**Contributor 1**	**Contributor 2**
Mix5 (20:1)	D1	0.96	0.04	0.95	0.05
	D2	0.96	0.04	0.95	0.05
Mix6 (10:1)	C1	0.93	0.07	0.91	0.09
	C2	0.92	0.08	0.91	0.09
Mix7 (5:1)	B1	0.87	0.13	0.83	0.17
	B2	0.85	0.15	0.83	0.17
Mix8 (3:1)	A1	0.79	0.21	0.75	0.25
	A2	0.81	0.19	0.75	0.25
Mix13 (1:1)	B1	0.52	0.48	0.50	0.50
	B2	0.52	0.48	0.50	0.50

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 Jefferson County's GlobalFiler™ mixtures was undertaken as per Taylor [6]. Departures from [6] include the use of average peak height rather than template amount and the use of apparent number of contributors (discussed further below).

With respect to interpretation methods, sensitivity is defined as the ability to reliably resolve the DNA profiles 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. 'Information' includes the amount of DNA from the contributor of interest, conditioning profiles (for example the complainant's profile on intimate samples), replicates, and decreasing numbers of contributors (i.e. reduced profile complexity).

Specificity is defined as the 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 be low (i.e. negative) and should trend upwards to 0 as less information is present within the profile.

Specificity and sensitivity were tested by calculating the *LR* for both known donors and non-contributors for a number of two-, three-, four-, and five-person mixed profiles that had been interpreted using STRmix™. The plots in [6] have been reproduced for Jefferson County's GlobalFiler™ data. Thirty-one two-person mixtures, twenty-one three-person mixtures, sixteen four-person mixtures, and four five-person mixtures were generated by the laboratory using known donors. Each sample was amplified in duplicate using the GlobalFiler™ multiplex to produce a total of 144 profiles. A summary of the profiles constructed is given in Table 5.

Table 5: Summary of experimental design for specificity and sensitivity tests

			Target DNA Amount (pg) of the Smallest Contributor					
Mixture	**N**	**Donor Ratio**	**A**	**B**	**C**	**D**	**E**	**F**
Mix 1	4	4:3:2:1	100	50	25	12.5	6.25	
Mix 2	4	10:5:2:1						
Mix 3	3	10:5:1						
Mix 4	3	3:2:1						
Mix 5	2	20:1						
Mix 6	2	10:1						
Mix 7	2	5:1						
Mix 8	2	3:1						
Mix 9†	2	3:1						
Mix 10†	3	10:5:1						
Mix 11	4	1:1:1:1	400	200	100	50	25	12.5
Mix 12	3	1:1:1						
Mix 13	2	1:1						
CM1	5	1:1:1:1:1	140					
CM2	5	1:5:1:10:1	39					
CM3	5	1:3:4:1:1	70					
CM4	5	2:4:5:1:3	47					

† Samples had a degraded component, where the major contributor was UV-degraded prior to use within the mixture.

These profiles represent a full range of sample types, including the 'worst' types of profiles likely to be encountered by the laboratory during casework analysis. The profiles are of varying DNA quantity, quality, and mixture proportions. The contributors include homozygous and heterozygous genotypes and there is varying amounts of allele sharing across the loci tested (standard 4.1.6.5). Given the template amounts, allele and/or locus dropout was expected to occur within profiles containing low DNA amounts (standard 4.1.7.1).

Before interpretation in STRmix™ the profiles were assessed 'blind' with regards to the true number of contributors (i.e. the number of contributors in the experimental set up). The apparent number of contributors (NoC) was assigned to each profile by an experienced analyst using the Maximum Allele Count (MAC) method supplemented with peak height information. The true number of contributors to a questioned/crime scene profile is always unknown and unknowable; the approach used in the present study ensures that the results produced are applicable to casework. Table 6 summarizes those profiles where there were discrepancies between the analyst-assigned NoC and experimental design NoC.

Table 6: Summary of profiles where apparent NoC differed from the experimental design NoC

Sample File	Apparent NoC	Design NoC
GF-MX_Mix3_B2_28c_15s_06_F04_04.hid	4	3
GF-MX_CM-Mix-4.1_28c_15s_07_G12_12.hid	4	5
GF-MX_CM-Mix-4.2_28c_15s_08_H12_12.hid	4	5
GF-MX_CM-Mix-2.1_28c_15s_03_C12_12.hid	4	5
GF-MX_CM-Mix-2.2_28c_15s_04_D12_12.hid	4	5
GF-MX_CM-Mix-3.1_28c_15s_05_E12_12.hid	4	5
GF-MX_Mix1_E1_28c_15s_02_B02_02.hid	3	4
GF-MX_Mix1_E2_28c_15s_03_C02_02.hid	3	4
GF-MX_Mix11_F1_28c_15s_01_A07_07.hid	3	4
GF-MX_Mix11_F2_28c_15s_02_B07_07.hid	3	4

Generally, where the NoC was under-assigned, the sample involved a high-order mixture where one or more contributors were present at low template and hence may not be apparent. The single profile where NoC was

over-assigned included peak height imbalances which led the analyst to consider the presence of an additional contributor.

Each profile was interpreted in STRmix™ using the apparent number of contributors assigned. The deconvolutions were then compared to a database containing profiles of the known contributors and 1,000 known non-contributors using the Database Search function within STRmix™. The non-contributors were artificially generated using the NIST Caucasian (July, 2017) allele frequency database using an MS Excel workbook available from the STRmix™ support team. An *LR* was calculated for each individual within the database.

To calculate the statistic, the propositions considered were:

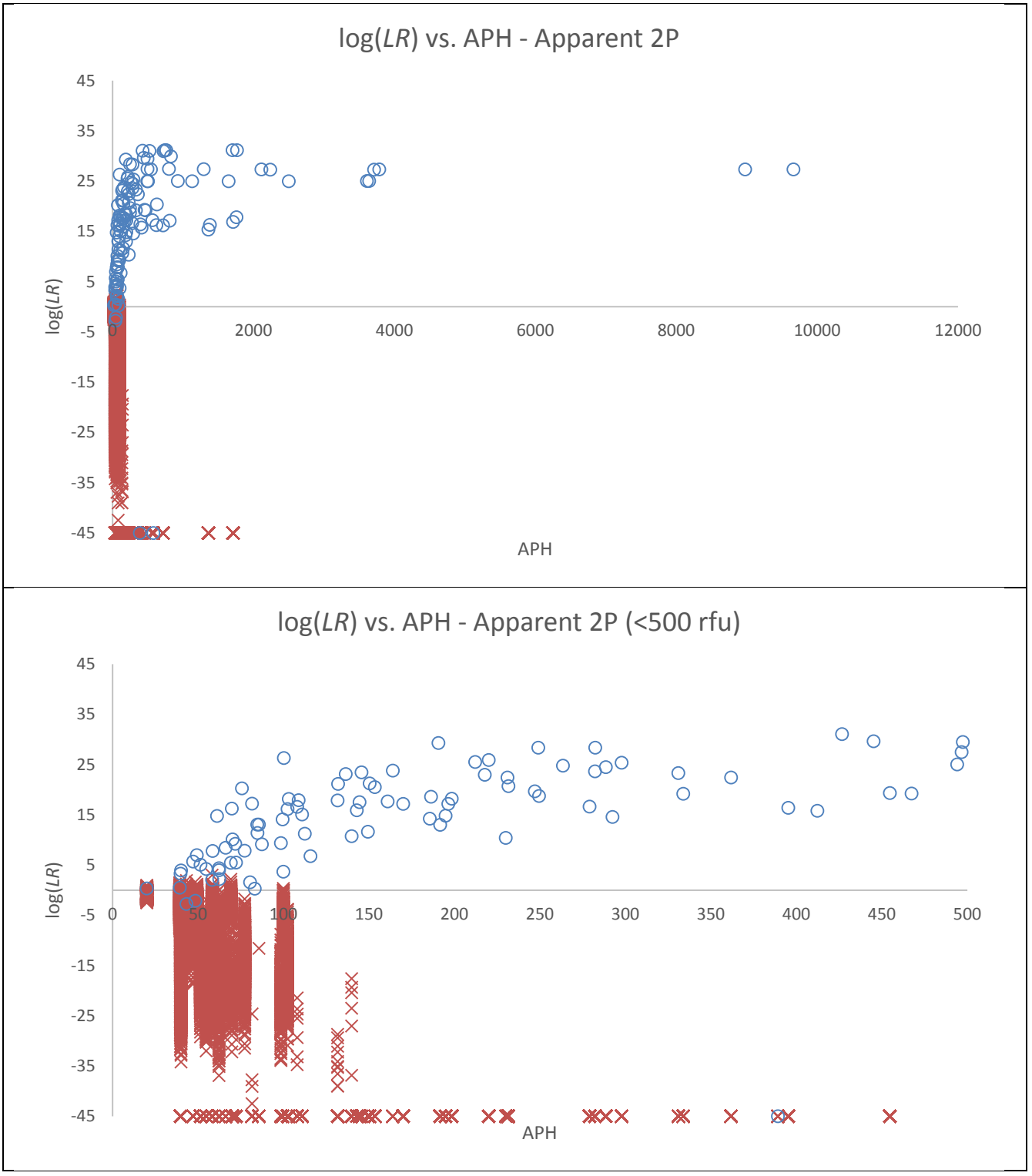
> H_p: The DNA originated from the database individual and *N*-1 unknown individuals
>
> H_d: The DNA originated from *N* unknown individuals

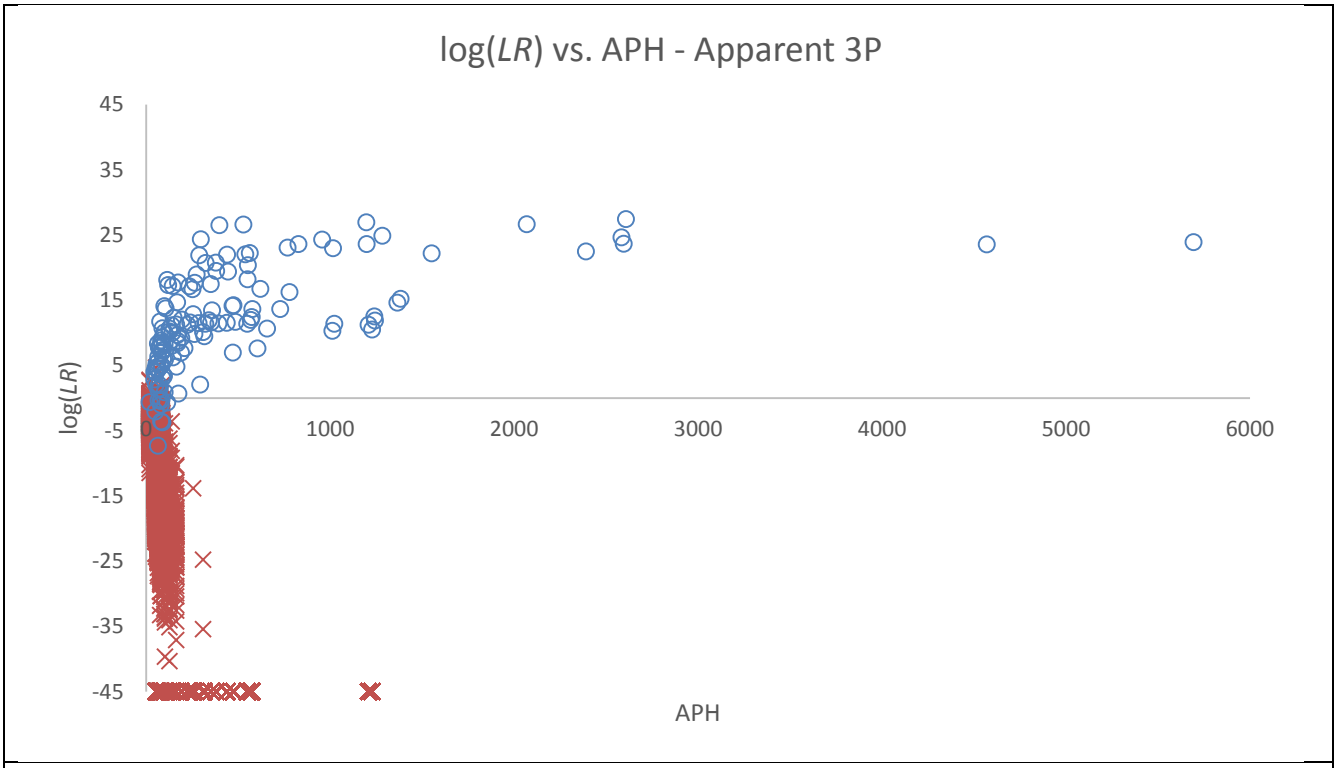
Where *N* is the apparent number of contributors assigned to the profile. *LR*s were calculated using the NIST Caucasian allele frequency database with an F_{ST} (θ) of 0.01 (1%). The 'Sub-source *LR*' calculated by STRmix™ was used as the point of comparison.

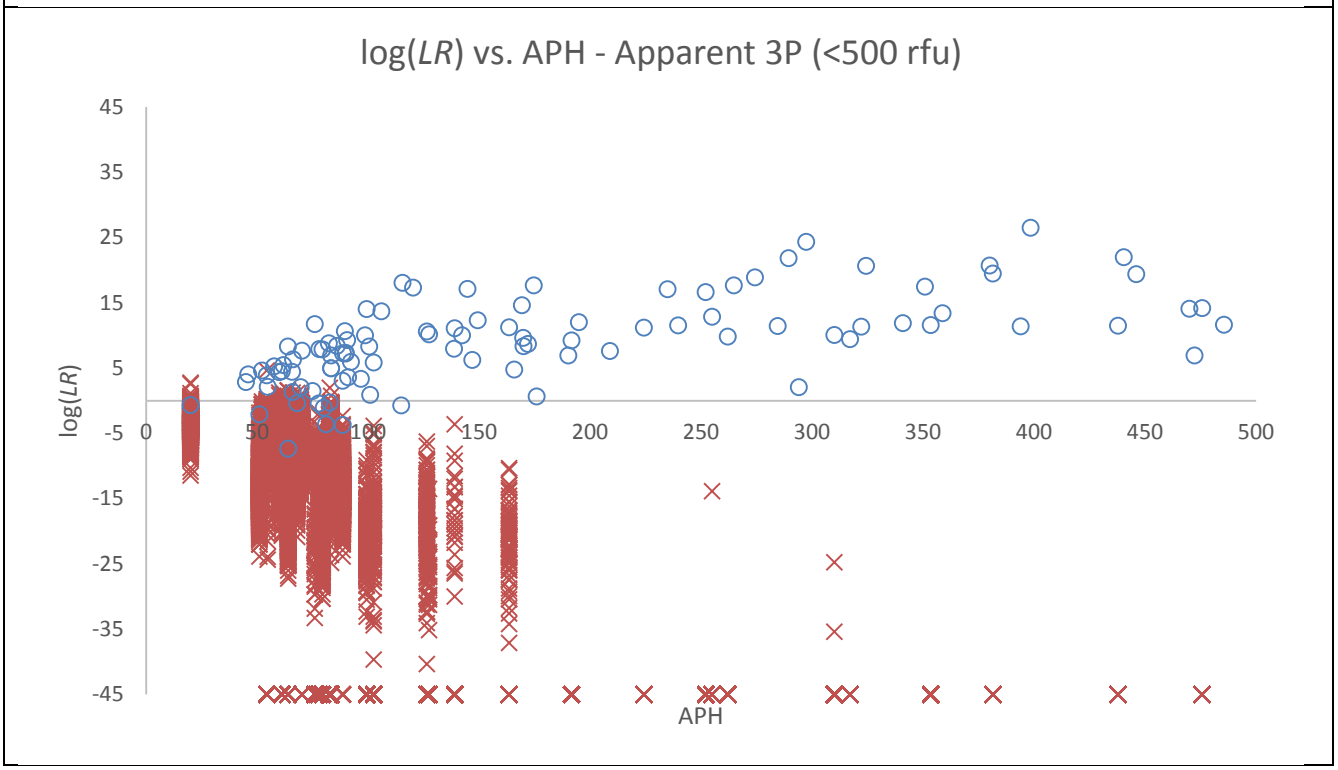
Plots of log(*LR*) versus the average peak height (*APH*) per contributor for the apparent two-, three-, four-, and five-contributor mixtures are given in Figure 4. *APH* was calculated using unmasked, unshared, and non-stutter-affected alleles for each contributor in the mixed profile. Where there were no such alleles, an *APH* of half the analytical threshold was assigned (in this case, 20rfu). The *APH* for the H_d true contributors was taken as the lowest *APH* across all known contributors.

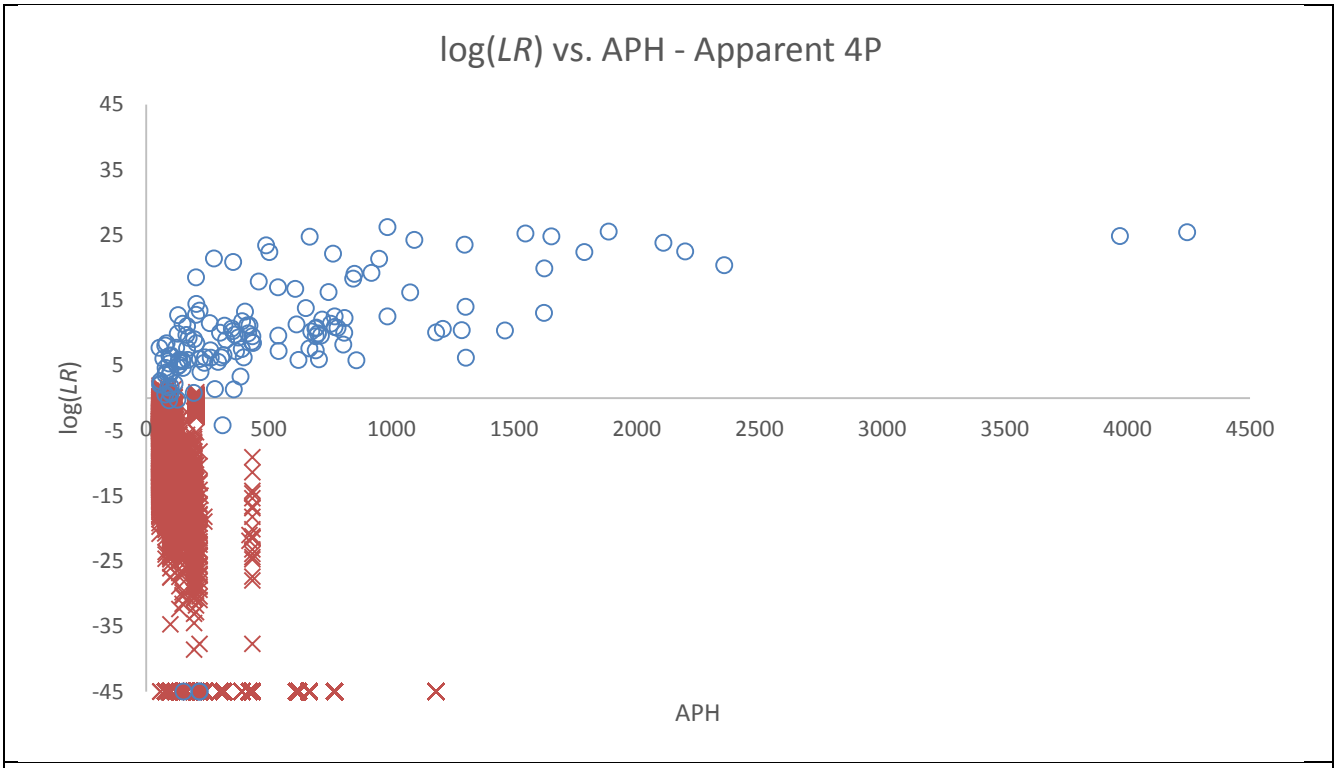
The results of all comparisons are provided in Figure 4. Exclusions (*LR* = 0) have arbitrarily been plotted as log(*LR*) = -45.

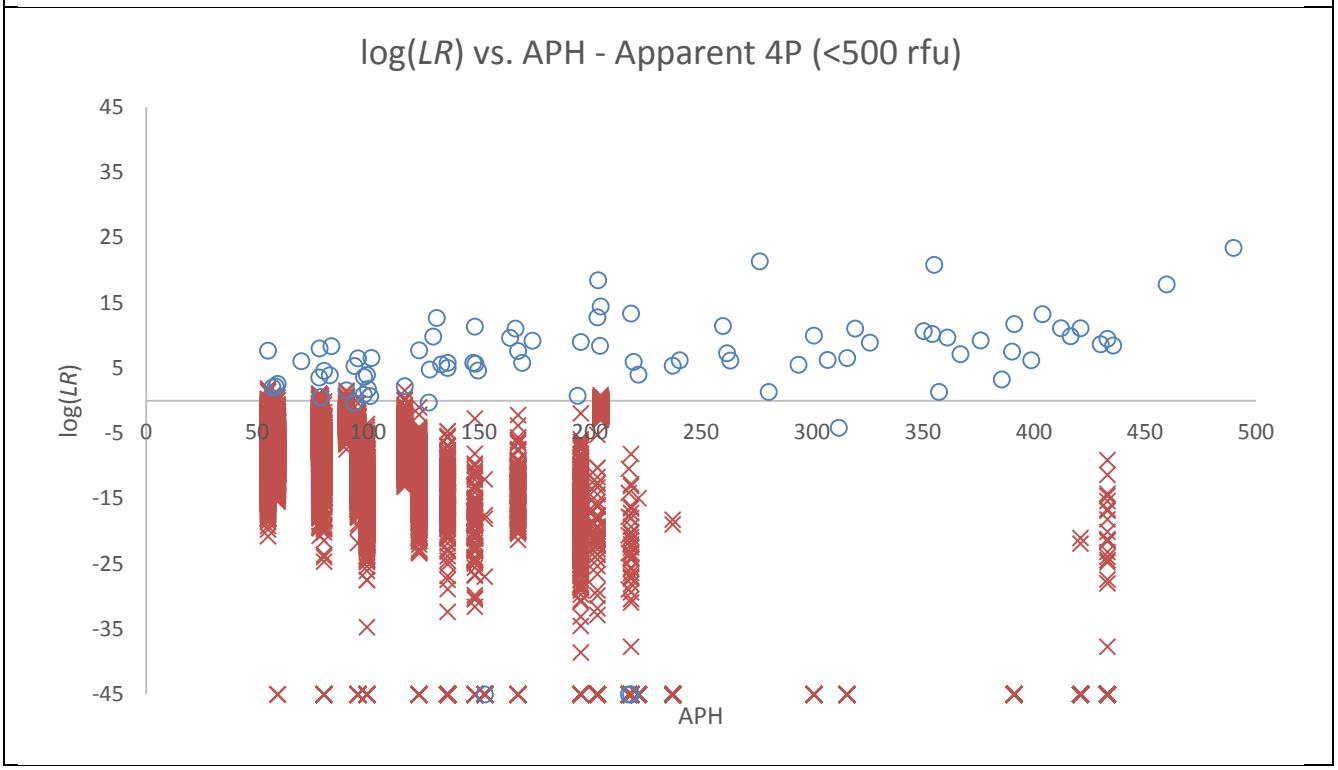
Figure 4: Log(LR) versus APH (rfu) for apparent two-, three-, four-, and five-person GlobalFiler™ mixtures. Each plot has been reproduced with the x-axis scaled to better show the results for low APH data (i.e. <500rfu). LRs for known donors are plotted as blue circles, LRs for non-contributors are plotted as red crosses

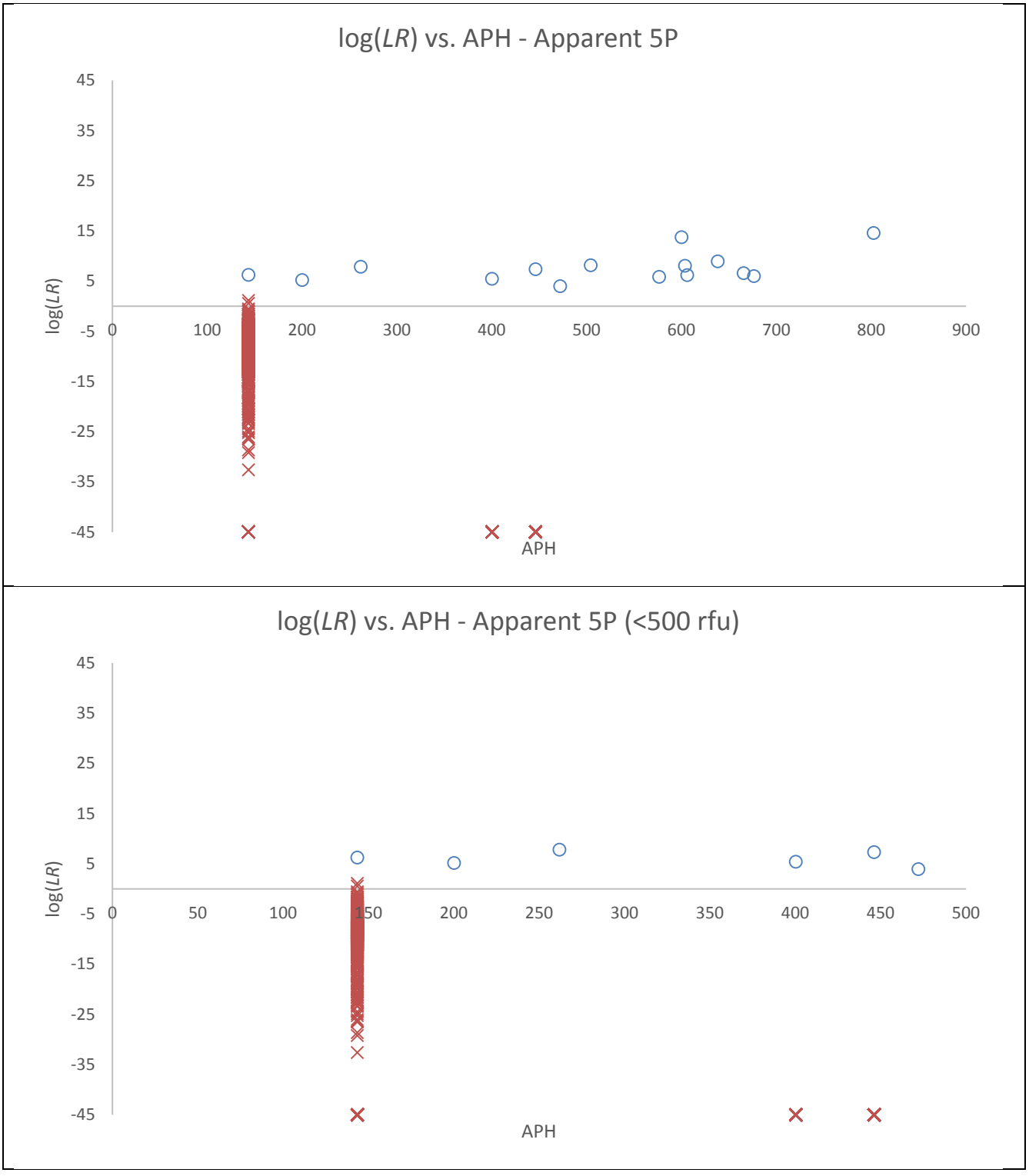



log(LR) vs. APH - Apparent 3P
45
35
25
15
5
-5
-15
-25
-35
-45
log(LR)
0
1000
2000
3000
4000
5000
6000
APH


log(LR) vs. APH - Apparent 3P (<500 rfu)
45
35
25
15
5
-5
-15
-25
-35
-45
log(LR)
0
50
100
150
200
250
300
350
400
450
500
APH


log(LR) vs. APH - Apparent 4P
log(LR)
APH
45
35
25
15
5
-5
-15
-25
-35
-45
0
500
1000
1500
2000
2500
3000
3500
4000
4500


log(LR) vs. APH - Apparent 4P (<500 rfu)
log(LR)
APH
45
35
25
15
5
-5
-15
-25
-35
-45
0
50
100
150
200
250
300
350
400
450
500



Inspection of Figure 4 reveals that there were a number of observations of interest where the *LR* calculated for a known contributor was exclusionary (*LR* < 1) or an exclusion (*LR* = 0). Each of these observations were investigated and are discussed in turn below.

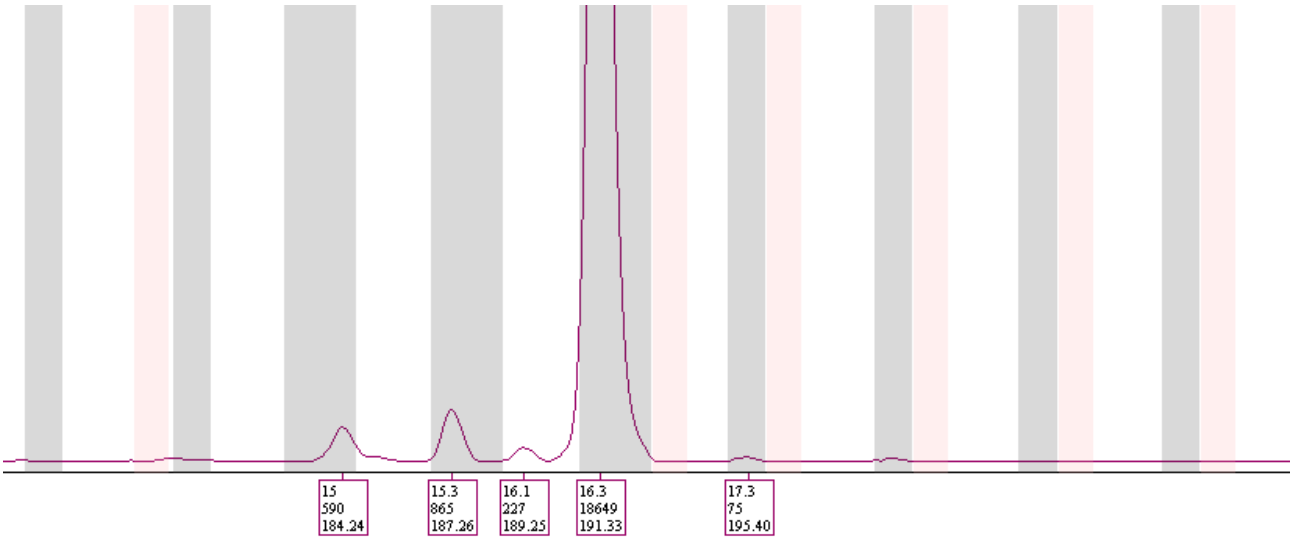
Under-assignment of NoC

Under-assignment of the number of contributors to the mixture appeared to be a likely cause for nine of these observations. This under-assignment in the number of contributors could be a result of allele-masking or the trace contributor dropping out of the profile. This can lead to false exclusions of true contributors as either the information from that contributor is not present (dropped out) or STRmix™ is restricted to proposing genotype combinations for one less contributor than is truly present. This precludes the acceptance of some genotype combinations of true contributors. These effects of under-assigning the number of contributors is further explored in Section F.

Single-locus Exclusion

Exclusions were observed for one of the known donors (reference 1224) to Mix5_A1 and Mix5_A2. Further investigation revealed that these were the result of a single-locus exclusion at locus D1S1656. In contrast, the remaining per-locus *LR*s were all inclusionary. The per-locus *LR*s are one of the primary diagnostics of STRmix™ and should be reviewed whenever an *LR* calculation is undertaken. Observation of an exclusion at a single locus with inclusionary *LR*s at all other loci may indicate an issue with the interpretation and should be investigated. The cause of the exclusions in this instance could be traced back to failure of the CE instrument to resolve the minor contributor's 17 allele. The genotypes of the known donors to Mix5_A1 and Mix5_A2 are 16.3,16.3 (major contributor) and 15,17 (minor contributor). A screenshot of the DNA data at locus D1S1656 from one of the affected samples is presented in Figure 5. While not obvious, it is likely that the minor contributor's 17 allele has fallen onto the shoulder of the major 16.3 allele and failed to be resolved. Furthermore, the template proposed for the minor contributor is quite high in both replicates (533rfu in Mix5_A1 and 384rfu in Mix5_A2). Therefore, it is reasonable that STRmix™ does not accept the genotype [15, Q] during MCMC. The difficulty in resolving peaks that are separated by one base pair is a known issue of the capillary electrophoresis process which can lead to issues when interpreting affected profiles using STRmix™. We advocate first trying to resolve the issue biologically (for example, by re-injecting the sample). Failing this, the locus can be ignored during STRmix™ interpretation. Whenever a locus is ignored, the analyst should first review the DNA data to ensure that a potentially exclusionary result is not being omitted.

Figure 5: Screenshot of locus D1S1656 for Mix5_A1. The minor contributor's 17 allele appears to have fallen in the shoulder of the major contributor's 16.3 allele and failed to be resolved during CE.



For demonstration purposes, *LR*s were re-calculated for both known contributors with locus D1S1656 ignored. The results are given in Table 7.

Table 7: Difference in LR when locus D1S1656 is ignored

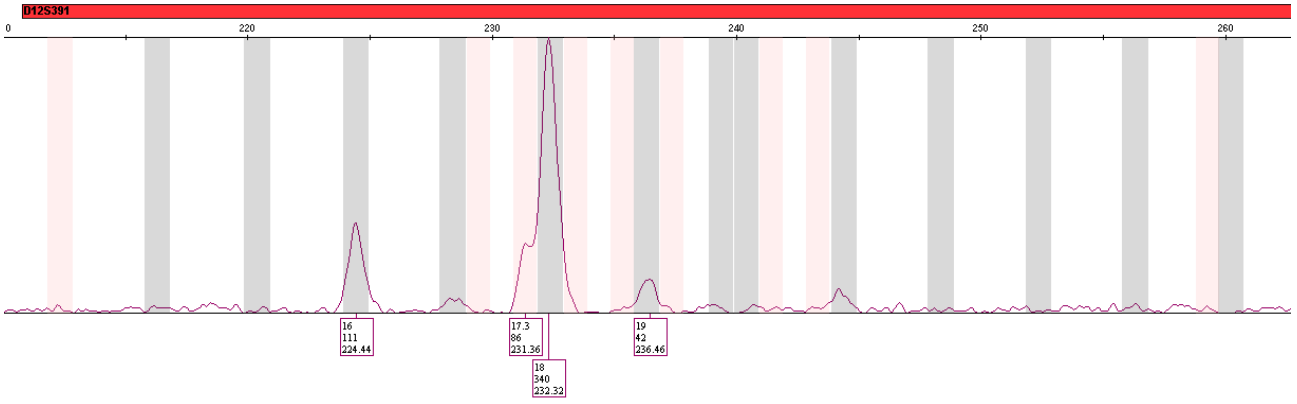
Sample	Component	*LR*	$LR_{\text{D1S1656 Ignored}}$
Mix5_A1	Major	2.1097×10^{27}	1.7450×10^{25}
	Minor	0	1.9762×10^{22}
Mix5_A2	Major	2.2703×10^{27}	1.8742×10^{25}
	Minor	0	2.0962×10^{22}

A similar issue was also observed in Mix 4_D2, resulting in *LR*s that supported exclusion being assigned to known contributors. This may be due in part to the non-resolution of peak information separated by a single base pair. An example of this issue is shown in Figure 6, where one of the replicates is able to resolve the 17.3 allele on the shoulder of the 18 allele at D12S391. Whereas, in the second replicate an indication of a peak is observed in the electropherogram but is unable to be resolved and is missing from the STRmix™ input file. Due to the low level nature of each profile, drop out combinations with the surviving resolved allele are able to be considered and accepted; therefore, an outright exclusion of the known contributor was not observed. The low level nature of the profile and the numerous genotype combinations considered at each locus also contributes to the low *LR* of a known contributor.

The use of replicate profiles from two or more amplifications may also be one way to approach this problem. As shown in Figure 6, if one of the replicates is able to resolve these single base pair differences, analysts can deconvolute the mixture using information from both replicates.

To demonstrate this use of replicates to resolve these issues with single base pair resolution differences, the two mixtures were re-interpreted with their corresponding PCR replicate. *LR*s were then recalculated to the known contributors where the *LR* previously supported exclusion of the known true contributor. The results, shown in Table 8, demonstrate the advantage of providing STRmix™ with additional information in the form of PCR replicates in this type of situation. However, it should be noted that some of the increase in the *LR* calculated using both replicates will be due to information from other loci, in addition to the resolution of the peaks a D12S391, also assisting the STRmix™ interpretation.

Figure 6: Screenshot of locus D12S391 from the two replicates of Mix4 dilution series D (top: replicate 1, bottom: replicate 2). Replicate 2 appears to have an unresolved 17.3 peak



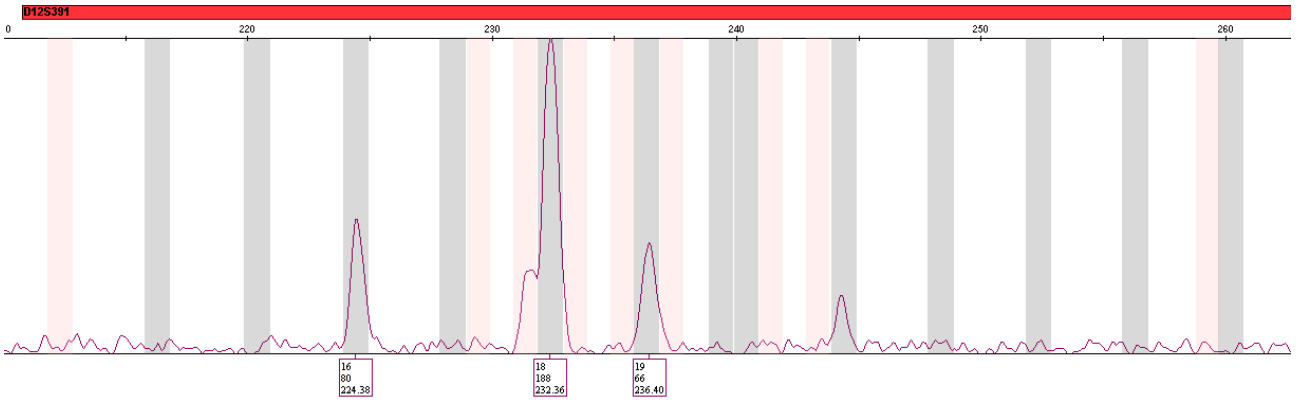


Table 8: Differences in LR when replicates (where issues with single base pair resolution is present in one replicate) are used in the STRmix™ deconvolution

Contributor	Sample	Replicate	*LR*	*LR*(with replicate)
1230	Mix4_D2	Mix4_D1	4.5255×10^{-1}	3.0555×10^{5}

Replicate PCR profiles were used in this example to demonstrate how relevant information can assist STRmix™ in the interpretation of low level profiles as exemplified by the increase in LR between interpretations displayed in Table 8. As for the previous example of unresolved peaks separated by one base pair, ignoring the locus following a review for potentially exculpatory information would be another approach to overcome this issue.

Low template mixtures

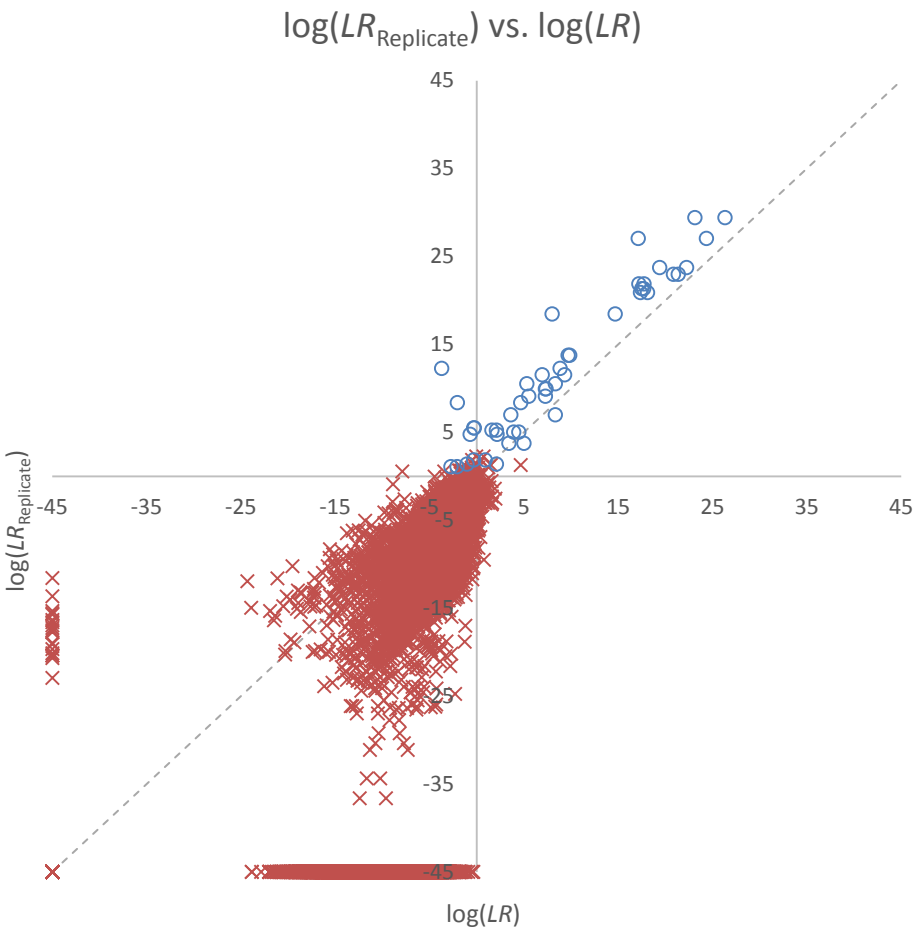
The remaining mixtures could be attributed to the low-level template of a minor or trace contributor. Due to the low-level nature of these contributors, there is increasing complexity leading to many genotype combinations being proposed and accepted. The weight is spread across these genotype combinations which leads to the slight exclusionary *LR*s calculated for the known contributors.

These low template profiles are examples where analysts can benefit from interpreting mixtures as replicates, by utilizing information present in replicate amplifications from the same extract. Therefore, the mixtures listed below were re-interpreted using both respective replicates.

- Mix 2 E replicates
- Mix 3 D replicates & E replicates
- Mix 4 D replicates & E replicates
- Mix 7 E replicates
- Mix 10 E replicates
- Mix 12 F replicates

As shown in Figure 7, it may be seen that with the addition of extra information the *LR* for some of the known contributors increased. Specifically, 45 of the 48 (93.75%) *LR*s for the known contributors increased. The largest difference involved an increase in one *LR* by almost 16 orders of magnitude. Use of replicates also caused the *LR*s calculated for non-contributors to decrease. In total, 14,248 of the 16,464 (86.54%) *LR*s for the known non-contributors decreased. This demonstrates the increased power to distinguish between known donors and non-contributors through the use of amplification replicates within STRmix™.

Figure 7: Comparison of log(LR) match statistics calculated for known donors and non-contributors following STRmix™ interpretation of a single PCR replicate in isolation (x-axis) and interpretation using both available PCR replicates for a given mixture (y-axis). LRs for known donors are plotted as blue circles, LRs for non-contributors are plotted as red crosses



Investigation of H_d true individuals with LRs greater than 100

Mixtures where a known non-contributor for whom an *LR* greater than 100 was calculated, when they were considered as a POI under H_p, were investigated and are discussed below. This *LR* value corresponds to the lower bound of the "moderate support" for H_p category of the SWGDAM verbal scale.

There were two known non-contributors who had an *LR* greater than 100 in the apparent two contributor mixtures. The mixtures where these were observed were Mix13_F1 (database Individual 770) and Mix 9_E1 (database individual 123). Both of these were low level mixtures with the maximum height observed in the input file being less than 220rfu in both cases. Following STRmix™ analysis, both had low log(likelihood) values; 1.41 and 1.16 respectively. Investigation of the weights section of each report indicated that many genotype sets had been considered and at a number of loci double drop out (Q,Q) options were assigned a high proportion of the weight enabling any individual to adventitiously link to these effective 'wild card'

designations. Comparison of the profiles of each of these individuals also showed a reasonable amount of correspondence. For mix 13_F1, 17 out of the 31 autosomal peaks present corresponded with database individual 770, and 14 out of the 22 autosomal peaks present in Mix 9_E1 corresponded with database individual 123. This relatively high degree of similarity and the high weightings of double drop out options leads to *LR*s that show moderate support for H_p.

There were three known non-contributors who had an *LR* greater than 100 in the apparent three contributor mixtures. The mixtures where these were observed were Mix 11_F2 (database individual 92) and Mix 8_E2 (database individual 387 and database individual 119). For Mix 8_E2, there was limited data available in the STRmix™ input file; the highest peak in the file was 252rfu and the 27 autosomal peaks were spread across 19 loci. In some instances, this lead to limited genotype combinations being considered for both contributors. For example, at loci where a single peak was observed below the drop-in cap STRmix™ could only consider the homozygote option, a heterozygote option with drop out, and then a double drop out option. The drop out combinations allow for adventitious matches and as the weight was fairly evenly spread across the limited combinations at some loci the LRs generated to the database individual mentioned above, that also shared a degree of similarity with the profile present, was relatively high. Mix11_F1 was again a low level mixture where the height of the tallest peak in the input file was 263rfu. When this mixture was analysed in STRmix™, many genotype combinations were considered at each locus including heterozygote drop options and double drop options. This allowed for adventitious matches to occur and high LRs were calculated when the aforementioned database individuals were considered as a POI under H_p due to the high degree of similarity between their profiles and the relatively low amount of information that was present in the input files. Individual 387 corresponded to 24 out of the 41 autosomal peaks and individual 119 corresponded to 28 out of 41 autosomal peaks.

Whilst there were known non-contributors that had *LR*s greater than log(LR) = 0 for the apparent four and five contributor mixtures, likely due to the high number of genotype combinations that needed to be considered for each contributor, there were no *LR*s calculated for known non-contributors that showed more than limited support for H_p.

Information displayed in the sensitivity and specificity plots; both the behaviour of the *LR*s for known contributors (H_p true) and known non-contributors over a range of average peak heights, can be helpful in considering whether or not to progress the interpretation of certain DNA profiles. The consideration of whether to progress an interpretation should also be made in combination with the confidence in the assignment of the number of contributors and other such information such as whether relatives may be involved. These factors are discussed later in the report.

Diagnostics

Within STRmix™ the primary diagnostics used to assess the appropriateness of the interpretation are the genotype weights, mixture proportions (M_x) and, where undertaken, the per-locus *LR*s. These values should be intuitive and align with a manual interpretation of the DNA profile.

STRmix™ outputs also contains a number of 'secondary diagnostics' and these should also be reviewed after each STRmix™ interpretation. These include the average log(likelihood), the Gelman-Rubin convergence diagnostic, and the posterior mean of the allele and stutter variances. A summary of the secondary diagnostics produced for the interpretations carried out within Section D can be found in Appendix 3.

With respect to the above diagnostics, STRmix™ appears to be performing as expected. Elevated or out of range diagnostics may indicate that the STRmix™ results require further scrutiny. In particular, results should be closely examined if the mixture proportions, genotype weights, or per-locus *LR*s are not intuitive.

Mixtures of related individuals

Two-, three-, four-, and five-person mixtures were also prepared by JCRCL using DNA from known individuals that were both related and unrelated. These mixtures were designed to mimic challenging casework scenarios that may involve family members. Each mixture was amplified in duplicate using the GlobalFiler™ kit, separated by CE, and analyzed in GeneMapper® *ID-X*. A summary of the mixtures prepared is provided in Table 9.

Table 9: Summary of experimental design for additional specificity and sensitivity test mixtures containing related individuals. Case circumstances provided by JCRCL are also presented.

Code	Mixture Ratio of Contributors					
	1	2	3	4	5	Scenario
T1	1	1	1			*Incest case:* Semen stain on parents' bedding; Father is suspected of abusing Daughter.
T2	3	1	1			
T3	5	1	1			
T4	9	1	1			
T5	1	1	1			*Homicide:* Bloodstain on victim's (Father) shirt; Mother and Child are suspects.
T6	1	3	1			
T7	1	5	1			
T8	1	9	1			
T9	1	1	1			*Vehicle Theft:* Steering wheel swabs; Child not supposed to drive car.
T10	1	1	3			
T11	1	1	5			
T12	1	1	9			
S13	3	1	1	1		*Burglary:* Swabs of rock used to break window; Brother is main suspect.
S14	5	1	1	1		
S15	9	1	1	1		
S16	3	1	3	1		*Assault:* Bloodstain on victim's (Brother) shirt; unknown assailant but Sister came to aid.
S17	5	1	3	1		
S18	9	1	3	1		
S19	9	5	1	1		*Burglary:* Swabs of gun cabinet in apartment shared by two Brothers.
S20	5	9	1	1		
S21	5	1	3	3		
D22	1	1	1	1		*Vehicle Theft:* Steering wheel swabs; Parent and Child drive car.
D23	3	1	1	1		
D24	5	1	1	1		
D25	9	1	1	1		
D26	1	1	1	1		*Sexual assault:* Neck swabs; Assailant is either Father or Son (unrelated to Victim).
D27	1	1	1	3		
D28	1	1	1	5		
D29	1	1	1	9		
D30	1	1	1	1		*Vehicle Theft:* Steering wheel swabs; Parent and Child drive car.
D31	1	3	1	1		
D32	1	5	1	1		
D33	1	9	1	1		
P34	5	1	1	1		*Criminal Trespass:* Swabs of jewellery box in bathroom shared by Mother, Son, and Daughter.
P35	1	5	1	1		*Criminal Trespass:* Swabs of back door handle of residence of Mother, Son, and Daughter.
P36	1	1	5	1		*Vehicle Theft:* Swabs of hat left on scene.
P37	4	2	1	1	1	*Vehicle Theft:* Steering wheel swabs; Suspect(s) unknown.
P38	9	5	5	1	1	
P39	1	9	5	3	3	
P40	9	1	1	3	3	

During analysis, the apparent number of contributors was assigned by an experienced analyst. Given the amount of allele sharing between related individuals, the task of correctly assigning the number of contributors to such mixtures can prove to be difficult. An analyst will often under-assign the number of contributors for mixtures of close relatives. This can have the effect of excluding known contributors to the mixture, as is further explained in Section F. In the present study, the number of contributors was under-assigned for many of the mixtures examined. As discussed below, the use of conditioning profiles can assist in this regard. Each mixture was interpreted in STRmix™ using the assigned number of contributors, following which *LR*s were calculated for known donors and known non-contributors to examine sensitivity and specificity. As above, the specificity and sensitivity was tested by calculating the *LR* to each database individual for a number of experimentally designed two-, three-, four-, and five-person mixtures. A summary of these mixtures are provided in Table 9. Furthermore, some of the non-contributors were also known relatives to donors to the mixture.

Nevertheless, each profile was interpreted in STRmix™ using the apparent number of contributors assigned. These deconvolutions were compared to the known contributors and 1,000 known non-contributors using the Database Search function within STRmix™.

To calculate an *LR*, the propositions considered were:

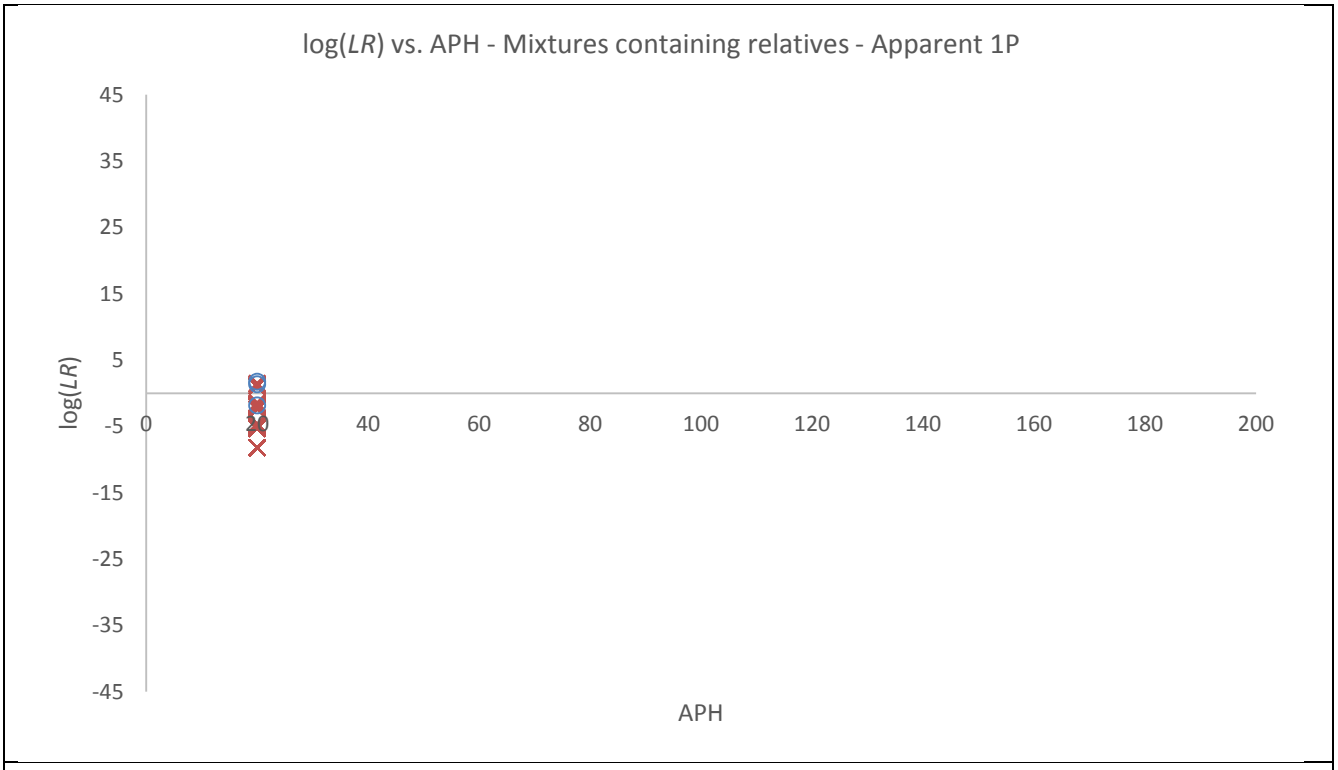
H_p: The DNA originated from the database individual and *N*-1 unknown individuals

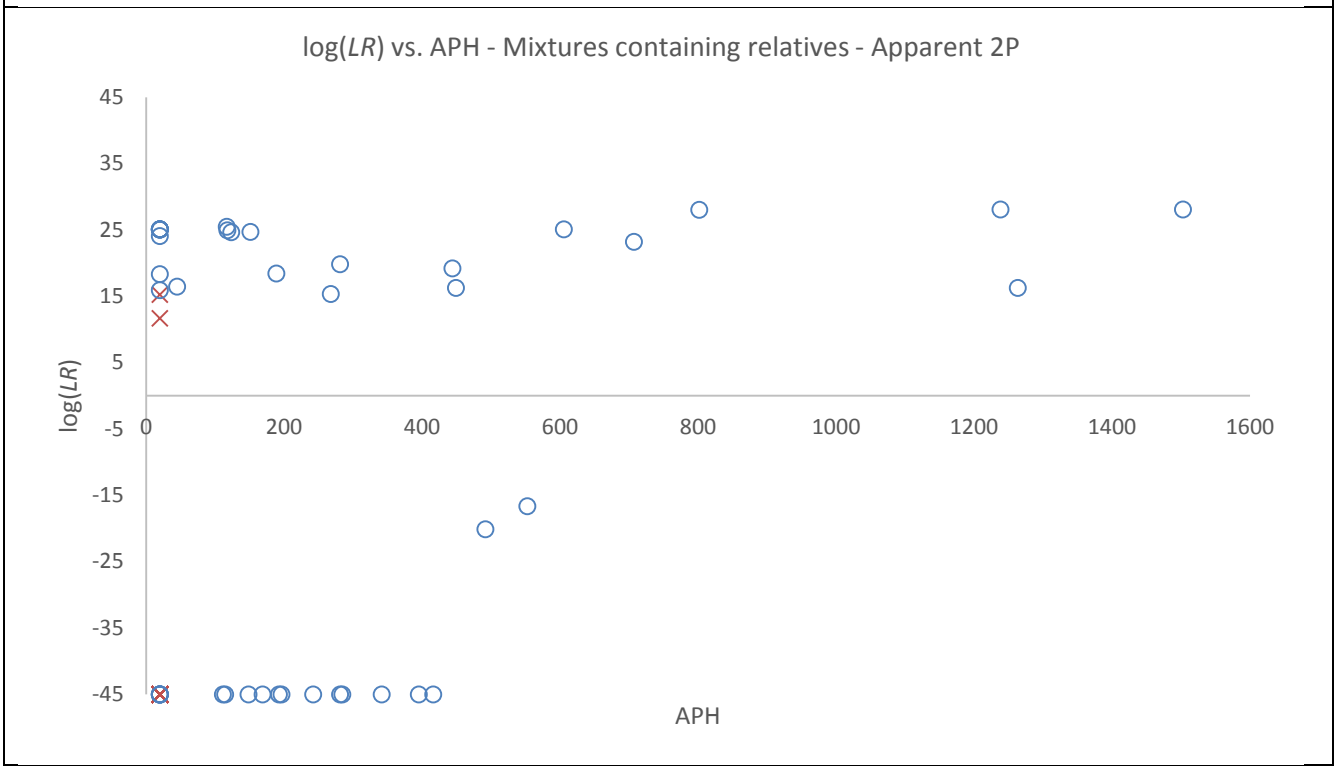
H_d: The DNA originated from *N* unknown individuals

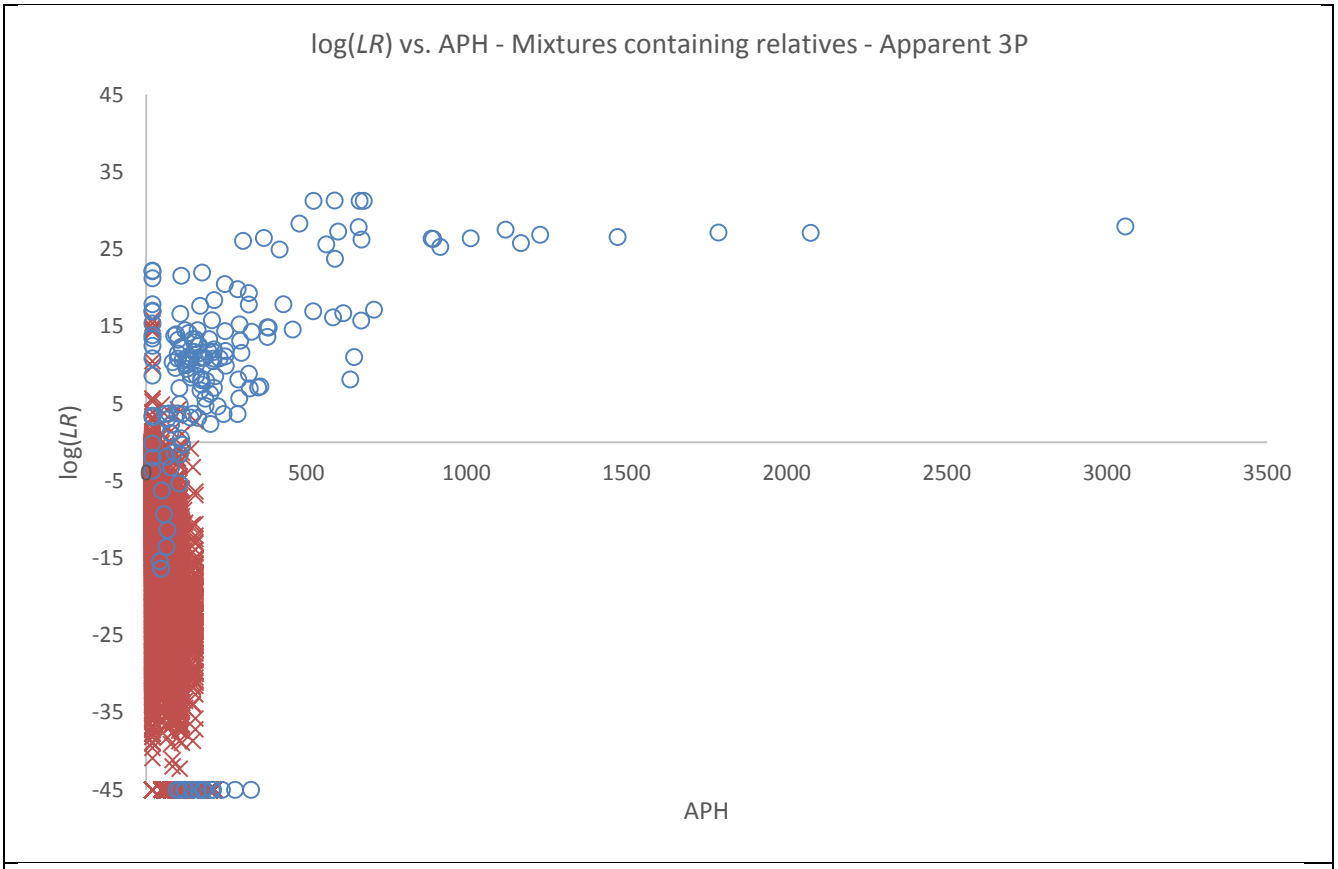
Where *N* is the number of contributors assigned to the profile. The *APH* for each contributor was again calculated, keeping in mind that it is likely that some of the *APH*s calculated will default to a value of half AT (20rfu) due to allele sharing between relatives. Plots of log(*LR*) versus the *APH* per contributor for the apparent one-, two-, three-, four-, and five-contributor profiles are provided in Figure 8.

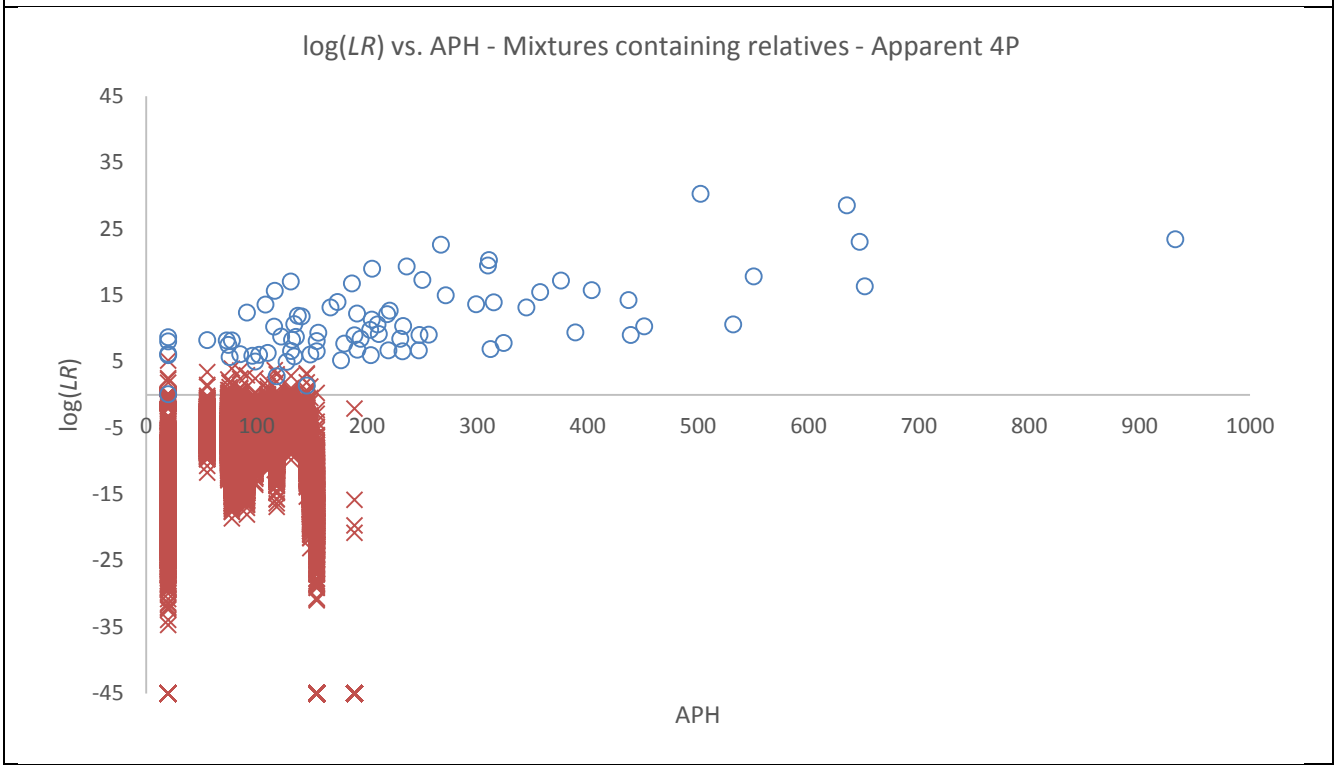
As shown in Figure 8, it may be seen that there is still good discriminatory power between the H_p and H_d true contributors. There were a number of false exclusions likely due to the under-assignment of the number of contributors to these complex mixtures of related individuals. There were also a number of non-contributors that gave *LR*s supporting H_p. *LR*s greater than 1,000 (log(*LR*) = 3) were investigated further for the apparent two, three, four and five-contributor mixtures. In most examples the non-contributor was found to be a close relative (mother, father, sibling) of at least one of the known contributors. These individuals would be expected to have very similar DNA profiles and the correspondence to components of the mixture is not unexpected.

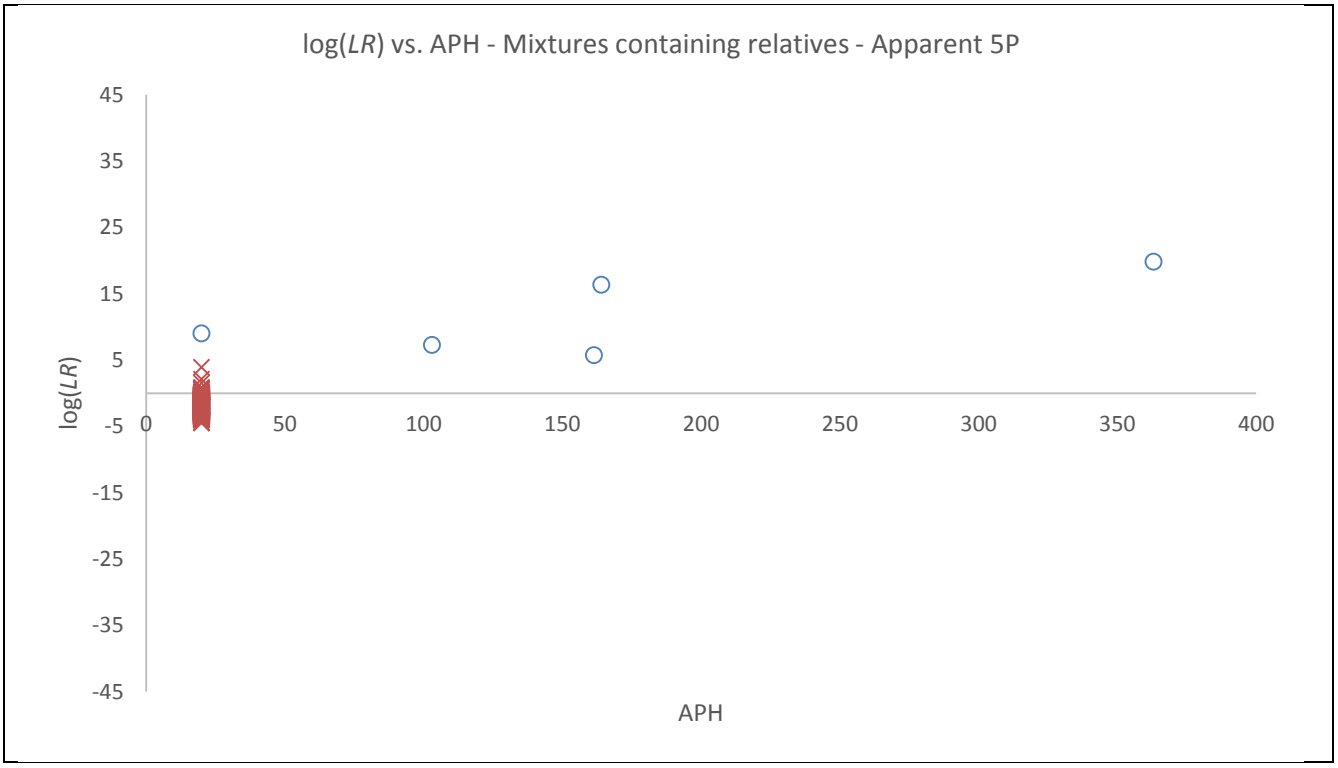
Figure 8: Log(LR) versus APH (rfu) for apparent one-, two-, three-, four-, and five-person GlobalFiler™ profiles containing related contributors. LRs for known donors are plotted as blue circles, LRs for non-contributors are plotted as red crosses


log(LR) vs. APH - Mixtures containing relatives - Apparent 1P
log(LR)
APH
45
35
25
15
5
-5
-15
-25
-35
-45
0
40
60
80
100
120
140
160
180
200


log(LR) vs. APH - Mixtures containing relatives - Apparent 2P
log(LR)
APH
45
35
25
15
5
-5
-15
-25
-35
-45
0
200
400
600
800
1000
1200
1400
1600


log(LR) vs. APH - Mixtures containing relatives - Apparent 3P
log(LR)
APH
45
35
25
15
5
-5
-15
-25
-35
-45
0
500
1000
1500
2000
2500
3000
3500


log(LR) vs. APH - Mixtures containing relatives - Apparent 4P
log(LR)
APH
45
35
25
15
5
-5
-15
-25
-35
-45
0
100
200
300
400
500
600
700
800
900
1000



In many of the mock case scenarios provided, it may be reasonable to condition the interpretation on one or more individuals. This involves interpreting the mixture under the assumption that the conditioned individual(s) is a contributor(s) under both H_p and H_d. Similar to the use of replicate amplifications, the use of conditioning profiles improves the ability of STRmix™ to reliably discriminate between true donors and non-contributors. Furthermore, consideration of the conditioning profiles can assist in assigning NoC. In the present study, the first replicate of each mixture was re-interpreted using relevant conditioning profiles to explore the effect on the interpretation and subsequent *LR* calculations. The conditioning profiles were also used to assist in the assignment of NoC. The effect on the *LR* when conditioning on a contributor is also more formerly explored in Section E.

In most circumstances investigated with the relatives dataset (with the exception of mixture P38.1), conditioning had an impact on the assignment of the number of contributors. For example, mixtures that presented as apparent two-contributor mixtures in the absence of conditioning information could be plausibly explained as originating from at least three individuals when the profile(s) of the assumed contributor(s) was considered. This change in NoC will in itself lead to changes in the *LR*. As discussed above and explained further in Section F, under-estimation of the true number of contributors can lead to false exclusions as seen in Figure 8 whereas the use of conditioning information can assist in the assessment of NoC, increase the NoC assigned relative to what may be estimated at first glance and this can lead to the appropriate genotype combinations given the true contribution being accepted.

Conditioning on an individual can also assist STRmix™ by providing it with relevant information which can refine the genotype combinations being considered. Without conditioning information, STRmix™ must consider many possible genotype combinations. However, when using the reference profiles of one or more individuals to condition on can lead to fewer combinations being accepted. This means that adventitious matches to non-contributors are less likely and discrimination between related individuals may be improved. The refined and

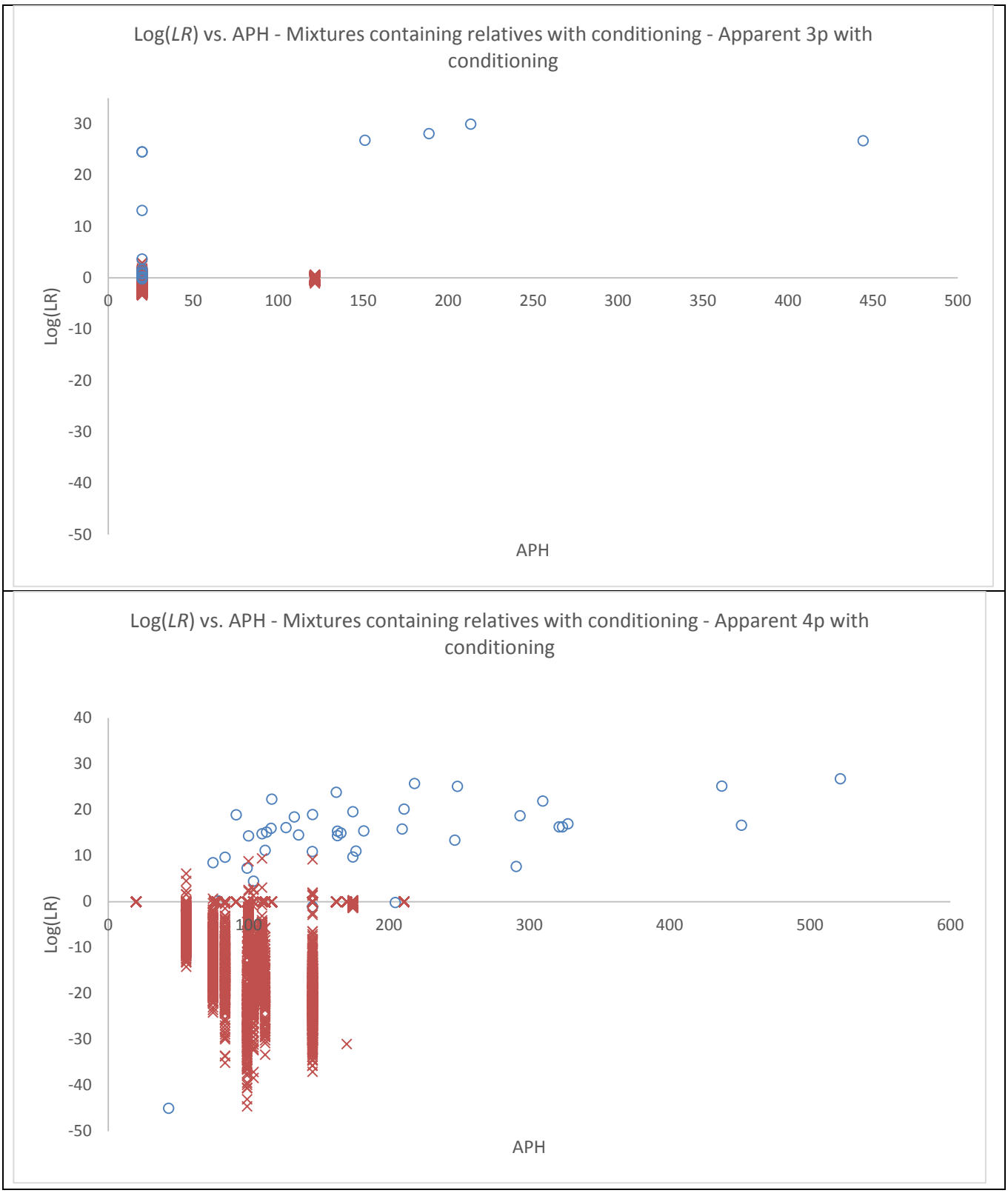
reduced number of possible genotype combinations also means that the *LR*s for true contributors may increase relative to the non-conditioned equivalent, again leading to better discrimination between true contributors and non-contributors.

The mixtures comprising DNA from related individuals were re-interpreted in STRmix™ using the conditioning information in the scenarios supplied (see Table 9). Each deconvolution was compared to the database used previously in Section D and an *LR* was calculated to each individual on this database, factoring in conditioning references into the propositions where appropriate.

The results of the testing carried out using conditioning profiles are displayed in Figure 9.

Based on the scenarios supplied, the mixtures could have been interpreted a number of ways with different individuals being selected to condition upon. In this study, STRmix™ was supplied with as much of the available information as possible (i.e. all relevant conditioning reference profiles used) with the aim of demonstrating that the provision of relevant information typically leads to better distinction between true contributors and non-contributors. One mixture was examined two ways to demonstrate this; sample P35 was interpreted conditioning on one of the individuals who could reasonably be assumed to be present (LH), and then re-interpreted conditioning on all of the individuals who could reasonably be assumed to be present (LH, BH, and KH). In the first scenario, an *LR* of 5.44E9 was calculated for true contributor JA but this increased to 3.86E19 when three conditioning reference profiles were used. The choice of which individuals to condition on (if any) must be informed by the available case circumstances. Studies examining the recovery of DNA under comparable circumstances may also assist the analyst in determining whether there is a reasonable expectation to recover an individual's DNA in the case at hand. There may be instances where there is ambiguity in whether an individual should be treated as a known contributor during interpretation; in such circumstances one approach may be to carry out a number of deconvolutions that explore the effect of conditioning on different individuals as was done in the present study for mixture P35. Ultimately, the analyst must be prepared to defend their decision to condition on an individual if challenged at court.

Figure 9: Log(LR) versus APH (rfu) for experimental design two-, three-, four-, and five-person GlobalFiler™ mixtures containing related contributors. The mixtures were interpreted in STRmix™ using relevant conditioning profiles, based on the case circumstances provided. LRs for known donors are plotted as blue circles, LRs for non-contributors are plotted as red crosses


Log(*LR*) vs. APH - Mixtures containing relatives with conditioning - Apparent 3p with conditioning
Log(LR)
APH
30
20
10
0
-10
-20
-30
-40
-50
0
50
100
150
200
250
300
350
400
450
500
Log(*LR*) vs. APH - Mixtures containing relatives with conditioning - Apparent 4p with conditioning
Log(LR)
APH
40
30
20
10
0
-10
-20
-30
-40
-50
0
100
200
300
400
500
600

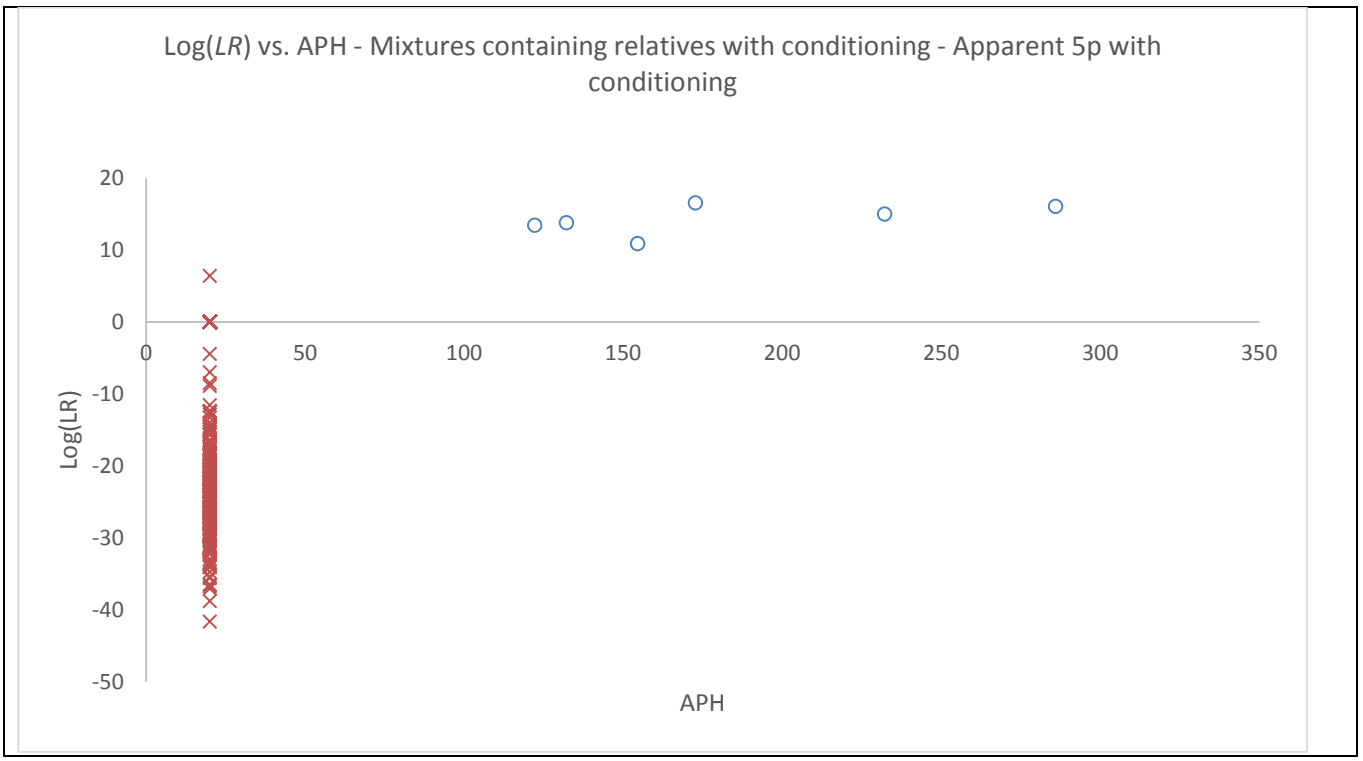
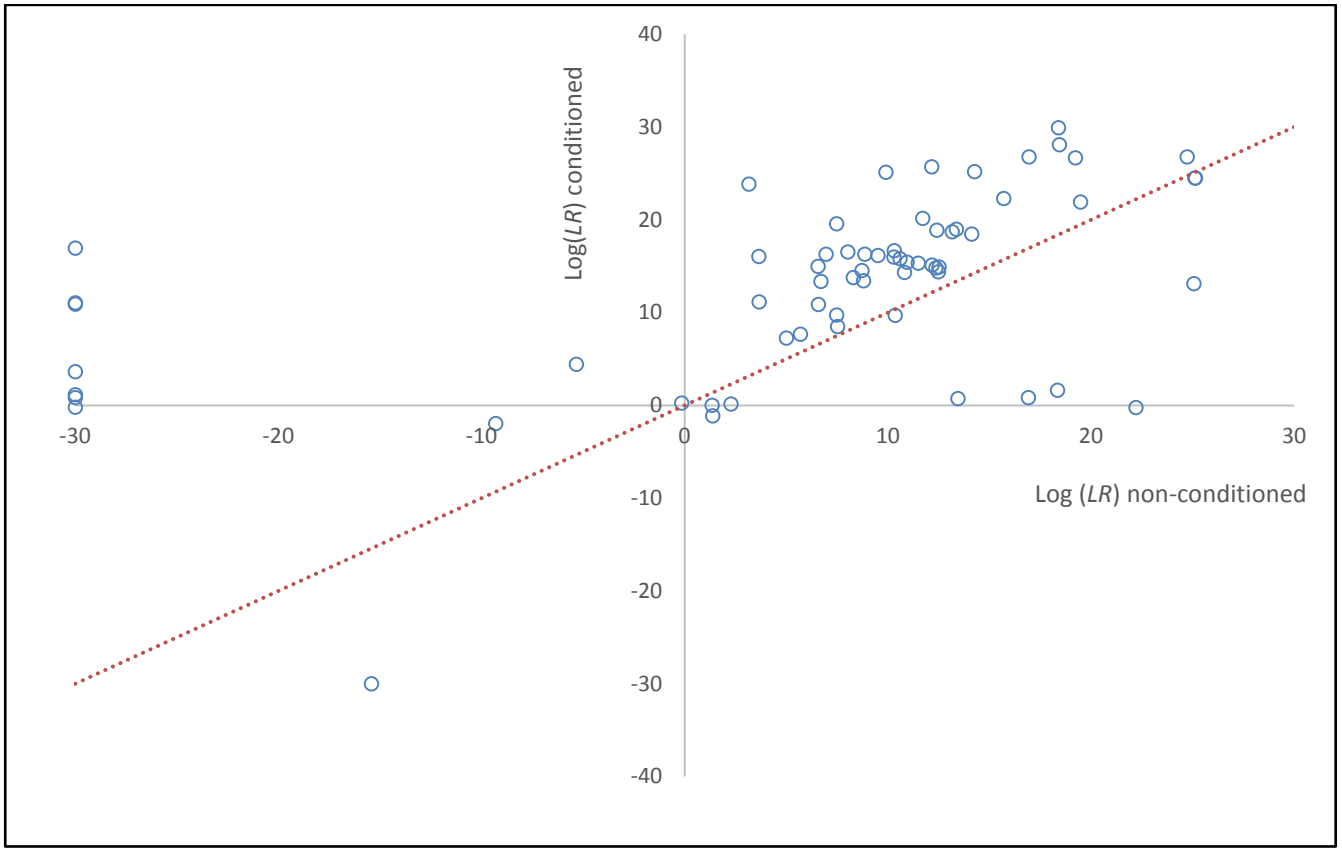


Figure 9a: Comparison of log(LR) values calculated for known donors to the related mixture dataset obtained from deconvolutions where conditioning references were not used and deconvolutions where conditioning reference samples were used.



By comparing the data in Figure 9 to the equivalent in Figure 8, it may be observed that in general when a conditioning profile is used adventitious matches to non-contributors are less likely and discrimination between related individuals may be improved. The *LR*s for true contributors may increase relative to the non-conditioned equivalent, again leading to better discrimination between true contributors and non-contributors. This is due to the generally refined and reduced number of possible genotype combinations considered and accepted when conditioning reference profiles are used in an interpretation.

However, it may be seen in the apparent four-person mixture data that even when conditioning information is used false exclusion of a true contributor and the high *LRs* in favour of inclusion for some non-contributors persist. These were explored further and are discussed below. Changes to the log(LR) of known contributors are explored further in Figure 9a, where the log(*LR)*s calculated for known contributors without using conditioning information are compared to the log(*LR)*s obtained when conditioning information is applied.

The false exclusion data point observed in the apparent four-person mixture data is known contributor PW-snr being excluded from mixture P38.1 (this data point sits below the *y=x* line in the bottom left hand quadrant of Figure 9a). The reference profiles of three of the other contributors were used to assist with the estimation of NoC; however, despite this and due to the relatively low-level nature of the profile an estimation of four rather than five contributors was made. The reduced NoC estimation and the use of the conditioning profiles restricts the range of genotype combinations that can be considered and accepted.

The false inclusions observed in the apparent four-person mixture plot displayed in Figure 9 all involved relatives of known donors to the mixtures examined. Reference BH falsely matched mixtures D28.1 (log(*LR*) =9.44), D27.1 (log(*LR*) =9.24), and D29.1 (log(*LR*) =8.79). Further inspection of the composition of these mixtures revealed that both the father (PW Sr) and brother of BH (PW Jr) are known contributors, likely leading to the large inclusionary *LR*s observed for BH. References PW Sr and LH both falsely matched mixture D33.1 with log(*LR*)s of 6.11 and 4.51, respectively. BH, who is the daughter of PW Sr and mother of LH, is one of the known contributors to this mixture. Finally, reference LH falsely matched mixture D28.1 with a log(*LR*) of 3.10. Both the maternal uncle (PW Jr) and maternal grandfather (PW Sr) of LH are contributors to this mixture. The remaining inclusionary *LR*s produced for non-contributors were all less than 1,000 and might reasonably be considered to be 'low-grade' adventitious matches.

False inclusion of a non-contributor was also observed with apparent five contributor mixture P37_1. Reference LR on the database falsely matched this mixture with a log(*LR*) of 6.43 being calculated; however, review of the mixture composition details revealed that this individual is the mother of a known contributor (BA) to the mixture.

Increases in the log(*LR*) when conditioning information is used may be observed as the data points in the top left hand quadrant of Figure 9a and above the y=x line in the top right quadrant in this plot. The individuals who correspond to the data points in the top left quadrant were excluded (plotted as log(*LR*)=-45) when initially interpreted without conditioning information but upon use of conditioning information a reassessment of the number of contributors was made and the increase in assigned NoC led to the acceptance of the genotypes of the true known contributors to these mixtures, hence a relatively high *LR* that in most cases supported inclusion.

Figure 9a supports the general trend of an increased *LR* for known contributors with the use of a conditioning profile relative to the *LR* calculated when conditioning is not performed. However, there are some outliers to this general trend where the *LR* calculated to a known contributor is lower when one or more of the

contributors was assumed to be present compared to the *LR* obtained when no conditioning information was used. These five data points sit below the line in the top right quadrant. These were investigated further and may be classified into two general groups.

Under-assignment of NoC

Upon further review of mixtures T10.1 and T6.1, it was found that these mixtures had been under-assigned in terms of NoC. The under-assignment results in a reduced and restricted number of genotype combinations but as these samples were mixtures of DNA from close family members (mother, father, and child) the genotypes of the unseen third contributor were still present, accepted by STRmix™, and led to a relatively high *LR*. However, when these mixtures were reviewed using the reference information of an individual assumed to be present the NoC estimate increased to three. Due to the peak sharing of relatives, STRmix™ struggles to see the third contributor. For mixture T6.1, STRmix™ gives little to no template to this third contributor i.e. it was placed in a trace contributor position which means that STRmix™ proposes and accepts many different combinations of the alleles present and includes drop out into the combinations. The weights get spread across these genotype combinations and therefore the *LR* is much reduced compared to the non-conditioned equivalent.

Following the interpretation of mixture T10.1 a high Gelman Rubin (GR) value was noted (>13) which could indicate non-convergence of the MCMC chains. It was also noted that the posterior mean allele variance value appeared elevated relative to the mode of the prior distribution. Re-interpretation of this sample was attempted with increased accepts (burn-in and post burn-in accepts were increased by a factor of ten) with the intention of allowing STRmix™ longer to explore the probability space and to see if this would encourage the chains to converge possibly leading to lower GR and posterior allele variance values. Whilst the GR did decrease slightly (GR = 11) it was still well in excess of the advised value of 1.2. Closer inspection of the input file in combination with the conditioning reference profiles also revealed that the mixture proportions determined by STRmix™ were not intuitive. Given the genotypes of the conditioning reference profiles and the peak heights present the mixture ratio appeared to be approximately 1:1:3 or 1:1:4 (these proportions were apparent at the vWA and SE33 loci).

When unintuitive mixture proportions are encountered such as those described for mixture T10.1, in addition to the assumption of the presence of certain individuals it may also be appropriate to consider what mixture proportions may be expected or evident in the profile. Due to the apparent difference between STRmix™ mixture proportion priors and analyst expectation, this mixture was investigated further using the M_x Priors function of STRmix™. By default, mixture proportions are uninformed in STRmix™; however, a user can elect to set priors on the mixture proportion when setting up a run. This could be based on case circumstance information combined with information in the electropherogram or based on the observation of sub-threshold information that is apparent to the analyst but not STRmix™. In the latter scenario, an additional contributor may be present but below AT, meaning that this information is not available to STRmix™ as it is not present in the input file. In the absence of this information STRmix™ could model the second contributor by splitting the template equally above threshold leading to inappropriate genotypes. However, the information visible to the analyst sub-threshold may indicate otherwise and can be useful in guiding STRmix™ to appropriate mixture proportions.

When mix T10.1 was reinterpreted using M_x priors of 0.16:0.16:0.66 and conditioning reference profiles, the GR was much reduced compared to the previous interpretations (1.05) and the posterior allele variance value sat

within the prior gamma distribution. When compared to the database, the LR calculated to the known third contributor was 3.3E24 which is much closer to the LR calculated for the original interpretation (1.2E25). This observation along with the case circumstances could prompt a user to explore the use of M_x Priors.

Relatedness

Similar to the deconvolutions discussed above as under-assigned, some mixtures where there are closely-related individuals contributing DNA in fairly equal amounts can also lead to some issues even when conditioning reference profiles are used. In the deconvolutions of mixtures T9.1, T5.1, and T2.1, the conditioning profiles could account for half of the peaks present meaning STRmix™ modelled the third contributor as a very trace contribution due to the high degree of peak sharing. This again means that many genotype combinations are accepted and the weight spread across these. In the non-conditioned runs, STRmix™ has a little more freedom to explore genotype combinations and more of the template is allowed for the third contributor leading to fewer genotype combinations and less dispersion of the weight.

Two of these mixtures, where it appeared that given the case circumstances of closely-related individuals could have donated DNA in approximately equal amounts were re-interpreted using apparent *N*, conditioning profiles, and M_x Priors. The *LR*s of the true contributors increased in comparison to the non-conditioned and conditioned (without M_x Priors) equivalents. The results produced are summarized in Table 10. During review of the interpretation of T9.1 it was also noted that there was a high posterior mean allele variance value likely due to apparent peak height differences caused by stacking of shared alleles due to relatedness. Elevated posterior variance values relative to the mode of their prior distribution can act as an indicator that the profile is not being explained well by STRmix™ and further scrutiny of interpretation criteria and results could be warranted.

Table 10: Comparison of the log(LR) values produced using different proposition sets and M_x priors

Sample name	Known contributor considered as POI	Log(*LR*) - no conditioning	Log(*LR*) - conditioning	M_x Priors used	Log(*LR*) - conditioning & M_x Priors
T5.1	RL	16.91775	0.865643	0.33 : 0.33 : 0.33	17.792
T9.1	EC	22.20615	-0.20786	0.33 : 0.33 : 0.33	24.422

To summarize the work carried out in Section D, good differentiation between true contributors and non-contributors has been observed at high APH for a range of mixtures, including mixtures containing related individuals. As average peak height diminishes or profile complexity increases it becomes more difficult to distinguish true contributors from non-contributors. It is highly recommended that relevant case circumstances be used to improve the outcome of STRmix™ interpretations. This can be in the form of assuming the presence of a contributor to assist with the assessment of the number of contributors and in the deconvolution. The use of replicate amplifications can also assist STRmix™ to deconvolute DNA profiles and, where appropriate, the use of M_x priors can also assist. In certain circumstances, however, conditioning on an expected contributor(s) may result on the false exclusion of a low level true contributor(s) if they are related to the assumed contributor because STRmix may not be able to differentiate allele sharing from normal allele height variance.

Section E: Alternative 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 selection of profiles from Section D were re-interpreted in STRmix™ using alternative propositions. In these interpretations, one of the contributors was assumed under both H_p and H_d. This would be akin to assuming a donor's DNA were present on, for example, an intimate sample. The following twenty samples were selected to cover a range of templates, mixture proportions, and complexity:

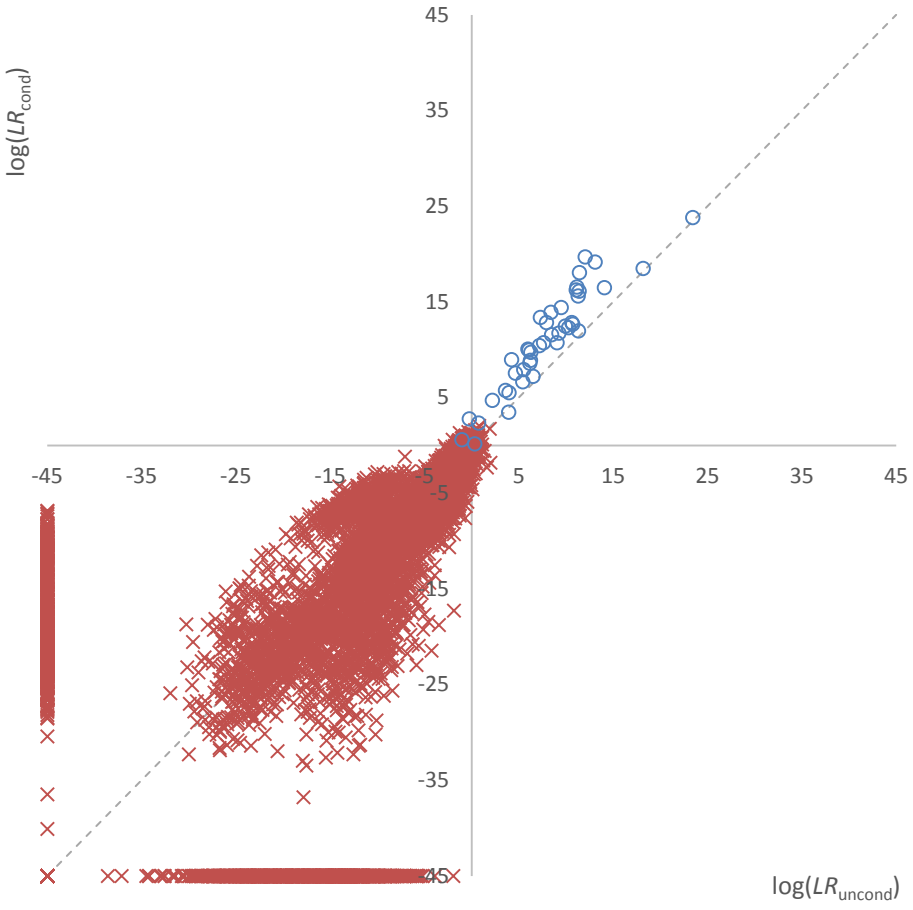
GF-MX_Mix1_C1_28c_15s_06_F01_01.hid
GF-MX_Mix1_D1_28c_15s_08_H01_01.hid
GF-MX_Mix4_B1_28c_15s_01_A06_06.hid
GF-MX_Mix4_D1_28c_15s_05_E06_06.hid
GF-MX_Mix4_E2_28c_15s_08_H06_06.hid
GF-MX_Mix7_B1_28c_15s_07_G10_10.hid
GF-MX_Mix7_C2_28c_15s_02_B11_11.hid
GF-MX_Mix7_D1_28c_15s_03_C11_11.hid
GF-MX_Mix8_C1_28c_15s_06_F01_01.hid
GF-MX_Mix8_D1_28c_15s_08_H01_01.hid
GF-MX_Mix11_B1_28c_15s_01_A06_06.hid
GF-MX_Mix8_E1_28c_15s_02_B02_02.hid
GF-MX_Mix11_C1_28c_15s_03_C06_06.hid
GF-MX_Mix11_D1_28c_15s_05_E06_06.hid
GF-MX_Mix11_E2_28c_15s_08_H06_06.hid
GF-MX_Mix12_C2_28c_15s_03_C08_08.hid
GF-MX_Mix12_D2_28c_15s_RT_04_D03_03.hid
GF-MX_Mix12_E2_28c_15s_07_G08_08.hid
GF-MX_Mix1_A1_28c_15s_02_B01_01.hid
GF-MX_Mix1_B1_28c_15s_04_D01_01.hid

Following interpretation, *LR*s were calculated for the remaining known donors and known non-contributors to examine sensitivity and specificity. The propositions considered were:

> H_p: The DNA originated from the assumed individual, the database individual and N-2 unknown individuals
>
> H_d: The DNA originated from the assumed individual and N-1 unknown individuals

The *LR*s calculated were compared with those produced in Section D where the same mixtures were interpreted without the use of conditioning profiles. The results obtained have been plotted in Figure 10.

Figure 10: Comparison of log(LRs) obtained from mixtures when assuming a contributor (conditioned) to when there is no person assumed (i.e. unconditioned, the original Section D result). LRs for known donors are plotted as blue circles, LRs for non-contributors are plotted as red crosses



Values above the line at $y=x$ for the H_p true *LR*s indicate that the *LR* generally increases when conditioning on, or assuming, a true contributor. In the H_d true data there is a spread of data points around the $y=x$ trend line, this is not unexpected as more variation is generally observed with low *LR*s (due to more uncertainty in the lower weighted genotypes). It is also possible to see in Figure 10 that a proportion of the data points are positioned below the $y=x$ trend line, at the bottom of the bottom left hand quadrant. These data points demonstrate that the use of a conditioning profile can also decrease the *LR* for non-contributors or else drive them, correctly, to outright exclusion. This shows that the addition of more relevant information (such as the addition of assumed contributors) typically improves the performance of STRmix™. MCMC variability is one likely explanation for the set of H_d true *LR*s that have gone from outright exclusion without conditioning to favouring exclusion (far left hand side of the bottom left quadrant). In the majority of instances, the change in numerical values equates to a move from exclusion to very strong support for the defence hypothesis.

In summary, the plots indicate that the addition of correct conditioning profiles (known contributors under both H_p and H_d) can improve the ability of STRmix™ to distinguish true contributors from non-contributors.

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.

In casework, the true number of contributors to a questioned profile is always unknown. Analysts are likely to add contributors in the presence of ambiguous peaks such as artefact peaks or inflated stutter peaks. The assumption of one fewer contributor 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 in family scenarios where mixtures of related individuals might reasonably be expected to be recovered.

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 [13]. The inclusion of an additional contributor beyond that 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.

This effect was investigated on JCRCL data by comparing STRmix™ interpretations where the apparent number of contributors to a mixture (as assigned in section D above) differs from the experimental design number of contributors to the mixture. For the purposes of Section F, *N* is defined as the experimental design number of contributors. *N*+1 indicates the false assumption of one additional contributor than *N* whereas *N*-1 indicates the false assumption of one fewer contributor than *N*.

Addition of one contributor

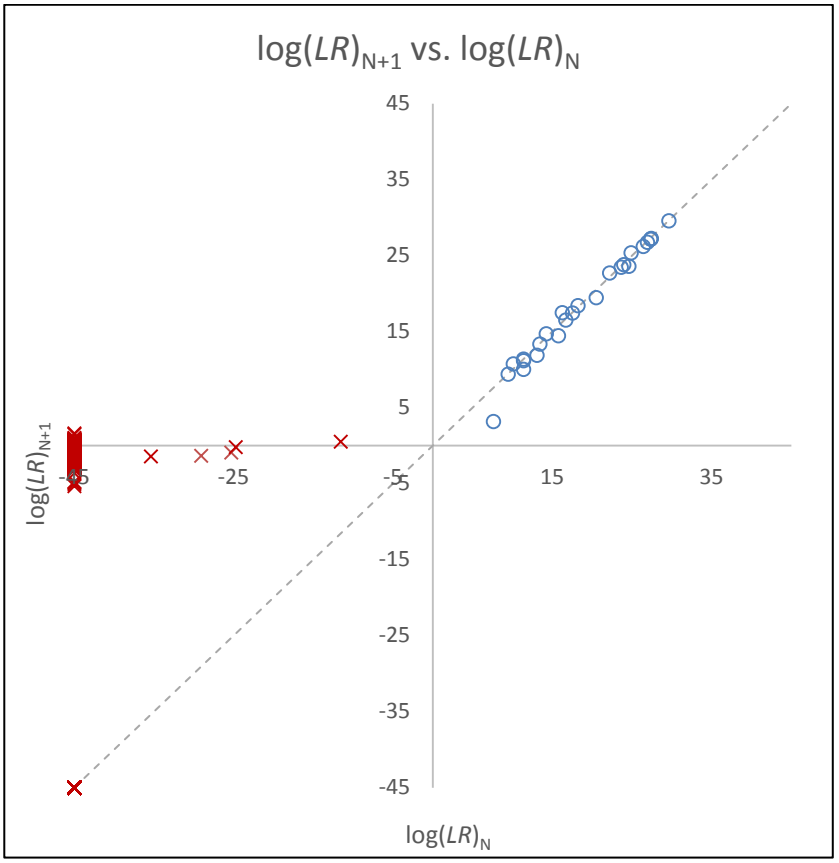
During the analysis of the profiles in Section D, there was one mixture where the experienced analyst assigned one more contributor than the experimental design number of contributors (hence, *N*+1). This mixture was Mix3_B2 and was overestimated due to potential imbalances with putative heterozygote pairs at some loci. Nine other mixtures were also selected to demonstrate this effect of over-assigning the number of contributors.

These mixtures were:

- Mix5_D1
- Mix10_A2
- Mix10_B1
- Mix8_B1
- Mix12_C2
- Mix3_B1
- Mix4_A1
- Mix5_B1
- Mix5_C1

Once again, the *LR*s for each database individual were calculated using the same parameters as in Section D. These *LR*s were compared to those calculated in Section D. The results produced are provided in Figure 11.

Figure 11: Comparison of log(LR) obtained when profiles were run in STRmix™ as N versus N+1 contributors for the known and non-contributors (where H_p true is represented by blue circle data points and H_d true are represented by red crosses)



The plot above demonstrates that there is no significant effect on the *LR* of the 'true' major or clear minor contributors when the number of contributors is overestimated. Some of the mixtures, for example Mix5 D1, Mix3 B1, and Mix5 B1 have a trace contributor with mixture proportion 5% or less. Even when an additional contributor was added, there was not a substantial effect on the *LR*s calculated to the known contributor. Even the known contributor 'outlier' that sits below the $y=x$ trendline was investigated further and found to be likely due to an unconverged MCMC chains in the original analysis. This was identified as a slightly elevated Gelman Rubin value and was subsequently re-run in STRmix™ using an increased number of accepts (per chain burn-in and post burn-in accepts were increased by a factor of 10). Once re-interpreted the deconvolution, which had a GR value less than 1.2, was then compared to the database and a likelihood ratio of 4.23E8 was calculated for known contributor 1230. This would bring this data point back around the $y=x$ trendline. The inclusion of an additional contributor beyond that present in the profile can have the effect of decreasing the log(*LR*) for known donors present at low levels although this was not observed in this study. This is because STRmix™ usually adds the additional (unseen) profile at low DNA amount (template) levels, diffusing the genotype probabilities. For H_d true comparisons, the inclusion of an additional contributor beyond that present in the profile has the effect of increasing the log(*LR*). This may be observed in the plot above with the data points on the far left of the *x*-axis (x= -45), here H_d true contributors that were excluded in the analyses under the assumption of *N* contributors either produced *LR*s that range from moderate support to for H_d to limited support for H_p using the proposed

SWGDAM verbal scale. Again, this is due to STRmix™ adding the additional (unseen) contribution at low DNA amount (template) levels. This leads to many genotype combinations being accepted at each locus thus allowing 'adventitious' matching of non-contributors. Overestimating the number of contributors did not result in any significant changes from *LR*s supporting exclusion to *LR*s supporting inclusion for any non-contributors in this limited selection of samples.

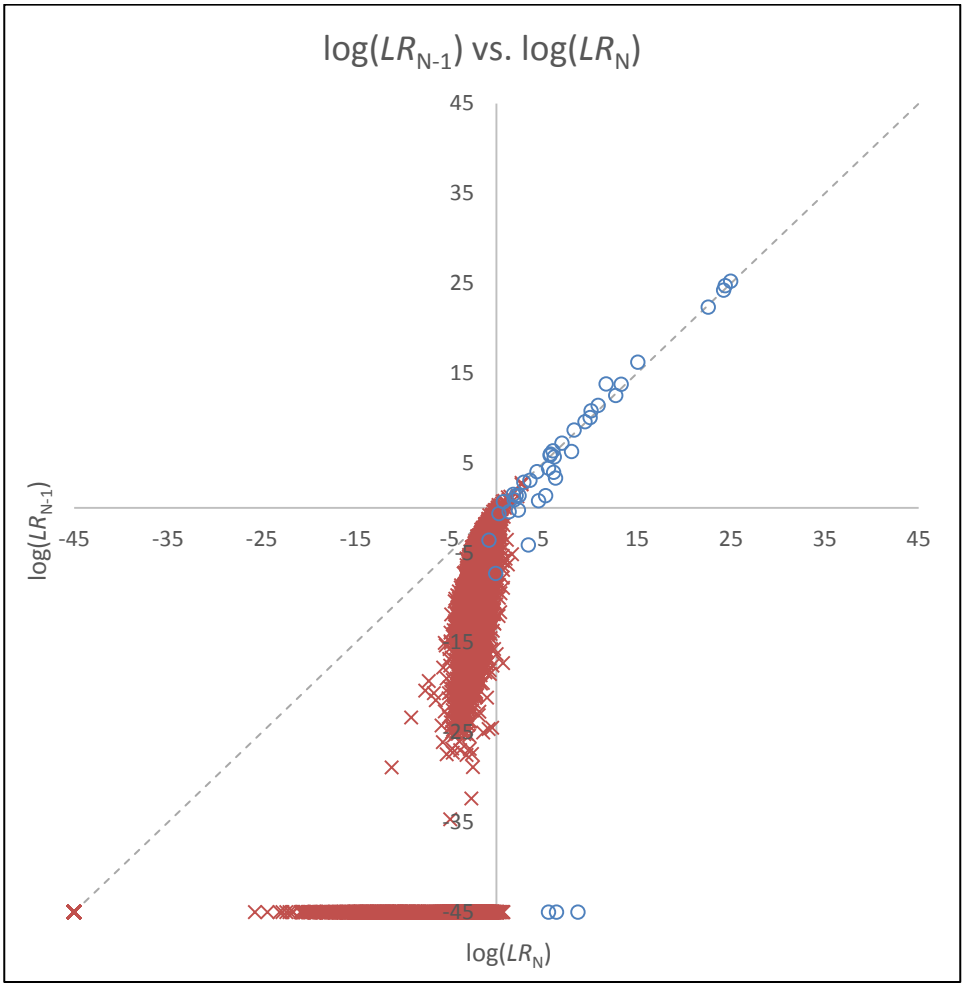
Subtraction of one contributor

During the analysis of the profiles in Section D, there were nine mixtures where the analyst assigned one less contributor than the experimentally designed number of contributors.

Each of these profiles was re-interpreted assuming the experimental design number of contributors (*N*). *LRs* for both the known contributors and known non-contributors (as 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* (Section D) and *N*-1 contributors. A plot of log(*LR*) for *N* and N-1 is provided in Figure 12 (where H_p true is represented by blue circle data points and H_d true is represented by red crosses).

Figure 12: Comparison of log(LR) values for true and non-contributors under the assumption of N and N-1 contributors, where N represents the experimental design number of contributors



The plot above demonstrates that, in general, there is no significant effect on the *LR* of the 'true' major or clear minor contributors to the mixture if the number of contributors is underestimated. However, for weaker contributors or profiles where there are close proportions and ambiguity, false exclusions or results that favour exclusion can occur. This is not unexpected as STRmix™ has to restrict the number and range of genotype combinations given fewer proposed contributors than are truly present and hence there is potential for some genotype combinations of the true contributors to not be accepted. For the same reason, underestimating the number of contributors can result in lower *LR*s for H_d true comparisons and can lead to outright exclusions of individuals that under the assumption of *N* contributors lead to non-zero *LR*s. Minor differences above and below the $y=x$ line are likely due to MCMC run variability.

Section G: Drop-in

This section covers the following standard:

4.1.8. Allele drop-in

Observed drop-in rates at JCRCL were modelled with the lab-specific parameters determined within the JCRCL STRmix™ kit file. Four experiments were undertaken to test these settings. In the first experiment, a realistically-sized (95rfu) drop-in peak was artificially added to locus D16S539 of profile 'GF_M3-B1_2ng_28c_15s_06_F09_09.hid', a *high template* single-source profile that had previously been interpreted using STRmix™. The height of the added peak is less than the JCRCL drop-in cap of 120rfu. The profile was interpreted as a single-source profile and an *LR* calculated for the known contributor using the propositions:

H_p: The DNA originated from the person of interest

H_d: The DNA originated from an unknown individual

As expected, STRmix™ modelled the additional peak as drop-in as its height was below the drop-in cap and because pairing it with the D16S539 12 allele detected (peak height: 7124rfu) would involve unreasonable allele variance. The resulting *LR* was identical to the *LR* calculated from the unmodified profile. This drop-in peak is written to the genotype probability txt output as drop-in.

In the second experiment, a realistically-sized (82rfu) drop-in peak was artificially added to the D19S433 locus (a 10 peak) of 'GF_M3-G1_0.063ng_28c_15s_06_F11_11.hid', a *low template* single-source profile that had previously been interpreted using STRmix™. The profile was interpreted as a single-source profile and an *LR* calculated using the same propositions as detailed above. In this example, STRmix™ accepted a number of genotypes and either considered the added peak to be allelic in origin or a drop-in peak. As expected STRmix™ modelled the additional peak as both drop-in and a true allele as it was of a similar height to the low template alleles at that heterozygote locus (<120 rfu). As the peak heights of the naturally observed peaks were lower than the drop-in cap used in the JCRCL STRmix™ kit, drop-in was naturally being considered with the peaks present. When the input file was edited to include the 10 peak at D19S433 this also was considered both allelic and as putative drop-in allele. This is the expected behaviour given the height of the peak. The resulting *LR* was slightly lower than the *LR* calculated from the unmodified profile; however, as drop-in was already being considered at this locus using the unmodified input this difference may be due to MCMC variation.

In the third experiment, a drop-in peak with height exceeding the JCRCL drop-in cap was added to a heterozygous locus (TPOX 12, height = 150rfu) of 'GF_M3-B1_2ng_28c_15s_06_F09_09.hid'. As the added peak was not in a stutter position and exceeded the drop-in cap it must be considered as part of the allelic

component of the profile. As expected, an interpretation could not be progressed by STRmix™ as the profile could no longer be explained by one contributor.

Finally, a 20 peak with a height of 55rfu (below the JCRCL drop-in cap) was added at locus D3S1358 of a two-person mixture with a mixture ratio of approximately 10:1. Prior to modification, D3S1358 presented with style AB:C, where AB represents the major contributor with genotype 15,17 and C represents a single 18 allelic peak from the minor contributor. During interpretation, STRmix™ assigned the most weight to a genotype of 18,20 for the minor contributor, but also accepted several other genotypes in which the 20 peak was modelled as drop-in. Given the DNA data at this locus, the genotype combinations accepted and their associated weights were intuitive. A summary of the results obtained in this section are displayed in Table 11.

Table 11: Summary of the tests carried out to investigate drop-in parameters in STRmix™

Experiment	Sample name	Locus edited	Allele(s) present (height in rfu)	Peak added	Height (rfu)	Accepted genotype(s)	Weight		Original *LR* from unedited input	*LR* when drop-in included
1	GF_M3-B1_2ng_28c_15s_06_F09_09.hid	D16S539	12	14	95	12,12	1.0		6.1885E27	6.1885E27
2	GF_M3-G1_0.063ng_28c_15s_06_F11_11.hid	D19S433	14 & 15.2	10	82	14,15.2 10,14 10,15.2 14,14	0.468 0.273 0.259 0.00002		6.0371E25	2.6969E25
3	GF_M3-B1_2ng_28c_15s_06_F09_09.hid	TPOX	8 & 10	12	150	Unable to progress interpretation under the assumption of two contributors due to three possible allelic peaks with height > drop-in cap				
4	GF_Mx_Mix6_C1_28c_15s_04_D09.09.Hid	D3S1358	15,17,18	20	55	Major	Minor			
						15,17	18,20	0.98405	1.57E16	3.9966E13
						15,17	18,18	0.00888		
						15,17	17,18	0.00350		
						15,17	15,18	0.00264		
						15,17	16,18	0.00047		
						15,17	14,18	0.00032		
						15,17	Q,18	0.00015		

Section H: Forward and reverse stutter

This section covers the following standard:

4.1.9. Forward and reverse stutter

STRmix™ models stutter peaks in a probabilistic fashion using per-allele stutter models. The models used are based on the allelic designation of the parent allele. Where allelic designation is found to be a poor descriptor for stutter ratios at a given locus, the longest uninterrupted stretch (LUS) of repeats within the parent allele can instead be used. Alternatively, per-allele or per-locus average observed stutter ratios may be utilized. Within STRmix™ V2.6, any type of stutter variant may be modelled, including back stutter, forward stutter, double back stutter, and partial stutter products such as the minus 2 base pair stutter variants commonly observed at the D1S1656 and SE33 loci. To model stutter appropriately, STRmix™ requires that stutter peak labels be retained during profile analysis within GeneMapper® *ID-X*. This ensures that stutter peaks are written to the STRmix™ input file and are available for consideration during profile interpretation.

As the height of a peak in a stutter position increases, we expect more weight to be placed on genotype combinations that consider the peak to be at least partly allelic in origin. The modelling of stutter is particularly important when interpreting mixed DNA profiles where one or more minor contributors are present at similar levels to stutter peaks from the major contributor(s). Throughout the present study, stutter appeared to be modelled appropriately using the parameters and models described within the document 'Estimation of STRmix™ parameters for Jefferson County Regional Crime Laboratory (GlobalFiler™ 3500)'. By way of example, the genotypes accepted and their associated weights are reproduced for one locus of a two-person mixture with ratio ~10:1 (Mix6 A1) (Table 12). The STRmix™ input file at this locus has also been provided (Table 13).

Table 12: Genotypes accepted at D13S317 in a two-person mixture (Mix6 A1) where the minor contributor is present at similar levels to the back stutter peaks from the major contributor. The mixture ratio is approximately 10:1

Major contributor	Minor contributor	Weight
9,11	9,12	0.39942
9,11	11,12	0.33723
9,11	12,12	0.26295
9,11	10,12	3.1496E-4
9,11	9,11	5.0788E-5
9,11	9,9	4.0923E-5

Table 13: STRmix™ input file for Mix6 A1 at locus D13S317

Allele	Height	Size
8	123	210.84
9	4898	214.87
10	204	218.95
11	4751	223.01
12	511	227.08

Inspection of Table 12 reveals that STRmix™ has accepted a range of genotypes for the minor contributor, with most of the weight assigned to genotypes that include a 12 allele. This is expected; although the 12 peak is in a

forward stutter position, the observed forward stutter ratio (10.76%) is well above the expected forward stutter ratio of 0.492%. Nevertheless, the observed forward stutter ratio is below the maximum allowable forward stutter ratio of 15%, causing STRmix™ to accept genotypes where the 12 peak has been modelled as originating solely from forward stutter, albeit with very low weight. Note that STRmix™ has also considered a minor genotype of 10,12 (i.e. proposing that the 10 peak, which is in a back stutter position, is allelic) however has assigned it a fairly low weight given the peak heights observed.

Section I: Intra locus peak height

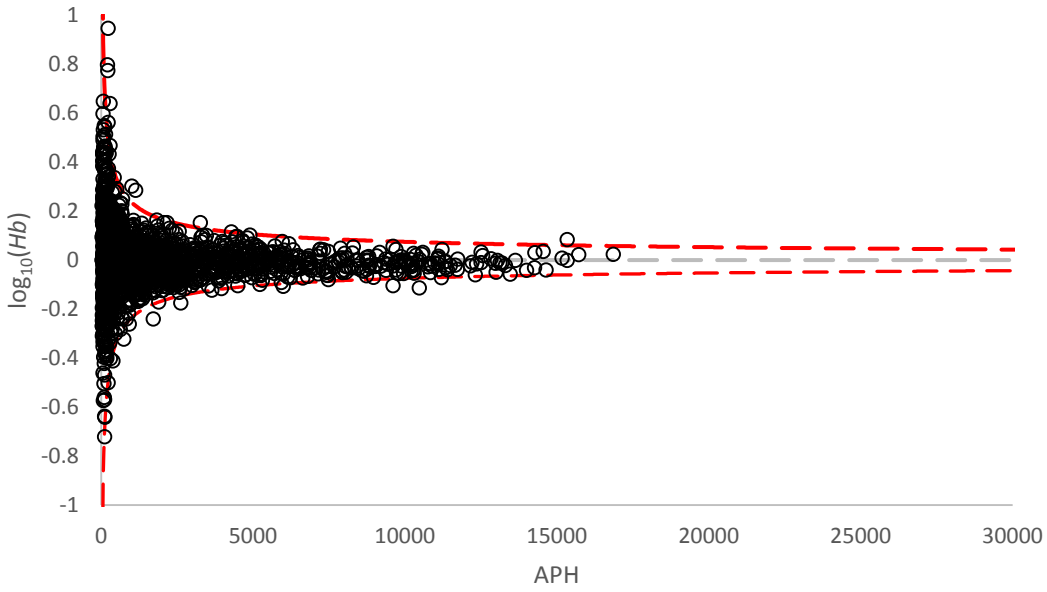
This section covers the following standard:

4.1.10. Intra-locus peak height variance

STRmix™ models the variability of single peaks. Peak height variance models are determined using laboratory data. This is undertaken within STRmix™ using the Model Maker function. Traditionally we investigate heterozygote balance (*Hb*), which can be thought of as the variability of two alleles at a heterozygous locus. A plot of log(*Hb*) versus average peak height (APH) of a locus demonstrates that the variability in *Hb* decreases as APH increases. The performance of Model Maker is checked by plotting the bounds informed by the Model Maker results (refer to the JCRCL Laboratory STRmix™ Implementation report for further details).

The plot of log(*Hb*) versus APH and the expected 95% bounds (plotted as dotted lines) calculated by $\pm\sqrt{2}\times 1.96\times\sqrt{\frac{c^2}{APH}}$ where c^2=7.23, the 75th percentile of the allele variance prior gamma distribution determined for JCRCL GlobalFiler™ data using Model Maker, is displayed in Figure 13. The plot in Figure 13 is an approximate check of the parameters derived through Model Maker.

Figure 13: Plot of log(Hb) versus APH for GlobalFiler™ data from the Jefferson County Regional Crime Laboratory



Section J: Inter-Locus peak heights

This section covers the following standard:

4.1.11. Inter-locus peak height variance

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, 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. Individual LSAE values for each locus are provided in the STRmix™ results following interpretation. LSAE and per-locus APH can be plotted to demonstrate the relationship between these two values. Similar trends should be observed between LSAE and APH. This was demonstrated by interpreting a single-source GlobalFiler™ profile (GF_M-D1_0.5ng) within STRmix™ and plotting the resulting LSAE values along with the per-locus APH (Figure 14). Individual loci have been plotted along the x-axis in order of increasing molecular weight. Inspection of Figure 14 demonstrates the expected relationship between APH and LSAE.

The same single-source GlobalFiler™ input file was then ***artificially inhibited*** in MS Excel by editing the peak heights at 4 loci. These 4 loci (D3S1358, D18S51, D13S317, and D2S1338) were 'inhibited' by reducing peak heights by 60%; the remaining loci were not edited. The edited profile was then interpreted within STRmix™ and the LSAE and APH values plotted as before (Figure 15). Inspection of Figure 15 demonstrates that the expected relationship between LSAE and APH holds for inhibited profiles.

Figure 14: Plot of APH and LSAE value for each locus for a single-source GlobalFiler™ profile

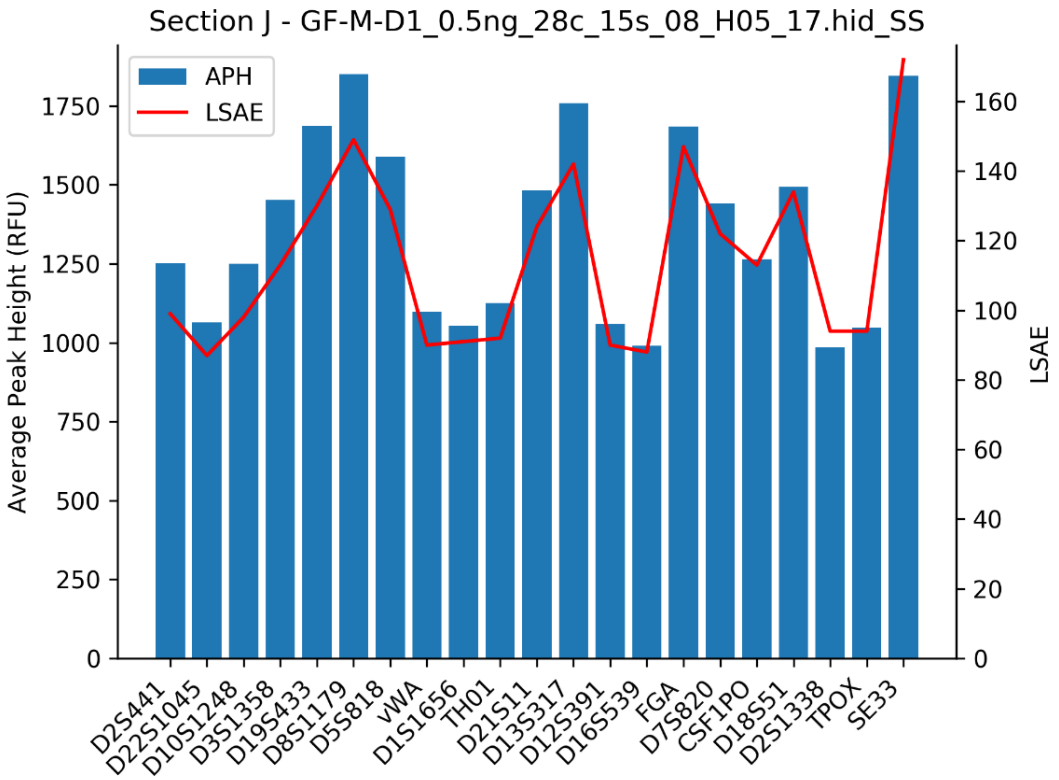
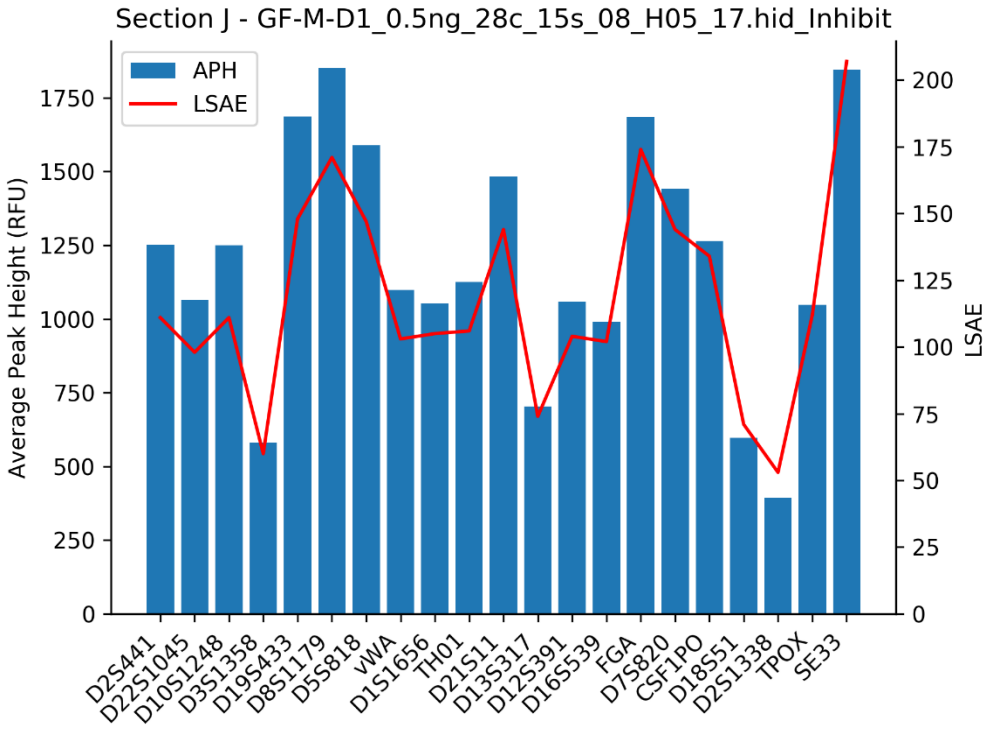
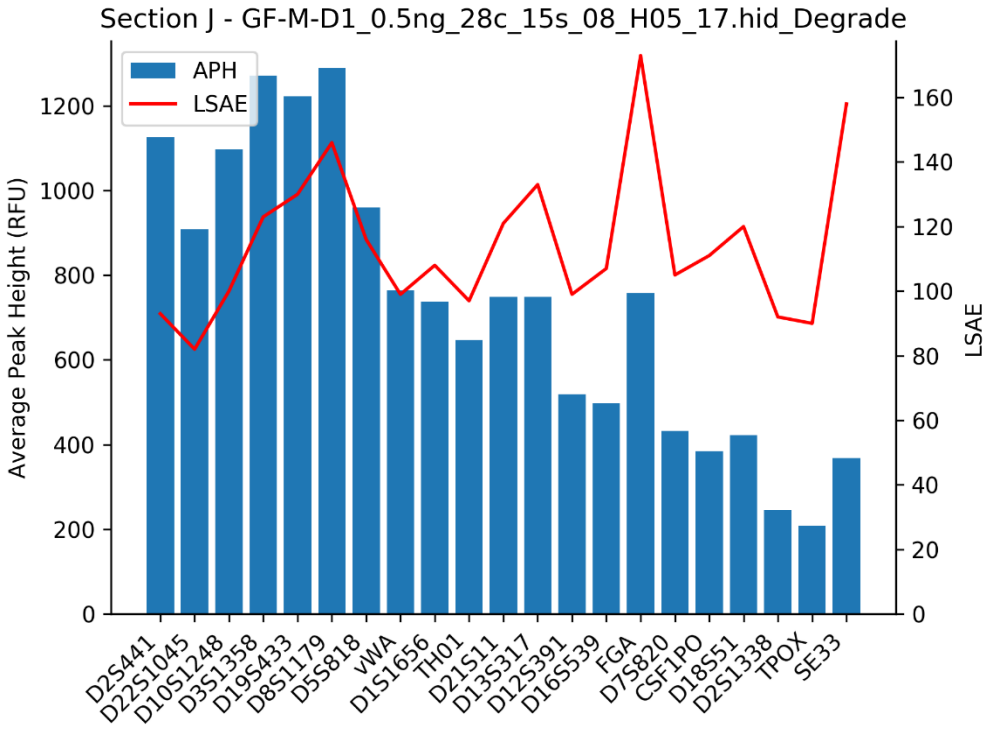


Figure 15: Plot of APH and LSAE value for each locus for a single-source GlobalFiler™ profile with artificial inhibition at D3S1358, D8S1179, D18S51, and D13S317



The original profile was ***artificially degraded*** by reducing the peak heights in MS Excel. High molecular weight loci were degraded by up to 80% whereas low molecular weight loci were only 'degraded' by 5%. As before, the edited profile was then interpreted using STRmix™ and the LSAE and APH values plotted (Figure 16). The effect of degradation is observable within Figure 16 by the decrease in APH with increasing marker size. LSAE values are independent of this degradation and the individual locus efficiencies can be seen by the LSAE values. This is most obvious at SE33.

Figure 16: Plot of APH and LSAE value for each locus for a single-source GlobalFiler™ with artificial degradation



Inspection of Figure 16 demonstrates that the expected relationship between APH and LSAE holds even with degraded samples. As degradation is also taken into account in the STRmix™ biological model the LSAE values in this example present differently to the APH values at some loci.

The average of the post burn-in degradation values (linear approximation) for the original profile was 0.773 rfu/bp. In contrast, the same parameter was 4.538 rfu/bp for the artificially degraded profile.

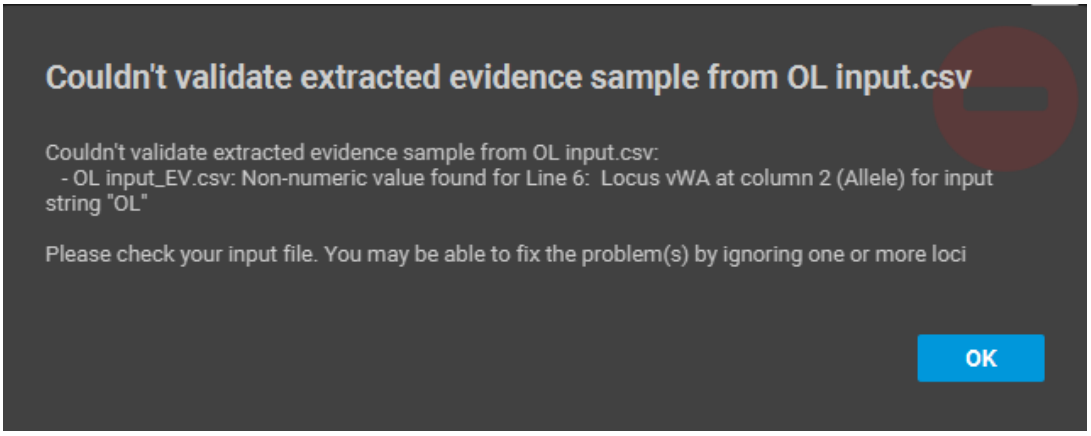
Section K: Challenge testing

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 peaks not removed during analysis) will cause STRmix™ to halt the interpretation. A sample (GF_M3_D1_0.5ng_28c_15s_05_E10_10.hid) was manually edited to include an 'OL' peak in an input file (renamed as OL input_EV.csv) and an attempt was made to run this in STRmix™. The following error message was obtained when the input file was added:

Figure 17: Error message produced by STRmix™ when attempting to add an input file containing an 'OL' peak

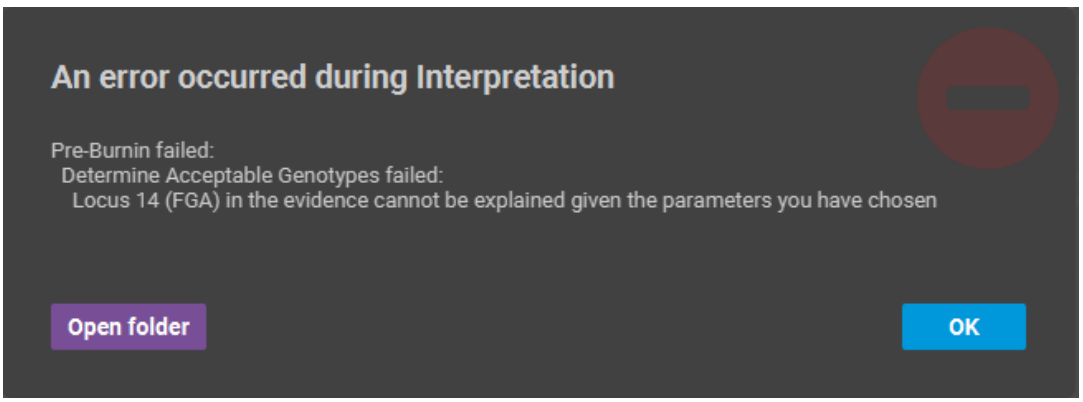


STRmix™ will not progress an interpretation until the issue is resolved. A review of the electropherogram should be undertaken and a determination made regarding whether the off-ladder peak originates from a microvariant allele not represented in the allelic ladder or is instead an artefact peak. If the analyst believes the peak to originate from an off-ladder allele, efforts should be made to type the allele based on its molecular weight relative to allelic ladders run with the sample. Note that allelic designations must be an integer value and designations such as 'R', '<', or '>' are not permitted within STRmix™ input files.

The presence of an artefact peak that has sized within an allelic bin and has been retained within the input file can have a number of effects on the STRmix™ interpretation. These include:

- No effect. If drop-in is observed within a laboratory and modelled within STRmix™, the artefact peak may be modelled as a drop-in peak if it less than the drop-in cap
- Failure to interpret. If an artefact is retained in a profile it may artificially increase the minimum number of contributors required to explain the profile. In the example shown below, peak data was inserted at the FGA locus to mimic a pull up event from a peak at the D18S51 locus in single-source sample GF_M3_D1_0.5ng_28c_15s_05_E10_10.hid. The FGA locus is a heterozygous locus and the additional 29.2 peak information means the profile cannot be explained as originating from a single contributor; this peak cannot be modelled as stutter or drop-in due to the position of the peak and its height, which exceeds the maximum drop-in height allowed in the JCRCL STRmix™ kit. Following a message such as the one displayed in Figure 18, it is suggested a user review the electropherogram to confirm calls and the number of contributors assigned.

Figure 18: Error message produced by STRmix™ when a profile cannot be explained by the number of contributors input by the user



- An exclusionary *LR*. If the artefact is modelled as having originated from the person of interest this may result in an exclusion. An example of this type of outcome has been tested here. A single-source profile (GF_M3_D1_0.5ng_28c_15s_05_E10_10.hid) was edited at the TH01 locus to include a 7 peak at 125rfu (above the drop-in cap). This was done to represent an artefact caused by pull-up from the D1S1656 locus. Under the assumption that the profile originates from a single contributor, this forced STRmix™ to assign all of its weight to a 7,9 genotype, causing the exclusion of the known donor (genotype= 9,9). Inspection of the per-locus *LR*s, displayed in Figure 19, shows inclusionary *LR*s at all other loci and an *LR* of 0 at TH01. This per-locus *LR* information is a useful diagnostic and a review of this locus would be warranted based on this result. An elevated variance value was also observed due to the large difference in peak height between the 7 and 9 peaks at the TH01 locus. This may also alert analysts to an unexpected peak becoming incorporated in the genotype combinations of a contributor. Once remedied, the affected profile can then be re-interpreted in STRmix™ and *LR*s that support inclusion at all loci would be expected. It may be harder for an analyst to spot a similar example in a multi contributor profile where the peak heights of the allelic component are similar to that of an artefact. However, careful review of the input file prior to interpretation in STRmix™ and review of the post-run diagnostics (e.g. Per locus LRs and weights) can assist with the identification of such issues.

Figure 19: Excerpt from the STRmix™ report of a run where LRs that support inclusion are observed at all but one locus (TH01)

PER LOCUS LIKELIHOOD RATIOS

	NIST1036_CAUC 0.01b(1.0, 1.0)		
LOCUS	Pr(E|Hp)	Pr(E|Hd)	LR
D3S1358	1	8.6620E-2	1.1545E1
vWA	1	4.6571E-2	2.1472E1
D16S539	1	1.0958E-1	9.1258E0
CSF1PO	1	1.0603E-1	9.4309E0
TPOX	1	6.1086E-2	1.6370E1
Yindel			
D8S1179	1	3.7750E-2	2.6490E1
D21S11	1	6.0292E-2	1.6586E1
D18S51	1	2.4181E-2	4.1354E1
DYS391			
D2S441	1	3.3934E-2	2.9469E1
D19S433	1	3.2598E-2	3.0677E1
TH01	0	5.1378E-2	0
FGA	1	3.1956E-2	3.1293E1
D22S1045	1	2.4701E-1	4.0484E0
D5S818	1	1.0627E-1	9.4100E0
D13S317	1	1.7775E-1	5.6259E0
D7S820	1	1.0885E-1	9.1869E0
SE33	1	9.9612E-3	1.0039E2
D10S1248	1	2.3782E-2	4.2049E1
D1S1656	1	7.0452E-3	1.4194E2
D12S391	1	3.5571E-2	2.8113E1
D2S1338	1	4.8576E-2	2.0586E1
SUB-SUB-SOURCE LR			0
SUB-SOURCE LR			0
99% 1-SIDED LOWER HPD INTERVAL			0

Section L: Comparison of STRmix™ to current interpretation and statistical methods

This section covers the following standards:

4.1.7. Partial profiles, to include the following:

4.1.7.2. DNA degradation

4.1.7.3. Inhibition

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.

This comparison study involves cases that had previously been interpreted manually (following CRIME-OPS 4-413) and statistically evaluated with the semi-continuous probabilistic genotyping software Lab Retriever (following CRIME-OPS 4-414), where applicable. It includes a range of data quality and comparison conclusions typically encountered in casework. Some of the selected cases did not originally involve a statistical calculation due to a manual exclusion or a determination of unsuitability. The study also provided a training exercise for the JCRCL DNA analysts prior to competency testing.

Multiple profile types and scenarios were identified in previously analysed casework. Each case and the associated samples were stripped of their case identifiers and assigned a code. Each coded case was contained within its own run folder(s) with the associated allelic ladders. The run folders were added to a GMID-X project for analysis. The allelic ladders and reference samples were analysed with the stutter filters. The evidence samples were analysed without stutter filters so that labels were assigned to peaks in stutter positions. These included back stutter and forward stutter at all STR loci except DYS391, two-base back stutter at SE33 and D1S1656, and double-back stutter at D10S1248. Other forms of stutter (if present) and all other artefacts such as spikes, noise, and pull-up were edited in the project. A STRmix™ table was exported from GMID-X. Lab Retriever statistical analysis was either previously performed as part of routine casework (and recalculated with the same drop-in probability as used in STRmix™) or performed specifically for this study on profiles that were initially deemed to be unsuitable for comparison and on comparisons that were manually determined to be exclusions. In some cases, the masking stutter approach was taken for the Lab Retriever calculation.

A total of 54 profiles were included in this comparison study. Table 14 summarizes the profile categories included. In addition to these categories, the profiles exhibited a range of degradation, dropout, and other sources of complexity.

Table 14 Summary of the profile categories included in this study

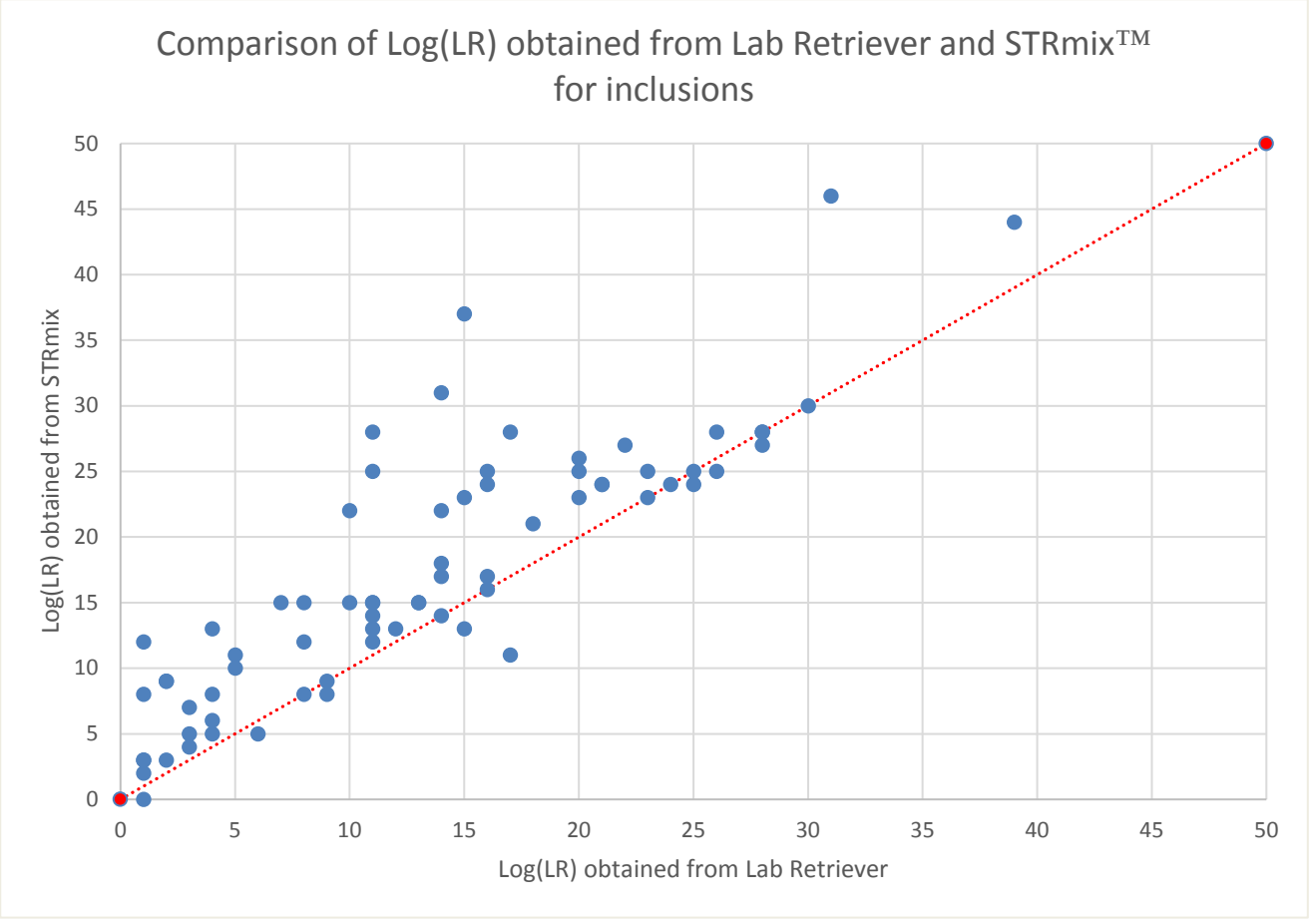
Single Source	Major/Minor Mix	2p Mix	3p Mix	4p Mix
6	12	7	17	12

Some of these profiles were interpreted in multiple ways, resulting in a total of 97 conclusions. The conclusions were separated into Inclusions, Exclusions, and Unsuitable for Comparison.

Inclusions

There were 71 inclusions among the 97 conclusions. The lowest likelihood ratio (LR) calculated by Lab Retriever for the African American, Caucasian, and Hispanic population groups was compared to the lowest sub-source LR calculated by STRmix™ for the same groups. In each case, both Lab Retriever and STRmix™ yielded an LR in support of inclusion of the person of interest. STRmix™ provided a higher LR value in all but eight instances. Figure 20 illustrates the comparison of the two software programs. The red dotted line indicates where the values on the X and Y axes are equal. The majority of the data are above this line, which is expected because STRmix™ can use more of the available information in the profile (i.e., peak heights) to assess the genotype weights to be factored into the LR given the proposed hypotheses. The average difference in LR values was four orders of magnitude. The largest log difference was 22 orders. This occurred when two persons of interest were included jointly, which is a scenario that is expected to yield higher LR values.

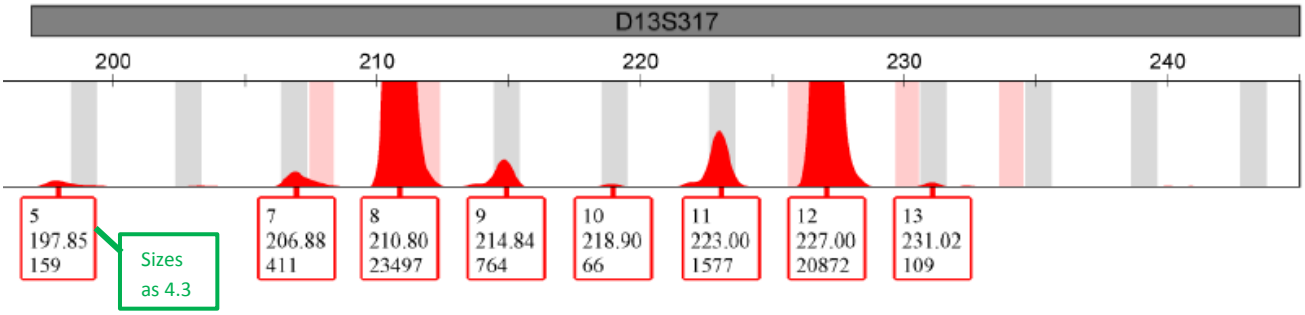
Figure 20: Plot of log(LR) obtained for Inclusion conclusions when calculated in Lab Retriever and STRmix™



If STRmix™ yielded a lower LR value than Lab Retriever, which occurred in eight of the 71 inclusions, the difference was typically one order of magnitude, which does not suggest a significant divergence. One case

presented an unusual result at the D13S317 locus. A peak was present in a bin less than the allelic ladder 5 position and was assigned a "5" label for the STRmix™ input file. Figure 21 is a screenshot of the locus.

Figure 21: Zoom view of the D13S317 locus of a profile used in the comparison study



The profile was interpreted under different numbers of contributors and with and without an assumed contributor. When no contributor was assumed, STRmix™ returned a higher LR for the person of interest (POI) (the lesser contributor with genotype 9,11) than did Lab Retriever. When one contributor was used to condition the interpretation (the greater contributor with genotype 8,12), however, the STRmix™ result was always lower than the Lab Retriever result. In fact, STRmix™ returned a LR=0 when the profile was interpreted as either a 2- or 3-person mixture with the assumed contributor. The Per Locus LR values were >1 for every locus except D13S317. Upon closer inspection of the genotype sets at D13S317, most of the weight was given to the option that included the "5" allele in the minor contributor genotypes. The 9,11 genotype was not assigned any weight, and consequently the POI in this case was excluded at this locus.

When the profile was interpreted as a 4-person mixture, the weight was spread more widely across different genotypes for the minor contributors. The POI's 9,11 genotype was considered possible and did not yield a Per Locus LR=0. The STRmix™ diagnostics (particularly the GR convergence, allele variance, and log(likelihood) values) provided indications that there was an issue with the modelling. Similarly, the Lab Retriever calculation provided a Per Locus LR <1 at D13S317.

The profile was additionally interpreted in three ways: a) without the D13S317 locus, b) without the "5" allele labelled at the D13S317 locus, and c) with the height of the "5" allele lowered to 120 RFU to allow it to be modelled as possible drop-in. In each of these instances, the STRmix™ LR values provided support for the inclusion of the POI but were 2-6 orders of magnitude lower than those obtained from Lab Retriever when the same adjustments were made to the input file. This difference may reflect the amount of uncertainty in the interpretation caused by this labelled peak in the profile. This profile behaved as expected in STRmix™ based on the rest of the validation data in this document and provided an opportunity to explore troubleshooting steps. As a side note, similar results were obtained when the peak was assigned a "4.3" label, which is how the peak sized compared to the allelic ladder.

Exclusions

In this study, there were 19 instances where a POI was manually excluded as a possible contributor. STRmix™ was used in all of these cases, whereas Lab Retriever was only performed in six cases. A LR=0 or <1 (support for exclusion) was obtained from both programs for 14 of these calculations. STRmix™ returned a LR=1 (uninformative) for one case involving a mid-level 4-person mixture. In this case, the uninformative LR applied to Contributor 4, who was attributed to 1% of the mixture. The STRmix™ report made no overt exclusion of the POI to Contributors 1-3, who constituted the other 99% of the mixture. Four other cases showed a LR>1 (between 3.4 and 760.) These cases involved partial and low-level 3- and 4-person mixtures. These results are

not unexpected due to the limited amount of information, the consideration of dropout, and the uncertainty in the genotype sets determined from the deconvoluted profile (the deconvolution report for one of these profiles was 916 pages). The outcome is consistent with those described in Section D.

Unsuitable for Comparison

The current interpretation procedure states that a profile is not interpretable if the analyst cannot determine whether there is a single or multiple contributors. It also states that a profile is not suitable for comparison if the number of contributors cannot be confidently assigned or if there appears to be more than five contributors. Six cases in this study involved a manual determination that the minor component of the profile was either uninterpretable or unsuitable for comparison. In all of these cases, the major component could be isolated and used for statistical analysis with Lab Retriever. This is not the case for STRmix™. Therefore, STRmix™ was performed by assuming different total numbers of contributors. There was no significant difference to the LR value for a POI who corresponded to the major component when the total number of contributors was changed. Table 15 provided an example of this from case kk1.

Table 15 Comparison of likelihood ratios for a mixed profile with a major contributor and an unknown number of minor contributors

	Lab Retriever	**STRmix™**
Major Component only	3.5E+20	
Mixture of 2		1.0E+24
Mixture of 3		6.6E+23
Mixture of 4		3.4E+23

In addition to a sub-source LR, STRmix™ can also calculate a sub-sub source LR. The sub-source LR reflects the whole mixture profile and is subject to the 'two-trace transfer' or the 'factor of N!' concept whereby the LR is adjusted by contributor order (influenced by the number of contributors assigned) and genotype set orders. In contrast, the sub-sub source LR value reflects a portion of a mixture profile (or a Contributor as deconvoluted by STRmix™). This value may be useful for profiles with a major component (or Contributor >75%) and an uninterpretable or unsuitable minor component. To illustrate how the sub-sub source LR may be utilized, five of the profiles in this study that exhibited major and minor components were selected. The sub-source LR was compared to the sub-sub source LR. Table 16 provides the results.

Table 16 Values reported for sub-source and sub-sub source LR for selected profiles with major and minor components

Case/Sample	**Profile description**	**Lowest sub-sub source LR**	**Lowest sub source LR**	**Fold-difference**
KDM2/007.A	Minor uninterpretable – total NOC as 2	2.5E+25	1.2E+25	~2x
KDM2/008.A	Minor unknown – total NOC as 3	2.2E+26	7.4E+25	~3x
KDM7/003.D.1	Mixed major with 2 POI Minor uninterpretable – total NOC as 3	5.4E+28	2.7E+28	~2x
BH4/002.A	Minor unknown – total NOC as 3	5.2E+16	1.9E+16	~3x
SS4/002.A	Minor interpreted as 3 – total NOC as 4	2.0E+29	5.2E+28	~4x

In general, the sub-source LR is lower than the sub-sub source LR by approximately an amount equivalent to the number of contributors assigned. This is intuitive because the sub-source LR is attempting to explain a

portion of the whole profile whereas the sub-sub source LR is attempting to explain only one portion while ignoring the rest of the profile.

Of interest is Item 008.A in case KDM2. The case involved two different POIs and a profile with a minor component that was manually determined to be unsuitable for comparison due to an unknown number of contributors. The total number of contributors was set to three for the STRmix™ deconvolution and LR calculation. While the major component was deconvoluted (and results are contained within Table 16), STRmix™ focused its comparison of the two POIs on the minor components that were deconvoluted. One POI was assigned an LR=320 (suggesting support for inclusion), and the other POI was assigned an LR=0.019 (suggesting limited support for exclusion). These results are expected based on the uncertainty in the interpretation of this profile. They are in the range of adventitious matches as expected based on the results of Section D. This case scenario also provides support for the continued manual interpretation to determine if a profile (or portion thereof) is suitable for comparison.

Summary

The comparison of STRmix™ to the current interpretation and Lab Retriever procedures showed sufficient concordance for inclusive conclusions. Options exist for treating mixture profiles with major components by either reporting the sub-source LR (which is not greatly influenced by the total number of contributors) or the sub-sub source LR. Where applicable, STRmix™ also provided statistical support for the current procedure for manually determining exclusions and inconclusive conclusions.

Section M: Precision

This section covers the following standard:

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

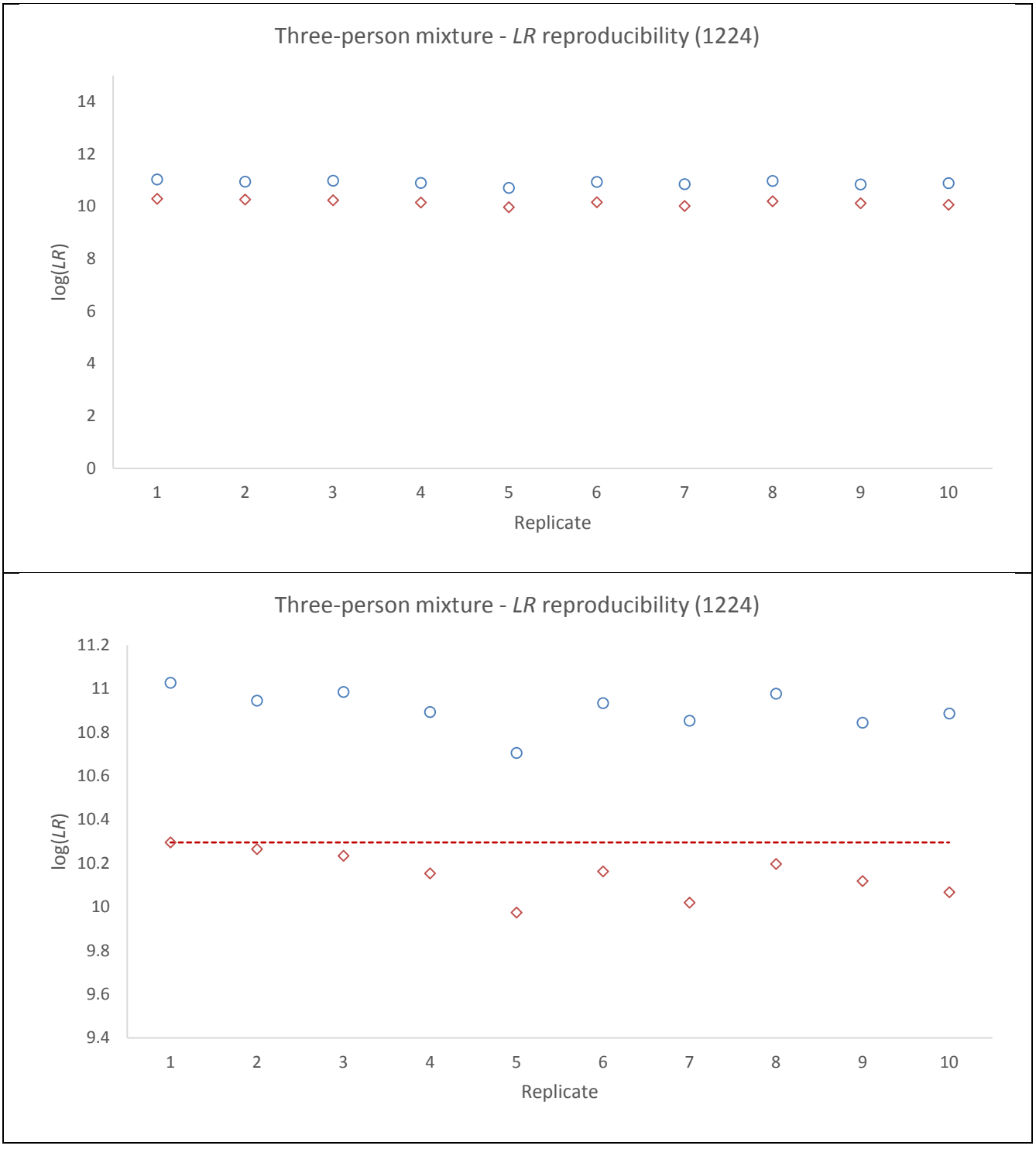
Refer to section D above for details of sensitivity and specificity tests.

STRmix™ uses MCMC to generate weights for different genotype combinations. This is a sampling procedure and therefore the weights will vary slightly if the interpretation is repeated. The variability in *LR*s between replicate interpretations has previously been explored [9]. The MCMC process was shown to be a small source of variability compared with other laboratory variables including the PCR and CE processes. Variability due to the size of the allele frequency database and the MCMC process is taken into account within STRmix™ V2.6 using the highest posterior density (HPD) method [10-12] (a type of confidence interval).

Run-to-run variability within STRmix™ was investigated by interpreting a three-person mixed DNA profile from Section D (GF-MX_Mix12_D1) ten times. The mixture examined is a complex three-person mixture with proportions of approximately 0.42, 0.33, and 0.24. Given the mixture proportions and low template amounts, STRmix™ has accepted numerous genotype combinations at each locus resulting in fairly diffuse weights. Following each interpretation, an *LR* was calculated for one of the known contributors (reference 1224). A plot of the log(*LR*) for each STRmix™ replicate is given in Figure 22. The blue circle data points indicate the sub-source *LRs* and the red rhombus data points are the 99% 1-sided lower HPD interval.

Inspection of Figure 22 shows that the *LR*s are highly reproducible and that the lower 99% bound of the HPD is always below the sub-source *LRs*, as shown by the red dashed line representing the largest HPD *LR* observed.

Figure 202: Plot of the sub-source LRs (blue circles) and 99% 1-sided lower HPD interval LRs (red rhombuses) calculated for a known contributor following ten replicate STRmix™ interpretations of the same profile. The plot has been reproduced with the y-axis expanded to better show the results



Parameters within STRmix™ that affect run-to-run variability include the number of iterations and the RWSD (random walk standard deviation). By default, number of MCMC accepts is set to 100,000 burn-in accepts and 50,000 post burn-in accepts per-chain. These values should be suitable for the majority of profiles interpreted. Decreasing the number of MCMC accepts may mean that the MCMC chains do not converge, resulting in

increased variability. Increasing the number of MCMC accepts improves precision at the expense of longer interpretation run times.

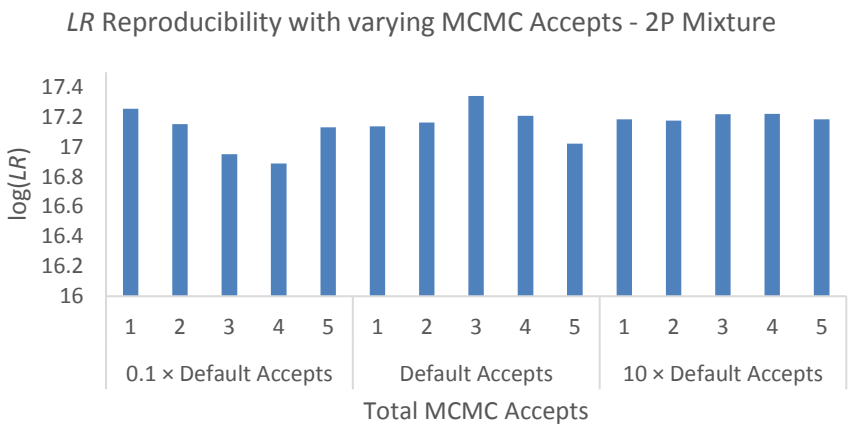
A two-, three-, and four-person mixture were each interpreted using a variety of MCMC conditions pertaining to the number of MCMC accepts required. Each profile was interpreted five times under each of the three conditions examined. Following interpretation, an *LR* was calculated for one of the known contributors to the mixture (reference 1224). A summary of the MCMC settings used are provided in Table 17.

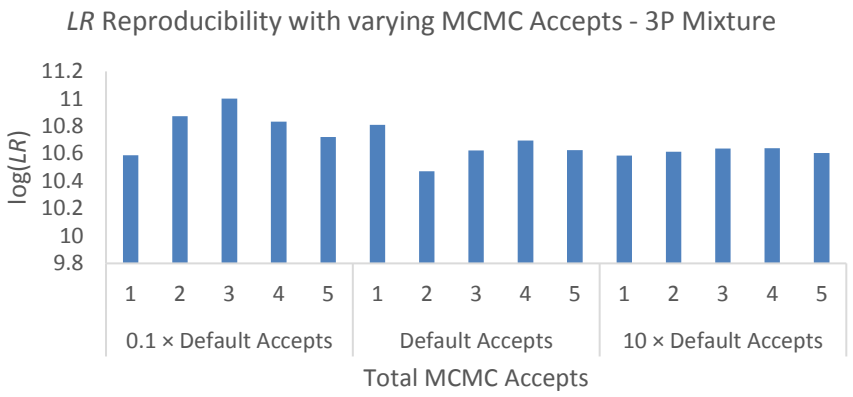
Table 17 Summary of MCMC conditions tested. The three mixtures examined were each deconvoluted five times under the MCMC conditions listed. Asterisks indicates the default settings within STRmix™

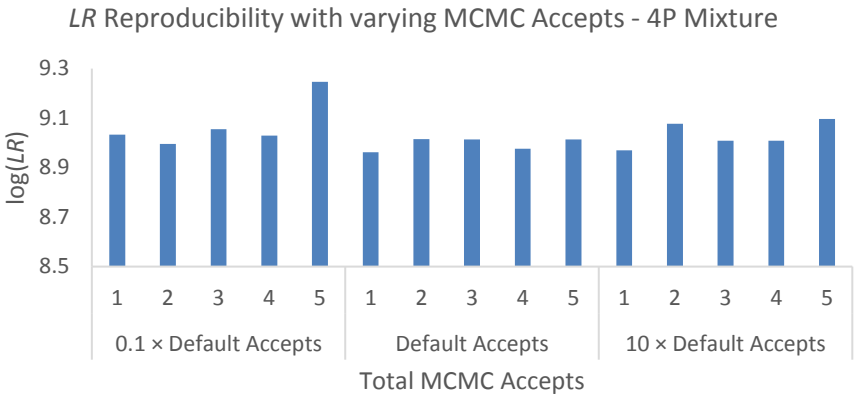
Sample	N	MCMC Chains	Burn-in Accepts/Chain	Post Burn-in Accepts/Chain
Mix11_D1	4	8	10,000	5,000
			100,000*	50,000*
			1,000,000	500,000
Mix12_D1	3		10,000	5,000
			100,000*	50,000*
			1,000,000	500,000
Mix13_D1	2		10,000	5,000
			100,000*	50,000*
			1,000,000	500,000

Figure 23 displays the Log(*LR*) calculated for the known contributor selected for each of the two-, three-, and four-person mixtures that were interpreted five times in STRmix™ using different numbers of MCMC accepts. Inspection of Figure 22, shows a reduction in run-to-run variability as MCMC accepts increase. The *LRs* calculated varied by less than one order of magnitude, even when using reduced MCMC accepts. Use of the default settings of 100,000 burn-in and 50,000 post burn-in accepts per-chain (using 8 chains) should be more than sufficient for the majority of casework profiles encountered. However, the option remains to increase this value as and when required depending on profile complexity.

Figure 213: Log(LR) calculated for a known contributor to a two-, three-, and four-person mixture interpreted five times in STRmix™ using different numbers of MCMC accepts. Default Accepts: 100,000 burn-in accepts/50,000 post burn-in accepts per-chain




LR Reproducibility with varying MCMC Accepts - 3P Mixture
log(*LR*)
11.2
11
10.8
10.6
10.4
10.2
10
9.8
1 2 3 4 5 1 2 3 4 5 1 2 3 4 5
0.1 × Default Accepts
Default Accepts
10 × Default Accepts
Total MCMC Accepts


LR Reproducibility with varying MCMC Accepts - 4P Mixture
log(*LR*)
9.3
9.1
8.9
8.7
8.5
1 2 3 4 5 1 2 3 4 5 1 2 3 4 5
0.1 × Default Accepts
Default Accepts
10 × Default Accepts
Total MCMC Accepts

Conclusion

This document describes the Jefferson County Regional Crime laboratory's internal validation activities for STRmix™ V2.6 using the GlobalFiler™ kit and a capillary electrophoresis injection setting of 15 seconds on a 3500 instrument. It has been shown that STRmix™ V2.6 is suited for its intended use at the JCRCL for the interpretation of profiles generated from crime scene samples.

The project work has been performed in accordance with the FBI Quality Assurance Standards for Forensic DNA Testing Laboratories, the ANAB Accreditation Requirements, and the SWGDAM Guidelines for the Validation of Probabilistic Genotyping.

Signatures

Beth Hewitt [signature redacted]

DNA Technical Leader

Jefferson County Regional Crime Laboratory

References

[1] FBI Quality Assurance Standards for Forensic DNA Testing Laboratories. 2011.

[2] Taylor D, Bright J-A, Buckleton JS. The interpretation of single source and mixed DNA profiles Forensic Science International: Genetics. 2013;7:516-28.

[3] Bright J-A, Evett IW, Taylor D, Curran JM, Buckleton J. A series of recommended tests when validating probabilistic DNA profile interpretation software. Forensic Science International: Genetics. 2015;14:125-31.

[4] Balding DJ, Nichols RA. DNA profile match probability calculation: how to allow for population stratification, relatedness, database selection and single bands. Forensic Science International. 1994;64:125-40.

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

[6] Taylor D. Using continuous DNA interpretation methods to revisit likelihood ratio behaviour. Forensic Science International: Genetics. 2014;11:144-53.

[7] Taylor D, Bright J-A, Buckleton J. The 'factor of two' issue in mixed DNA profiles. Journal of Theoretical Biology. 2014;363:300-6.

[8] Budowle B, Moretti TR, Baumstark AL, Defenbaugh DA, Keys KM. Population data on the thirteen CODIS core short tandem repeat loci in African Americans, US Caucasians, Hispanics, Bahamanians, Jamaicans and Trinidadians. Journal of Forensic Sciences. 1999;44:1277-86.

[9] Bright J-A, Stevenson KE, Curran JM, Buckleton JS. The variability in likelihood ratios due to different mechanisms. Forensic Science International: Genetics. 2015;14:187-90.

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

[11] Triggs CM, Curran JM. The sensitivity of the Bayesian HPD method to the choice of prior. Science & Justice. 2006;46:169-78.

[12] Curran JM, Buckleton JS. An investigation into the performance of methods for adjusting for sampling uncertainty in DNA likelihood ratio calculations. Forensic Science International: Genetics. 2011;5:512-6.

[13] 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

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 and 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 and 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, Characterising the STR locus D6S1043 and examination of its effect on stutter rates. Forensic Science International: Genetics, 2014. 8(1): 20-23

6. D. Taylor and J.S. Buckleton, Do low template DNA profiles have useful quantitative data? Forensic Science International: Genetics, 2015. 16: 13-16
7. D. Taylor, J. Buckleton and J.-A. Bright, 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 and 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 and J.S. Buckleton, Investigating a common approach to DNA profile interpretation using probabilistic software. Forensic Science International: Genetics, 2014. 16: 121-131
4. T.R. Moretti, R.S. Just, S.C Kehl, L.E. Willis, J.S. Buckleton, J.-A. Bright 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. J.-A. Bright, D. Taylor, C.E. McGovern, S. Cooper, L. Russell, D. Abarno et al., Developmental validation of STRmix™, expert software for the interpretation of forensic DNA profiles. Forensic Science International: Genetics, 2016. 23: 226-39
6. D. Taylor, J.-A. Bright, C. McGovern, C. Hefford, T. Kalafut and J. Buckleton, 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

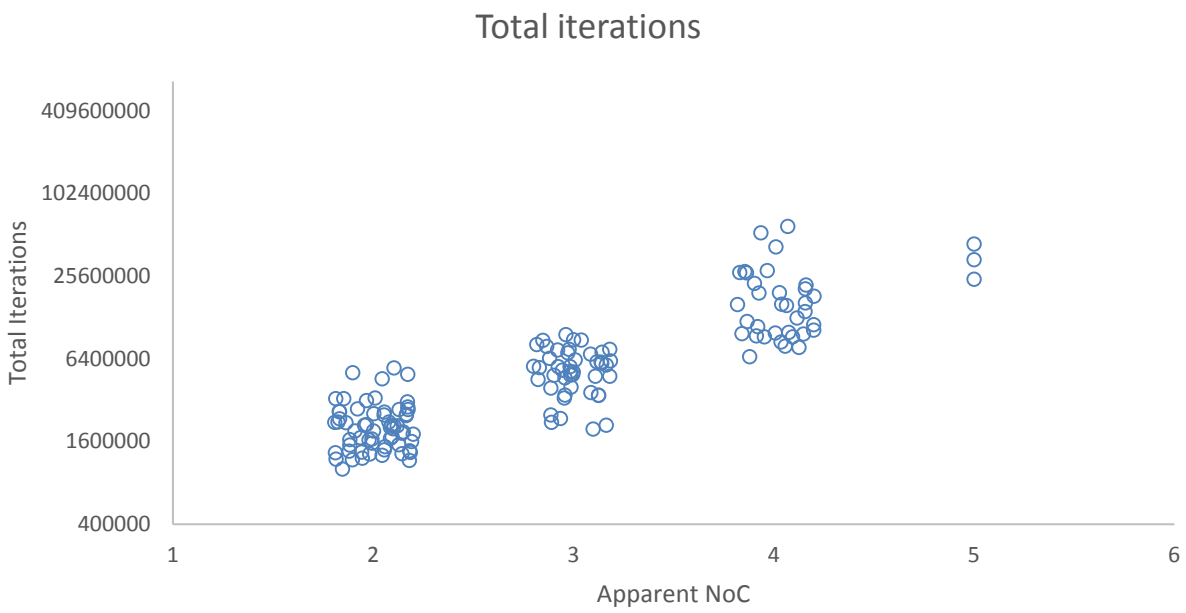
Appendix 3: Review of Secondary Diagnostics of Section D

This section reviews the secondary diagnostics for each mixture in Section D. These 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 that the interpretation has progressed as expected. Individual secondary diagnostics may indicate whether further review of the results is warranted; however, analysts should not rely on these diagnostics alone. Elevated values for one of these diagnostics alone does not necessarily mean the results are not fit for purpose. Further review of the other diagnostics and the profile itself could indicate that STRmix™ is performing as expected.

Total Number of Iterations

The total number of iterations shows the total number of iterations that were required to achieve 400,000 accepts during the MCMC process (50,000 post burn-in accepts per chain with 8 Markov chains used). We expect the number of iterations to increase as DNA profiles become more complex. The total iterations required to interpret the mixtures examined within Section D has been plotted in Figure 22. As expected, the number of iterations required to interpret a DNA profile increased in line with the complexity of the profile. An excessive number of iterations or low acceptance rate could indicate that STRmix™ could not converge on a good probability space during MCMC, perhaps due to the complexity of the profile.

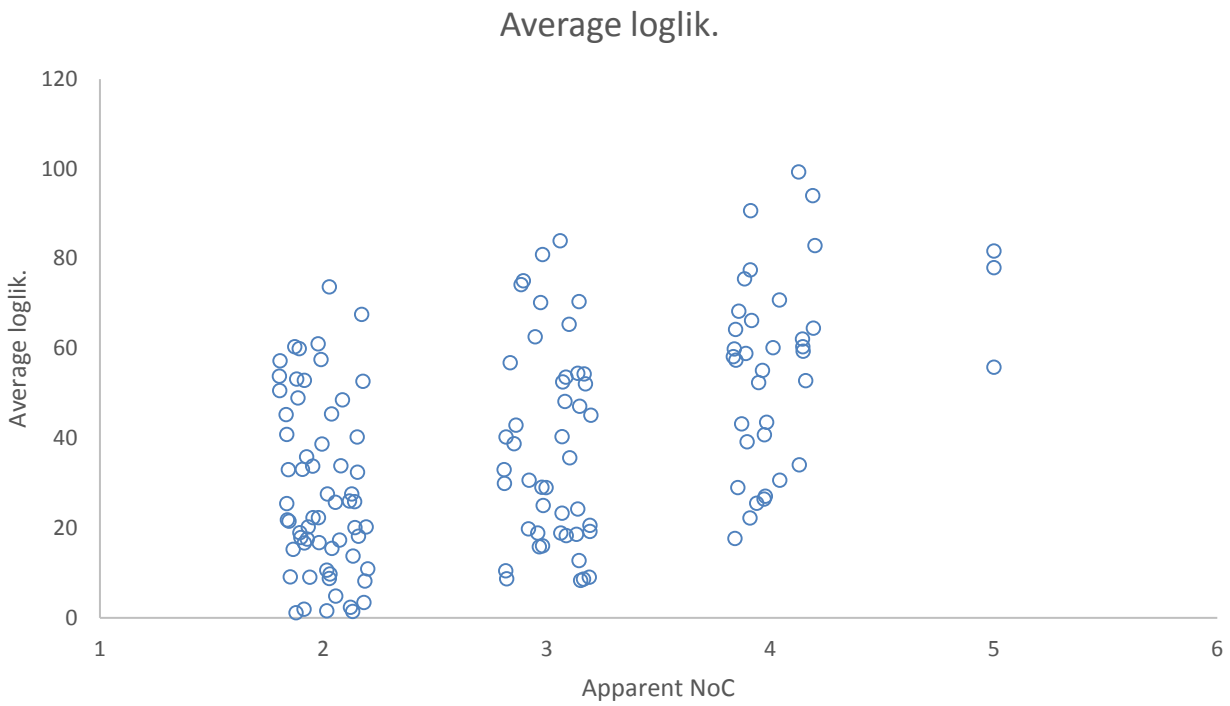
Figure 22: The total number of iterations required for each mixture to achieve 400,000 post burn-in accepts



Average log(likelihood)

The average log(likelihood) can be described as the average of the post burn-in probability density (or likelihood) values across the chains used in a deconvolution. Generally, a high average log(likelihood) is considered desirable as it indicates that STRmix™ has been able to describe the observed profile better. Low log(likelihoods) may be obtained due to STRmix™ being unable to describe the data well due to large forced stochastic events resulting from an incorrect assessment on the number of contributors, absence of expected data (labels may have been inadvertently removed during analysis), presence of an unexpected peak (for example, a pull up peak label being retained at analysis), or due to data from replicates being disparate in nature. However, low or even negative values, do not necessarily indicate an issue. Low-level profiles with very little information present may naturally yield low average log(likelihood) values as there are few peaks contributing to the probability value calculated at each iteration. When low or negative log(likelihood) values are observed, other diagnostics such as mixture proportion and weights should be considered carefully before progressing the interpretation or the reporting of a result. The average log(likelihood) values produced following STRmix™ interpretation of the 144 mixtures examined within Section D is plotted in Figure 23. In the present study, no negative log(likelihood) diagnostics were observed, indicating that STRmix™ was able to explain the observed profiles well.

Figure 23: The average log(likelihood) output for each mixture deconvolution undertaken within Section D

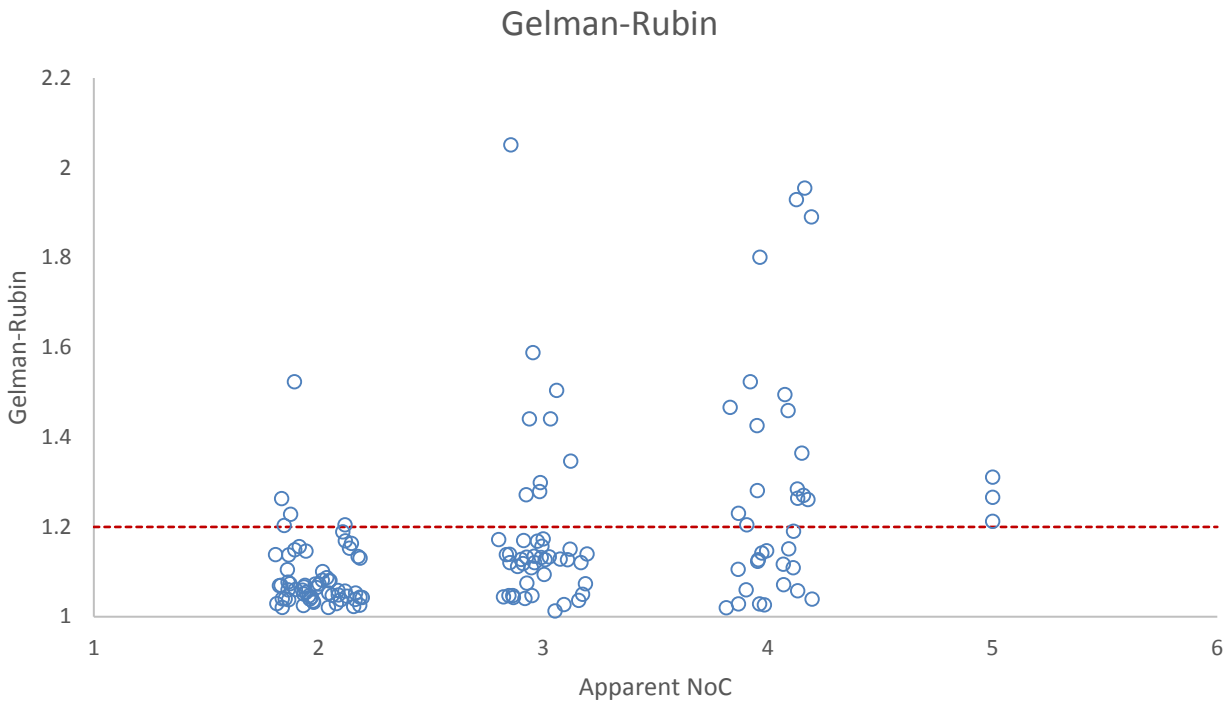


Gelman-Rubin (GR) convergence diagnostic

The Gelman-Rubin (GR) convergence diagnostic is a diagnostic value that indicates whether there is likely convergence of the MCMC chains. This value is a comparison of the within chain and between chain variances. A value of 1.2 or less typically indicates likely convergence of the MCMC chains. Figure 24 shows a spread of GR values with the majority of the points (76.4%) below 1.2 (shown as the dashed line).

However, GR values greater than 1.2 do not necessarily indicate that the deconvolution is unsuitable for use. This value could simply be indicative of 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, potentially allowing STRmix™ to investigate the mixture more thoroughly. If this does not decrease the GR value, it may indicate a further review of the input file or reassessment of *N* could be warranted.

Figure 24: The GR value for each mixture deconvolution undertaken within Section D



From Figure 24, there are 5 observations where the GR was greater than 1.8. Upon review of these profiles, there was no obvious indication as to why these GR values were inflated, as the primary and other secondary diagnostics appeared reasonable. The inflated GRs could simply be a testament to the complexity of these mixtures.

These five mixtures were re-interpreted with a ten-fold increase in the number of burn-in and post burn-in accepts per chain. In all but one example the GR following reinterpretation using increased number of accepts was less than 1.2. The results are shown in Table 15.

Table 14: Summary of GR changes increasing the number of post burn-in accepts

Sample	GR	GR (Extended Accepts)
GF-MX_Mix4_A1	2.051197633	1.615924234
GF-MX_Mix1_B2	1.955022324	1.024489754
GF-MX_Mix2_B1	1.928813065	1.016773863
GF-MX_Mix2_C1	1.890298119	1.022113557
GF-MX_Mix1_A2	1.800885558	1.027869599

Posterior Mean Allele Variance

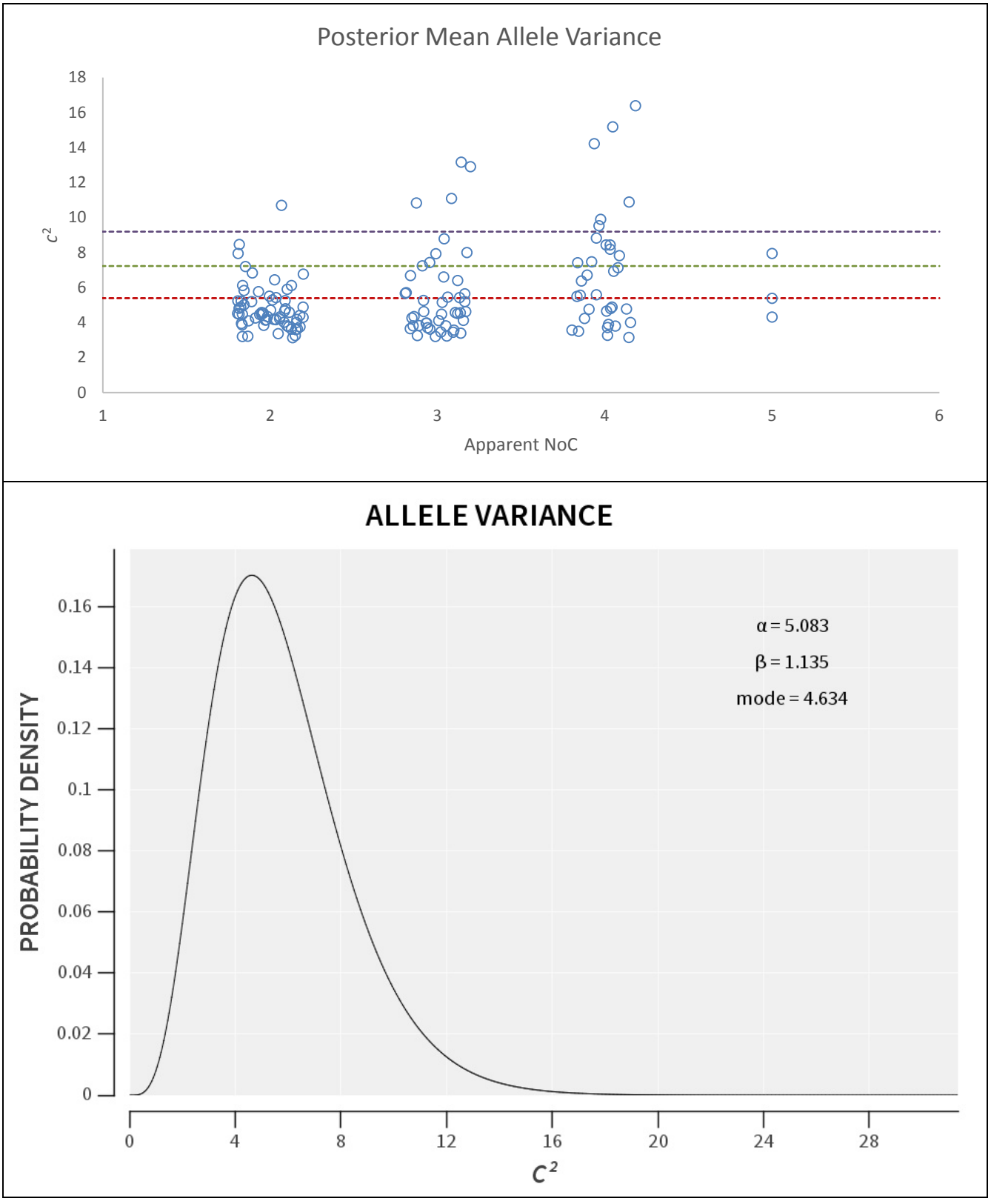
Figure 25 shows the spread of the allele variance values from the STRmix™ outputs for each sample run from section D. This is the posterior mean across the entire post burn-in analysis. The red (lower), green (middle), and purple (upper) dotted lines represents the 50^{th}, 75^{th}, and 90^{th} percentile of the stutter variance prior distribution; however, the mode is also a useful point of reference as well. It can be helpful to consider where the value sits compared to the prior distribution, determined during parameter setting, which is provided in the lower panel of Figure 25.

From Figure 25, it can be seen that the majority of the posterior mean allele variances for each sample are situated around the body of the allele variance prior distribution. The prior distribution is modelled by a gamma distribution, $\Gamma(5.083, 1.135)$.

Occasionally a STRmix™ deconvolution of a DNA profile will show an inflated posterior mean allele variance, in that for whatever reason the profile requires a high allele variance. This alone does not invalidate a run; however, excessively high values, when compared to the main body of the prior distribution, may warrant further review.

For example, the three highest posterior mean allele variance values originated from complex four-person mixtures that were of high template values. However, all three of these mixtures demonstrated *LR*s that supported inclusion for the known contributors and reliably excluded known non-contributors, where the highest *LR* for a non-contributor was 8.219×10^{-10}.

Figure 25: Plot of the average posterior mean allele variance values for each DNA profile deconvoluted as part of section D (upper panel). Red, green, and purple dashed lines represent the 50th, 75th, and 90th percentile values of the prior distribution modelled by a gamma distribution (lower panel)



Posterior Mean Stutter Variances

Figures 26 through 29 show the spread of the stutter variance values from the STRmix™ outputs for each sample run from section D, for each stutter type being modelled. This is the posterior mean across the entire post burn-in analysis. The red (lower), green (middle), and purple (upper) dotted lines represent the 50^{th}, 75^{th}, and 90^{th} percentile of the stutter variance prior distribution, however the mode is also a useful point of reference. Again, the prior distribution is also provided as a point of reference. As seen in the figures below, the majority of the posterior mean stutter variances appear to be situated within the body of the stutter variance prior distribution for each stutter type. The prior distributions are modelled by a gamma distribution: $\Gamma(1.509, 10.85)$ for back stutter, $\Gamma(1.592, 7.041)$ for forward stutter, $\Gamma(1.775, 1.856)$ for two base-pair back stutter, and $\Gamma(3.963, 2.366)$ for double back stutter.

Figure 26: Plot of the average posterior mean (-1,0) stutter variance values for each DNA profile (upper panel). Red, green, and purple dashed lines represents the 50th, 75th, and 90th percentile values of the prior distribution modelled by a gamma distribution (lower panel)

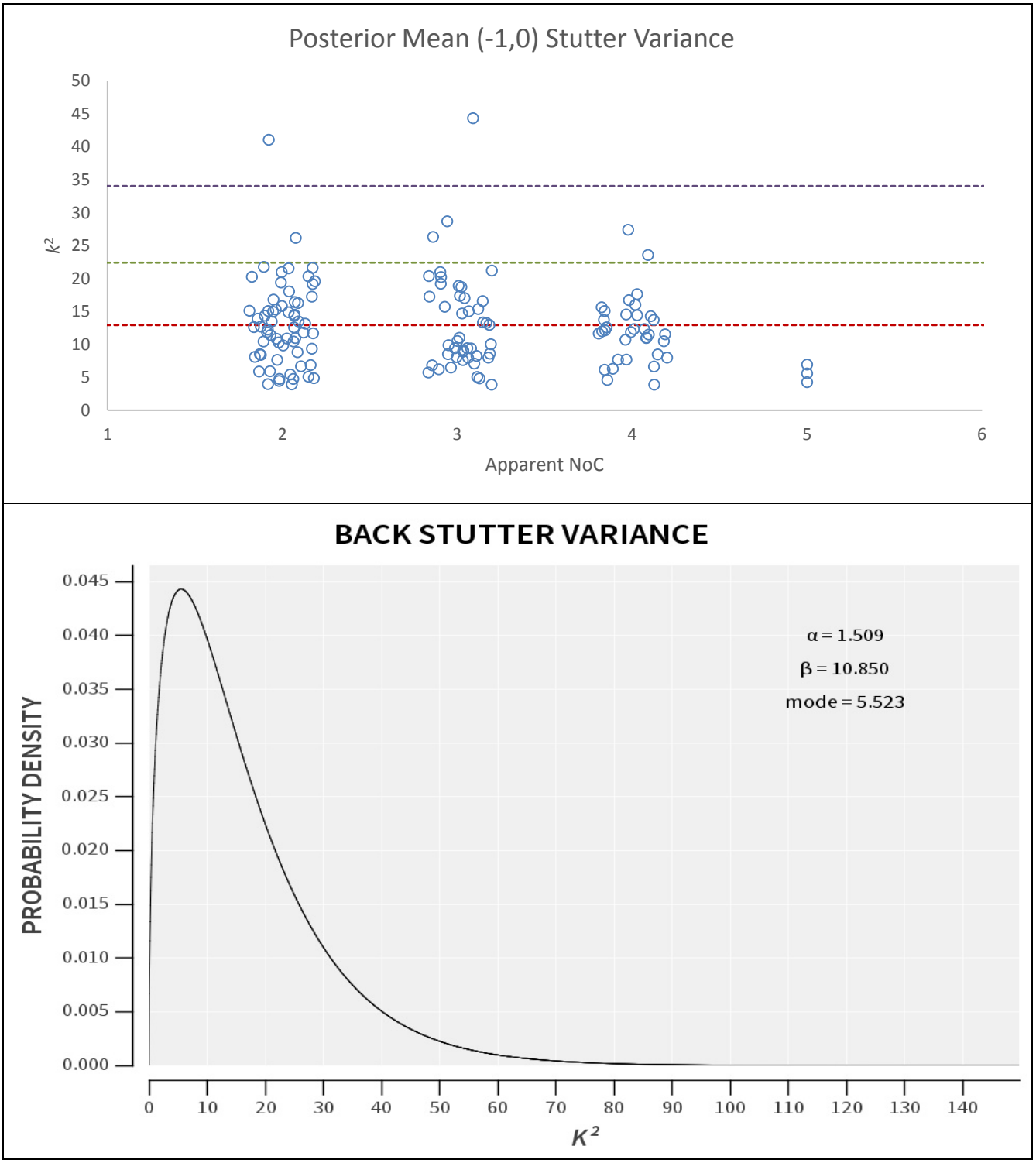


Figure 27: Plot of the average posterior mean (1,0) stutter variance values for each DNA profile (upper panel). Red, green, and purple dashed lines represents the 50th, 75th, and 90th percentile values of the prior distribution modelled by a gamma distribution (lower panel).

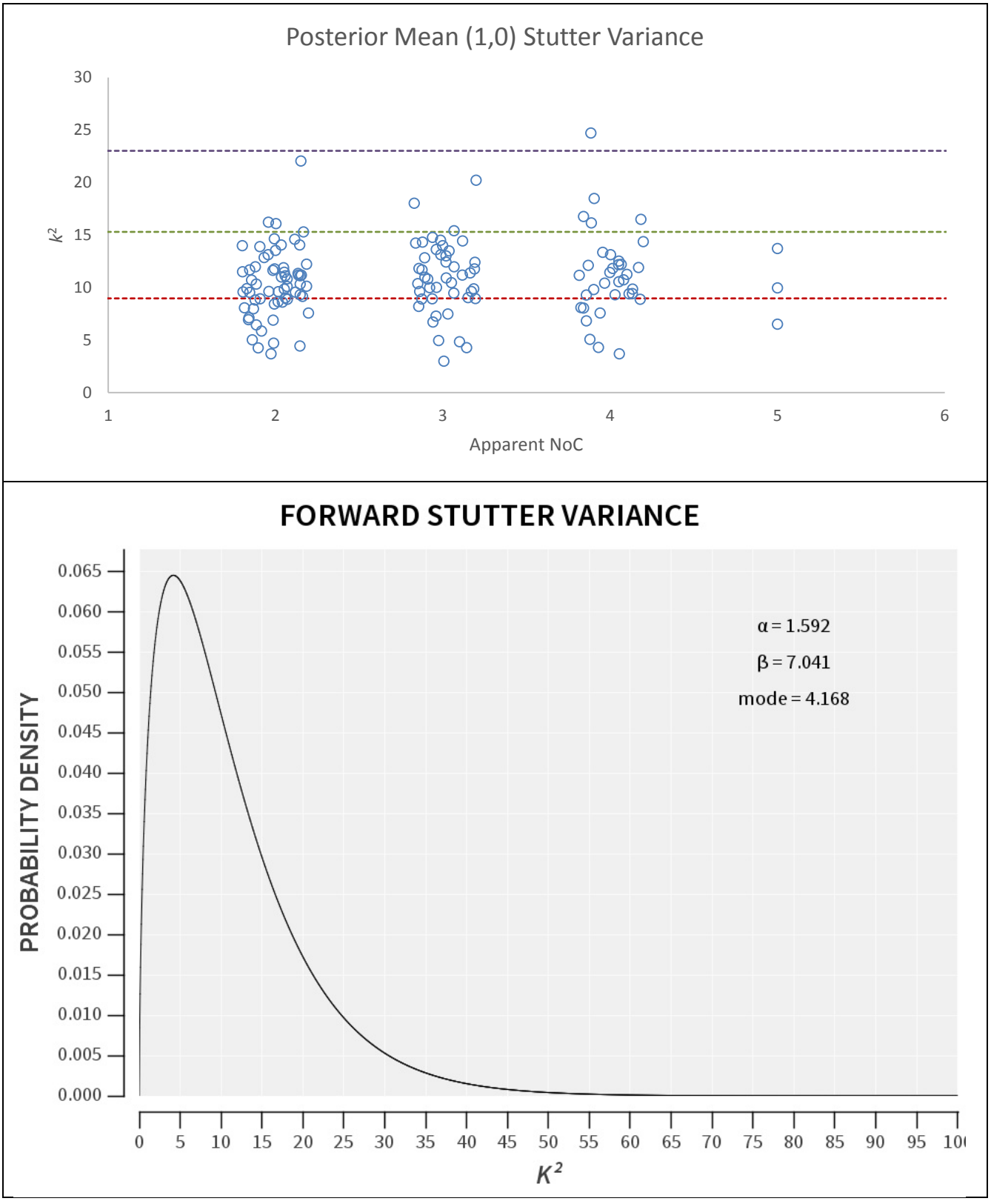


Figure 28: Plot of the average posterior mean (0,-2) stutter variance values for each DNA profile (upper panel). Red, green, and purple dashed lines represents the 50th, 75th, and 90th percentile values of the prior distribution modelled by a gamma distribution (lower panel).

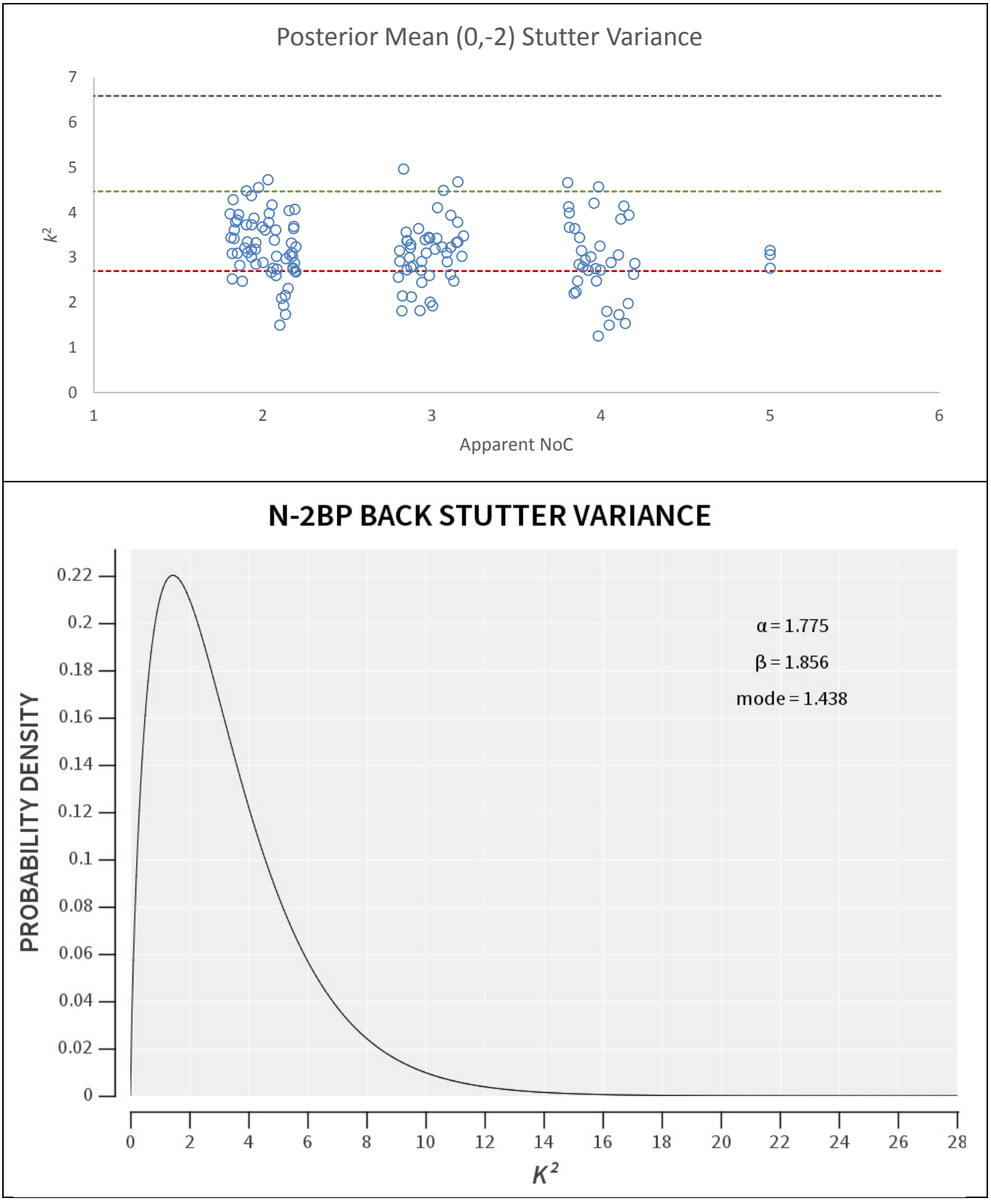
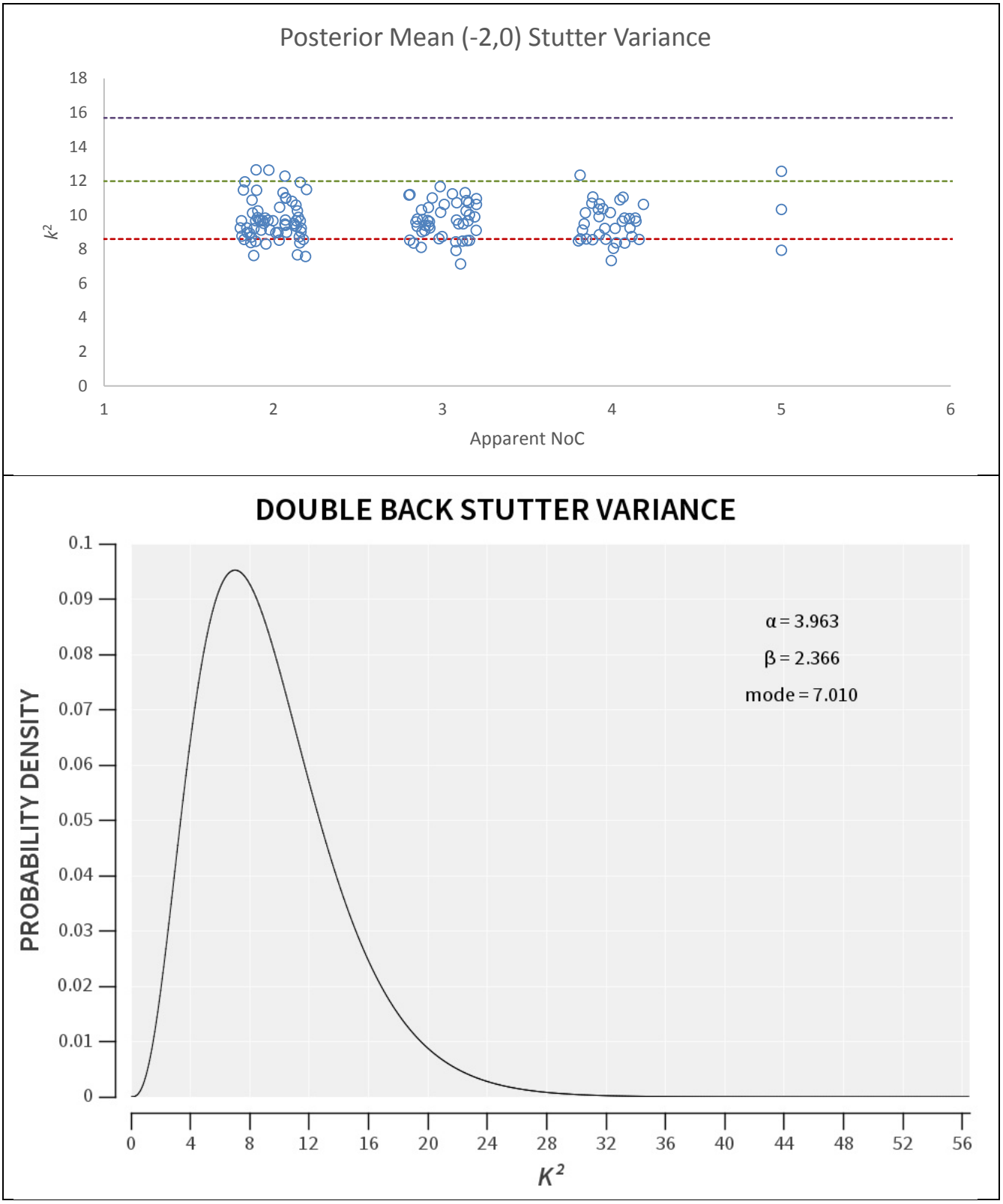


Figure 29: Plot of the average posterior mean (-2,0) stutter variance values for each DNA profile (upper panel). Red, green, and purple dashed lines represents the 50th, 75th, and 90th percentile values of the prior distribution modelled by a gamma distribution (lower panel).



Addendum: Upgrade to STRmix™ V2.6.3 and validation of the VarNOC feature of STRmix™ for GlobalFiler™ data produced within the Jefferson County Regional Crime Laboratory

Addendum: Upgrade to STRmix™ Version 2.6.3

In the period between the Jefferson County Regional Crime Laboratory (JCRCL hereafter) completing validation studies (in V2.6.0) and preparing to use the software on casework, a number of point releases of the STRmix™ software were made. STRmix™ V2.6.1 was released to address user licensing issues and use of the software that did not affect the STRmix™ calculations. However, mathematical changes were made in versions 2.6.2 and 2.6.3. These changes are detailed in the Release and Testing Report for each version of the software[1]. In both of these reports it is recommended that laboratories who have validated V2.6 repeat Likelihood Ratio calculations and compare the values obtained in the new version to those obtained in previous versions of the software. Given the changes made it was anticipated that there should be no difference in the likelihood ratios (*LR*s) for analyses performed with the same deconvolution seed and the same *LR* seed. Prior to going live with STRmix™ in casework JCRCL would like to demonstrate that STRmix™ V2.6.3 behaves as expected and the software is fit for purpose with the respects to analysing GlobalFiler™ data produced at the laboratory. At this time, JCRCL also wished to explore the option to use the VarNOC functionality within STRmix™.

This addendum displays comparisons of the alpha and beta prior gamma distributions for variance for each of the peak types modelled within the JCRCL STRmix™ kit. This report also details work that was carried out to demonstrate that there is no effect on the standard STRmix™ *LR* using STRmix™ V2.6.3 compared to the previously validated version V2.6.0 and that the work done to date to demonstrate the JCRCL STRmix™ kit is suitable for use on GlobalFiler™ profiles developed within the laboratory is valid. This addendum then goes on to demonstrate how *LR*s may behave using the VarNOC feature with JCRCL GlobalFiler profiles.

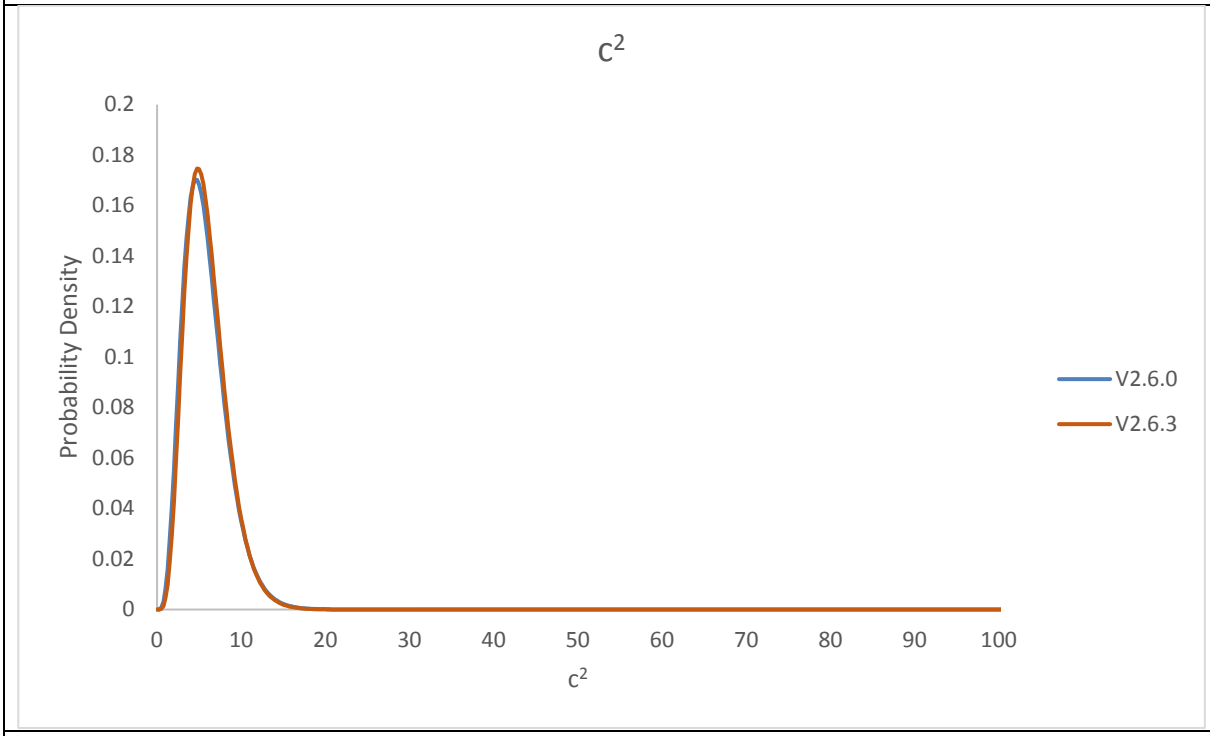
Comparison of Model Maker results obtained from STRmix™ V2.6.0 and V2.6.3

The JCRCL Model Maker data that was used in the initial estimation of parameters for STRmix™ v2.6.0 (see document 'Estimation of STRmix™ Parameters for Jefferson County Regional Crime Laboratory (GlobalFiler™ 3500)') was reanalysed in STRmix™ V2.6.3. A comparison of the prior gamma variance distributions for each peak type modelled within the JCRCL GlobalFiler™ kit are displayed in Figure 1.

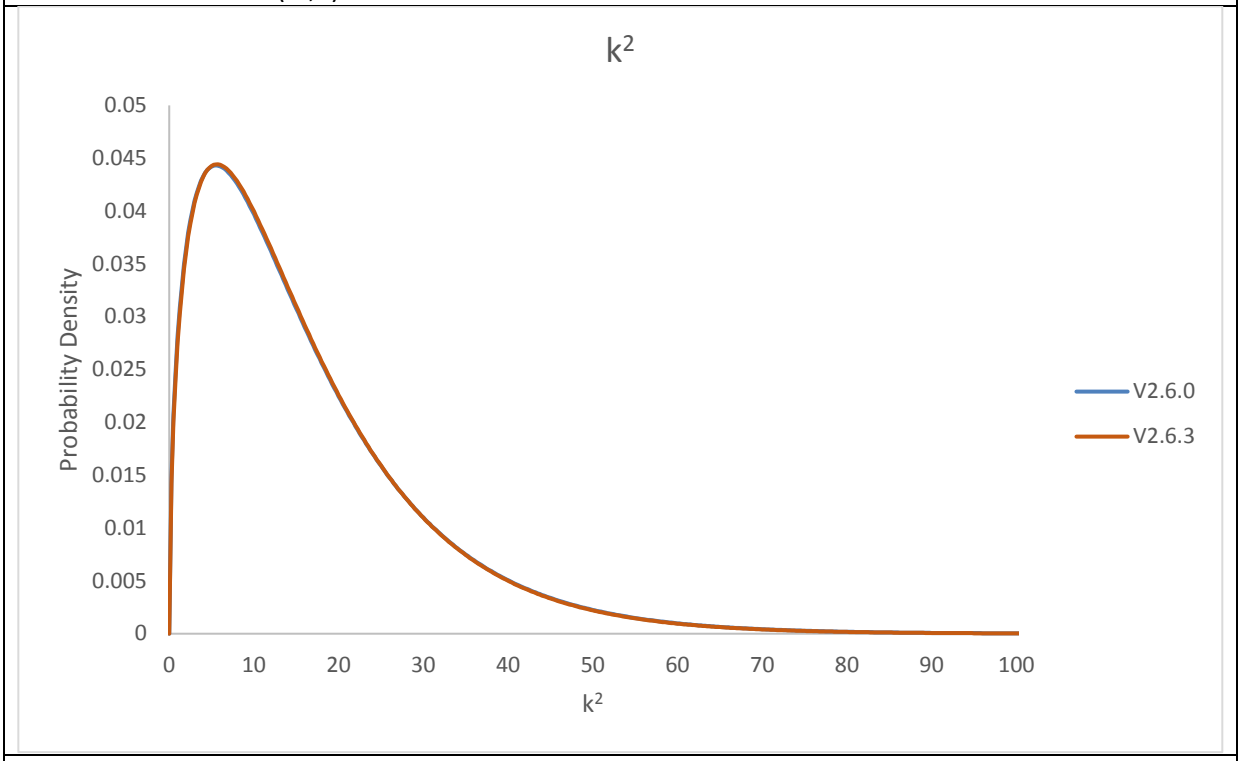
Figure 1: *Comparisons of the prior gamma variance distributions arrived at using Model Maker in STRmix™ V2.6.0 and V2.6.3 for each peak type modelled within the JCRCL GlobalFiler™ kit. Note: the Y-axis scale changes per plot.*

[1] These can be found on the www.support.strmix.com site, within the implementation & validation folder

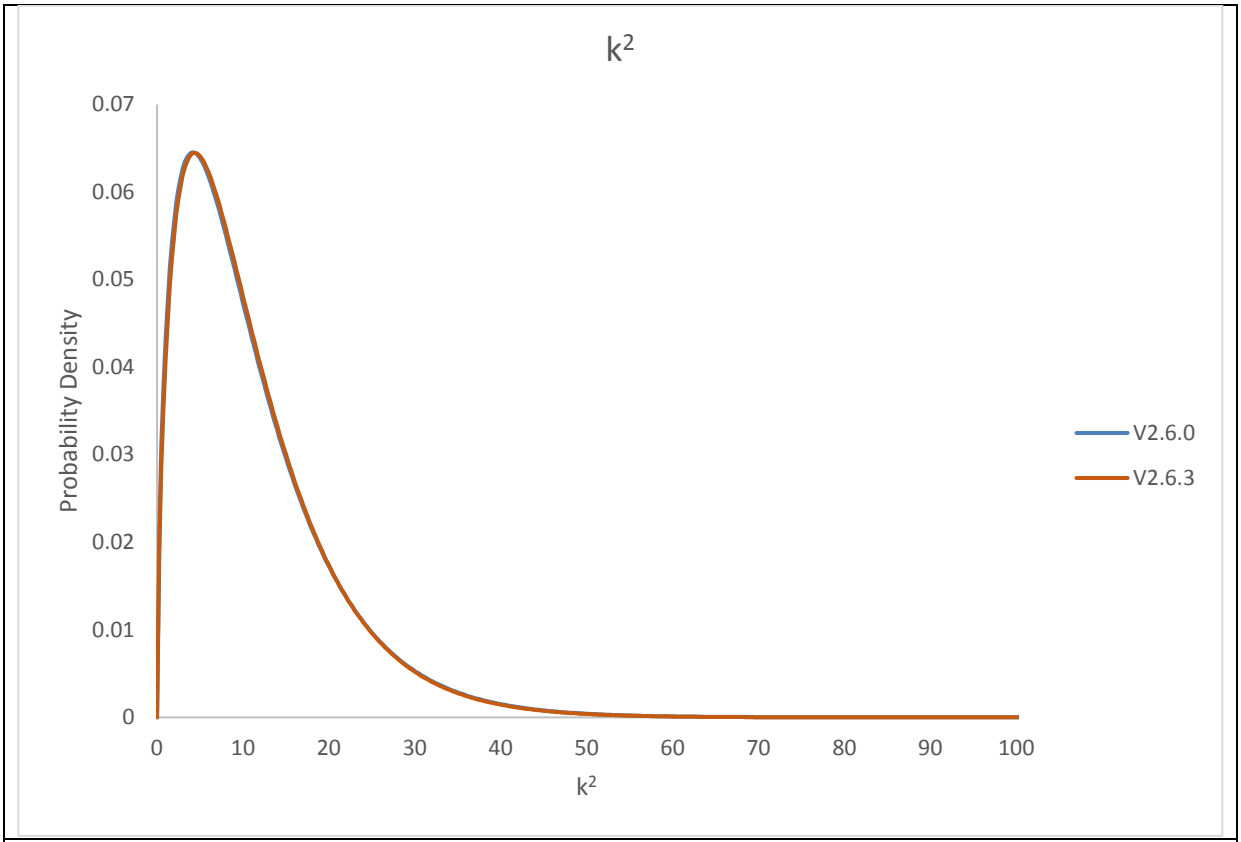
Allele variance



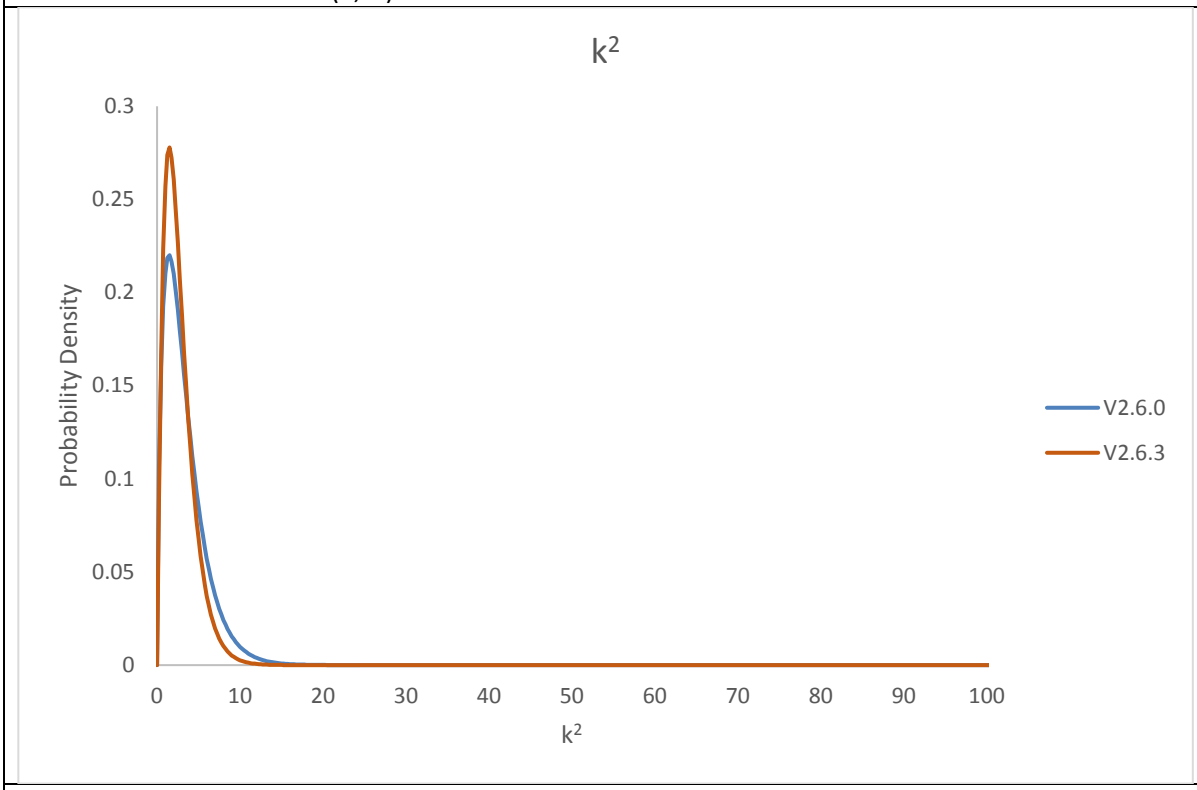
Back stutter variance (-1,0)



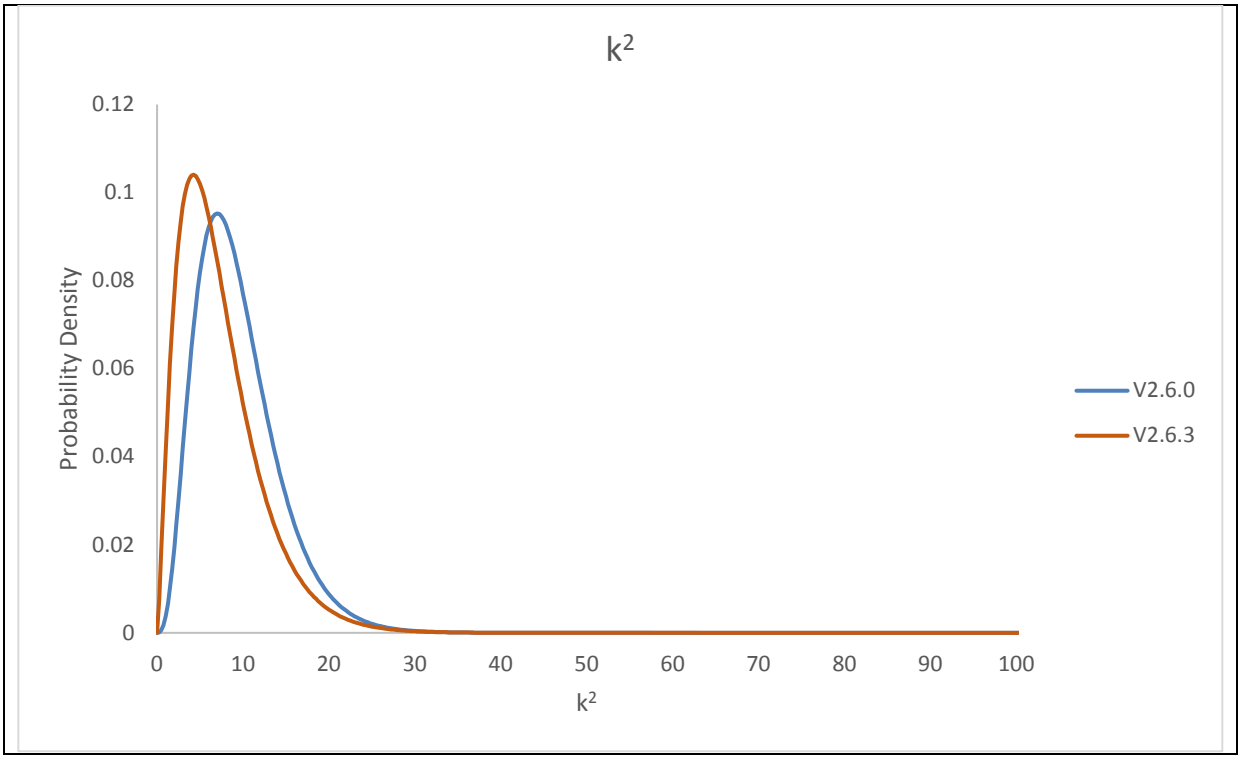
Forward stutter variance (1,0)



Half back stutter variance (0,-2)



Double back stutter (-2,0)

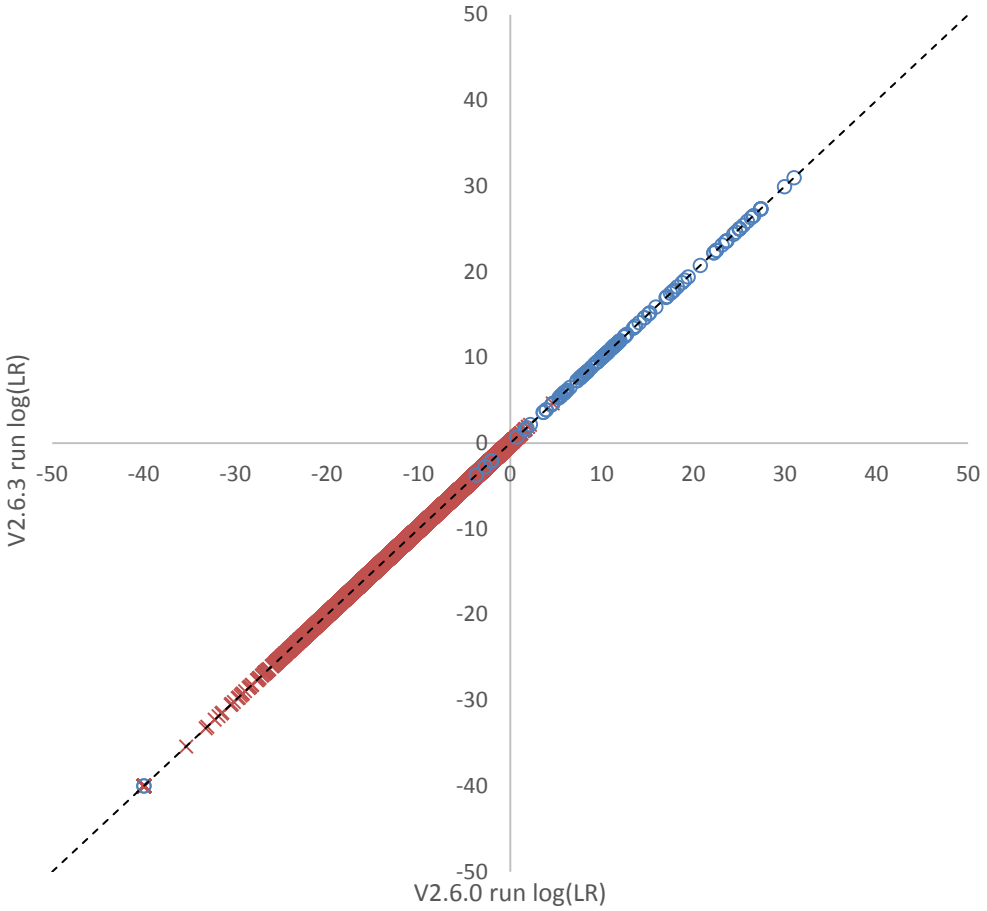


A comparison of the mode and the shape of each distribution demonstrates that there is very little difference between the prior distributions for each peak variant. Any differences observed are likely due to natural MCMC variability. Based on the Model Maker comparisons displayed in Figure 1, it appears appropriate to continue with the same prior variance values as were established within V2.6.0 of STRmix™ and simply migrate these through to V2.6.3. Little to no difference in the outcome of deconvolutions would be anticipated purely from changing these variance values.

Comparison of results obtained in STRmix™ v2.6.0 to v2.6.3

To assess the impact of the changes made to version V2.6.2 and V2.6.3, forty-five mixtures were selected from the 144 studied previously in V2.6.0 (described in section D of the initial JCRCL validation document). This included two-, three- and four contributor mixtures. The process of deconvolution and *LR* generation was repeated using STRmix™ V2.6.3 with the same starting position (seed) as the V2.6.0 runs. The (non-VarNOC) *LR*s of both known contributors and non-contributors were compared to the *LR*s calculated using v2.6.0 (original results displayed in Section D of the JCRCL validation document), the outcome of this is displayed in Figure 2.

Figure 2: *Comparison of the log(LR)s obtained from the analysis of forty-five randomly selected profiles in STRmix™ V2.6.0 and v2.6.3. The H_p true LR data points are represented by blue circles and the H_d true data points are represented by red crosses. The dashed line shows a y = x trend line.*



The 46,400 *LR*s calculated in V2.6.3 are identical to the values calculated in V2.6.0. This may be seen in Figure 2 and the associated data work-up where both the H_p true *LR* data points and the H_d true data points are the same and hence sit on the y = x trend line.

The minimal difference in the Model Maker prior gamma variance distributions and the identical values in the above comparison of the calculation of 'standard' *LR*s indicates that despite the validation work discussed in the main JCRCL document being completed in STRmix™ v2.6.0, the STRmix™ kit and STRmix™ v2.6.3 are suitable for use with Jefferson County Regional Crime Laboratory GlobalFiler™ profiles.

Validation of the VarNOC feature of STRmix™ v2.6.3

The variable number of contributors (hereafter varNOC) function is a new feature in STRmix™ v2.6 onwards. Conventionally, when setting up an analysis in STRmix™, the user is required to input the apparent number of contributors (NOC) to the profile being interpreted. There may be occasions where NOC cannot be assigned with confidence; in these situations, STRmix™ allows for a profile to be interpreted using a range of values for *N*.

However, it is recommended that the varNOC function is only used in casework after all other avenues to reduce the uncertainty in assigning *N* have been explored. Whilst any range can be entered, developmental validation of the varNOC function has only been carried out for a contributor range of (+/-) one. If an increased range is required, it could be argued that too much uncertainty exists in the profile to progress a meaningful interpretation.

In this section the effect of varNOC interpretation on subsequent *LR*s was examined. Nineteen mixtures, where there was some ambiguity in the NOC, including some where there was a difference between the analyst assigned NOC and experimentally designed NOC, were re-interpreted using a contributor range NOC->NOC+1. Here NOC indicates the lower estimate number. The mixtures selected, the experimental design and apparent number of contributors, and the range of contributors used in varNOC calculations are displayed in Table 1.

Table 1*: Summary of mixtures interpreted using varNoC. Apparent NOC indicates the assumption made for NOC in Section D.*

Sample File	Apparent NOC	Design NOC	varNOC Range
GF-MX_CM-Mix-2.1_28c_15s_03_C12_12.hid	4	5	4-5
GF-MX_CM-Mix-2.2_28c_15s_04_D12_12.hid	4	5	4-5
GF-MX_CM-Mix-3.1_28c_15s_05_E12_12.hid	4	5	4-5
GF-MX_CM-Mix-4.1_28c_15s_07_G12_12.hid	4	5	4-5
GF-MX_CM-Mix-4.2_28c_15s_08_H12_12.hid	4	5	4-5
GF-MX_Mix1_E1_28c_15s_02_B02_02.hid	3	4	3-4
GF-MX_Mix1_E2_28c_15s_03_C02_02.hid	3	4	3-4
GF-MX_Mix11_F1_28c_15s_01_A07_07.hid	3	4	3-4
GF-MX_Mix11_F2_28c_15s_02_B07_07.hid	3	4	3-4
GF-MX_Mix3_B1_28c_15s_05_E04_04.hid	3	3	3-4
GF-MX_Mix4_A1_28c_15s_07_G05_05.hid	3	3	3-4
GF-MX_Mix5_B1_28c_15s_06_F07_07.hid	2	2	2-3
GF-MX_Mix5_C1_28c_15s_08_H07_07.hid	2	2	2-3
GF-MX_Mix5_D1_28c_15s_02_B08_08.hid	2	2	2-3
GF-MX_Mix8_B1_28c_15s_04_D01_01.hid	2	2	2-3
GF-MX_Mix10_A2_28c_15s_04_D04_04.hid	3	3	3-4
GF-MX_Mix10_B1_28c_15s_05_E04_04.hid	3	3	3-4
GF-MX_Mix12_C2_28c_15s_03_C08_08.hid	3	3	3-4
GF-MX_Mix3_B2_28c_15s_06_F04_04.hid	4	3	3-4

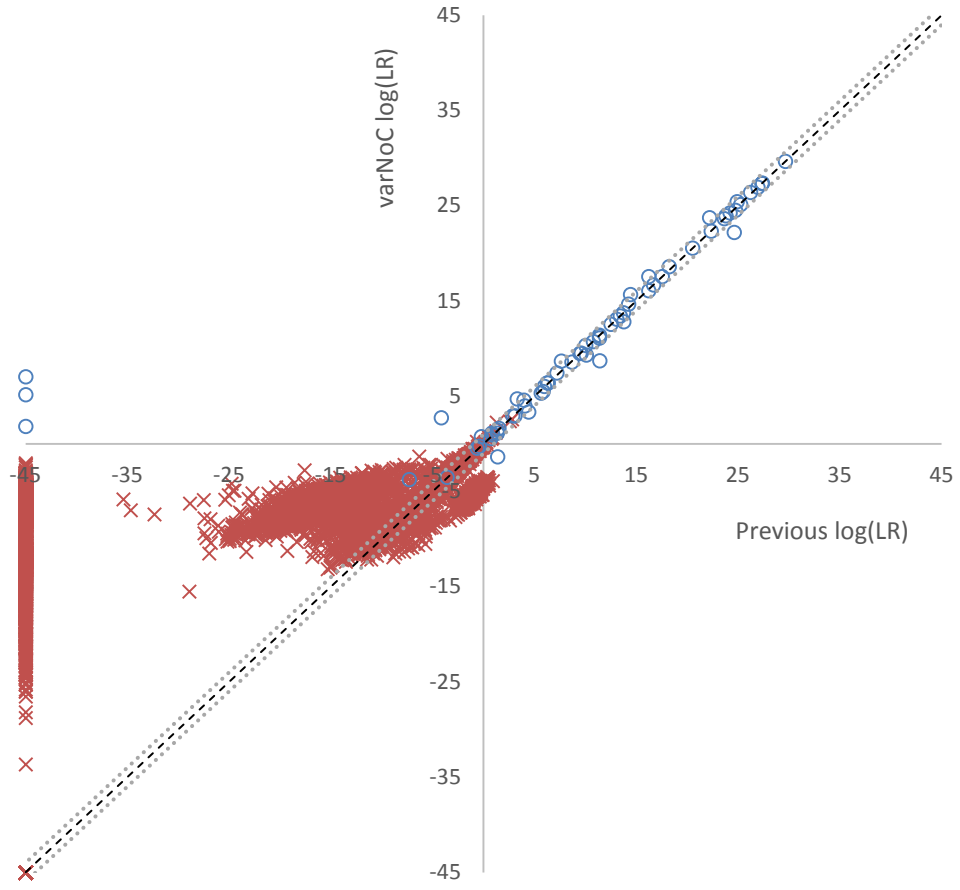
The database search function was used in the same manner as section D to calculate an *LR* for each individual on the database; both contributors and non-contributors, with an *LR* threshold of 0, however a varNOC stratified *LR* was calculated rather than a 'standard' *LR*. The proposition sets used in this section of the analysis were:

H_p: The DNA originated from the database individual and *the varNOC range of NOC minus 1* unknown individuals

H_d: The DNA originated from the *varNOC range of NOC* unknown individuals

The NIST Caucasian allele frequencies with a theta (F_{ST}) of 0.01 (1%) were used for Database Search *LR* calculations. The original (standard, non-varNOC) *LR*s calculated in STRmix™ v2.6.0 were compared with the varNOC *LR*s (V2.6.3) and are plotted in Figure 3.

***Figure 3**: Comparison of the log(varNoC LR) using STRmix™ v2.6.3 and the log(LR) using STRmix™ v2.6.0. True contributors are shown as blue circles and known non-contributors are shown as red crosses. The dashed line indicates a $y = x$ trend, and dotted lines indicate a ±1 order of magnitude from the $y = x$ trend line.*



As shown in Figure 3, the results of the known contributors show good concordance between the varNOC *LR*s (v2.6.3) and the *LR*s (v2.6.0) calculated using a single value for the NOC assignment. Figure 3 shows that 52 of the 67 observations (78%) for the true contributors exhibited *LR*s that were within one-order of magnitude of the *LR* obtained from varNOC deconvolution methods. This is shown by the majority of the blue circles falling in-between the grey dotted lines that indicate ±1 order of magnitude difference from the $y = x$ trend line.

It also may be noted from Figure 3, that 3 of the 67 observations of true contributor data points changed from outright exclusions with the initial work to *LR*s that favoured inclusion, when the varNOC feature was used. These values were investigated and were found to be instances where the assigned number of contributors was an underestimation with regards to the experimental design. The under assumption of one less contributor to a

mixture can affect the *LR* of true contributors and lead to exclusions. The consideration of a range of contributors within this interpretation entails that true contributors are not excluded.

There is more pronounced variation observed with the non-contributor data points; however, this is expected due to the likely low weights of the genotype combinations that non-contributors associate with and due to the difference in approach to dealing with the number of contributors. Typically, the trend is similar to that observed from section F, as most deconvolutions have included a NOC+1 consideration and hence the genotypes and weights can be spread amongst more contributors.

Overall, these results demonstrate the robustness of the *LR*s calculated for the known contributors to each mixture when deconvoluted using a range of contributors. Any variability can be attributed to not only MCMC variation, but also a varNOC deconvolution process.

Conclusion

This Addendum describes the Jefferson County Crime laboratory's internal validation activities for STRmix™ V2.6.3 using the GlobalFiler™ kit on a 3500 instrument. It has been shown that STRmix™ v2.6.3 is suited for its intended use at the JCRCL for the interpretation of profiles generated from crime scene samples. This addendum also details work to demonstrate that the VarNOC feature of the software is fit for purpose on casework samples.

The project work has been performed in accordance with the FBI Quality Assurance Standards for Forensic DNA Testing Laboratories, the ANAB Accreditation Requirements, and the SWGDAM Guidelines for the Validation of Probabilistic Genotyping.

Signatures:

Beth Hewitt [signature redacted]

DNA Technical Leader

Jefferson County Regional Crime Laboratory