Internal Validation of STRmix™ V2.4 Sacramento County District Attorney's Crime Laboratory

STRmix™ internal validation

This document describes the internal validation of STRmix™ V2.4 at the Sacramento County District Attorney's Crime Laboratory (SacDA). STRmix™ V2.4 is a fully continuous probabilistic genotyping program for the interpretation of autosomal STR profiles. While STRmix™ was designed to interpret evidence profiles for any number of contributors, the SacDA protocol has been validated only for one to four contributors. The validation was specific to Promega® PowerPlex® Fusion 6C Amplification Kit results from 3500xL capillary electrophoresis instruments. The population databases used throughout the validation were taken from four populations: African-American, Caucasian, Hispanic, and Asian. [1].

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 SacDA has undertaken in-house before the implementation of STRmix™ into routine casework. This document follows the internal validation section of the SWGDAM Guidelines for the Validation of Probabilistic Genotyping Systems [3] and satisfies Standard 8.7 of the FBI Quality Assurance Standards for Forensic DNA Testing Laboratories (September 1, 2011) [4]. The internal validation includes the examination of known and non-probative evidence samples, investigations into reproducibility and precision, sensitivity and stochastic studies, and mixture studies. The section where specific SWGDAM guidelines are discussed in this document is cross referenced in Appendix 2.

The data and results of all experiments related to the internal validation of STRmix™ at the Sacramento County District Attorney's Crime Laboratory are retained electronically and in hard copy.

STRmix™ parameters

The parameters described in the document *Estimation of STRmix™ parameters for SacDA* 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 SWGDAM guidelines:

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

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

A dilution series of three single source profiles were constructed where the peak heights ranged above and below 450 rfu (stochastic threshold). Note: the stochastic threshold will only apply to the interpretation of casework reference samples. The template DNA in picograms for the serial dilutions were: 1000, 200, 100, 50, 25, and 10 pg. The profiles were analyzed with an analytical threshold (AT) = 100 rfu.

The profiles were interpreted in STRmix™ using the propositions:

H_1: The DNA originated from the person of interest

H_2: The DNA originated from an unknown individual

The Likelihood Ratio (*LR)* was calculated for the true contributor for three population groups (African-American, Caucasian, and Hispanic) with an F_{ST} (θ) of 0.01. A plot of log(*LR*) versus input DNA is provided in Figures 1-3.

Figure 1: Plot of log(*LR*) versus input amount (pg) for African-American population

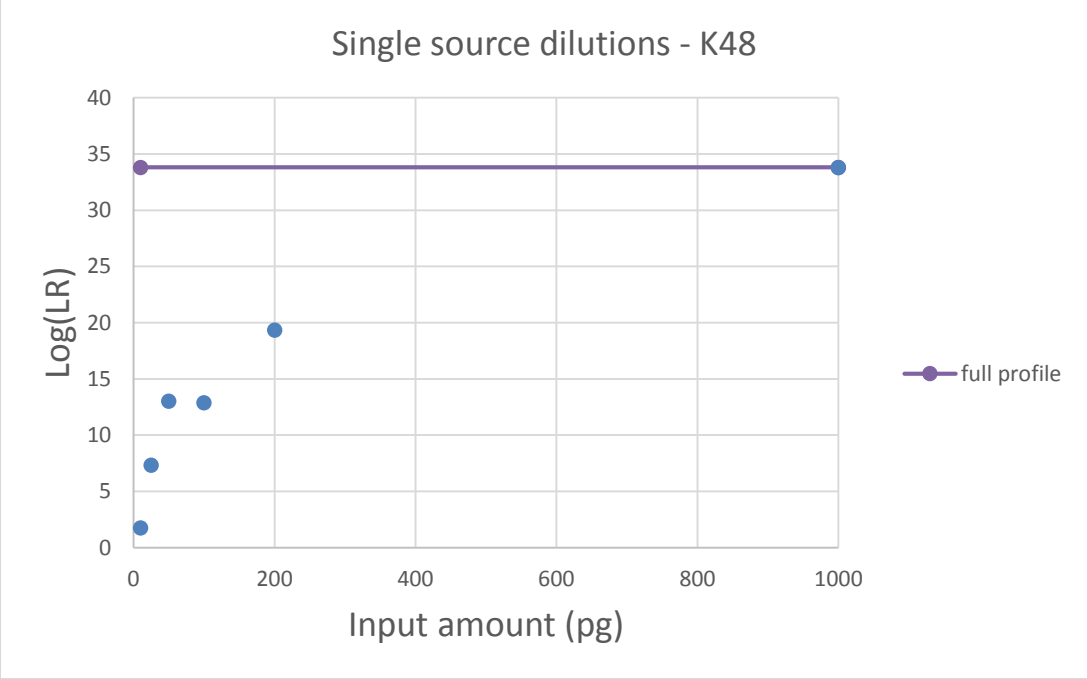


Figure 2: Plot of log(*LR*) versus input amount (pg) for African-American population

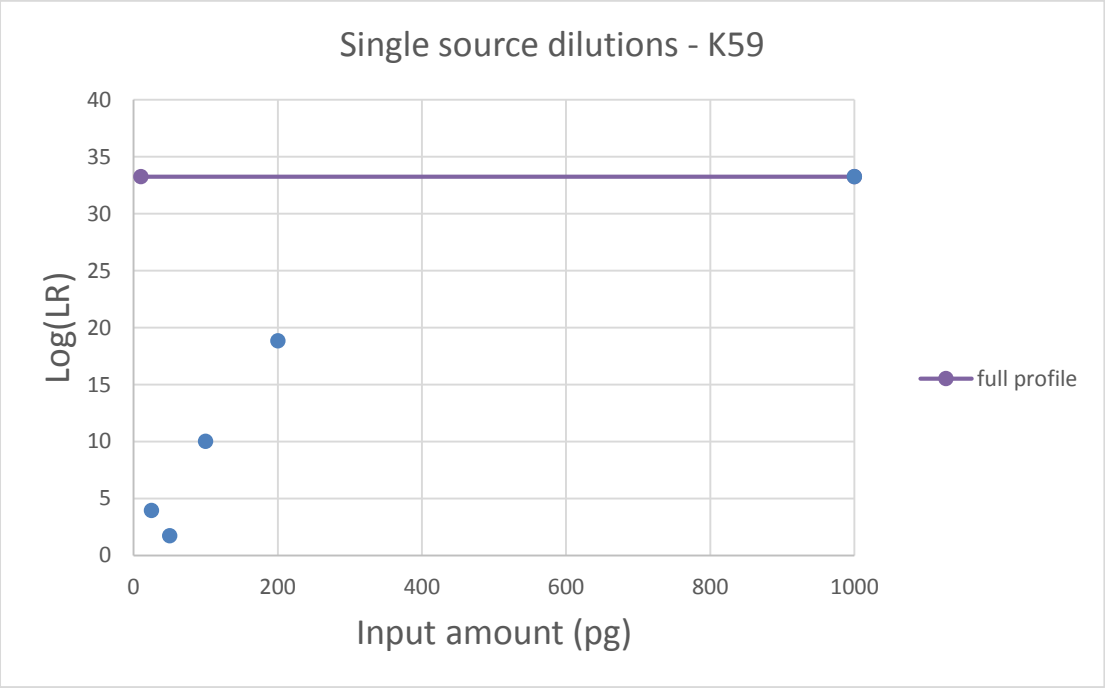
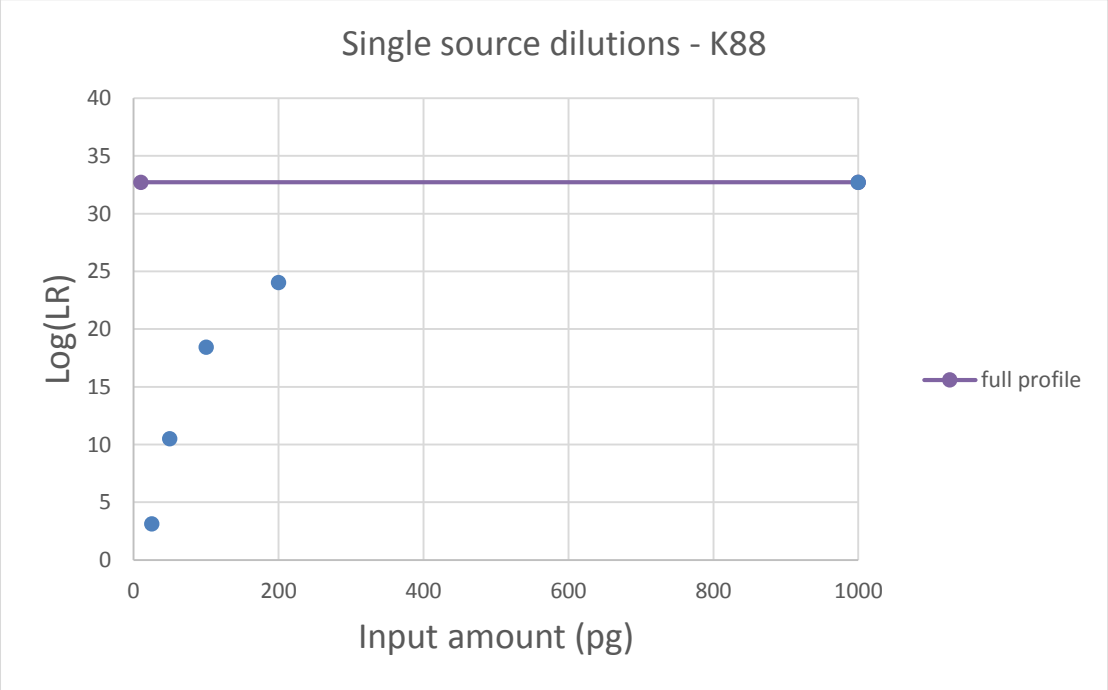


Figure 3: Plot of log(*LR*) versus input amount (pg) for African-American population



Inspection of the plot shows the log(*LR)* progressing from the value for the single source log(*LR)* calculated for a full profile towards log(*LR)* = 0 as the DNA template decreases. As expected, the weights for genotypes considering dropout increased as template drops. In addition, the DNA amounts from the STRmix™ output (*t* or template mass parameter) decline steadily as peak heights decrease.

Reproduction of single source LR

There is a small subset of profiles where the 'answer' is known or can be estimated easily [5]. These include single source profiles where the weight is one (or 100%) for one genotype at each locus. The point estimate *LR* was calculated 'by hand' at each locus for ten single source profiles and the individual locus *LR*s compared with the STRmix™ results. The 'by hand' calculated and STRmix™ results for the single source profiles are given in Table 1 (African-American population provided).

Table 1: 'By hand' (Excel) calculation of *LR* versus STRmix™ results for ten single source profiles

Sample	STRmix™	Excel
K88	1.04E+33	1.04E+33
K89	3.06E+31	3.06E+31
K90	1.69E+32	1.69E+32
K91	8.45E+34	8.45E+34
K92	2.33E+33	2.33E+33
K96	3.53E+32	3.53E+32
K97	7.72E+32	7.72E+32
K98	2.88E+35	2.88E+35
K99	5.02E+33	5.03E+33
K100	2.68E+32	2.68E+32

This was undertaken twice for one of the samples (K88); once using an F_{ST} (or θ) value = 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 [6] formulae (or equations 4.10 from NRCII [7]) are applied. For single source profiles:

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

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

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

$$\frac{x_i+\frac{1}{k}}{N_a+1} \qquad [3]$$

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

The 'by hand' calculated and STRmix™ results for a single source profile for θ=0 and θ=0.01 are given in Table 2 (African-American population provided).

Table 2: 'By hand' (Excel) calculation of *LR* versus STRmix™ results for one single source profile (K88) with varying F_{ST} values

Locus	Locus *LR*s		Locus *LR*s	
	Excel Θ = 0	STRmix Θ = 0	Excel Θ = 0.01	STRmix Θ = 0.01
D3S1358	26	26	23.5	23.5
D1S1656	156	156	118	118
D2S441	22.1	22.1	20	20
D10S1248	7.75	7.75	7.53	7.53
D13S317	8.53	8.53	8.16	8.16
Penta E	177	177	131	131
D16S539	8.72	8.72	8.41	8.41
D18S51	41.5	41.5	36.6	36.6
D2S1338	34.7	34.7	31	31
CSF1PO	42.7	42.7	35.5	35.5
Penta D	35.4	35.4	28.2	28.2
TH01	107	107	71	71
vWA	16	16	13.9	13.9
D21S11	34.8	34.8	27.8	27.8
D7S820	17	17	15.5	15.5
D5S818	9.57	9.57	9.23	9.23
TPOX	21.3	21.3	18	18
D8S1179	7.77	7.77	7.55	7.55
D12S391	46.7	46.7	40	40
D19S433	9.69	9.69	9.33	9.33
SE33	2580	2580	671	671
D22S1045	33.7	33.7	29.8	29.8
FGA	75.4	75.4	61.7	61.7
Total	**9.55E+34**	**9.55E+34**	**1.04E+33**	**1.04E+33**

The results in Tables 1 and 2 show that STRmix™ is giving the expected answer based on the population genetic model being used and that *LR* calculations between STRmix™ and Excel are identical.

Section B: Use of peak heights

This section covers the following SWGDAM guideline:

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. As the weights for genotypes considering dropout increase, the weights for genotype combinations for the *true* contributors decrease and subsequently the *LR* decreases. This can be observed in Figures 1-3. This is the expected result.

STRmix™ treats all peaks that are greater than the saturation threshold (calculated as 30,000 rfu) qualitatively and not quantitatively. Saturated profiles should not be interpreted within STRmix™, as a profile that exceeds the saturation threshold is likely to have higher stutter peak heights than expected by STRmix™. The effect of higher stutter values can be seen by reviewing the stutter variance values which deviated significantly from the stutter variance mode.

Seven single source samples were amplified with deliberately high input amounts of DNA (6 – 8 ng). The profiles were interpreted in STRmix™ and the weights were reviewed. All profiles were interpreted correctly, with weights = 1 for the known genotype combination.

Section C: Weights

This section covers the following SWGDAM guideline:

> 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.

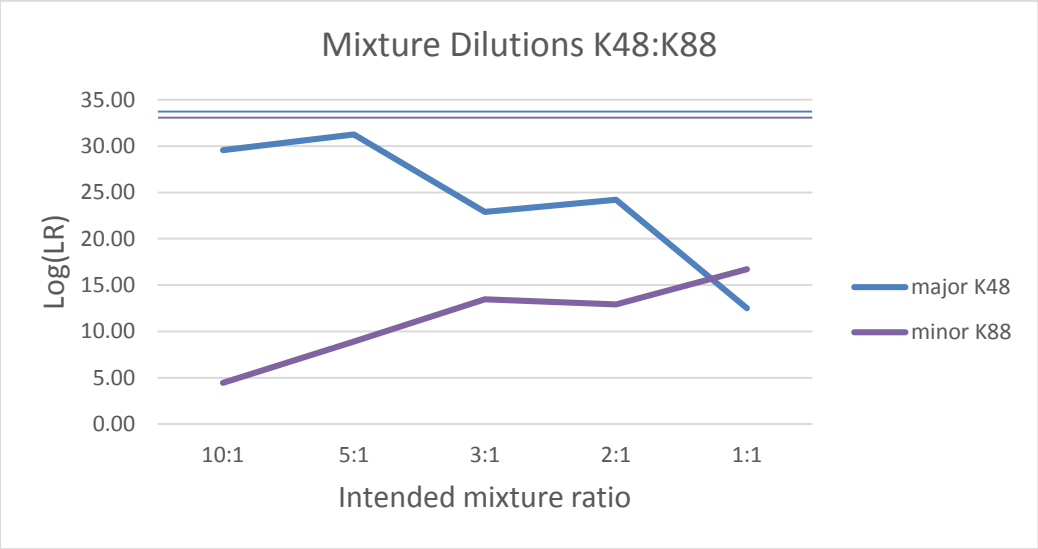
The weights are described as the primary output from STRmix™. They can be used as a diagnostic of the deconvolution process and should be intuitively correct, where the most supported genotypes have the highest weights.

A two person mixture series was constructed in the following ratios 10:1, 5:1, 3:1, 2:1 and 1:1. The total amount of DNA in the profiles was approximately 500 pg DNA. The profiles were interpreted in STRmix™ under the following propositions and an LR was calculated:

> H_1: The DNA originated from the person of interest (known major or minor) and an unknown individual
>
> H_2: The DNA originated from two unknown individuals

A plot of Log(*LR*) for the major (blue) and minor (purple) contributors for the mixture series is provided in Figure 4. (*LR* values for full profiles found as horizontal lines at top of plot, blue=major, purple=minor.) Data is for the African-American population and based on the 99% 1-sided lower HPD value.

Figure 4: Log(*LR*) for the major (blue) and minor (purple) contributors for the mixture series



Inspection of Figure 4 shows that the *LR* decreases by approximately half (10 orders of magnitude) for the 1:1 mixture when compared to the single source *LR* calculated for the major or minor contributor. The decrease starts at ratios where it is reasonable for the assignment of alleles from a major and minor contributor to be more ambiguous (approximately 5:1 or 3:1). The *LR* for the minor contributor increases as their DNA contribution increases and the ratio between the two contributors becomes more similar. In addition, the mixture proportions in the STRmix™ output changed appropriately as the mixture ratios varied.

Section D: Sensitivity and specificity and mixtures

This section covers the following SWGDAM guidelines:

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 SacDA Fusion 6C mixtures was undertaken as per Taylor [8]. With respect to interpretation methods, sensitivity is defined as the ability of the software to reliably resolve the DNA profile of known contributors within a mixed DNA profile for a range of starting DNA templates. The log(*LR*) for known contributors (H_1 true) should be high and should trend to 0 as less information is present within the profile. Information includes amount of DNA from the contributor of interest, conditioning profiles (for example the victim's profile on intimate samples), replicates and decreasing numbers of contributors. Specificity is defined as ability of the software to reliably exclude known non contributors (H_2 true) within a mixed DNA profile for a range of starting DNA templates. The log(*LR*) should trend upwards to 0 as less information is present within the profile.

Specificity and sensitivity were tested by calculating the *LR* for 81 two, three and four person profiles for both known contributors and known non-contributors. Thirty-one 2-person, twenty-seven 3-person, and twenty-three 4-person mixed DNA profiles were generated at various mixture ratios by the laboratory. Additionally, six 5-person mixtures were generated. However, two attempts to deconvolute 5-person mixtures resulted in STRmix™ run time errors. SacDA will not interpret 5-person mixtures. A summary of the DNA mixed profiles is given in Table 3.

Each profile was interpreted in STRmix™ and compared to 200 known non-contributors and 27 known contributors (2-4 of which were present in each mixture) using the Database Search function within STRmix™ for a total of 18,160 *LR*s. The profiles from non-contributors were artificially generated using the NIST allele frequency database.

These profiles represent a wide range of profiles likely to be encountered by the laboratory. The profiles are of varying DNA quantity and mixture proportions. The contributors include homozygote and heterozygote alleles, and there is varying amounts of allele sharing across the different loci (guideline 4.1.6.5). Given the template amounts, allele and/or locus dropout was expected to occur within the profiles containing the lower DNA amounts (guideline 4.1.7.1).

The propositions considered were:

> H_1: The DNA originated from the true contributor and *N*-1 unknown individuals
>
> H_2: The DNA originated from *N* unknown individuals

Plots of log(*LR*) versus the average peak height (*APH*) per contributor for the two, three and four contributor mixtures are given in Figure 5. Exclusions (*LR*=0) are plotted as log(*LR*)=-30. The per contributor amount of DNA for H_2 true contributors is taken from the lowest of the known contributors. The *APH* per known contributor is calculated using the unmasked and unshared alleles. The lowest contributor *APH* for each profile was used for the H_2 true contributors. The results of all comparisons are provided in Figure 5.

Table 3: Summary of experimental design for specificity and sensitivity tests

Two person

Mixture ratio	Number of samples	Range of DNA (ng) for smallest contributor
1:1	7	0.025 – 0.500
2:1	3	0.250 – 0.833
3:1	6	0.025 – 0.625
5:1	6	0.050 – 0.417
10:1	6	0.068 – 0.300
20:1	3	0.120 – 0.500

Three person

Mixture ratio	Number of samples	Range of DNA (ng) for smallest contributor
1:1:1	10	0.040 – 0.833
3:2:1	4	0.040 – 0.300
3:3:1	3	0.107 – 0.357
4:1:1	3	0.125 – 0.417
4:2:1	3	0.107 – 0.357
10:5:1	4	0.100 – 0.800

Four person

Mixture ratio	Number of samples	Range of DNA (ng) for smallest contributor
1:1:1:1	5	0.050 – 0.625
4:3:2:1	6	0.040 – 0.500
4:4:1:1	3	0.075 – 0.250
5:1:1:1	3	0.094 – 0.313
5:3:1:1	3	0.075 – 0.250
10:5:2:1	3	0.060 – 0.220

Five person

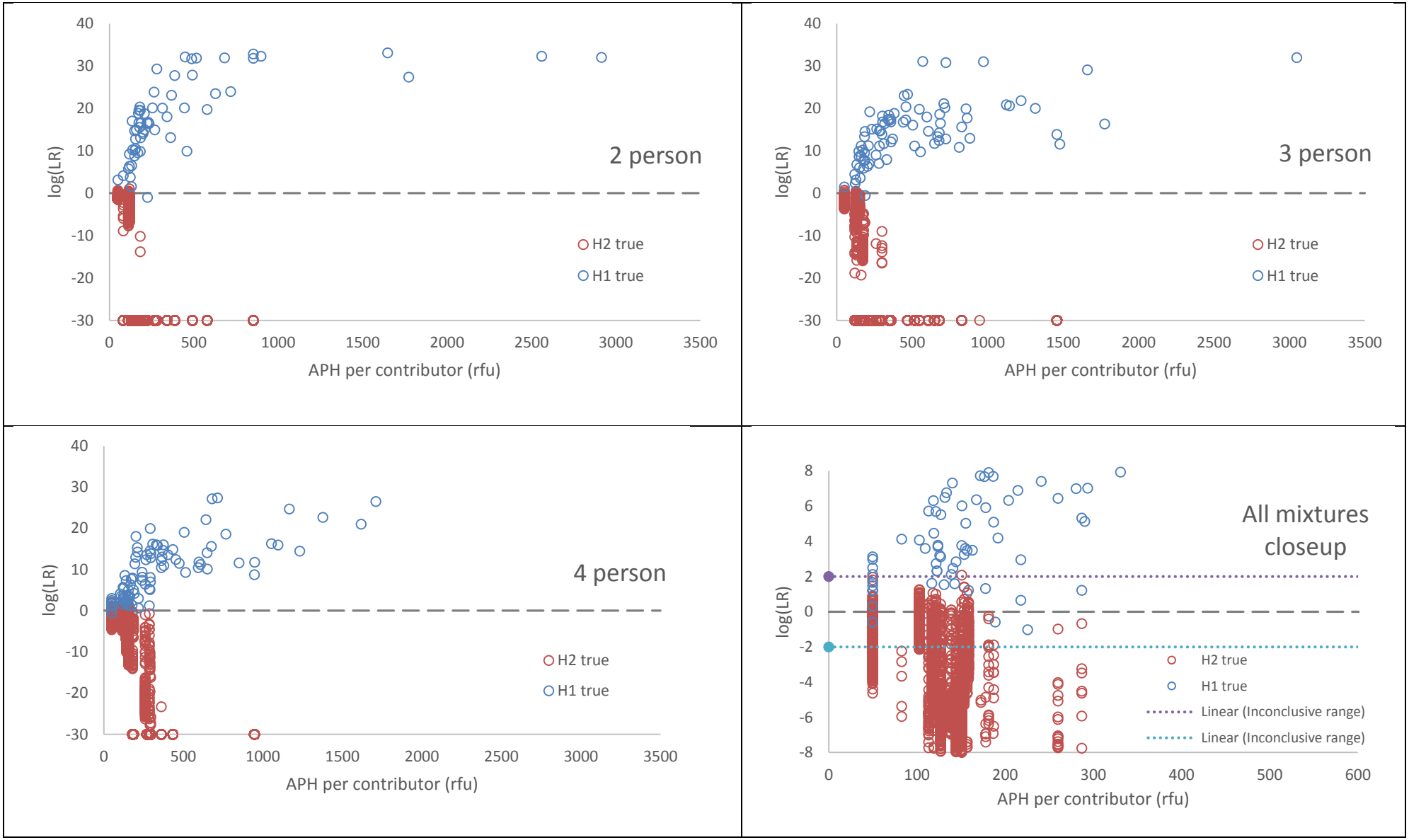
Mixture ratio	Number of samples	Range of DNA (ng) for smallest contributor
10:1:1:1:1	3	0.038 – 0.077
15:1:1:1:1	3	0.011 – 0.042

Inspection of the plots in Figure 5 indicates that, as expected, at high template STRmix™ correctly and reliably resulted in high *LRs* for true contributors and low *LRs* for false contributors. At low template or high contributor number STRmix™ correctly and reliably reported that the analysis of the sample tends towards an uninformative or inconclusive *LR*.

To determine an inconclusive range, the plots in Figure 5 can help inform the limits of STRmix™, particularly the lower limit of *LR*s where an H_1 true hypothesis results in an LR less than 1 and the upper limit where an H_2 true hypothesis results in an *LR* greater than 1. The experiment results show that false positives (an LR greater than 1 where H_2 is true) and false negatives (an LR less than 1 where H_1 is true) generally occur between an *LR* of .01 and 100 (or Log(*LR*) between -2 and 2) and when the APH for the contributor is <250 rfu. Upon examining the *LR*s when H_1 is true, fifteen gave

*LR*s in the inconclusive range (between 0.01 and 100), with the lowest at 0.10. All of these profiles exhibited a lot of allele dropout and often an ambiguous number of contributors. An inconclusive result for these profiles is a reasonable interpretation. Upon examining the *LR*s when H_2 is true, 1,375 gave *LR*s in the inconclusive range (between 0.01 and 100). There was only one profile that gave an *LR* greater than 100 (121.1), and this false contributor is a parent/child of one of the true contributors. The profiles from the next highest ten *LR*s were further examined, and an inconclusive result was found to be a reasonable interpretation and consistent with qualitative expectations.

Figure 5: Log(*LR*) versus APH for four, three and two person mixtures amplified by the SacDA laboratory.



Section E: Alternate propositions

This section covers the following SWGDAM guideline:

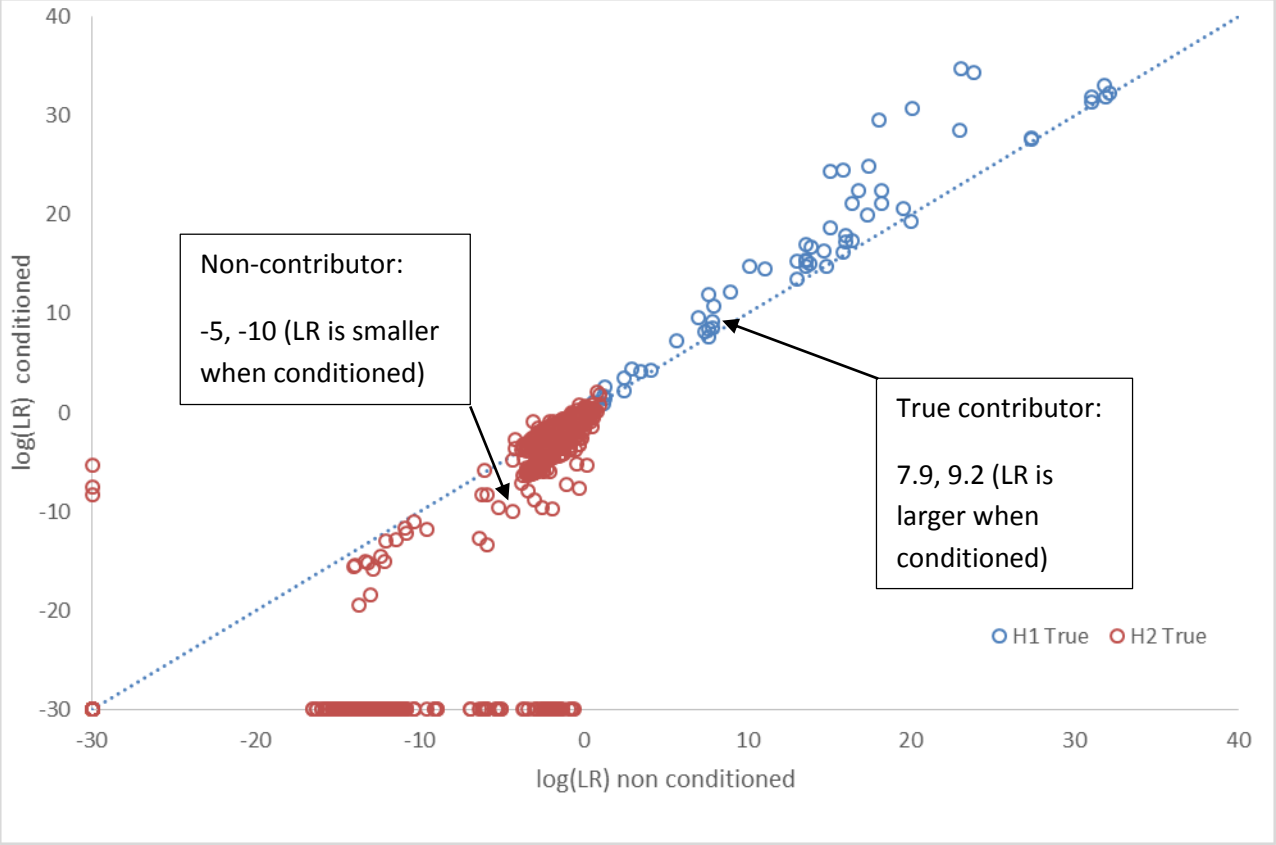
> 4.1.2.1. The laboratory should evaluate more than one set of hypotheses for individual evidentiary profiles to aid in the development of policies regarding the formulation of hypotheses. For example, if there are two persons of interest, they may be evaluated as co-contributors and, alternatively, as each contributing with an unknown individual. The hypotheses used for evaluation of casework profiles can have a significant impact on the results obtained.

Fifteen 2, 3, and 4 person mixtures were taken from Section D (non-conditioned) and reinterpreted in STRmix™ with alternate propositions. Each mixture was deconvoluted twice in STRmix™, once assuming one known contributor, and again assuming another known contributor. In these interpretations assume one of the contributors is a known under both H_1 and H_2. The different propositions being considered are:

> H_1: The DNA originated from the known individual, the true contributor, and N-2 unknown individuals
>
> H_2: The DNA originated from the known individual and N-1 unknown individuals

A plot of the log(*LR*) with the conditioned contributor plotted against the log(*LR*) without the conditioned contributor is provided in Figure 6.

Figure 6: Log(*LR*) with the conditioned contributor plotted against log(*LR*) without the conditioned contributor



Inspection of the plot in Figure 6 shows that the addition of more relevant information, such as the addition of assumed contributors, improves the performance of STRmix™. Under H_1 true proposition, the *LR*s trend

upward with the addition of an assumed contributor. Under H_2 true proposition, the *LR*s trend downward with the addition of an assumed contributor. It makes sense that, with more information, the *LR*s for true contributors will get stronger, and the *LR*s for non-contributors will get weaker.

The addition of correct conditioning profiles (known contributors under both H_1 and H_2) improved the performance.

Section F: Assigning number of contributors

This section covers the following SWGDAM guideline:

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

The effect of the uncertainty in the number of contributors was tested by both increasing and decreasing the number of contributors compared with the known (*N*+1 and *N*-1 trials). The true number of contributors to a profile is always unknown. Analysts are likely to add contributors in the presence of:

- an artifact
- high stutter
- forward stutter peaks.

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. The assumption of one fewer contributor than that actually present may be made when:

- contributors are at very low levels and dropping out (or visible below the analytical threshold)
- mixture profiles are consisting of DNA from individuals with similar profiles, such as relatives.

Addition of one contributor

Six single source profiles, five 2-person profiles, and five 3-person profiles were interpreted as 2, 3 and 4 person profiles, respectively (*N*+1). The *LR* for both the known contributors and 200 known non-contributors (see the specificity and sensitivity studies, Section D) were calculated. The propositions considered were:

H_1: The DNA originated from the true contributor and *N* unknown individuals

H_2: The DNA originated from *N*-1 unknown individuals

The *LR* was compared for the known contributors and known non-contributors under the assumption of *N* and *N*+1 contributors. Table 4 summarizes the results of adding a contributor. Under the correct number of contributors, all of the true contributors are included, and all of the non-contributors are excluded. The addition of a contributor still leads to an inclusion of the true contributors, but more than half of the non-contributors are now in the inconclusive range. Only one sample gave an *LR* greater than 100 (123.4) for a non-contributor when assuming *N*+1 contributors.

Table 4: Summary of *LR* results comparing contributors to non-contributors

Sample	Contributors			Non-contributors		
	Exclusion *LR*<0.01	Inconclusive 0.01<*LR*<100	Included *LR*>100	Exclusion *LR*<0.01	Inconclusive 0.01<*LR*<100	Included *LR*>100
N contributors	0	0	31	3600	0	0
N+1 contributors	0	0	31	1442	2157	1

A plot of log(*LR*) (assuming *N*) vs log(*LR*) (assuming *N*+1) is provided in Figure 7. A summary of the original log(*LR*) assuming the correct number of contributors (*N*) and after assuming *N*+1 is given in Table 5. Inspection of the values in Table 5 shows that as expected there is no significant effect on the *LR*, and any effect is to lower the *LR*.

Figure 7: plot of log(*LR*) (assuming *N*) vs log(*LR*) (assuming *N*+1)

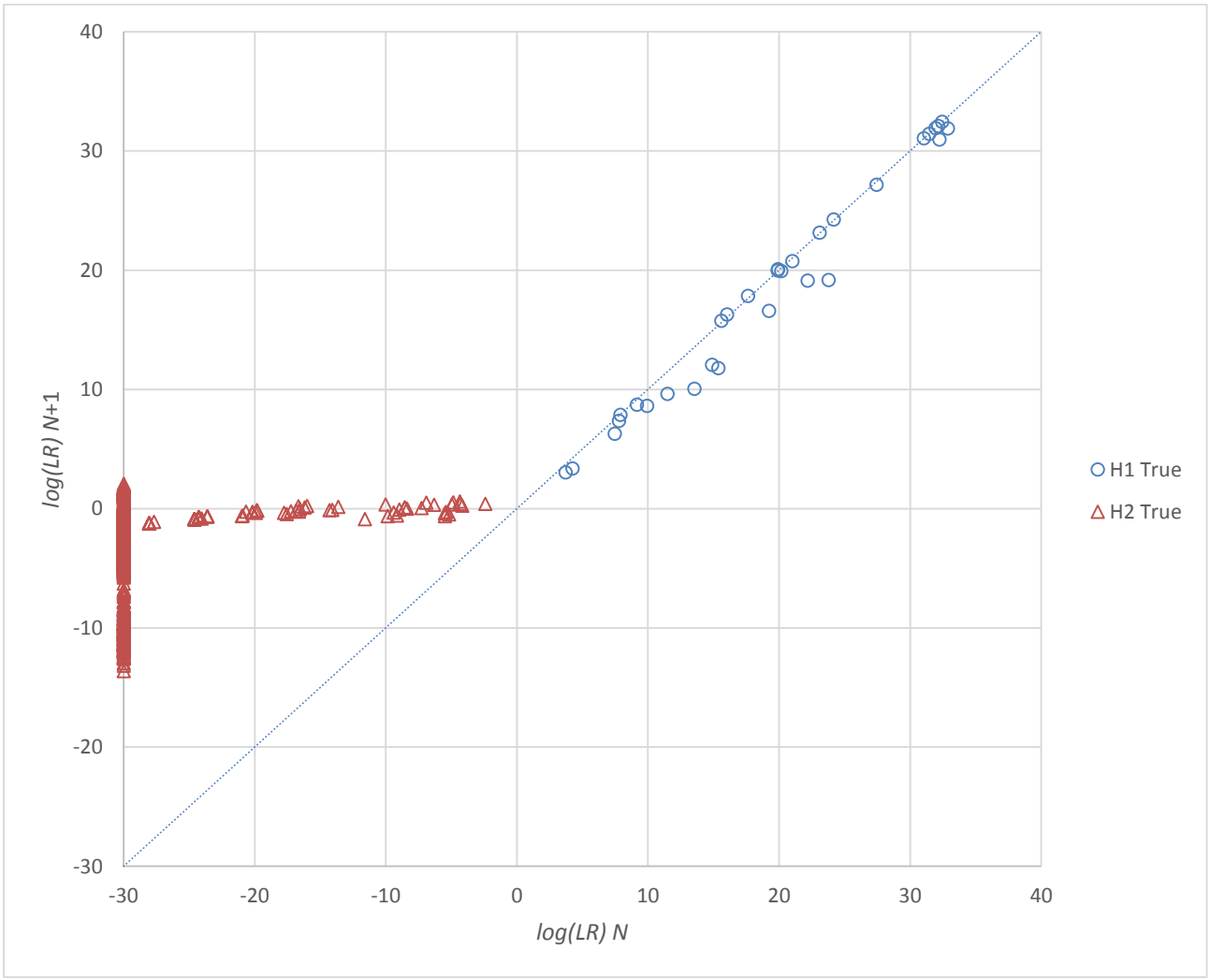


Table 5: Log(*LR*) values (H_1 true) for 1, 2, and 3 person mixtures assuming *N* and *N*+1 contributors

Sample	*N*	Cont.	*N* log(*LR*)	*N*+1 log(*LR*)
K48_25pg	1	K48	7.46	6.27
K48-1ng	1	K48	31.45	31.45
K59_1ng	1	K59	32.44	32.44
K59_200pg	1	K59	19.23	16.57
K88_200pg	1	K88	22.17	19.13
K88_25pg	1	K88	3.71	3.02
1-1	2	K55	24.16	24.25
		K66	19.88	19.98
2-1	2	K53	27.43	27.16
		K55	32.86	31.90
3-2	2	K66	32.24	30.95
		K74	13.54	10.06
4-3	2	K44	31.93	31.91
		K53	14.88	12.07
5-3	2	K74	32.14	32.09
		K87	4.24	3.36
6-1	3	K46	17.63	17.83
		K50	15.59	15.75
		K55	19.91	20.07
7-3	3	K54	16.02	16.27
		K74	7.89	7.87
		K75	9.14	8.71
8-1	3	K85	20.17	19.94
		K86	21.00	20.74
		K87	23.08	23.14
9-3	3	K85	23.78	19.18
		K89	11.48	9.62
		K91	15.36	11.78
10-3	3	K44	31.04	31.07
		K75	9.93	8.62
		K86	7.77	7.35

Subtraction of one contributor

Five 2 person profiles were selected that had no more than two alleles per locus. Six 3 person profiles were selected with no more than four alleles at a locus, and three 4 person profiles were selected with no more than six alleles at a locus. Each of these profiles were interpreted assuming one, two or three contributors, respectively (*N*-1). The *LR* for both the known contributors and 200 known non-contributors (see the specificity and sensitivity studies, Section D) were calculated. The propositions considered were:

H_1: The DNA originated from the true contributor and *N*-2 unknown individuals

H_2: The DNA originated from *N*-1 unknown individuals

Table 6 summarizes the results of subtracting a contributor. Under the correct number of contributors, all of the true contributors are included. 473 non-contributors fell in the inconclusive range, and 2 non-contributors had inclusionary *LR*s (111 and 120). It should be noted that the same non-contributor gave both of these inclusionary *LR*s and that this person is related to a true contributor in the mixture. The subtraction of a contributor leads to an exclusion of about half of the true contributors, and all of the non-contributors.

Table 6: Summary of *LR* results comparing contributors to non-contributors

Sample	Contributors			Non-contributors		
	Exclusion *LR*<0.01	Inconclusive 0.01<*LR*<100	Included *LR*>100	Exclusion *LR*<0.01	Inconclusive 0.01<*LR*<100	Included *LR*>100
N contributors	0	0	40	2663	473	2*
N-1 contributors	20	1	19	3138	0	0

**non-contributor related to true contributor*

A plot of log(*LR*) (assuming *N*) vs log(*LR*) (assuming *N*-1) is provided in Figure 8. A summary of the original log(*LR*) assuming the correct number of contributors (*N*) and after assuming *N*-1 is given in Table 7. Differences in the log(*LR*) of true contributors resulting in an inclusion (under *N* contributors) to an exclusion (under *N*-1 contributors) have been highlighted in gray on the table. Differences in the log(*LR*) resulting in an inconclusive (under *N* contributors) to an exclusion (under *N*-1 contributors) have been highlighted in blue on the table. Inspection of the values in Table 6 shows that underestimating the number of contributors can lead to false exclusions of true contributors. It should be noted that these profiles were chosen because they had allele counts that allowed STRmix™ to deconvolute given an incorrect number of contributors (*N*-1) (eg, no more than four alleles in a 3 person mixture). However, most of these profiles have indications, either in peak height imbalances or peaks below the analytical threshold, that the number of contributors is incorrect.

Figure 8: plot of log(*LR*) (assuming *N*) vs log(*LR*) (assuming *N*-1)

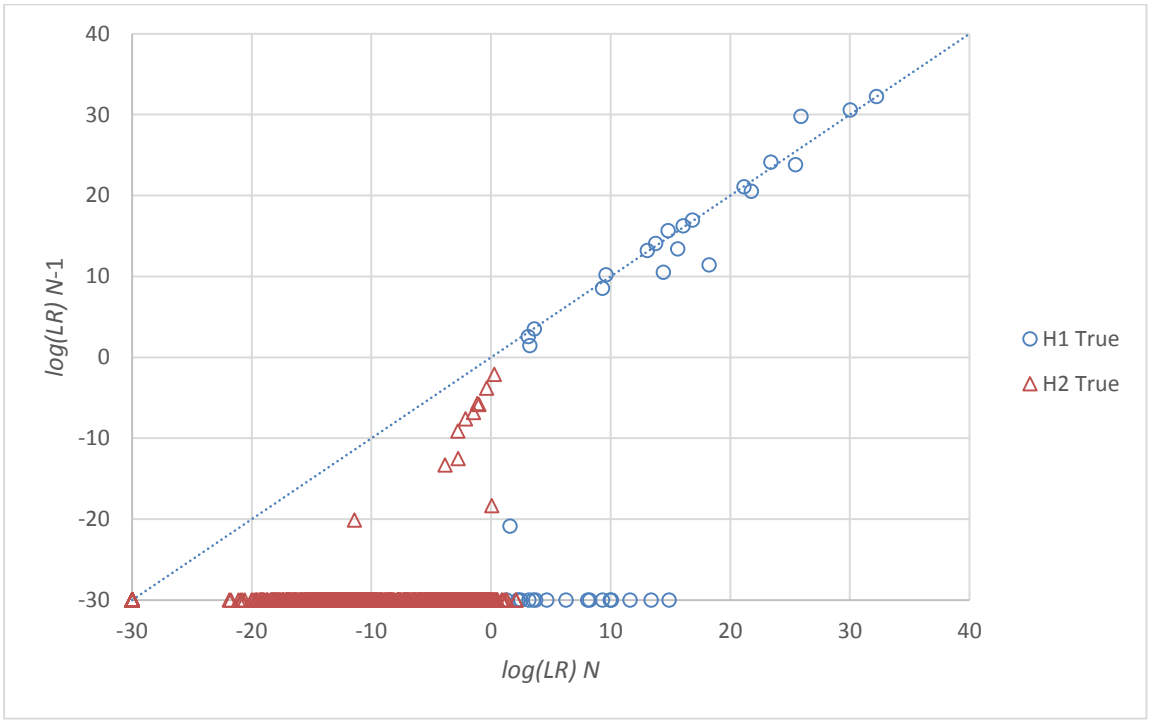


Table 7: Log(*LR*) values (H_1 true) for 2, 3, and 4 person mixtures assuming *N* and *N*-1 contributors

Sample	*N*	Cont.	*N* log(*LR*)	*N*-1 log(*LR*)
44-100_3.1_.05	2	K44	6.27	-30.00
		K100	9.96	-30.00
44-100_3.1_.025	2	K44	3.23	1.47
		K100	3.73	-30.00
53-74_1.1_.025	2	K74	9.31	8.54
		K53	1.58	-20.85
48-88_1-10	2	K48	1.27	-30.00
		K88	23.39	24.10
1-1 C9	2	K65	16.05	16.23
		K69	1.58	-30.00
57-74-100_3.2.1_.08	3	K57	3.11	2.58
		K74	3.60	3.50
		K100	9.61	10.18
86-53-100_1.1.1_.08	3	K53	2.27	-30.00
		K86	14.87	-30.00
		K100	10.06	-30.00
86-53-100_1.1.1_.04	3	K53	-0.05	-30.00
		K86	14.80	15.64
		K100	4.65	-30.00
3-2-1 C3	3	K41	21.12	21.07
		K65	3.17	-30.00
		K58	3.51	-30.00
4-2-1 32.5 PG	3	K41	25.90	29.78
		K65	8.08	-30.00
		K58	8.25	-30.00
10-5-1 C4	3	K41	30.02	30.58
		K65	2.49	-30.00
		K58	15.58	13.39
1-1-1-1 C7	4	K41	16.82	16.95
		K65	13.75	14.06
		K69	2.21	-30.00
		K58	13.06	13.19
4-3-2-1 C1	4	K41	25.44	23.82
		K65	13.38	-30.00
		K69	11.61	-30.00
		K58	14.40	10.50
10-5-2-1 C1	4	K41	32.23	32.21
		K65	18.23	11.42
		K69	9.31	-30.00
		K58	21.76	20.52

Gray shaded: LR went from included to excluded with N-1

Blue shaded: LR went from inconclusive to excluded with N-1

Section G: Drop-in

This section covers the following SWGDAM guideline:

4.1.8. Allele drop-in

Observed drop-in rates at the SacDA Laboratory have been modelled and the appropriate parameters are within STRmix™. To test these settings seven experiments were undertaken. In the first five experiments (K88 – K92), a realistically sized (≤150 rfu) drop-in peak was artificially added to *high template* single source STRmix™ input files that had been previously interpreted within STRmix™. The profiles were interpreted as single source profiles. As expected STRmix™ completely modelled the additional peak as drop-in because it could not pair with the high template alleles (>1000 rfu). The resulting *LRs* were identical to the original profile *LRs*.

In the sixth experiment (K59), a realistically sized (≤150 rfu) drop-in peak was artificially added to a *low template* single source STRmix™ input file that had been previously interpreted within STRmix™. The profile was interpreted as a single source profile. 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 (≤150 rfu). The resulting *LR* was less than the original profile *LR*.

In the last experiment (K96), a 'drop-in' allele was added to a heterozygote locus outside SacDA Laboratory's parameters (>150 rfu) in a single source profile. As expected, the interpretation halted with an error because the profile could no longer be explained by one contributor. This error can be used as a diagnostic tool in evaluating the profile for possible artifacts/contamination/tri-alleles.

Section H: Forward and reverse stutter

This section covers the following SWGDAM guideline:

4.1.9. Forward and reverse stutter

STRmix™ implements a 'per allele' back stutter model. This is alternatively based on the longest uninterrupted stretch (LUS) of common repeats in the allele or the allele designation itself. STRmix™ V2.4 also models forward stutter using a per locus model. Stutter peak labels are retained at analysis and within the STRmix™ input file. The modelling of stutter peaks can be seen in the interpretation of single source profiles where stutter peaks are retained at interpretation. As part of the MCMC process they are considered as alleles in the genotype but those combinations are not accepted and therefore receive no weight. In mixed DNA profiles, when the minor contributor is of a similar height as the stutter peaks they start to be considered as minor alleles.

Seven experiments were conducted for the modelling of forward stutter. In the first five experiments (K88 – K92), a realistically sized forward stutter peak (≤10% of the parent peak) was artificially added to *high template* single source STRmix™ input files that had been previously interpreted within STRmix™. The profiles were interpreted as single source profiles. As expected STRmix™ completely modelled the additional peak as forward stutter because it could not pair with the high template alleles (>1000 rfu). The resulting *LRs* were identical to the original profile *LRs*.

In the sixth experiment (K96), a forward stutter peak (>10% of the parent peak) was artificially added to a *high template* single source STRmix™ input file that had previously been interpreted within STRmix™. The profile was interpreted as a single source profile. As expected STRmix™ modelled the additional peak as an allele since it did not fit within the forward stutter parameters. This additional allele created a genotype

that was not the same as the true contributor's genotype, and consequently, the true contributor gave an *LR* of 0.

In the last experiment (K97), a forward stutter peak (>10% of the parent peak) was added to a heterozygote locus in a single source profile. As expected, the interpretation halted with an error because the profile could no longer be explained by one contributor. This error can be used as a diagnostic tool in evaluating the profile for possible artifacts/contamination/tri-alleles.

Section I: Intra locus peak height

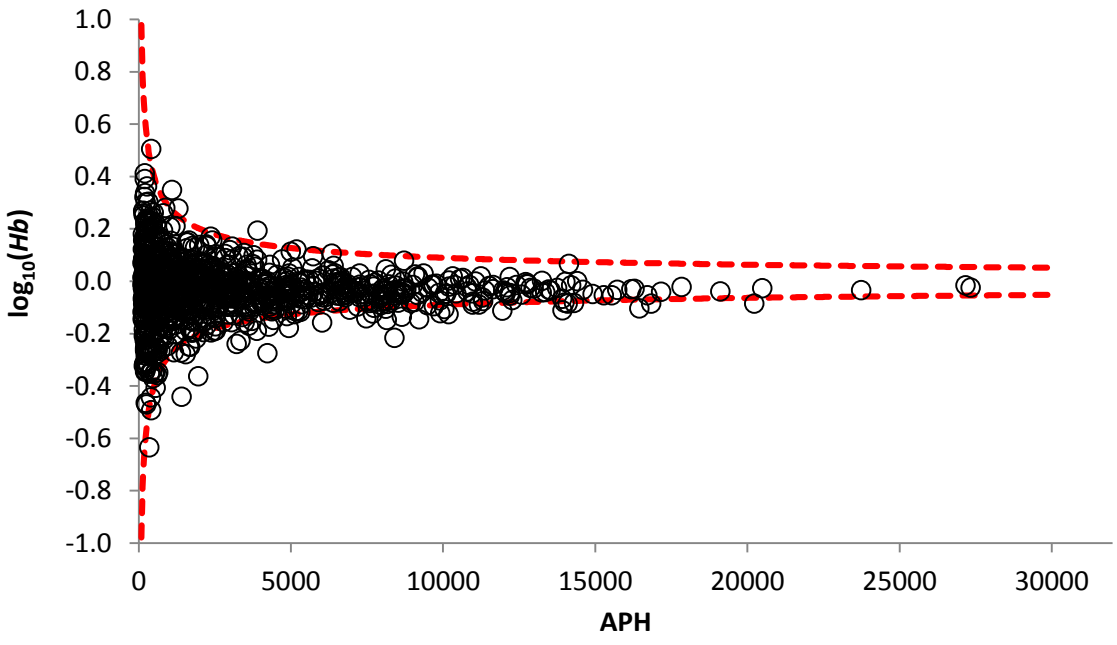
This section covers the following SWGDAM guideline:

4.1.10. Intra-locus peak height variance

STRmix™ models the variability of single peaks. The variance of this model is determined by directly modelling laboratory data. This is undertaken within STRmix™ using the Model Maker 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. Using an Excel worksheet provided by the STRmix™ validation team, the performance of Model Maker is checked by plotting the bounds informed by the Model Maker results (refer to the SacDA Laboratory STRmix™ Parameter 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 = 10.48, the 75th percentile from the gamma distribution from the combination data set. The 95% bounds encapsulate sufficient data as demonstrated in the graphs (coverage = 94.6%) demonstrating that the values for variance are sufficiently optimised. The plot of log(*Hb*) versus APH is given in Figure 9.

Figure 9: Plot of log(*Hb*) versus APH For SacDA Laboratory Fusion 6C data



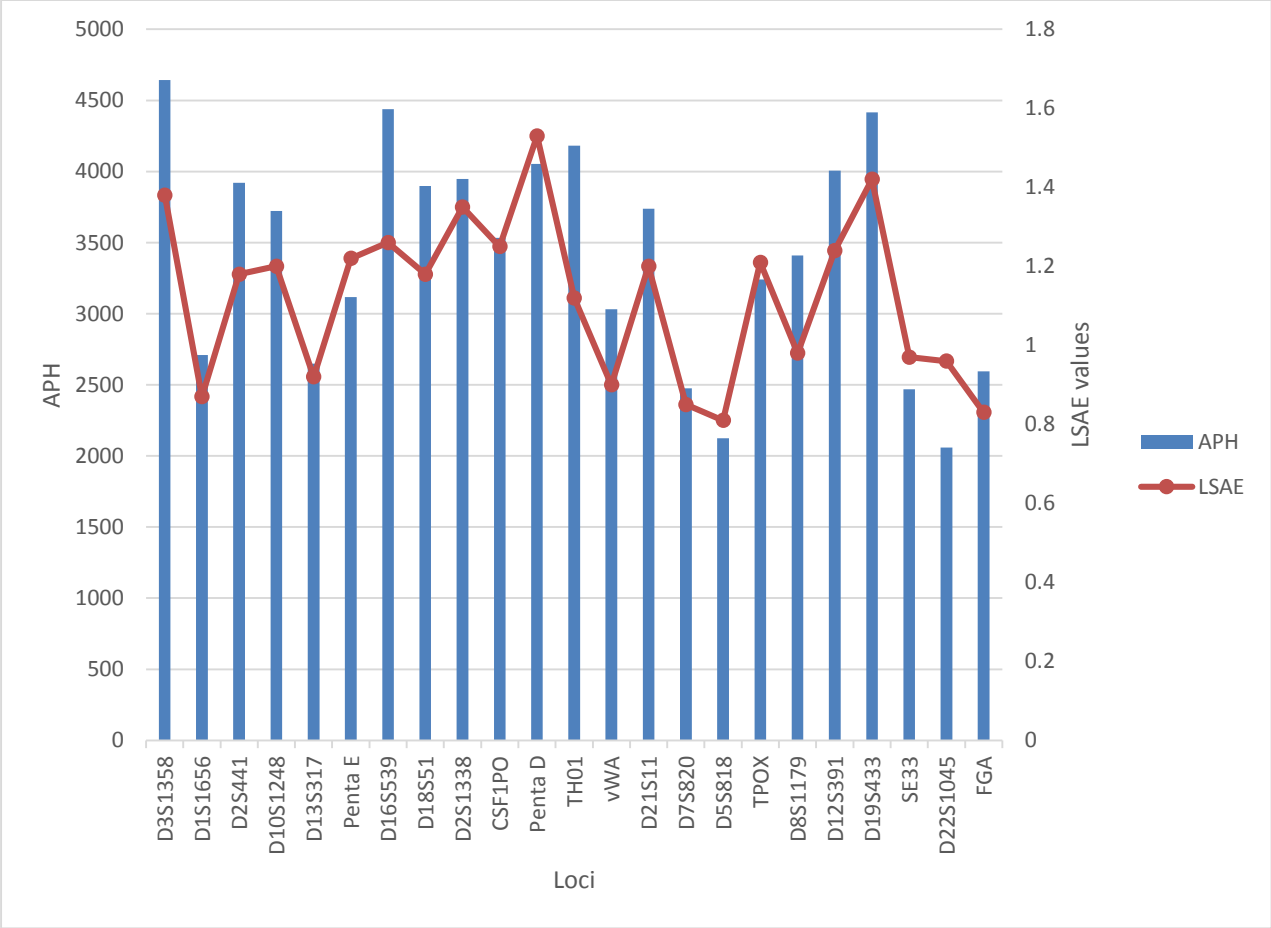
Section J: Inter-Locus peak heights

This section covers the following SWGDAM guideline:

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, the peak heights between loci are still more variable than predicted. The variance of this model is determined by directly modelling laboratory data. LSAE values for each STRmix™ interpretation appear within the results. We can demonstrate the relationship of LSAE values to average peak heights (APH) via a simple plot. The LSAE values should mimic the average peaks heights of the locus. This is demonstrated for one single source profile (K48) in Figure 10.

Figure 10: Plot of APH and LSAE value for each locus for a single source profile



Section K: Challenge testing

This section covers the following SWGDAM guideline:

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

STRmix™ requires that only numeric values are retained within the input file. Any values that are not numeric (such as OL alleles not removed at analysis) will prohibit STRmix™ from completing the

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

- A decreased *LR*. Analysis proceeds normally until it reaches the locus containing the artifact (OL). No downstream loci can be processed.
- An exclusionary *LR*. If the artifact is modelled as having originated from the POI (for example the artifact is paired with a homozygote allele to create a heterozygous genotype that then excludes the POI), this may result in an exclusion.
- No effect. If drop-in is observed within a laboratory, the artifact may be modelled as a drop-in peak if it is less than the drop-in height threshold.
- Failure to interpret. If an artifact within an allelic bin is retained in a profile it may artificially increase the minimum number of contributors within the profile. For example an artifact at a heterozygous locus in a single source profile (not modelled as stutter or drop-in) will increase the minimum number of contributors by one. STRmix™ will not proceed assuming only one contributor, and the interpretation will result in an error.

Each of these expected outcomes was demonstrated by editing five previously examined single source input files (K88 – K92) and calculating an LR within STRmix™. In addition, it was observed that these artifacts caused suboptimal run diagnostics e.g. large variance values for the profile, low or negative average log(likelihood) values and low acceptance rates.

Section L: Casework profiles

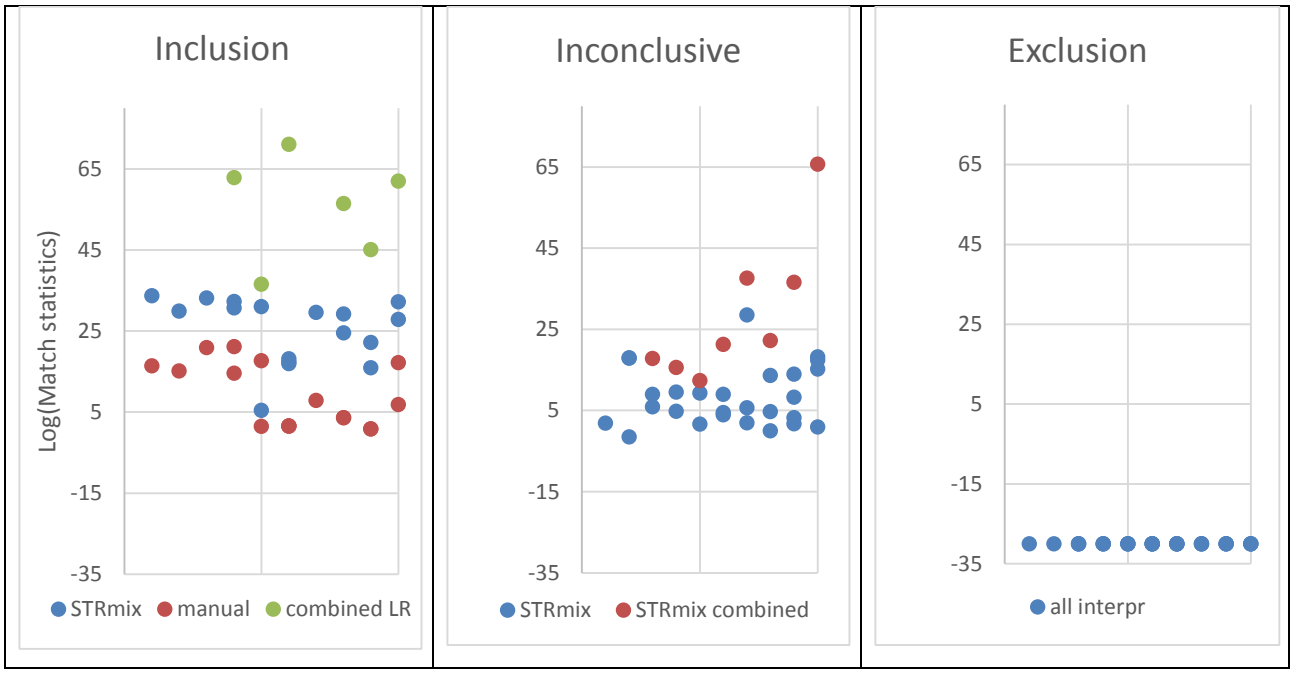
This section covers the following SWGDAM guidelines:

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.

Thirty profiles covering a range of numbers of contributors (1, 2, 3, and 4 person) and profile types (some with an assumed contributor, and some without) were interpreted in STRmix™ where the person of interest (POI) was previously considered to be excluded, inconclusive, or included using our traditional RMP statistical methods. Traditional methods were mimicked by amplifying the interpretable profiles with Identifiler™ Plus, and performing a simplified interpretation. A plot of the log(*match statistics*) versus a qualitative assessment is provided in Figure 11.

Figure 11: Plot of log(*match statistics*) as calculated in STRmix™ versus traditional assessment of the profile



For the inclusion section, ten profiles were examined, and 17 *LR*s were calculated for the POIs. All seventeen POIs would have been included using traditional methods and are still included using STRmix™. The STRmix™ *LR*s calculated for a sample tended to be larger than the random match calculations. This is in part due to the additional loci included in the Fusion 6C kit versus the Identifiler™ Plus kit.

For the inconclusive section, ten mixtures were examined, and 26 *LR*s were calculated for the POIs. Using the manual interpretation method, all ten profiles would have been inconclusive and therefore not compared to reference profiles. Of the 26 *LR*s calculated after STRmix™ deconvolution, six POIs gave *LR*s in the inconclusive range (.01 – 100), and one POI was excluded with an *LR* of 0.029. Upon further examination of this profile, the POI had dropout at all but one locus.

Ten profiles were examined for the exclusion section. Using both traditional interpretation methods and STRmix™ interpretation, all POIs were excluded.

As an extension of this experiment, an additional *LR* was calculated for all mixtures in the "inclusion" and "inconclusive" sections under the following proposition:

H_1: The DNA originated from POI 1 + POI 2 + POI *N*…

H_2: The DNA originated from *N* unknown individuals

If both contributors *together* were a good fit for the profile, then the resulting *LR* should be approximately additive (and often additive plus more). If both contributors together are not a good fit for the profile, then the resulting *LR* will not be approximately additive and will sometimes give an *LR* of 0 if the two contributors cannot be present at the same time in a profile. A summary of the separate and combined *LR*s for each mixture is given in Table 8.

Table 8: Summary of the separate and combined *LR*s

Sample	N	*LR*: POI 1	*LR*: POI 2	*LR*: POI 3	*LR*: POI 4	Combined *LR*
1-1 C9	2	7.89E+17	2.88E-02			9.03E+17
4-1	2	5.28E+30	1.93E+32			1.02E+63
5-1 C2	2	2.80E+05	1.02E+31			2.86E+36
44-100 .05	2	7.52E+05	9.53E+08			7.07E+17
46-86 .08	2	3.26E+09	6.39E+04			3.67E+15
53-74 .025	2	4.64E+01	1.87E+09			2.29E+12
6-2	3	9.64E+16	2.47E+17	1.39E+18		3.31E+52
9-3	3	3.49E+24	1.53E+29	*assumed*		5.34E+53
4:2:1 15.625pg	3	3.20E+28	4.82E+05	8.58E+01		3.75E+37
57-74-100 .08	3	8.10E+03	3.16E+04	9.86E+08		1.80E+21
86-53-100 .04	3	9.02E-01	4.19E+13	5.22E+04		1.69E+22
14-1	4	8.17E+15	1.62E+22	*assumed*	*assumed*	1.32E+38
15-1	4	7.54E+27	1.77E+32	*assumed*	*assumed*	1.33E+60
44-53-74-86 .15	4	5.59E+01	1.69E+03	1.91E+08	8.51E+13	3.52E+36
1:1:1:1 C7	4	1.57E+18	2.65E+17	1.64E+15	8.73E+00	5.30E+65

Section M: Precision

This section covers the following SWGDAM guideline:

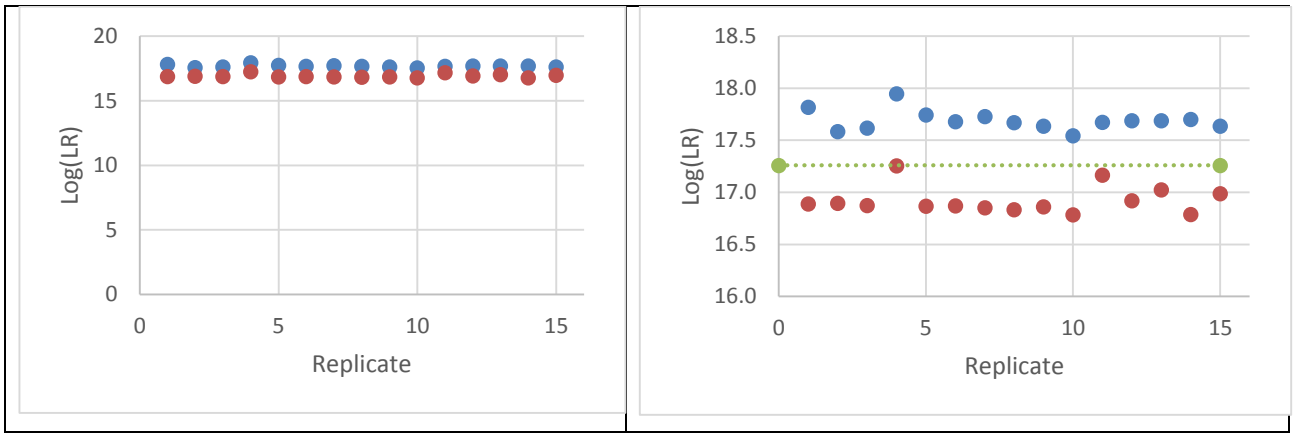
> 4.1.13. Sensitivity, specificity and *precision*, as described for Developmental Validation

Refer to Section D for details of sensitivity and specificity tests.

The MCMC process is used to generate the weights within STRmix™ for different genotype combinations. This is a sampling procedure and therefore the weights will vary slightly between each run. 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 lab variables including the PCR and CE process. The variability due to the size of the allele frequency database and the MCMC process is taken into account within STRmix™ V2.4 using the highest posterior density (HPD) method [10-12] (which defines the interval most likely to contain the true value (e.g. allele frequencies or *LR*s).

The extent of STRmix™ run variability was investigated by SacDA Laboratory by interpreting one of the mixed DNA profiles from Section D (sample 6-3) where there was ambiguity in the genotype combinations, fifteen times. A plot of log(*LR*) for each replicate is given in Figure 12. The blue dots indicate the point estimate *LR* values and the red dots are the lower 99% bound of the HPD. The green line shows the value of the highest lower 99% bound of the HPD (Log(*LR*) = 17.26).

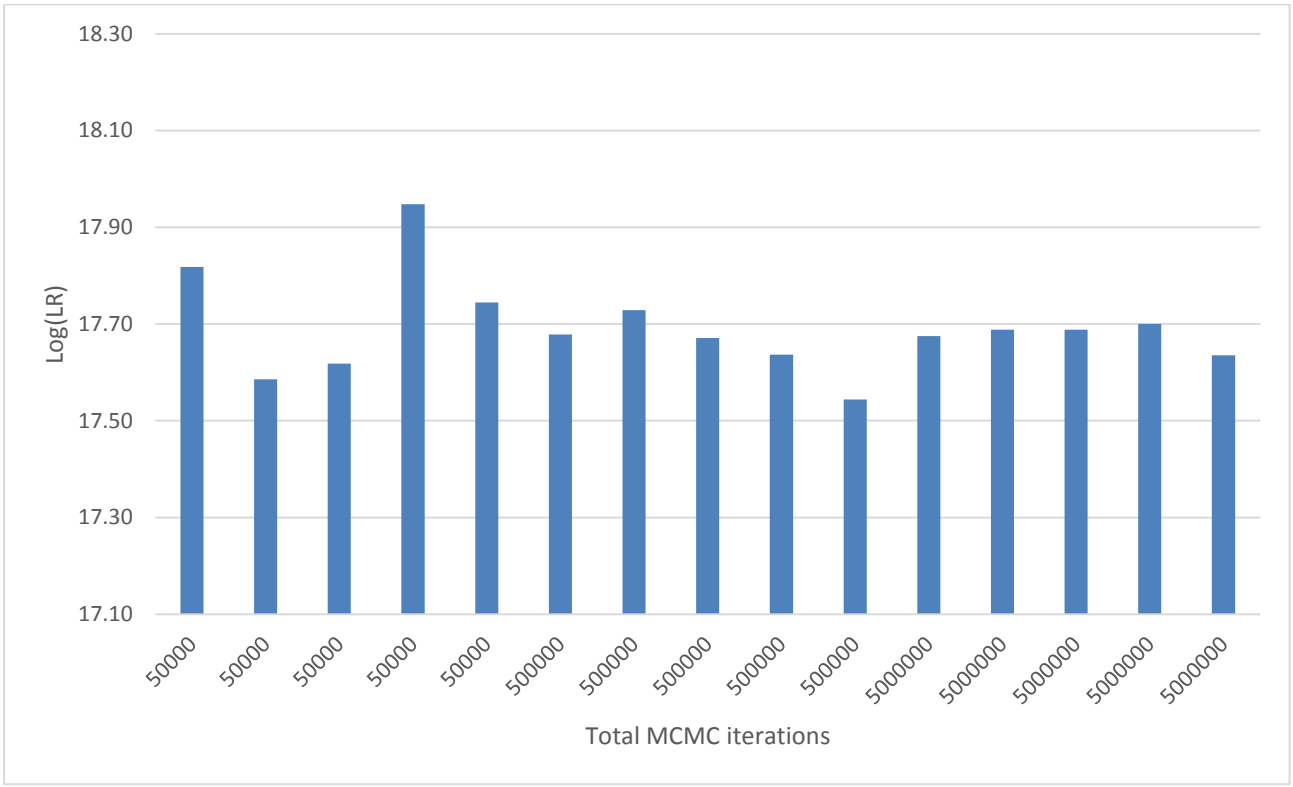
Figure 12: Plot of replicate log(*LR*) demonstrating reproducibility of STRmix™ (pane 1) and zoom of Y axis (pane 2)



Inspection of Figure 12 shows that the *LR*s are very reproducible and that the lower 99% bound of the HPD is always below the point estimate *LR* values.

Parameters within STRmix™ that affect run variability include the number of iterations and the RWSD (random walk standard deviation). The default number of iterations is set to 100,000 burn-in and 400,000 post burn-in. These will be suitable for many different types of profiles. Decreasing the number of iterations may mean that STRmix™ has not converged and more variability is expected. Increasing the number of iterations may mean convergence is achieved (if it hasn't already) and will certainly mean higher run times. One three person mixture was interpreted using three different sets of iterations (total 50,000, 500,000 and 5,000,000) five times each. A plot of log(*LR*) for each replicate is given in Figure 13.

Figure 13: Log(*LR*) of complex three person mixture interpreted five times in STRmix™ using different numbers of accepts



Results from Figure 13 show that the overall Log(*LR)* was highly reproducible across 15 replicates and that Log(*LR)* variability decreased as the number of iterations increased.

Section N: Degraded/inhibited samples

This section covers the following SWGDAM guidelines:

4.1.7. Partial profiles, to include the following:

4.1.7.2. DNA degradation

4.1.7.3. Inhibition

Five previously analyzed casework samples with degradation or differential degradation (one single source and four 2 person mixtures) were amplified with Fusion 6C and deconvoluted with STRmix™. Five victim references and two suspect references were also available and amplified with Fusion 6C. Three of the five victim reference sample profiles showed signs of degradation. The degradation quantities calculated by STRmix™ ranged from 0.3 to 23.4 rfu/bp. All of the casework deconvolutions gave inclusionary *LR*s for both the victims and the suspects. The *LR* quantities, mixture ratios, and genotype weightings were consistent with qualitative expectations.

Thirty single-source samples were subjected to an inhibiting substance. Because Fusion 6C is good at overcoming inhibition, only three of these samples showed signs of inhibition on the electropherograms. These three samples were deconvoluted with STRmix™ and gave genotype weightings that were consistent with qualitative expectations. These experiments demonstrate that STRmix™ can reliably be used to analyze profiles exhibiting degradation or inhibition.

Section O: Informed priors

The default settings in STRmix™ use uninformed priors for DNA amount and hence mixture proportions. STRmix™ has the ability for the user to provide informed priors by approximating the mixture proportions and variance for each contributor before deconvoluting the mixture. There are some instances when a user may wish to provide informed priors. One instance would be when information is present in an electropherogram that is below the detection threshold, suggesting a low level contributor. In these circumstances, as the information has not been detected, STRmix™ will not be able to make use of that information. The user can overcome this by choosing an informed mixture proportion for one contributor at trace levels. Another instance where informed priors may be used is if the user has scenario-specific information that multiple relatives have contributed DNA to a sample. In this situation profiles will have significant allele sharing, and there may be a number of mixture proportions that fit the profile. In this instance the user may wish to use the case context to provide informed priors.

Upon reviewing the deconvolutions of 132 two, three, and four person mixtures (see Section R), none of the mixtures would benefit from informed priors. For validation purposes, one mixture was chosen to experiment with informed priors. Sample 20-1 C3 is a two person major/minor mixture. The original STRmix™ deconvolution gave mixture proportions of 98%-2%, which were appropriate for the mixture. This mixture was deconvoluted two additional times using informed priors, once at 90%-10% and once at 50%-50% to calculate an LR for the minor contributor. Table 9 shows the *LR* for the compared individual gets lower as the mixture proportion moves away from the true proportion. This is expected.

Table 9: Informed priors experiment

Sample name	Conditioned on	Compared to	Informed priors	STRmix mix ratio	*LR*
20-1 C3	-	K69	none	98:2	21.5
20-1 C3	-	K69	90:10	97:3	4.35
20-1 C3	-	K69	50:50	71:29	1.67E-21

Section P: Assuming contributors

This section investigates the scenario when a person is assumed to be present in a mixture, but some of their alleles are dropping out of the evidence profile. In the genotype probability distribution table, these "dropped out" alleles from the assumed contributor will be used to create genotypes for the remaining contributors. During the MCMC iterations, STRmix™ uses alleles from an assumed contributor when making the genotype combinations, whether or not these alleles have dropped out. By telling STRmix™ that a contributor is present, STRmix™ assumes their alleles are present, and when it does not observe the alleles, it assumes that they have dropped out. The only alleles that are covered by "Q" are those that have not been seen yet.

Four 2 person mixtures and one 3 person mixture, with varying levels of dropout observed for the assumed contributor, were deconvoluted in STRmix™. The genotype probability distributions were examined and genotypes containing dropped alleles from the assumed contributor were noted. While this does give slightly higher weights to genotypes containing alleles that were not observed in the evidence profile, it has very little effect on the overall *LR* results because these weights are either very small or the remaining profile has very little information.

Section Q: Relatives

Relatives are expected to share more alleles with each other than with unrelated people. STRmix™ automatically outputs the *LR*s for alternative hypotheses involving relatives. An example of a hypotheses used if the brother of the person of interest (POI) is in question is:

H_1: The DNA originated from the POI

H_2: The DNA originated from a brother of the POI

To better understand the effect that relatives have on comparisons, two parent/child relationships were examined and one sibling relationship was examined. Table 10 shows nineteen 2 and 3 person mixtures. In each mixture, an *LR* was calculated for POI 1, who is a true contributor. An *LR* was also calculated for POI 2, who is not a contributor, but a parent/child of POI 1. The *LR*s obtained for POI 2 are higher than would be expected from an unrelated population. One of the 3 person mixtures gave an inclusionary *LR* when both POI 1 and POI 2 were in the numerator of the *LR*. However, the combined *LR* is not additive, and this may be used as a diagnostic that these two contributors do not explain the mixture very well.

Table 10: Comparison of *LR*s between true contributors and parent/child non-contributors

Sample	**N**	***LR*: POI 1** ***(true contributor)***	***LR*: POI 2** ***(non-contributor)***	**Combined *LR***
3-3	2	9.65E+21	0	0
5-3	2	1.24E+34	0	0
1-3	2	1.07E+18	0	0
3-2	2	4.74E+31	0	0
7-2	3	1.39E+12	2.60E+00	0
7-3	3	8.51E+09	4.69E+02	0
10-3	3	7.85E+10	4.23E+03	0
1-1-1 C9	3	5.03E+14	6.88E+02	0
1-1-1 C10	3	8.06E+08	9.30E+02	4.49E+08
3-2-1 C1	3	2.76E+16	6.07E+04	0
3-2-1 C2	3	1.28E+11	5.99E+00	0
3-2-1 C3	3	3.59E+04	3.66E+02	0
4-2-1 15.625PG	3	2.05E+02	5.29E+01	0
4-2-1 32.5PG	3	7.05E+08	5.93E+03	0
4-2-1 62.5PG	3	4.14E+09	5.90E+00	0
4-2-1 125PG	3	1.35E+20	1.63E-09	0
4-2-1 250PG	3	9.35E+19	0	0
10-5-1 C3	3	7.54E+05	2.90E+02	0
10-5-1 C4	3	5.59E+02	3.39E+01	0

Table 11 shows twenty-nine 2, 3, and 4 person mixtures. In each mixture, an *LR* was calculated for POI 1, who is a true contributor. An *LR* was also calculated for POI 2, who is not a contributor, but a sibling of POI 1. The *LR*s obtained for POI 2 are higher than would be expected from an unrelated population. Six of the mixtures gave an inclusionary *LR* when both POI 1 and POI 2 were in the numerator of the *LR*. However, the combined *LR* is not additive, and this may be used as a diagnostic that these two contributors do not explain the mixture very well.

Table 11: Comparison of *LR*s between true contributors and sibling non-contributors

Sample	**N**	***LR*: POI 1** ***(true contributor)***	***LR*: POI 2** ***(non-contributor)***	**Combined *LR***
78-41 .5	2	3.37E+12	0	0
78-41 .25	2	8.10E+15	0	0
78-41 .12	2	4.47E+00	0	0
11-1	3	1.19E+15	0	0
11-2	3	1.52E+15	0	0
11-3	3	8.58E+17	0	0
78-53-41 .8	3	5.27E+14	0	0
78-53-41 .4	3	6.36E+15	0	0
78-53-41 .2	3	1.17E+16	0	0
78-53-41 .1	3	4.67E+10	2.30E+04	0
1-1-1 C6	3	1.40E+28	0	0
1-1-1 C7	3	9.85E+20	0	0
1-1-1 C8	3	4.29E+19	2.34E+09	0
1-1-1 C9	3	9.64E+19	5.00E+13	7.26E+19
1-1-1 C10	3	6.44E+08	4.11E+06	1.03E+11
3-2-1 C1	3	3.94E+35	2.91E+01	0
3-2-1 C2	3	1.16E+29	8.03E+10	0
3-2-1 C3	3	1.35E+25	2.52E+09	4.71E+24
4-2-1 250PG	3	4.10E+35	0	0
4-2-1 125PG	3	3.43E+35	0	0
4-2-1 62.5PG	3	3.73E+35	3.34E-01	0
4-2-1 32.5PG	3	1.72E+29	5.95E+13	0
4-2-1 15.625PG	3	2.70E+30	6.42E+11	0
10-5-1 C1	3	4.09E+35	0	0
10-5-1 C2	3	3.97E+00	0	0
10-5-1 C3	3	2.03E+35	1.47E+03	4.72E+34
10-5-1 C4	3	1.54E+33	4.95E+09	2.27E+32
41-74-86-91 .1	4	6.87E+14	1.01E+05	0
41-74-86-91 .05	4	4.40E+14	3.18E+07	5.50E+15

Section R: Extreme dropout

As the amount of DNA increases, the probability of dropout decreases. STRmix™ uses this feature during the deconvolution. Occasionally dropout occurs, and the remaining peak is higher than expected. Out of the 96 single-source Model Maker samples, five instances of extreme dropout were observed. Each of these profiles was run with STRmix™. The true contributor gave inclusionary *LR*s for each profile, even though the individual *LR* at the dropout locus was small. See Table 12.

Table 12: Summary of results for profiles with extreme dropout

Sample Name	Remaining allele (rfu)	Visible peak under threshold?	Weight for dropout genotype	POI included?
15.625pg DNA1	426	no	0.047	yes
31.25pg DNA3	347	yes	0.026	yes
62.5pg DNA2	513	yes	0.002	yes
31.25pg DNA8	329	yes	0.059	yes
250pg DNA10	597	yes	0.000*	yes

* rounded to 0

Section S: Mixture summary

STRmix™ results for 132 two, three, and four person mixtures were carefully scrutinized. Gelman-Rubin values and the ability of STRmix™ to capture true contributor genotypes in the presence of allele dropout were evaluated.

2 person mixtures (Tables 13-15):
Of the 53 two person mixtures, 45 of them had alleles from at least one contributor dropping out. Twenty-three mixtures were evaluated for correct genotype calls at 99%, and 22 mixtures were evaluated at 99.5%. All of the Gelman-Rubin Convergence numbers were appropriate, with the highest at 1.31. There was only one mixture (1-3) in which one of the contributor's genotypes at one locus was not in the top 99%. This was a mid-level balanced mixture where each contributor was contributing ~50%. Upon further inspection of the genotypes, both contributors were fully represented, however at D2S1338, one of the contributors had an allele dropping out and partial dropout was considered, but only with a weighting of 0.048%, which did not make the top 99% cutoff that was investigated in this study. All other mixtures (including low level, balanced and imbalanced mixtures) were deconvoluted by STRmix™ in a way that was intuitive and genotypes from the known contributors fell in the top 99% of weights.

Table 13: 2 person mixture results

Sample Name	N	Dropout ?	Gelman-Rubin	Reference IDs	Manual % contribution	STRmix % contribution	Correct genotype in 99%	STRmix genotype possibilties intuitive?
1-1	2	N	1.04	K55	56	53	yes	yes
				K66	44	47	yes	yes
1-2	2	N	1.02	K55	52	53	yes	yes
				K66	48	47	yes	yes
1-3	2	Y	1.00	K55	52	54	yes	yes
				K66	48	46	**no**	yes
2-1	2	N	1.01	K53	68	69	yes	yes
				K55	32	31	yes	yes
2-2	2	N	1.04	K53	65	63	yes	yes
				K55	35	37	yes	yes
2-3	2	N	1.02	K53	63	56	yes	yes
				K55	37	44	yes	yes
3-1	2	N	1.01	K66	85	87	yes	yes
				K74	15	13	yes	yes
3-2	2	Y	1.19	K66	73	80	yes	yes
				K74	27	20	yes	yes
3-3	2	Y	1.01	K66	74	78	yes	yes
				K74	26	22	yes	yes
4-1	2	N	1.01	K44	86	86	yes	yes
				K53	14	14	yes	yes
4-2	2	Y	1.04	K44	83	87	yes	yes
				K53	17	13	yes	yes
4-3	2	Y	1.04	K44	81	87	yes	yes
				K53	19	13	yes	yes
5-1	2	Y	1.02	K74	94	93	yes	yes
				K87	6	7	yes	yes
5-2	2	Y	1.02	K74	87	93	yes	yes
				K87	13	7	yes	yes
5-3	2	Y	1.06	K74	84	93	yes	yes
				K87	16	7	yes	yes

Table 14: 2 person mixture results

Sample Name	N	Dropout ?	Gelman-Rubin	Reference IDs	Manual % contribution	STRmix % contribution	Correct genotype in 99%	STRmix genotype possibilties intuitive?
78-41 .5	2	Y	1.05	K78	73	60	yes	yes
				K41	27	40	yes	yes
78-41 .25	2	Y	1.03	K78	67	63	yes	yes
				K41	33	37	yes	yes
78-41 .12	2	Y	1.03	K78	62	63	yes	yes
				K41	38	38	yes	yes
46-86 .3	2	Y	1.00	K46	51	55	yes	yes
				K86	49	45	yes	yes
46-86 .15	2	Y	1.01	K46	51	55	yes	yes
				K86	49	45	yes	yes
46-86 .08	2	Y	1.05	K46	57	65	yes	yes
				K86	43	35	yes	yes
53-91 .18	2	Y	1.01	K53	53	56	yes	yes
				K91	47	44	yes	yes
53-91 .09	2	Y	1.01	K53	61	59	yes	yes
				K91	39	41	yes	yes
53-91 .05	2	Y	1.02	K53	53	56	yes	yes
				K91	47	44	yes	yes
44-100 .1	2	Y	1.02	K44	51	58	yes	yes
				K100	49	42	yes	yes
44-100 .05	2	Y	1.03	K44	54	62	yes	yes
				K100	46	38	yes	yes
44-100 .025	2	Y	1.15	K44	71	68	yes	yes
				K100	29	32	yes	yes
53-74 .2	2	Y	1.05	K53	54	55	yes	yes
				K74	46	45	yes	yes
53-74 .1	2	Y	1.02	K53	55	56	yes	yes
				K74	45	44	yes	yes
53-74 .05	2	Y	1.01	K53	50	59	yes	yes
				K74	50	41	yes	yes
53-74 .025	2	Y	1.02	K53	53	67	yes	yes
				K74	47	33	yes	yes

Table 15: 2 person mixture results

Sample Name	N	Dropout ?	Gelman-Rubin	Reference IDs	Manual % contribution	STRmix % contribution	Correct genotype in 99.5%	STRmix genotype possibilties intuitive?
1-1 C6	2	N	1.07	K65	-	88	yes	yes
				K69	-	12	yes	yes
1-1 C7	2	Y	1.05	K65	-	86	yes	yes
				K69	-	14	yes	yes
1-1 C8	2	Y	1.31	K65	-	74	yes	yes
				K69	-	26	yes	yes
1-1 C9	2	Y	1.06	K65	-	60	yes	yes
				K69	-	40	yes	yes
1-1 C10	2	Y	1.17	K65	-	68	yes	yes
				K69	-	32	yes	yes
3-1 C1	2	Y	1.03	K65	-	94	yes	yes
				K69	-	6	yes	yes
3-1 C2	2	Y	1.09	K65	-	87	yes	yes
				K69	-	13	yes	yes
3-1 C3	2	Y	1.05	K65	-	99	yes	yes
				K69	-	1	yes	yes
3-1 C4	2	Y	1.04	K65	-	67	yes	yes
				K69	-	33	yes	yes
3-1 C5	2	Y	1.19	K65	-	70	yes	yes
				K69	-	30	yes	yes
5-1 C1	2	Y	1.04	K65	-	95	yes	yes
				K69	-	5	yes	yes
5-1 C2	2	Y	1.02	K65	-	95	yes	yes
				K69	-	5	yes	yes
5-1 C3	2	Y	1.02	K65	-	100	yes	yes
				K69	-	0	yes	yes
5-1 C4	2	Y	1.09	K65	-	82	yes	yes
				K69	-	18	yes	yes
5-1 C5	2	Y	1.21	K65	-	65	yes	yes
				K69	-	35	yes	yes
10-1 C1	2	Y	1.04	K65	-	95	yes	yes
				K69	-	5	yes	yes
10-1 C2	2	Y	1.05	K65	-	100	yes	yes
				K69	-	0	yes	yes
10-1 C3	2	Y	1.04	K65	-	99	yes	yes
				K69	-	1	yes	yes
20-1 C1	2	Y	1.02	K65	-	99	yes	yes
				K69	-	1	yes	yes
20-1 C2	2	Y	1.07	K65	-	93	yes	yes
				K69	-	7	yes	yes
20-1 C3	2	Y	1.06	K65	-	98	yes	yes
				K69	-	2	yes	yes
20-1 C4	2	Y	1.10	K65	-	100	yes	yes
				K69	-	0	yes	yes

3 person mixtures (Tables 16-18):

Of the 44 three-person mixtures, 26 of them had alleles from at least one contributor dropping out.

All of the Gelman-Rubin Convergence numbers were appropriate, with the highest at 1.25. Twenty-seven mixtures were evaluated for correct genotype calls at 99%, and 26 mixtures were evaluated at 99.5%. There was only one mixture (10-5-1 C1) in which one of the contributor's genotypes at one locus was not in the top 99%. This was a major/minor/trace mixture, and upon further inspection of the genotypes, all contributors were fully represented, except at one locus, the trace contributor had one allele dropping out. At D22S1045, there was an allele in a forward stutter position which was probably elevated by stutter such that STRmix™ gave a weighting of 99.52% for the minor contributor to have this allele. The minor contributor, however, does not have this allele. The trace contributor is homozygous with this allele.

All other mixtures (including low level, balanced and imbalanced mixtures) were deconvoluted by STRmix™ in a way that was intuitive and genotypes from the known contributors fell in the top 99% of weights.

Table 16: 3 person mixture results

Sample Name	N	Dropout ?	Gelman-Rubin	Reference IDs	Manual % contribution	STRmix % contribution	Correct genotypes in 99%	STRmix genotype possibilties intuitive?
6-1	3	N	1.01	K46	34	36	yes	yes
				K50	34	33	yes	yes
				K55	32	31	yes	yes
6-2	3	N	1.01	K46	39	38	yes	yes
				K50	32	33	yes	yes
				K55	29	28	yes	yes
6-3	3	N	1.04	K46	36	40	yes	yes
				K50	32	32	yes	yes
				K55	32	28	yes	yes
7-1	3	N	1.01	K54	38	36	yes	yes
				K74	31	34	yes	yes
				K75	31	30	yes	yes
7-2	3	N	1.01	K54	39	39	yes	yes
				K74	32	33	yes	yes
				K75	29	29	yes	yes
7-3	3	Y	1.02	K54	36	41	yes	yes
				K74	35	32	yes	yes
				K75	29	27	yes	yes
8-1	3	N	1.01	K85	45	44	yes	yes
				K86	39	36	yes	yes
				K87	16	20	yes	yes
8-2	3	N	1.01	K85	47	44	yes	yes
				K86	39	37	yes	yes
				K87	14	19	yes	yes
8-3	3	Y	1.07	K85	40	42	yes	yes
				K86	40	34	yes	yes
				K87	21	24	yes	yes
9-1	3	N	1.02	K85	55	55	yes	yes
				K89	23	24	yes	yes
				K91	22	21	yes	yes
9-2	3	Y	1.02	K85	62	67	yes	yes
				K89	19	18	yes	yes
				K91	18	15	yes	yes
9-3	3	Y	1.17	K85	46	47	yes	yes
				K89	30	30	yes	yes
				K91	24	23	yes	yes
10-1	3	N	1.02	K44	70	70	yes	yes
				K75	16	17	yes	yes
				K86	14	13	yes	yes
10-2	3	Y	1.01	K44	66	66	yes	yes
				K75	19	21	yes	yes
				K86	15	13	yes	yes
10-3	3	Y	1.01	K44	62	70	yes	yes
				K75	19	18	yes	yes
				K86	19	12	yes	yes

Table 17: 3 person mixture results

Sample Name	N	Dropout ?	Gelman-Rubin	Reference IDs	Manual % contribution	STRmix % contribution	Correct genotypes in 99%	STRmix genotype possibilties intuitive?
57-74-100 .3	3	Y	1.01	K57	39	40	yes	yes
				K74	34	33	yes	yes
				K100	28	27	yes	yes
57-74-100 .15	3	Y	1.01	K57	44	42	yes	yes
				K74	28	32	yes	yes
				K100	28	26	yes	yes
57-74-100 .08	3	Y	1.04	K57	36	51	yes	yes
				K74	34	33	yes	yes
				K100	30	16	yes	yes
57-74-100 .04	3	Y	1.03	K57	46	50	yes	yes
				K74	38	33	yes	yes
				K100	16	17	yes	yes
78-53-41 .8	3	N	1.00	K78	40	40	yes	yes
				K53	30	32	yes	yes
				K41	29	28	yes	yes
78-53-41 .4	3	N	1.01	K78	37	41	yes	yes
				K53	33	32	yes	yes
				K41	31	27	yes	yes
78-53-41 .2	3	Y	1.02	K78	39	41	yes	yes
				K53	32	33	yes	yes
				K41	29	26	yes	yes
78-53-41 .1	3	Y	1.03	K78	41	44	yes	yes
				K53	32	32	yes	yes
				K41	27	24	yes	yes
86-53-100 .3	3	N	1.00	K86	42	45	yes	yes
				K53	39	38	yes	yes
				K100	19	17	yes	yes
86-53-100 .15	3	Y	1.06	K86	45	44	yes	yes
				K53	34	35	yes	yes
				K100	21	21	yes	yes
86-53-100 .08	3	Y	1.02	K86	47	43	yes	yes
				K53	32	34	yes	yes
				K100	21	23	yes	yes
86-53-100 .04	3	Y	1.01	K86	36	44	yes	yes
				K53	36	33	yes	yes
				K100	29	22	yes	yes

Table 18: 3 person mixture results

Sample Name	N	Dropout ?	Gelman-Rubin	Reference IDs	Manual % contribution	STRmix % contribution	Correct genotypes in 99.5%	STRmix genotype possibilties intuitive?
1-1-1 C6	3	N	1.25	K41		42	yes	yes
				K58		33	yes	yes
				K65		25	yes	yes
1-1-1 C7	3	N	1.16	K41		44	yes	yes
				K58		35	yes	yes
				K65		21	yes	yes
1-1-1 C8	3	Y	1.04	K41		40	yes	yes
				K58		33	yes	yes
				K65		27	yes	yes
1-1-1 C9	3	Y	1.05	K41		42	yes	yes
				K58		33	yes	yes
				K65		25	yes	yes
1-1-1 C10	3	Y	1.04	K41		44	yes	yes
				K58		33	yes	yes
				K65		22	yes	yes
3-2-1 C1	3	Y	1.02	K41		72	yes	yes
				K58		16	yes	yes
				K65		12	yes	yes
3-2-1 C2	3	Y	1.03	K41		45	yes	yes
				K58		33	yes	yes
				K65		22	yes	yes
3-2-1 C3	3	Y	1.05	K41		52	yes	yes
				K58		34	yes	yes
				K65		14	yes	yes
4-2-1 15.625PG	3	Y	1.05	K41		60	yes	yes
				K58		31	yes	yes
				K65		10	yes	yes
4-2-1 32.5PG	3	Y	1.10	K41		52	yes	yes
				K58		32	yes	yes
				K65		16	yes	yes
4-2-1 62.5PG	3	Y	1.23	K41		71	yes	yes
				K58		20	yes	yes
				K65		8	yes	yes
4-2-1 125PG	3	N	1.16	K41		69	yes	yes
				K58		25	yes	yes
				K65		6	yes	yes
4-2-1 250PG	3	N	1.12	K41		70	yes	yes
				K58		26	yes	yes
				K65		4	yes	yes
10-5-1 C1	3	Y	1.04	K41		68	yes	yes
				K58		24	**no**	yes
				K65		8	yes	yes
10-5-1 C2	3	Y	1.02	K41		71	yes	yes
				K58		25	yes	yes
				K65		5	yes	yes
10-5-1 C3	3	Y	1.04	K41		70	yes	yes
				K58		26	yes	yes
				K65		4	yes	yes
10-5-1 C4	3	Y	1.02	K41		68	yes	yes
				K58		27	yes	yes
				K65		4	yes	yes

4 person mixtures (Tables 19-21):

Of the 35 four-person mixtures, 27 of them had alleles from at least one contributor dropping out. Twenty-three mixtures were evaluated for correct genotype calls at 99%, and 12 mixtures were evaluated at 99.5%. Only two of these mixtures (13-2 and 1-1-1-1 C6) had a diagnostic value that warranted a closer look. The Gelman-Rubin Convergence number for sample 13-2 was 1.50. This mixture showed some dropout of three of the four contributors, but full assessment of this sample did not indicate any other problem. Each known contributor's genotype fell into the top 99% of weights in the Component Interpretation section. The Gelman-Rubin Convergence number for sample 1-1-1-1 C6 was 1.70. This mixture did not have any dropout, and a full assessment did not indicate any problems. Each known contributor's genotype fell into the top 99.5% of weights in the Component Interpretation section.

All of the mixtures (including low level, balanced and imbalanced mixtures) were deconvoluted by STRmix™ in a way that was intuitive and genotypes from the known contributors fell in the top 99% of weights.

Table 19: 4 person mixture results

Sample Name	N	Dropout ?	Gelman-Rubin	Reference IDs	Manual % contribution	STRmix % contribution	Correct genotypes in 99%	STRmix genotype possibilties intuitive?
11-1	4	N	1.02	K41	29	28	yes	yes
				K45	25	26	yes	yes
				K49	23	24	yes	yes
				K55	23	22	yes	yes
11-2	4	N	1.06	K41	29	32	yes	yes
				K45	26	27	yes	yes
				K49	26	23	yes	yes
				K55	19	18	yes	yes
11-3	4	N	1.05	K41	30	33	yes	yes
				K45	28	27	yes	yes
				K49	22	22	yes	yes
				K55	20	18	yes	yes
12-1	4	N	1.04	K50	43	45	yes	yes
				K66	37	34	yes	yes
				K69	12	12	yes	yes
				K85	8	9	yes	yes
12-2	4	Y	1.03	K50	47	48	yes	yes
				K66	31	42	yes	yes
				K69	12	7	yes	yes
				K85	9	3	yes	yes
12-3	4	Y	1.09	K50	46	46	yes	yes
				K66	40	37	yes	yes
				K69	7	14	yes	yes
				K85	7	4	yes	yes
13-1	4	N	1.02	K51	64	64	yes	yes
				K55	14	15	yes	yes
				K57	12	12	yes	yes
				K65	10	9	yes	yes
13-2	4	Y	1.50	K51	54	49	yes	yes
				K55	20	24	yes	yes
				K57	13	16	yes	yes

				K65	13	11	yes	yes
13-3	4	Y	1.03	K51	57	77	yes	yes
				K55	19	11	yes	yes
				K57	14	8	yes	yes
				K65	10	4	yes	yes
14-1	4	N	1.05	K65	37	32	yes	yes
				K74	29	27	yes	yes
				K78	22	23	yes	yes
				K85	12	18	yes	yes
14-2	4	Y	1.11	K65	36	38	yes	yes
				K74	33	30	yes	yes
				K78	16	20	yes	yes
				K85	15	12	yes	yes
14-3	4	Y	1.03	K65	34	35	yes	yes
				K74	28	27	yes	yes
				K78	21	22	yes	yes
				K85	17	16	yes	yes
15-1	4	N	1.20	K69	48	36	yes	yes
				K86	21	26	yes	yes
				K89	19	22	yes	yes
				K100	12	17	yes	yes
15-2	4	Y	1.04	K69	50	47	yes	yes
				K86	27	26	yes	yes
				K89	12	17	yes	yes
				K100	11	10	yes	yes
15-3	4	Y	1.13	K69	41	42	yes	yes
				K86	24	26	yes	yes
				K89	18	19	yes	yes
				K100	17	13	yes	yes

Table 20: 4 person mixture results

Sample Name	N	Drop out?	Gelman-Rubin	Reference IDs	Manual % contribution	STRmix % contribution	Correct genotypes in 99%	STRmix genotype possibilties intuitive?
41-74-86-91 .1	4	Y	1.04	K41	32	35	yes	yes
				K74	28	27	yes	yes
				K86	20	21	yes	yes
				K91	20	17	yes	yes
41-74-86-91 .05	4	Y	1.02	K41	39	40	yes	yes
				K74	26	30	yes	yes
				K86	25	20	yes	yes
				K91	10	10	yes	yes
44-53-74-86 .15	4	Y	1.09	K44	41	39	yes	yes
				K53	20	30	yes	yes
				K74	20	20	yes	yes
				K86	20	10	yes	yes
44-53-74-86 .08	4	Y	1.02	K44	35	38	yes	yes
				K53	24	28	yes	yes
				K74	21	23	yes	yes
				K86	20	11	yes	yes
44-53-74-86 .04	4	Y	1.04	K44	45	48	yes	yes
				K53	18	26	yes	yes
				K74	18	24	yes	yes
				K86	18	3	yes	yes
46-57-100-91 .22	4	Y	1.03	K46	31	35	yes	yes
				K57	27	27	yes	yes
				K100	22	22	yes	yes
				K91	20	16	yes	yes
46-57-100-91 .12	4	Y	1.02	K46	49	43	yes	yes
				K57	34	27	yes	yes
				K100	8	18	yes	yes
				K91	8	12	yes	yes
46-57-100-91 .06	4	Y	1.06	K46	27	42	yes	yes
				K57	26	29	yes	yes
				K100	24	15	yes	yes
				K91	22	15	yes	yes

Table 21: 4 person mixture results

Sample Name	N	Dropout ?	Gelman-Rubin	Reference IDs	Manual % contribution	STRmix % contribution	Correct genotypes in 99.5%	STRmix genotype possibilties intuitive?
1-1-1-1 C6	4	N	1.70	K41	-	38	yes	yes
				K58	-	32	yes	yes
				K65	-	24	yes	yes
				K69	-	6	yes	yes
1-1-1-1 C7	4	Y	1.06	K41	-	36	yes	yes
				K58	-	32	yes	yes
				K65	-	29	yes	yes
				K69	-	3	yes	yes
1-1-1-1 C8	4	Y	1.07	K41	-	45	yes	yes
				K58	-	30	yes	yes
				K65	-	20	yes	yes
				K69	-	5	yes	yes
1-1-1-1 C9	4	Y	1.11	K41	-	45	yes	yes
				K58	-	31	yes	yes
				K65	-	17	yes	yes
				K69	-	7	yes	yes
1-1-1-1 C10	4	Y	1.01	K41	-	38	yes	yes
				K58	-	25	yes	yes
				K65	-	25	yes	yes
				K69	-	13	yes	yes
4-3-2-1 C1	4	Y	1.27	K41	-	45	yes	yes
				K58	-	28	yes	yes
				K65	-	25	yes	yes
				K69	-	2	yes	yes
4-3-2-1 C2	4	Y	1.17	K41	-	43	yes	yes
				K58	-	32	yes	yes
				K65	-	24	yes	yes
				K69	-	2	yes	yes
4-3-2-1 C3	4	Y	1.23	K41	-	41	yes	yes
				K58	-	32	yes	yes
				K65	-	25	yes	yes
				K69	-	1	yes	yes
10-5-2-1 C1	4	Y	1.03	K41	-	64	yes	yes
				K58	-	20	yes	yes
				K65	-	14	yes	yes
				K69	-	2	yes	yes
10-5-2-1 C2	4	Y	1.18	K41	-	58	yes	yes
				K58	-	25	yes	yes
				K65	-	16	yes	yes
				K69	-	2	yes	yes
10-5-2-1 C3	4	Y	1.21	K41	-	61	yes	yes
				K58	-	21	yes	yes
				K65	-	14	yes	yes
				K69	-	3	yes	yes
10-5-2-1 C4	4	Y	1.06	K41	-	63	yes	yes
				K58	-	19	yes	yes
				K65	-	13	yes	yes
				K69	-	5	yes	yes

Conclusion

This document describes the SacDA laboratory's internal validation activities for STRmix™ V2.4. STRmix™ is suited for its intended use for the interpretation of profiles generated from one to four contributors using Promega's PowerPlex® Fusion 6C Amplification Kit and 3500xL Genetic Analyzers. This document follows the internal validation section of the SWGDAM Guidelines for the Validation of Probabilistic Genotyping Systems [3] and satisfies Standard 8.7 of the FBI Quality Assurance Standards for Forensic DNA Testing Laboratories (September 1, 2011) [4].

References

[1] NIST US1036 Allele Frequencies, Hill et al. Forensic Science International: Genetics 7 (2013) e82–e83.

[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] Scientific Working Group of DNA Analysis Methods (2015) Guidelines for Validation of Probabilistic Genotyping Sysptems.

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

[5] 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.

[6] 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.

[7] 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.

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

[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.

APPENDIX 1: List of papers that support STRmix™

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

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

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

1. J.-A. Bright, J.M. Curran. Investigation into stutter ratio variability between different laboratories. Forensic Science International: Genetics, 2014. 13: 79-81
2. C. Brookes, J.-A. Bright, S.A. Harbison, and J.S. Buckleton, Characterising stutter in forensic STR multiplexes. Forensic Science International: Genetics, 2012. 6(1): 58-63
3. H. Kelly, J.-A. Bright, J.M. Curran, and J.S. Buckleton Identifying and modelling the drivers of stutter in forensic DNA profiles. Australian Journal of Forensic Sciences, 2014. 46(2): 194-203
4. J.-A. Bright, S. Neville, J.M. Curran, and J.S. Buckleton. Variability of mixed DNA profiles separated on a 3130 and 3500 capillary electrophoresis instrument. Australian Journal of Forensic Sciences, 2014. 46(3): 304-312
5. J.-A. Bright, K.E. Stevenson, M.D. Coble, C.R. Hill, J.M. Curran, and J.S. Buckleton Bright, Characterising the STR locus D6S1043 and examination of its effect on stutter rates. Forensic Science International: Genetics, 2014. 8(1): p. 20-23.
6. D. Taylor, J.S. Buckleton. Do low template DNA profiles have useful quantitative data? Forensic Science International: Genetics, 2015. 16: 13-16.

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

1. J.-A. Bright, I.W. Evett, D. Taylor, J.M. Curran and J.S. Buckleton, A series of recommended tests when validating probabilistic DNA profile interpretation software. Forensic Science International: Genetics, 2015. 14: 125-131
2. T.W. Bille, S.M. Weitz, M.D. Coble, J.S. Buckleton, J.-A. Bright. Comparison of the performance of different models for the interpretation of low level mixed DNA profiles. ELECTROPHORESIS. 2014;35:3125-33.
3. S.J. Cooper, C.E. McGovern, J.-A. Bright, D. Taylor, J.S. Buckleton. Investigating a common approach to DNA profile interpretation using probabilistic software. Forensic Science International: Genetics, 2014. 16: 121-131.

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