Department of Forensic Biology
Charles S. Hirsch Center for Forensic Sciences
421 East 26th Street, New York, NY 10016
Telephone: 212-323-1200 Fax: 212-323-1590
Email: DNALab@ocme.nyc.gov
Official Website: www.nyc.gov/ocme

Internal Validation of STRmix™ v2.7 for Fusion 5C/3500*xL* Data

STRmix™ is a probabilistic genotyping software that utilizes a fully continuous approach to DNA sample interpretation. It was developed by the Institute of Environmental Science and Research (ESR) in New Zealand, the New Zealand Crown Research Institute, and Forensic Science South Australia (FSSA). By using various biological parameters, STRmix™ aids in interpretation and deconvolution of DNA samples by providing weight to each possible genotype combination. Subsequently, the software can perform statistical analysis in the form of a likelihood ratio (LR) for the comparison of reference samples to evidence samples (1-3).

The New York City Office of Chief Medical Examiner (referred to as the NYC OCME hereafter) performed an internal validation of STRmix™ v2.4.05 (4) and began using it in casework in January 2017. Subsequently, a performance check of STRmix™ v2.4.08 was performed and implemented into casework in June 2019 (5). Since then, newer versions with software updates have been released. The NYC OCME has validated and plans to integrate the use of STRmix™ v2.7 for use in casework. While there have been various small user interface and technical changes made with the release of newer versions, some of the more noticeable advancements that have been integrated into subsequent versions and STRmix™ v2.7 include but are not limited to: implementation of "Smart Start" allowing the MCMC to begin in an informed starting position for each contributor within the profile, decrease in default burn-in accepts to 10,000, ability to model any type of stutter (referred to as generalized stutter), addition of checks on input files to note any peaks below the analytical threshold and potential missing stutter, modeling improvements for composite peaks, drop-in, and LSAE (locus specific amplification efficiency), and general report template updates (6-7).

The following document describes the internal validation of STRmix™ version 2.7 for Promega PowerPlex® Fusion 5C data run on Applied Biosystems™ 3500*xL* Genetic Analyzers for the NYC OCME Department of Forensic Biology. The experiments detailed within this document were performed prior to the use and implementation of STRmix™ v2.7 in casework. This internal validation was conducted using known and mock evidence single source and mixture samples to investigate various aspects of STRmix™ including but not limited to variable conditions and template amounts, sensitivity, specificity, precision, and variable peak height and stutter variances. The experiments were implemented following the Scientific Working Group on DNA Analysis Methods (SWGDAM) "Guidelines for Validation of Probabilistic Genotyping Systems", the FBI's "Quality Assurance Standards for Forensic DNA Testing Laboratories", and the STRmix™ v2.7

"Implementation and Validation Guide" (8-10). (Note: All numerical designations refer to specific sections of the SWGDAM "Guidelines for the Validation of Probabilistic Genotyping Systems" that is addressed for each experiment. Unless otherwise noted, all references to STRmix™ are to version 2.7.)

Sample Description and STRmix™ v2.7 Parameters

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

4.1.3. Variable DNA typing conditions (e.g., any variations in the amplification and/or electrophoresis parameters used by the laboratory to increase or decrease the detection of alleles and/or artifacts)

All single source and mixtures samples were prepared in-house using extracted buccal swabs. Additionally, old proficiency test and mock casework samples were extracted using various methods. The single source and various combinations of two-person, three-person, four-person, and five-person mixture samples were prepared at various template amounts and ratios. For analysis, all samples were amplified using the Promega PowerPlex® Fusion 5C amplification kit at half reactions at 29 cycles, according to current NYC OCME protocols. Electrophoresis was performed using two 3500*xL* genetic analyzers run at standard conditions of 13kV for 1500 seconds, as recommended by the manufacturer (11). Sample data was analyzed using GeneMarker® version 3.0.0 and subsequently exported for use by STRmix™ (12). Because STRmix™ can model various types of stutter, GeneMarker® stutter filters were turned off prior to generation in GeneMarker® of the STRmix™ input data.

Prior to use of STRmix™ and completion of the following validation experiments, laboratory-specific parameters had to be optimized. Each parameter is dependent on the STR amplification kit and capillary electrophoresis (CE) platform/protocols used and were determined through analysis of empirical data or modeled within STRmix™ v2.7 using Model Maker. The parameters include:

a. analytical threshold (AT),
b. CE saturation limit,
c. expected stutter ratios,
d. drop-in parameters,
e. allelic and stutter peak height variance prior distributions,
f. the hyper-parameter for the variance of locus specific amplification effects (LSAE), and
g. population settings including allele frequencies and Theta (θ) values.

These laboratory-specific parameters are described and defined in detail in the "NYC OCME Internal Validation of STRmix™ v2.7 for Fusion 5C/3500*xL* Data - STRmix™ Parameters" document (11, 13-14). Unless otherwise noted, all samples were analyzed using the laboratory's ATs of 85 RFU (Blue), 120 RFU (Green), 130 RFU (Yellow), and 160 RFU (Red) with GeneMarker® HID v3.0.0 (11). These parameters were subsequently used for all STRmix™ v2.7 analysis within this internal validation. Additionally, all LR calculations presented in this document use the NIST amended frequencies for the Caucasian population database (15).

Experiment 1-3: Single Source Specimens

4.1.5 Single-source specimens

Experiment 1:

A dilution series of six single source profiles were constructed where the peak heights varied above and below the AT. The six samples were amplified at input amounts ranging from 750 pg to 3.25 pg. The samples were interpreted using STRmix™ and LRs were calculated for the known contributors where H_p = known contributor and H_d = unknown person. Figure 1 depicts the resultant LRs. Note: Sample 12M at 750 pg is not included due to elevated stutter and other artifacts that increased the apparent number of contributors (NOC) to 2. Additionally, sample 34F at 15 pg, 25 pg, and 37.5 pg were removed due to higher than expected peak heights cause by possible dilution and/or pipetting error.

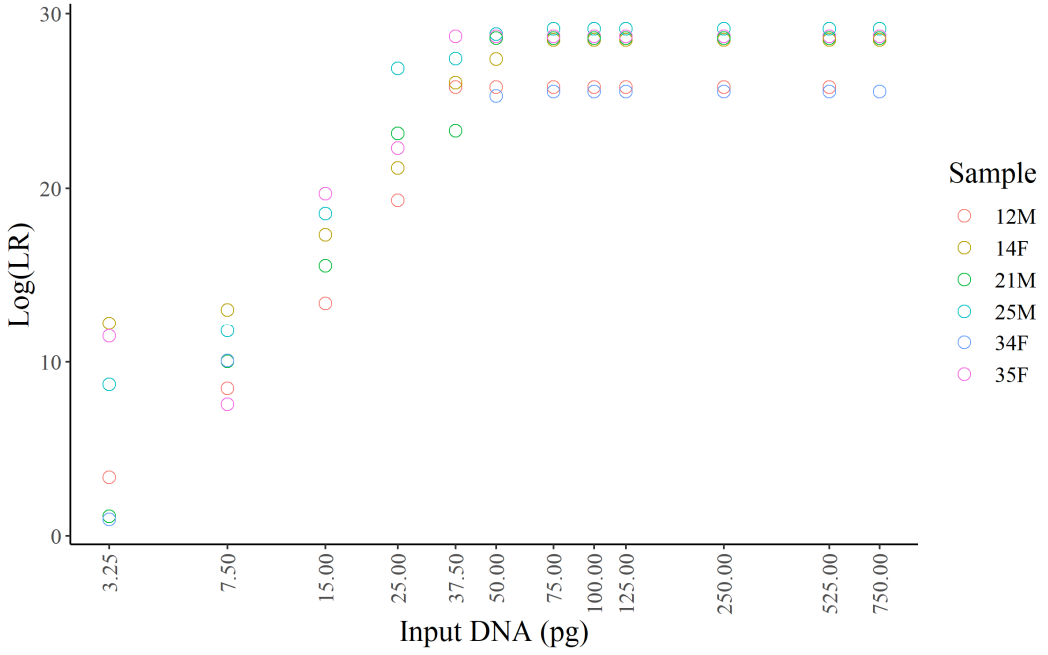


Figure 1. Log(LR) versus input DNA template (pg) for six single source samples.

Figure 1 shows that STRmix™ gives expectedly high LRs for known contributors to a single source sample. Additionally, as the template amount increases the LRs also increase. This is expected, as the weights for the genotypes increase as template amount increases due to more information available for comparison (i.e. less drop-out, higher peak heights, less stochastic effects). The direct correlation of the template amount and LR generally holds true when there is more information available and will continue to trend upwards until full profiles are deconvoluted, at which point the LR values will begin to plateau as observed in Figure 1.

Experiment 2:

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 the probabilistic methods.

For a subset of samples that may be encountered in casework, the genotypes for the profile and the comparison LRs can be easily estimated. This includes single source samples where the genotype at each locus is assigned a weight of 1 (or 100%). For a set of six single source profiles and their corresponding known contributor, a sub-sub-source LR (also called a point estimate LR) for all loci was calculated 'by-hand' (in Microsoft® Excel) using the Balding and Nichols formulas (recommendation 4.2 of NRCII and equations 1.10a and 3.10b) (16-17) and using STRmix™. The LRs were calculated two ways, once where F_{ST} (or θ) was 0 and once where F_{ST} was 0.03. Setting θ to zero returns the product rule:

$$2p_i p_j \quad \text{for heterozygote loci}$$
$$p_i^2 \quad \text{for homozygote loci}$$

Where p_i is the allele frequency for allele i and p_j is the allele frequency for allele j. When θ > 0, the Balding and Nichols formulae is applied.

For single source samples:

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

Where p_i is the allele frequency for allele i, p_j is 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:

$$\frac{x_i + {}^{1}/_{k}}{N_a + 1}$$

Where 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.

The single source profiles were amplified at the laboratory's target DNA amount of 525 pg and the LRs were calculated using the following hypotheses:

H_p: The DNA originated from a known contributor

H_d: The DNA originated from an unknown individual.

All LRs calculated 'by-hand' were calculated using the above equations with the aid of a Microsoft® Excel spreadsheet provided by the STRmix™ developers[1]. The LRs calculated 'by-hand' and using STRmix™ for one of the single source samples are presented in Table 1.

Table 1. Calculation of sub-sub-source LRs 'by-hand' (Microsoft® Excel) and using STRmix™ for single source profile 12M using the NIST amended Caucasian frequency population database for varying F_{ST} (θ) values.

	$\theta = 0$		$\theta = 0.03$	
Locus	**Microsoft® Excel**	**STRmix™ v2.7**	**Microsoft® Excel**	**STRmix™ v2.7**
D3S1358	1.3454E+01	1.3454E+01	9.4938E+00	9.4938E+00
D1S1656	6.3356E+01	6.3356E+01	3.9089E+01	3.9089E+01
D2S441	8.4929E+00	8.4929E+00	6.5726E+00	6.5726E+00
D10S1248	1.1297E+01	1.1297E+01	8.2725E+00	8.2725E+00
D13S317	9.6382E+01	9.6382E+01	5.1790E+01	5.1790E+01
Penta E	8.0028E+04	8.0028E+04	1.6030E+02	1.6030E+02
D16S539	5.0658E+00	5.0658E+00	4.8718E+00	4.8718E+00
D18S51	2.5879E+01	2.5879E+01	1.9970E+01	1.9970E+01
D2S1338	2.9075E+01	2.9075E+01	1.6863E+01	1.6863E+01
CSF1PO	6.3132E+00	6.3132E+00	5.9148E+00	5.9148E+00
Penta D	7.1376E+01	7.1376E+01	4.3295E+01	4.3295E+01
TH01	8.4225E+00	8.4225E+00	6.5280E+00	6.5280E+00
vWA	2.3670E+01	2.3670E+01	1.8394E+01	1.8394E+01
D21S11	1.2553E+01	1.2553E+01	8.9919E+00	8.9919E+00
D7S820	9.5329E+00	9.5329E+00	8.5750E+00	8.5750E+00
D5S818	6.6622E+00	6.6622E+00	5.3783E+00	5.3783E+00
TPOX	3.6367E+00	3.6367E+00	3.2072E+00	3.2072E+00
D8S1179	2.8978E+01	2.8978E+01	2.1843E+01	2.1843E+01
D12S391	2.2637E+02	2.2637E+02	9.2191E+01	9.2191E+01
D19S433	1.5429E+01	1.5429E+01	1.0556E+01	1.0556E+01
FGA	5.6138E+01	5.6138E+01	3.4275E+01	3.4275E+01
D22S1045	9.3601E+00	9.3601E+00	8.2282E+00	8.2282E+00
Total	3.3304E+31	3.3304E+31	6.7382E+25	6.7382E+25

As shown in Table 1, the sub-sub-source LR calculated at each locus within STRmix™ is the same as the 'by-hand' calculation for both $\theta = 0$ and $\theta = 0.03$. All sub-sub-source LRs calculated 'by hand' for the remaining five single source profiles were the same as those calculated by STRmix™ (Appendix A). These results serve as a check of the LR calculation within the STRmix™ software and the results are as expected.

[1] http://support.STRmix™.com/support/solutions/articles/1000219464-single-source-lr-calculator

Experiment 3:

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

Thirty-four single source profiles were amplified at above-optimal DNA input amounts (750 pg, 1 ng, 2 ng) to obtain saturated peaks (i.e. ≥ 30,000 RFU). These profiles were then interpreted through STRmix™ using their true number of contributor (NOC) of 1. Additionally, LRs of the known contributors were also calculated in STRmix™. For 22 of the samples run, all interpretations resulted in intuitive genotypes where the weight was equal to 1 (i.e. 100%) for the known contributor's genotype. For four of the samples, all interpretations resulted in intuitive genotypes except for locus D19S433. These runs were all amplifications from sample 12M where most of the weight was given to genotype 13,13 (92.60%, 69.51%, 84.94%, and 68.53%) and some weight was given to genotype 12,13 (7.40%, 30.49%, 15.06%, and 31.47%). For this location, the 13 allele was oversaturated leading to higher than expected back stutter. This led to the 12 stutter peak not only being modeled as stutter but also as a potential true allele peak. While 100% of the weight was not given to the expected genotype of 13,13 (the profile of 12M), most of the weight for this location was given to this intuitive genotype which allowed for the comparison to result in an intuitive inclusionary LR.

Five of the 34 samples resulted in LRs of 0 due to a non-intuitive genotype at a single locus having a 100% weight. Additionally, three samples failed to run due to a NOC of 1 not being a good fit for the data imported. Figure 2 provides a summary of the log(LR) versus input DNA (ng) for each sample (not including samples from failed runs or LRs of 0).

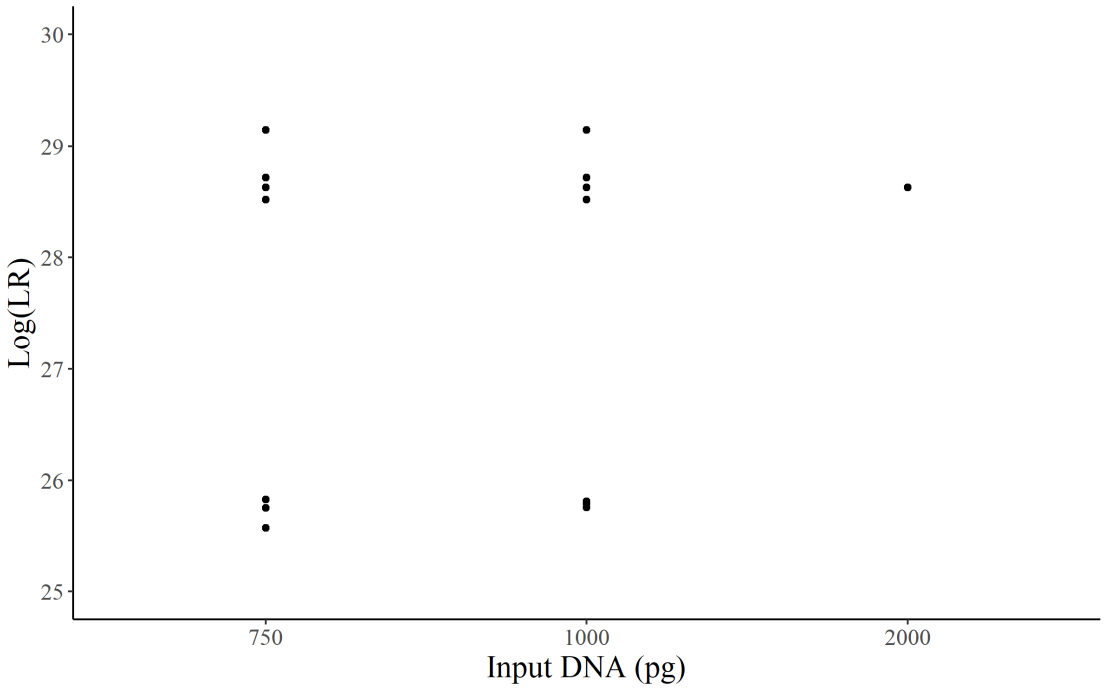


Figure 2. Log(LR) versus input DNA (pg) for 26 profiles amplified at above-optimal DNA amounts.

As seen in Figure 2, the 26 samples that gave intuitive genotypes resulted in high LRs for the comparison of the known contributor. For one of the samples with an LR of 0, G07_FBRE18-00004_OS_25M_1ng_rep.hid (Carmody011119 120-122), D5S818 had a non-intuitive genotype of 12.3,12.3 called instead of the intuitive and true genotype of 13,13. After further investigation, wide peak morphology in the sample and the internal lane standard at this location showed shifting and, thus, a possible migration issue during CE. The 4 remaining samples with an LR of 0 were all from 1 ng and 2 ng amplifications of sample 12M. At locus D19S433, STRmix™ gave a weight of 1 (100%) to a non-intuitive genotype of 12,13 instead of the intuitive and true genotype of 13,13. For each of these profiles, the 13 allele had heights of 32,510 RFU – 32,628 RFU which is well above the saturation point of the 3500*xLs*. The 12 allele for these profiles was 9.28% - 13.82% of the corresponding 13 allele. While these stutter ratios are below the maximum allowable stutter for back stutter (30%), they are above the expected stutter ratio for that particular allele (8.78%) (14). Additionally, oversaturated peaks can overwhelm the CCD camera of the CE instrument leading to inaccurate allelic peak height information, which subsequently affects the relationship between the allelic peak height and the expected stutter peak height. Oversaturation also becomes more likely at homozygote locations for samples at high template amounts. This shows how overly saturated peaks can lead to an unexpected genotype; in this case, the overly saturated homozygote peak is paired with its own back stutter peak. Results like these, with unintuitive allele pairings, are recognizable by trained analysts and the sample should be re-run at a dilution.

Additionally, Figure 3-6 provide summaries of each profiles stutter variance (k^2) for each type of stutter modeled versus the template RFU.

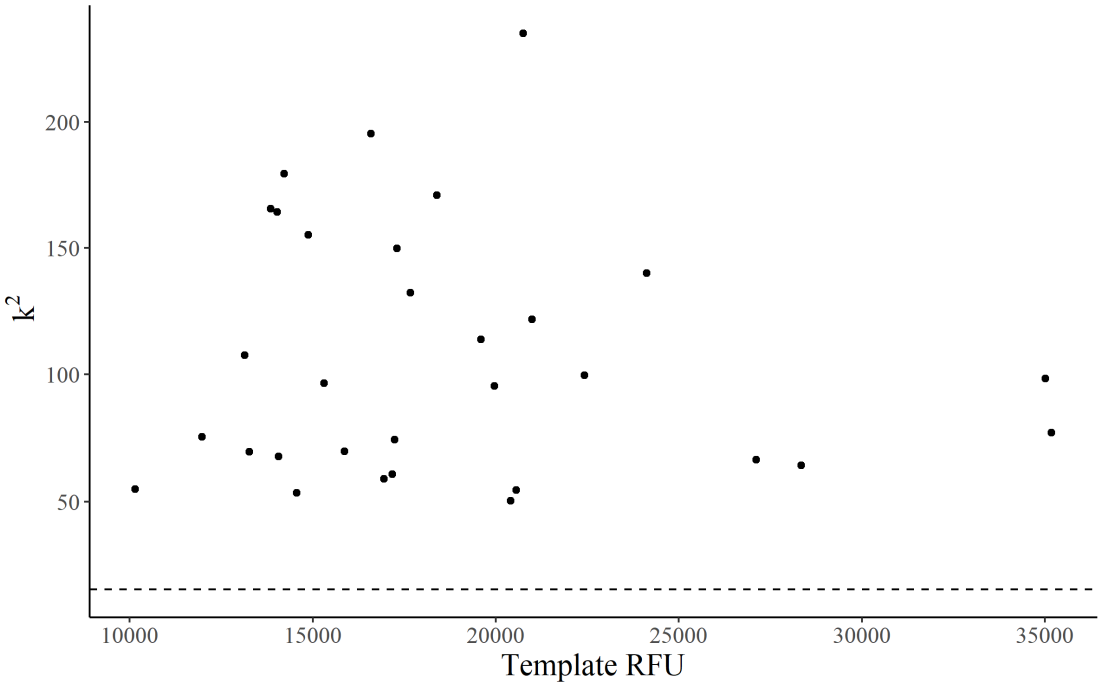


Figure 3. Per profile back stutter variance (k^2) versus template (mean over the post burn-in accepts) for samples amplified at above-optimal DNA amounts. The horizontal line corresponds to the mode (15.223) of the back stutter variance parameters (1.799, 19.052).

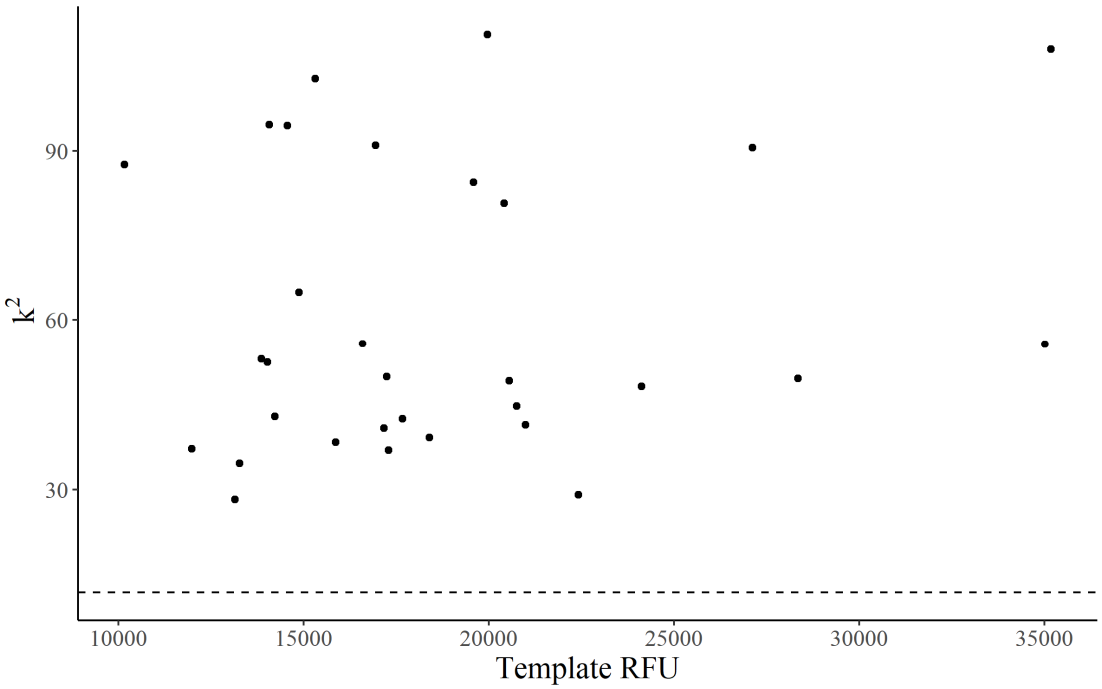


Figure 4. Per profile forward stutter variance (k^2) versus template (mean over the post burn-in accepts) for samples amplified at above-optimal DNA amounts. The horizontal line corresponds to the mode (11.691) of the front stutter variance parameters (1.999, 11.703).

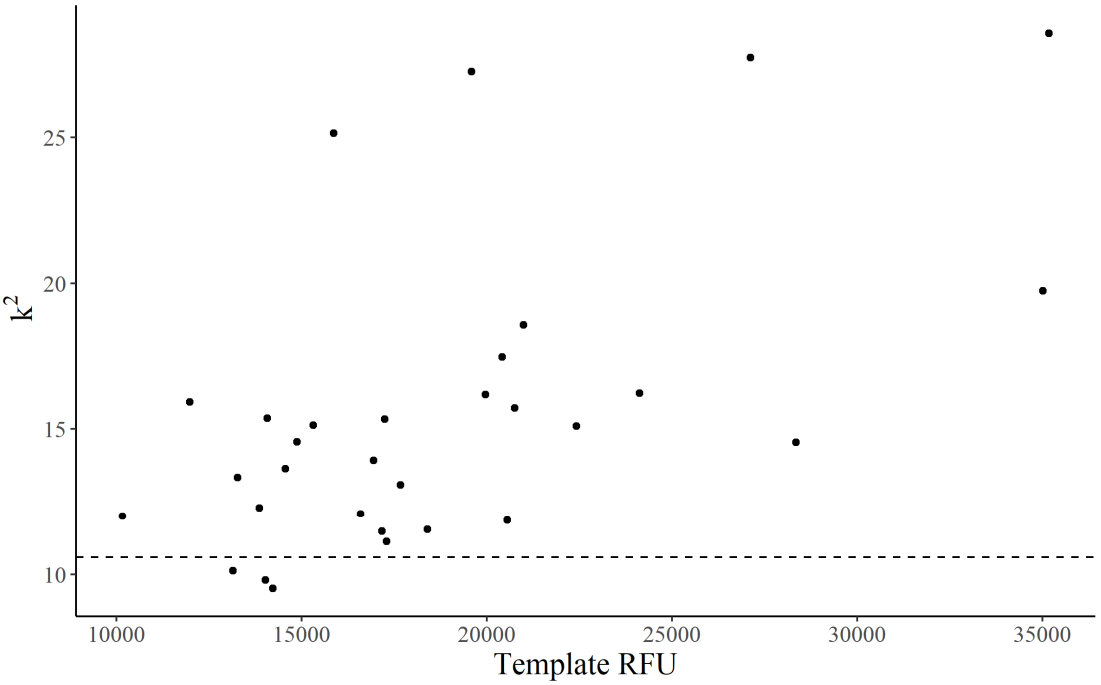


Figure 5. Per profile half back stutter variance (k^2) versus template (mean over the post burn-in accepts) for samples amplified at above-optimal DNA amounts. The horizontal line corresponds to the mode (10.591) of the half back stutter variance parameters (2.597, 6.632).

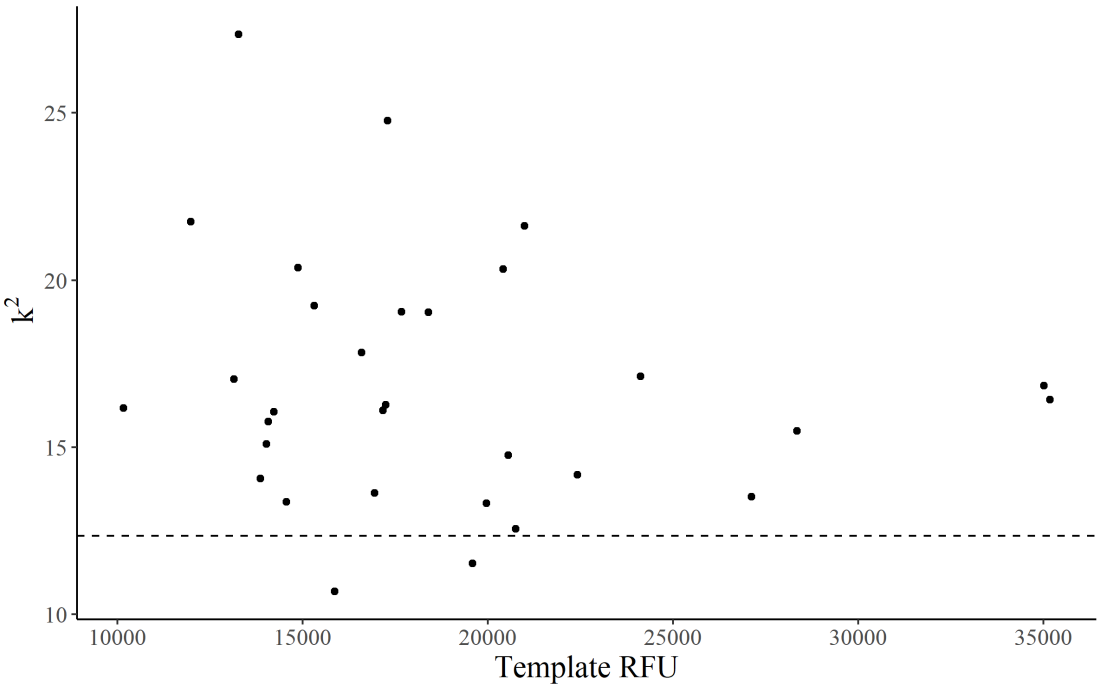
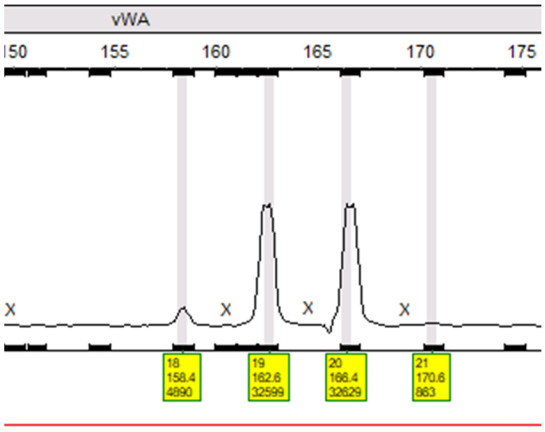


Figure 6. Per profile double back stutter variance (k^2) versus template (mean over the post burn-in accepts) for samples amplified at above-optimal DNA amounts. The horizontal line corresponds to the mode (12.334) of the double back stutter variance parameters (2.816, 6.792).

As observed in Figures 3-6, the stutter variance of the oversaturated profiles is generally higher than the mode. Due to the saturation of the peaks within the profiles, the expected height of the stutter peak is calculated based on the expected height of the parent allele and not the observed height of the parent allele. Ultimately, the observed corresponding stutter peak height ratios to the saturated parent peaks are larger than expected which leads to higher-than-expected variance between the observed and expected stutter peaks. Additionally, there were instances in which STRmix™ generated a message prior to running indicating potential missing stutter peaks for alleles where they were expected to appear. These messages were noted but the analyses were continued, and results were evaluated - these runs also generally correlated to higher than expected stutter variance values.

All three samples that failed to be run to completion in STRmix™ due to the NOC determination were amplified at 2 ng: D03_FBRE18-00004_OS_25M_2ng, E01_FBRE18-00004_OS_21M_2ng, F05_FBRE18-00004_OS_21M_2ng_rep (all from Carmody011119 120-122). For all three of these samples, there was one location in which the evidence could not be explained by the NOC parameter of 1. After further investigation, the locations marked as 'cannot be explained' by STRmix™ were heterozygote locations where the parent allele peak heights were >31,000 RFU. Figure 7 is an example of one of the failed samples at locus vWA where (a) is the locus with stutter filters on and (b) is the locus with stutter filters off within GeneMarker®.

(a)



(b)

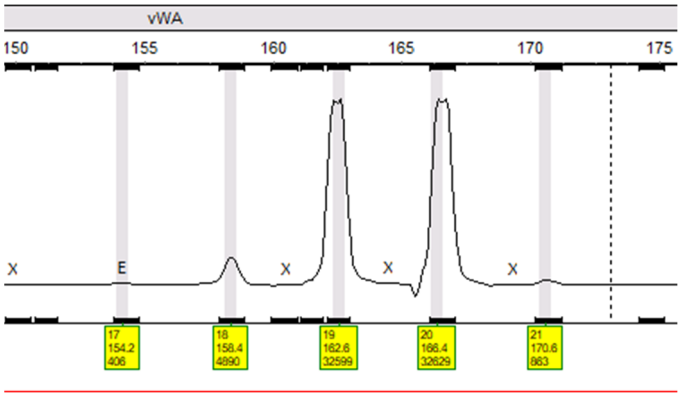


Figure 7. Locus vWA of saturated sample E01_FBRE18-00004_OS_21M_2ng that failed to run in STRmix™ 2.7 with stutter filters (a) and without stutter filters (b).

Double back stutter is defined as two repeat units shorter than the parent allele. Locus vWA, in Figure 7, does not have a double back stutter filter in GeneMarker® and double back stutter is not modeled at this location in STRmix™ (13-14). Because STRmix™ models stutter, the data is imported without applying the stutter filters, as shown in (b). Due to the oversaturation of the 19 true parent allele, both the 18 allele (stutter of the 19) and the 17 allele (stutter of the 18) are elevated. During STRmix™ interpretation, this 17 allele is above the drop-in cap and could not be modeled as double back stutter and, instead, had to be considered a true allele; the run could not complete as the data could not be explained with a NOC of 1. In cases like these, this result would then be reviewed by the analyst to determine whether the NOC needs to be reconsidered or rerun at a dilution to account for the sample's oversaturation.

While samples with oversaturated peaks can be accurately analyzed, this experiment exemplifies the importance of how saturated peak heights may lead to elevated stutter or other potential artifacts like elevated baseline and pull-up that could lead to unintuitive STRmix™ analysis results. Only one 2 ng sample tested in this experiment resulted in intuitive results while the rest were all 1 ng and 750 pg. Additionally, only samples amplified at 1 ng or 2 ng resulted in non-intuitive results with an LR = 0 or failure to run due to NOC. This shows that samples amplified at or above 750 pg are more affected by saturation, meaning that caution should be taken when deciding on how much to amplify above the optimal amount (525 pg for the NYC OCME). If a sample does have excessive oversaturation, a re-amplification at a dilution should be considered.

Experiments 4-7: Specificity, Sensitivity, and Precision

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

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

4.1.6.5. Sharing of alleles among contributors

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

Experiment 4:

Various hypotheses were tested using two-person, three-person, four-person, and five-person mixture sets based on the determined apparent NOC. The contributor profiles contained both homozygous and heterozygous loci, variable amounts of allele sharing between contributors, and a minor component of each mixture that experienced dropout as the DNA template amount decreased. The following hypotheses were tested:

H_p: the DNA originated from a known contributor and N-1 unknown individuals

H_d: the DNA originated from N unknown contributors

where N is the apparent NOC based on what is observed in the sample data. The apparent NOC was used to ensure that the results mimic casework in which the true number of contributors is never known. Determining the NOC is complicated by allele sharing and various artifacts that may be present from amplification and/or capillary electrophoresis. With a consensus of five trained analysts, the NOC of each sample was determined through scrutiny of the DNA profiles and assessment of various aspects within each sample. The characteristics that are evaluated in a

sample for determination of NOC include, but are not limited to, a combination of: (a) using the maximum allele count method as an initial estimate (maximum number of alleles at a locus divided by two); (b) peak height balance/imbalance; (c) amount of DNA amplified; (d) presence of peaks below analytical or stochastic threshold; (e) possible degradation; (f) mixture ratio; (g) potential allele sharing. For more information on the NYC OCME's protocol for determining NOC, see the NYC OCME's STR Analysis and Interpretation Manual for PowerPlex® Fusion 5C and STRmix™ v2.7 on 3500*xLs* (18).

The mixture samples were deconvoluted in STRmix™ and compared to their known contributors as well as 10,000 non-contributor profiles using the Database Search function found within the 'Investigation' module of STRmix™. The NIST amended frequencies for the Caucasian population database were used for LR calculations with theta ($\theta = 0.03$) and sub-source LR applied (Factor of N!). The non-contributor profiles were generated using a spreadsheet provided by the STRmix™ developers that can simulate a database of profiles based on a given set of allele frequencies. The goal of this experiment was two-fold:

1. To test the limits of STRmix™ v2.7 to see the range of LRs for true contributors and non-contributor profiles, and

2. To see the effects of DNA template amount on LR calculations.

Table 2 provides a summary of the number of profiles with their known and apparent NOC.

Table 2. Summary of profiles, true and apparent NOC for samples amplified at 37.5 pg – 750 pg.

		Apparent NOC of Profiles				
True NOC of Profiles		**1**	**2**	**3**	**4**	**5**
	1	-	5	-	-	-
	2	15	127	27	-	-
	3	-	9	75	6	-
	4	-	1	36	43	-
	5	-	-	7	18	15

As observed, there are instances in which the true NOC and the apparent NOC are not the same. While the true NOC to evidence profiles is never truly known, there are various reasons as to why the true and apparent NOC may not align, leading to an under- or over-estimation of the true NOC. For instance, an increased number of artifacts or elevated stutter peaks may present as an additional contributor, while low template amounts, degradation, trace contributors with drop-out, or potential allele sharing may lead to a determination of fewer contributors.

To investigate the range of LRs for both known and non-contributor comparison profiles, log(LR) values were plotted against the template amount in RFU per contributor. The template RFU was designated as the minimum average peak height (APH) among the contributors as assigned by the STRmix™ analysis. Figures 8-12 show this comparison for all apparent two, three, four, and five-person mixtures with any LRs of 0 plotted at Log(LR) = -50.

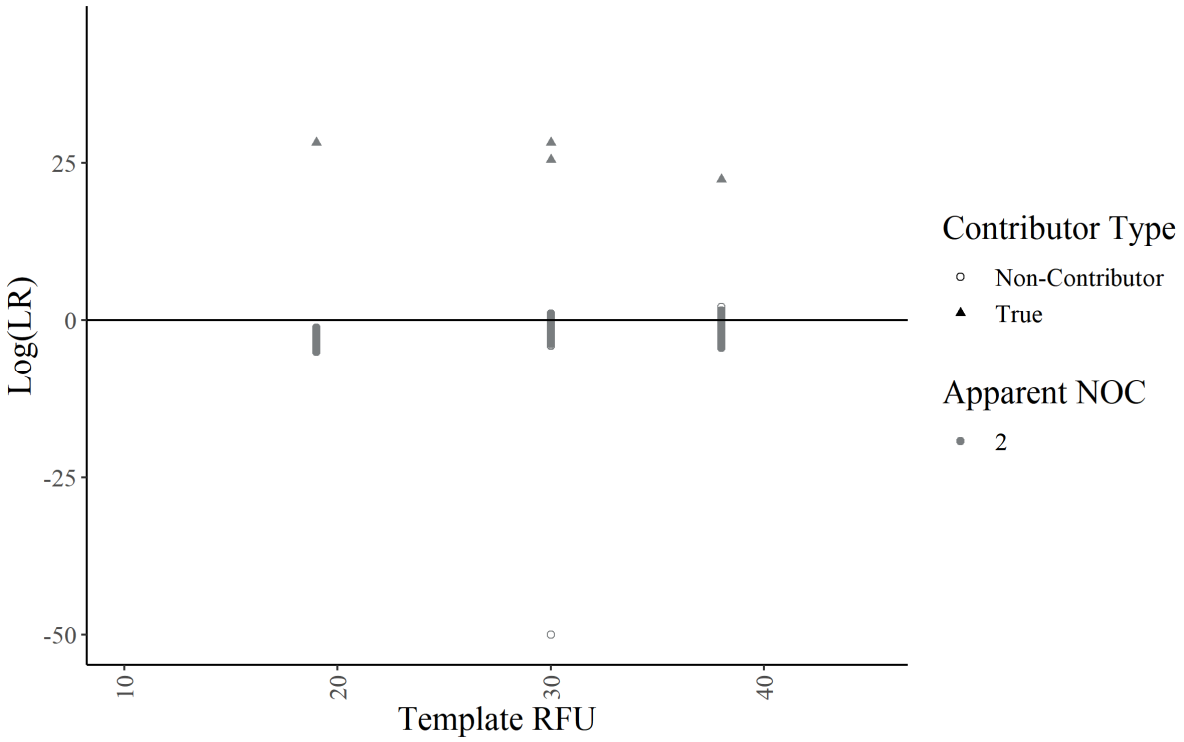


Figure 8. Log(LR) versus template RFU per contributor for true one-person samples of variable ratios and amplified amounts.

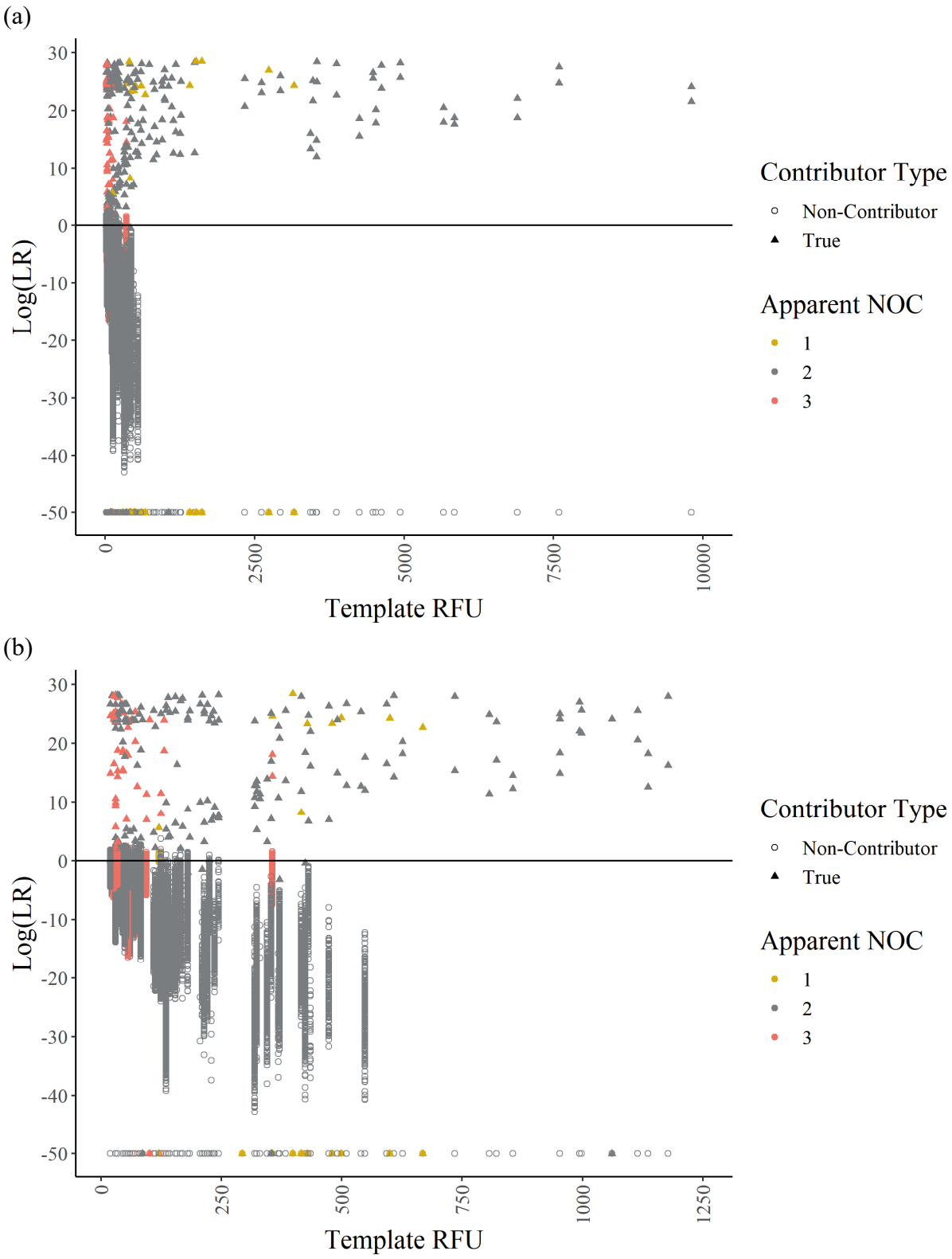


Figure 9. Log(LR) versus template RFU per contributor for true two-person mixtures of variable ratios and amplified amounts. (a) all data plotted; (b) zoomed in x-axis.

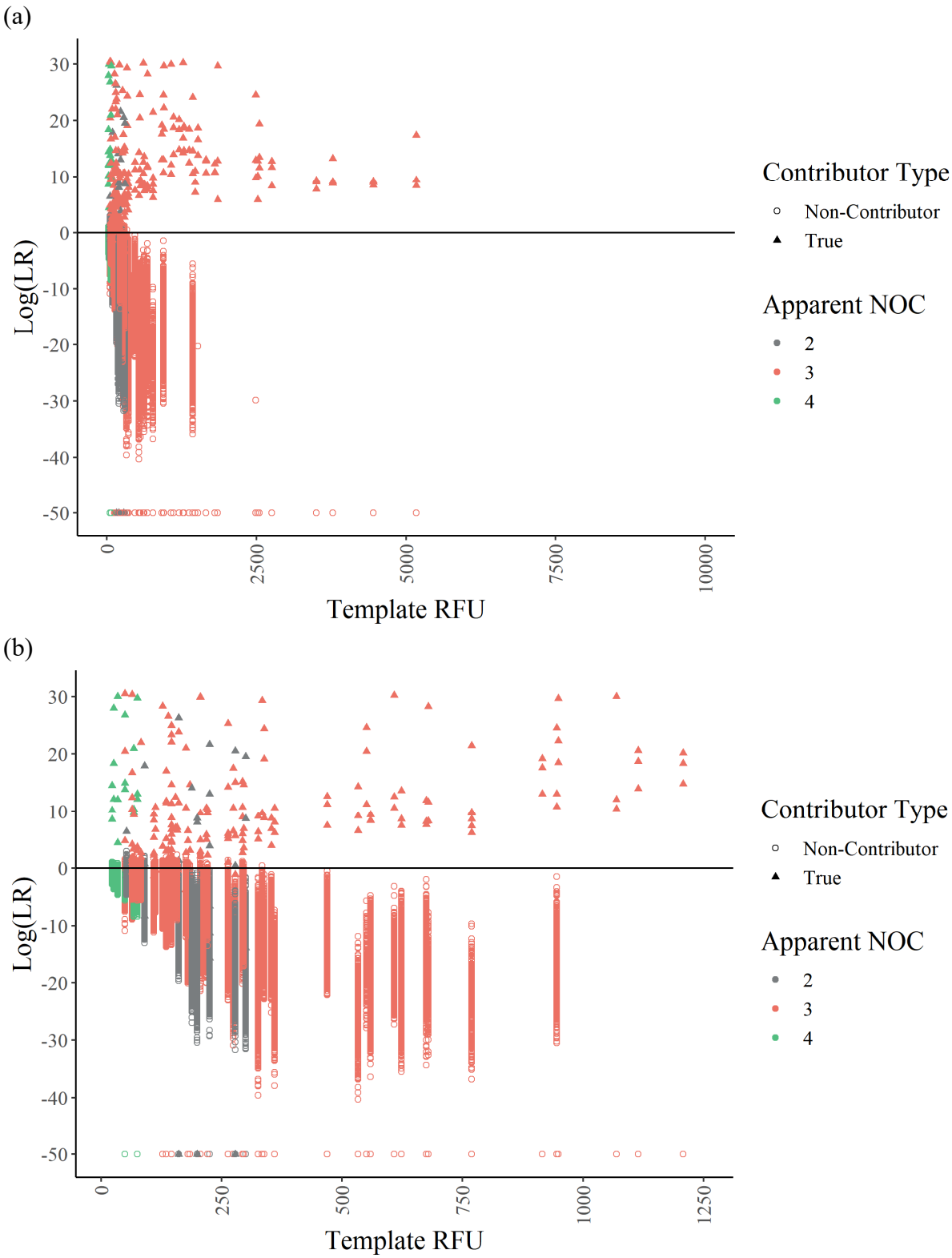


Figure 10. Log(LR) versus template RFU per contributor for true three-person mixtures of variable ratios and amplified amounts. (a) all data plotted; (b) zoomed in x-axis.

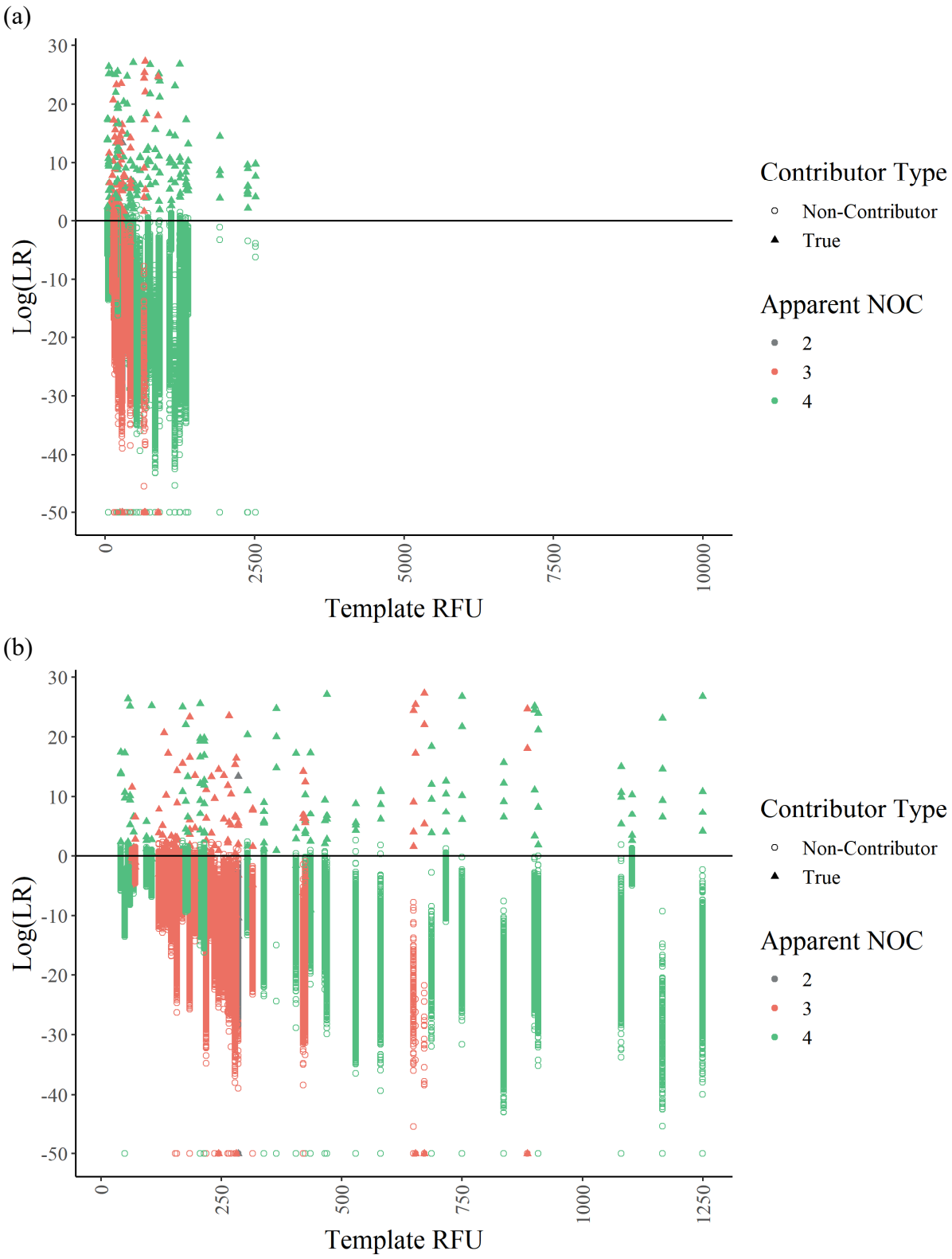


Figure 11. Log(LR) versus template RFU per contributor for true four-person mixtures of variable ratios and amplified amounts. (a) all data plotted; (b) zoomed in x-axis.

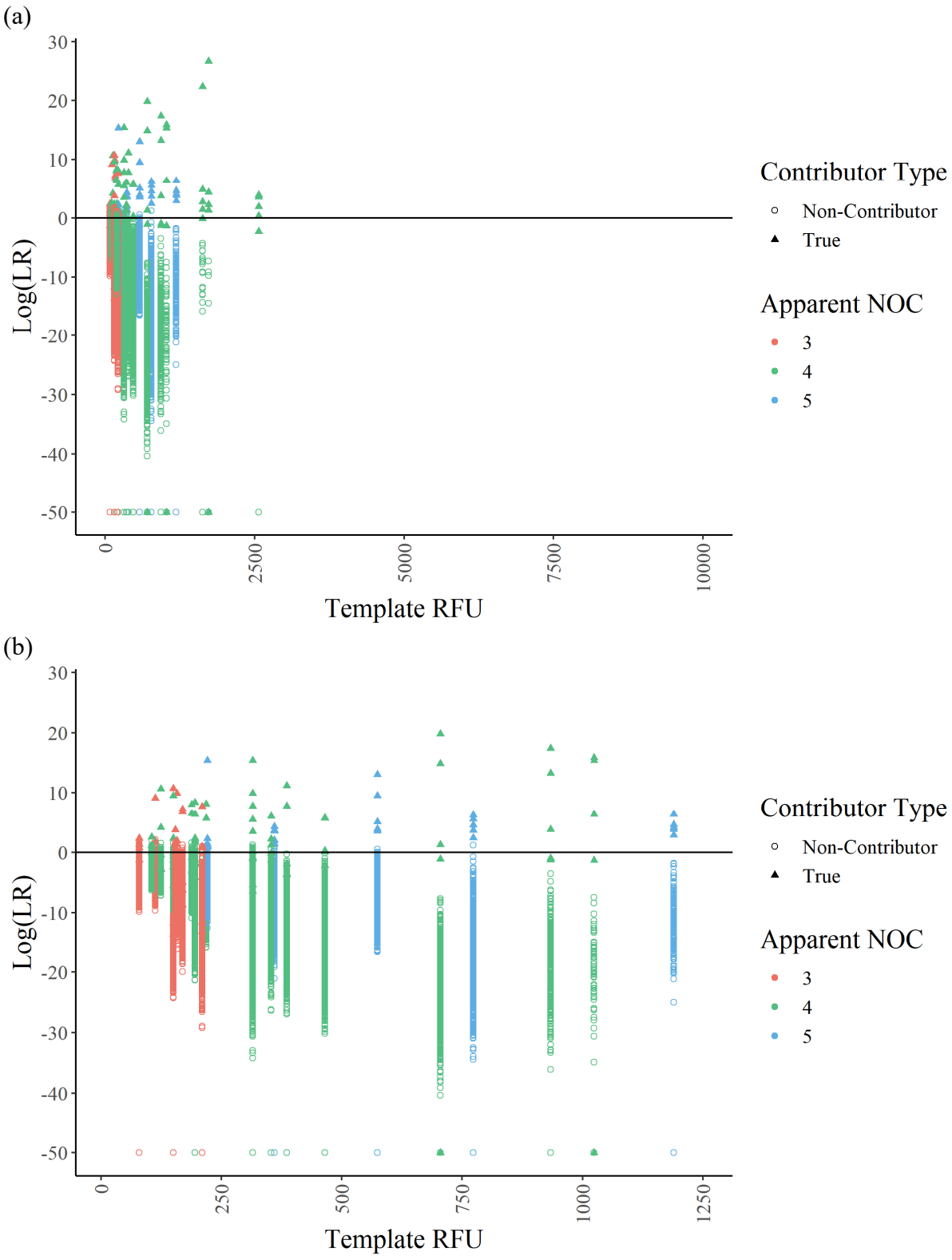


Figure 12. Log(LR) versus template RFU per contributor for true five-person mixtures of variable ratios and amplified amounts. (a) all data plotted; (b) zoomed in x-axis.

Figures 8-12 demonstrate that true contributors generally gave high LRs and the non-contributor profiles gave low LRs. This was especially true for high template samples. In some instances, it was observed that true contributors resulted in LRs in support of exclusion. These LRs were due to limited data observed for that contributor (especially in extreme mixture ratios) and/or when the apparent NOC was less than the true NOC. Although the true NOC can never be known, when applying a NOC that is less than that of the true, this can result in counter-intuitive genotype pairings for a minor or trace contributor. Ultimately, this may result in false support for exclusion for a low-level true contributor.

There were a few instances of high template RFUs for a minor contributor in which an LR supporting exclusion was obtained for a true contributor. Some were caused by artifacts falling within an allele bin inadvertently being left in the data and, subsequently, the STRmix™ import file. While these artifacts are well characterized as non-DNA peaks that may occur with the PowerPlex® Fusion amplification kit (19), they were being considered as true alleles by STRmix™, causing genotype possibilities to be generated with this artifact allele, that then subsequently did not align with the true minor contributor profile. These results highlight the importance of removing artifacts at the analysis stage to ensure accurate import into STRmix™. The remaining higher template samples resulting in false support of exclusion of a true contributor occurred with higher order mixtures, i.e. 4- and 5-p mixtures, where the apparent NOC was less than that of the true NOC and where the minor contributor(s) resulted in an exclusion. In these cases, there was limited data representing these lower-level contributors, thus, resulting in an LR = 0 for the comparison of the true contributors to the samples.

Table 3 tabulates the graphical information shown in Tables 8-12. Table 3a summarizes the number of LRs that were observed for both true contributors and non-contributor profiles falling below and above 1 for samples amplified at 37.5 pg – 750 pg based on their true NOC. An LR <1 indicates support for exclusion and an LR >1 indicates support for inclusion. Additionally, Tables 3b and 3c further break down the LRs based on the apparent NOC of the mixtures.

Table 3. Number of LRs less than 1 and greater than 1 for *true contributors and non-contributors* for samples amplified at 37.5 pg – 750 pg based on true NOC; (b) Number of LRs less than 1 and greater than 1 for *true contributors* compared to apparent NOC; (c) Number of LRs less than 1 and greater than 1 for *non-contributors profiles* compared to apparent NOC.

True NOC	True-Contributor LRs		Non-Contributor LRs	
	# LRs <1	# LRs >1	# LRs <1	# LRs >1
1	0	4	39065	1271
2	40	296	1675239	18705
3	23	247	905010	2370
4	60	256	795133	1266
5	54	96	302130	270

True NOC	True-Contributor LRs					
	# LRs <1			# LRs >1		
	Apparent < True NOC	Apparent = True NOC	Apparent > True NOC	Apparent < True NOC	Apparent = True NOC	Apparent > True NOC
1	n/a	0	0	n/a	0	4
2	16	22	2	14	232	50
3	12	11	0	15	214	18
4	54	6	0	90	166	0
5	54	0	0	71	25	0
Sum/Total per True NOC	76.8%	22.0%	1.1%	21.1%	70.9%	8.0%

True NOC	Non-Contributor LRs					
	# LRs <1			# LRs >1		
	Apparent < True NOC	Apparent = True NOC	Apparent > True NOC	Apparent < True NOC	Apparent = True NOC	Apparent > True NOC
1	n/a	0	39065	n/a	0	1271
2	151237	1265340	258662	8	15201	3496
3	90368	754503	60139	370	1647	353
4	362421	432712	0	495	771	0
5	251748	50382	0	252	18	0
Sum/Total per True NOC	23.0%	67.3%	9.6%	4.7%	73.9%	21.4%

Over 3 million comparisons were made to apparent 2-, 3-, and 4-person mixtures. For the database comparisons to non-contributor profiles, where the LRs were calculated using theta ($\theta = 0.03$) and sub-source LR applied, the rate of false support for inclusion of a non-contributor profile was ~0.65%. The non-contributor profile comparisons are further broken down by given LR ranges in Table 4.

Table 4. Total number and percentage of LRs within a given range for the comparison of >10,000 non-contributor profiles to 2-, 3- and 4-person mixtures amplified at 37.5 pg – 750 pg.

Given LR	# of LRs < Given LR	% of LRs < Given LR	# of LRs >= Given LR	% of LRs >= Given LR
1	3666195	99.3533%	23864	0.6467%
10	3688417	99.9555%	1642	0.0445%
100	3689989	99.9981%	70	0.0019%
1000	3690058	99.9999%	1	0.0001%
10000	3690059	100.0000%	0	0.0000%
100000	3690059	100.0000%	0	0.0000%
1000000	3690059	100.0000%	0	0.0000%

Additional calculations were performed for the one non-contributor profile comparison that resulted in an LR above 1000. This additional calculation was performed to apply the highest posterior density (HPD) to integrate another level of conservatism specifically related to the MCMC uncertainty (which would be used in case work). Table 5 outlines the LRs calculated both before and after re-running the comparison sample to the mixture.

Table 5. Non-contributor profile comparison with a database LR >1000 re-calculated with HPD applied. *Database LR was calculated using a theta of $\theta = 0.03$ and sub-source LR (Factor of N!) applied.

Sample	Apparent NOC	True NOC	Non-contributor Database Profile ID	Database LR*	HPD LR
M2_75-1_525_P 2p	2	2	9376	5471.8	2030.5

While the resulting LR is lower upon re-calculation to apply the HPD LR, the result for the comparison of this non-contributor profile remains above 1000. Review of the STRmix™ interpretation report shows that the major contributor to the sample deconvoluted out with a weight of 1 (100%) for each genotype. For the minor contributor, with which the non-contributor profile

comparison aligns, drop-out was observed throughout the profile. Additionally, the mixture proportion was 99.04% for contributor 1 (major) and 0.96% for contributor 2 (minor). This ratio conforms to the qualitative expectations based on the observed data and is consistent with the mixture ratio, 75:1. Due to the limited data for the minor contributor, this left room for ambiguity within the genotypes of the profile as is shown through the spread of genotype choices, including "Q" alleles representing any allele, for the minor contributor. Additionally, the non-contributor profile shares some alleles with both the known major and known minor contributor as well as has some of the same alleles as the stutter of the major contributor. Because of this, within the possible genotypes given for the minor contributor by STRmix™, the first or second highest weighted genotype combination for the minor aligned with the non-contributor profile. This ultimately resulted in an LR that supported inclusion for that non-contributor profile. Ultimately, this false support for inclusion of the non-contributor profile was due to the fortuitous allele sharing between the non-contributor profile and the true contributors and was not a failure of STRmix™.

Current OCME protocol is to report the lowest unified LR of the Caucasian, Asian, African American, or Hispanic population. The unified LR applies another level of conservatism that can be used when calculating LRs by considering relatives of the known contributor within the unknown population. Upon re-calculation of the above sample, the unified LRs were also calculated for all four populations. While comparisons were made using LRs from the Caucasian population, the LRs for the other populations were slightly different due to varying allele frequencies between the different populations. The unified LR for the Caucasian population was 2030; however, this was not the lowest unified LR. The lowest unified LR was 247 for the African American population which falls below 1000.

Overall, the vast majority of comparisons in this experiment resulted in high LRs for true contributors and low LRs (or an LR of 0) for non-contributor profiles. Although a small number of comparisons gave false support for exclusion of a true contributor or false support for inclusion of a non-contributor profile, these results are not unexpected. This is due to the characteristics of mixture samples including, but not limited to, the presence of extreme mixture ratios, an increase in artifacts or elevated stutter, drop-out, and overall limited data for minor and/or trace contributors where genotype weights are ultimately affected. These characteristics may lead to a true contributor not aligning to genotypes being given weight by STRmix™ or to a non-contributor's profile aligning to genotypes being given more weight by STRmix™, with the latter referred to as an "adventitious match" (20). One way to aid in mitigating these types of situations is to compare samples to a database of non-contributor profiles, creating a non-donor distribution of LRs, to establish an uninformative range. An uninformative range is determined from the analysis of these comparisons which can then define LR thresholds for support for inclusion, support for exclusion, and no support for either. The uninformative range for the NYC OCME will be set at 1000 as 99.9999% of the LRs for non-contributor profile comparisons fell below this value. For any comparison, if the LR is 0.001 – 1000, no support will be given to either hypothesis.

Additional figures consisting of a plot of the log(LR) versus total input DNA for all mixtures and summaries of the secondary diagnostics per apparent NOC including total iterations, average log(LR), Gelman Rubin (GR), effective sample size (ESS), allele variance, and stutter variances can be found in Appendix B and C.

Experiment 5:

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.

Mixtures were assessed through STRmix™ with known contributors as assumed (or conditioned) contributors. A selection of 2-, 3-, 4-, and 5-person mixtures of varying ratios and amplified amounts were chosen for analysis from the set of mixtures from experiment 4. The mixtures were re-interpreted through STRmix™ with a true, major contributor, set under both hypotheses. If a mixture had equal proportions for all contributors, one true contributor was selected for conditioning. The propositions are as follows for a mixture with N contributors:

> H_p: The DNA originated from the conditioned contributor, a known contributor, and N-2 unknown individual(s)
>
> H_d: The DNA originated from the conditioned contributor and N-1 unknown individual(s)

Figure 13 is a comparison of the log(LRs) for the minor contributor(s) prior to conditioning and after conditioning on a true major contributor. The LRs were calculated using a theta of 0.03 and the sub-source LR (factor of N!) applied.

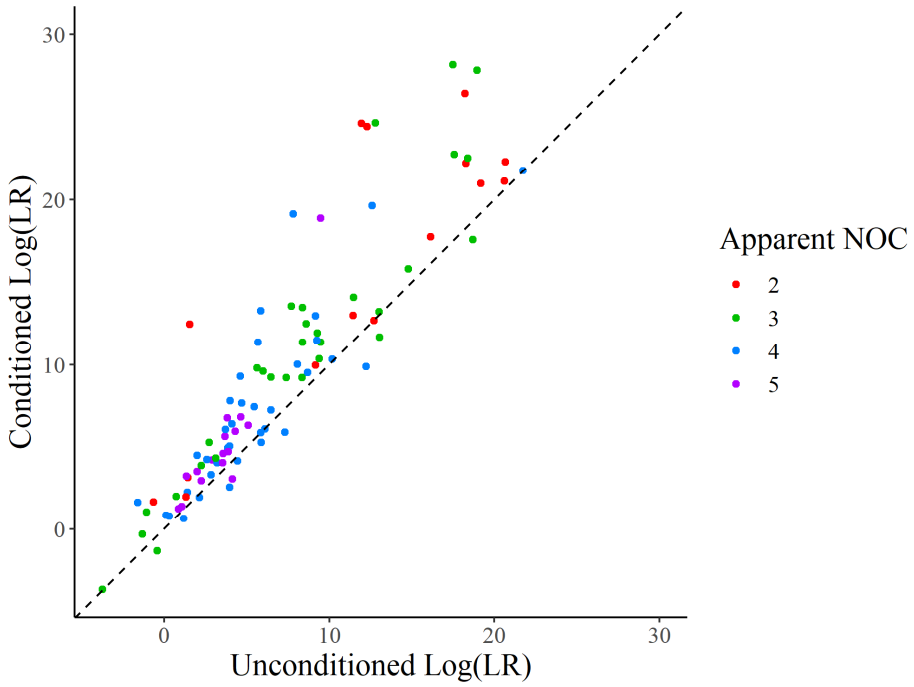


Figure 13. Log(LR) with no conditioning versus log(LR) with conditioning of a true contributor. The diagonal line indicates x=y.

As shown in Figure 13, there was an overall increase in the LR values after conditioning on a true contributor as indicated by the values above the x=y line. With the addition of more relevant information, such as conditioning on a true contributor, it better informs the deconvolution and LR calculation provided by STRmix™ (10). Providing information about a known contributor allows STRmix™ to use that information to better inform the genotypes, genotype weights, and mixture proportions for additional contributors. In general, LRs for known contributors (H_p true) will increase and the LRs for non-contributors (H_d true) will decrease (21-22).

There were 11 comparisons in which the LR did not increase but instead decreased slightly with the addition of more relevant information. An example of this is sample M2_10-5-3-1_525_C, where the LR of a minor contributor without a conditioned contributor was 1.76E+12 and the LR with a conditioned contributor was 7.63E+09. The comparison sample giving this LR was the known third contributor. When looking at the deconvolutions, this decrease in LR is likely associated with the slight shift in mixture proportions reported by STRmix™, going from 48.18%:26.99%:18.38%:6.44% for the unconditioned run to 42.77%:34.95%:16.24%:6.05% for the conditioned run. By conditioning and assigning genotypes to the major contributor, the ratio and available alleles for possible genotypes of the minor contributor are affected. This sample represents the largest decrease for samples falling into this category, with a log(LR) difference of ~2.3. Although this is a decrease, the qualitative interpretation remained the same for the comparisons to this sample with both LRs showing support for inclusion.

Prior to conditioning, six comparisons resulted in LRs < 1. After conditioning, five of the six comparison LRs slightly increased with three resulting in LRs > 1. For these samples, the difference between the log(LR) calculations were < 3.2, and all LRs maintained the same qualitative interpretation; all were within the uninformative range of 0.001-1000, as determined in experiment 4. When comparing all the LRs to the NYC OCME's uninformative range, 11 comparisons resulted in different conclusions between the non-conditioned run and the conditioned run. Nine of these comparisons resulted in LRs within the uninformative range for the non-conditioned run and support for inclusion with the conditioned run, with all log(LR) differences < 2.54. These results are expected as conditioning on a known contributor generally increases the LR for other true contributors. For one of the remaining two comparisons, sample M2_1-3-3-1_525_P, the LR decreased from 9.55E+03 with no conditioning to 3.70E+02 with conditioning. While this small difference shifted the overall conclusion, the comparison profile was that of a minor contributor and the LR did not change enough to push this sample in support of exclusion.

The last comparison resulting in different conclusions was for the comparison of the minor contributor to sample M2_10-1_37.5_C. The LR was 3.54E+01 prior to conditioning and 2.88E+12 after conditioning. Upon looking at the deconvolutions and the data observed for the sample, the low template amount not only caused drop-out of the minor contributor but also caused drop-out for the major contributor, in addition to other stochastic affects such as inconsistent peak height ratios. Without a conditioned contributor, this ultimately caused much more ambiguity for both the major and minor contributor genotypes with a wide spread of weight for various possible genotype combinations. When the major contributor was conditioned upon, there was less ambiguity in the genotypes and genotype weights for the minor contributor, going from 1 of 24 alleles ≥ 99% without conditioning to 15 of 24 alleles ≥ 99% with conditioning. This supports that conditioning on a known contributor can aid in modeling the genotypes and genotype weights for other contributors.

Additionally, there were six 5p samples that completed running with a conditioned contributor that initially did not run in experiment 4 without a conditioned contributor. By conditioning on a known contributor, the number of contributors to be deconvoluted was effectively lowered by one which allowed STRmix™ to require less computing power to calculate the genotype weights of those contributors. Again, this further indicates that adding more information about a sample may aid in the STRmix™ deconvolution for contributors within that sample.

Experiment 6:

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

STRmix™ uses the Markov chain Monte Carlo (MCMC) method to generate weights for different genotype combinations. This sampling procedure will result in slightly varied weights between each run. The variability within this process is taken into account within the LR calculation through the use of the highest posterior density (HPD) method (2, 23-26).

To assess the variability between runs, two samples of variable ratios and template amounts were chosen for each of the 2-, 3-, 4-, and 5-person mixtures and run through STRmix™ 10 times each. LRs were also calculated for a true contributor. These runs were performed with the default MCMC setting of 10,000 burn-in accepts per chain (80,000 burn-in accepts total) and 50,000 post burn-in accepts per chain (400,000 post burn-in accepts total), as STRmix™ will yield results with acceptable diagnostics in most instances at these settings. For some cases in which mixture samples do not yield acceptable diagnostics (high Gelman Rubin, non-intuitive mixture proportions and/or genotype weights) a re-run with an increased number of accepts in the MCMC process may aid in giving the STRmix™ chains more time to converge. To demonstrate this, one 3-, 4-, and 5-person mixture from the above set were run through STRmix™ an additional 10 times with an increased number of accepts - 100,000 burn-in accepts per chain (800,000 burn-in accepts total) and 500,000 post burn-in accepts per chain (4,000,000 post burn-in accepts total). Figures 14 and 15 show the sub-sub-source LRs and the 99% 1-sided HPD LR for each run of the samples.

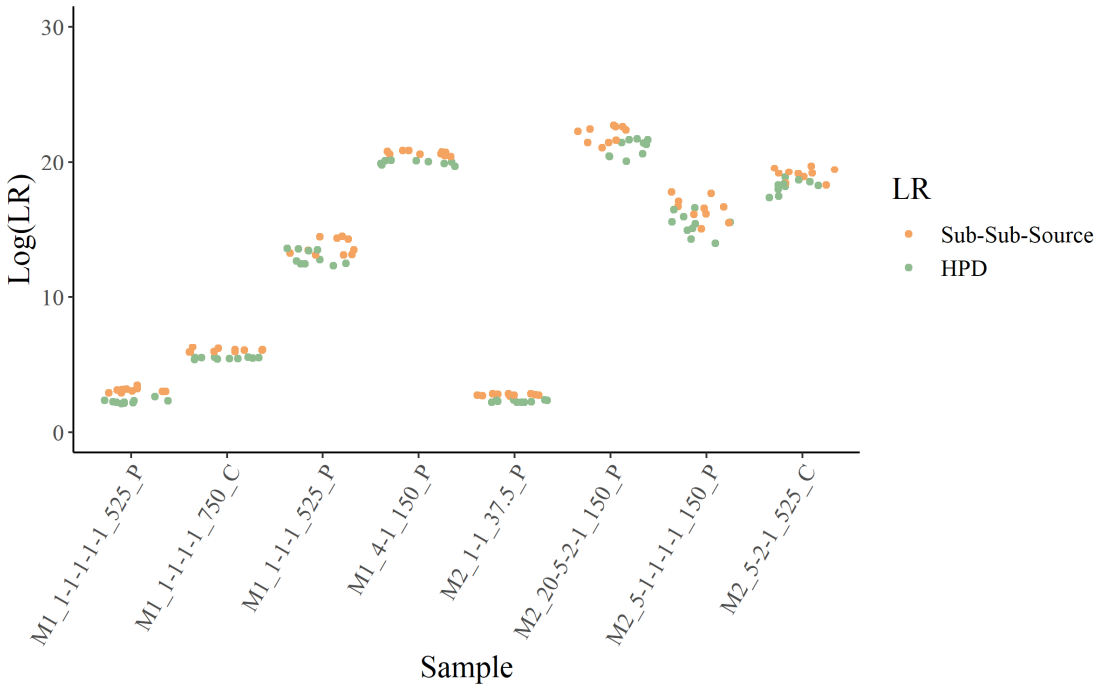


Figure 14. Plot of log(LRs) for 10 replicate interpretation and known comparison runs per sample. Data is jittered to better see the range of LRs.

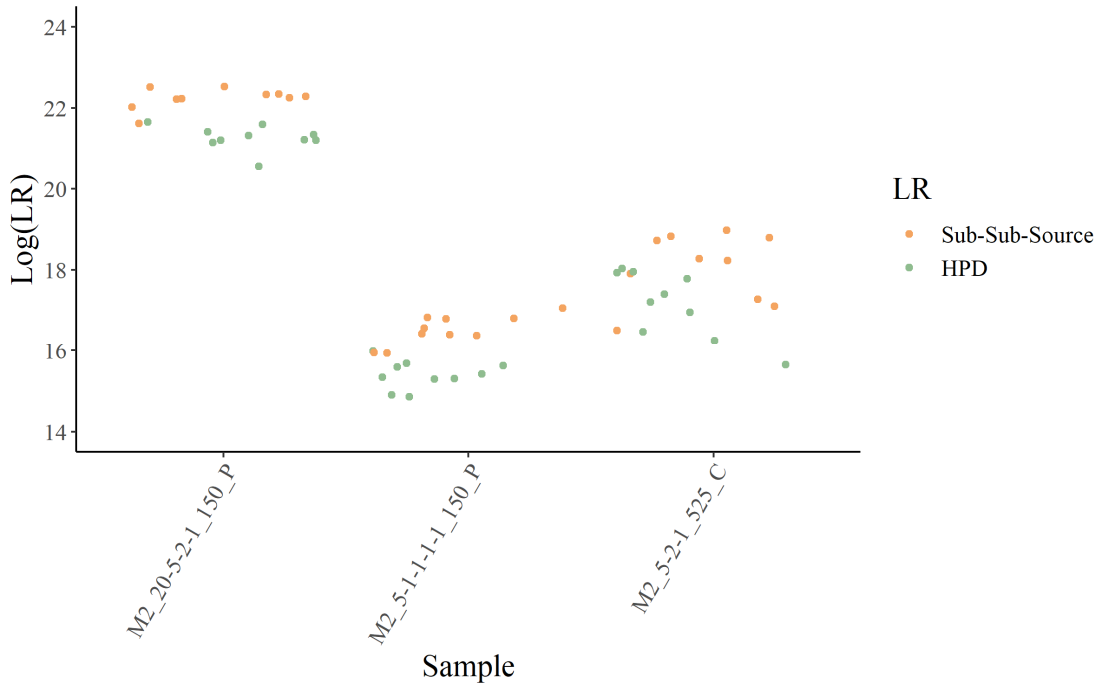


Figure 15. Plot of log(LRs) for 10 replicate interpretation and known comparison runs per sample with additional accepts. Data is jittered to better see the range of LRs.

Figures 14 and 15 demonstrate that, while there is inherent variability within the MCMC process, all sub-sub-source LRs and HPD LRs are within 1-4 orders of magnitude of each other. Additionally, the overall conclusions of support for inclusion, support for exclusion, or within the uninformative range did not change for the HPD LRs based on the NYC OCME's uninformative range of 0.001 – 1000 (experiment 4).

Experiment 7

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

Thirty-three mock evidence samples were assessed through STRmix™ v2.7. The mock evidence samples came from a variety of sources including the following:

- 9 mock sexual assault samples including single source and mixtures created using DNA from one female and/or one male
- 2 old proficiency test samples cut from bloodstain cards
- 8 drinking containers/straw samples taken from donated items from laboratory staff
- 10 touched item samples taken from various areas around the laboratory office space
- 4 NIST standard reference material (SRM) samples

All samples were assigned an apparent NOC according to the same procedure outlined in experiment 4. Each sample was interpreted using STRmix™ and compared to their presumed contributors and the 10,000 non-contributor profile database used in previous experiments. Presumed contributors refers specifically to the mock samples where there is an assumed contributor based on a reasonable expectation that the contributor's DNA would be on the item (i.e., the item belonged to a specific person). Figure 16 is a summary of the LRs for the presumed contributors.

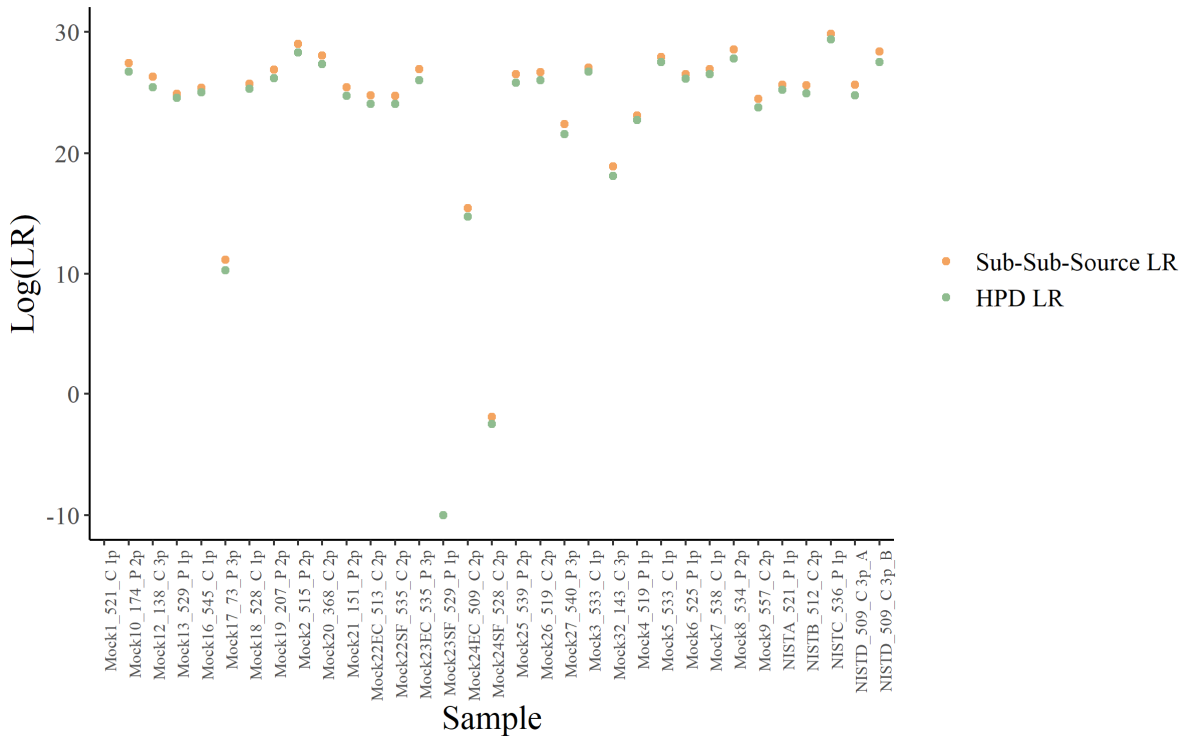


Figure 16. Log(LRs) for 33 mock evidence samples. LR of 0 plotted at Log(LR) = -10.

The comparisons of the presumed contributors (H_p true) were as expected resulting in high LRs for most samples. There were two instances where comparisons using the H_p true presumed contributors did not result in high LRs for two mock sexual assault samples. As a note, the apparent NOC was used for STRmix™; however, the true NOC for the mock sexual assault samples is known as they were prepared in-house.

- Mock23SF_529_P (created with female/male DNA) had an apparent NOC of 1 and a true NOC of 2. The comparison of the true minor contributor to the sample, who had dropped out at nearly every locus, resulted in an LR of 0.
- Mock24SF_528_C (created with male/male DNA) had an apparent and true NOC of 2 with an extreme mixture proportion of 99.7%:0.3% as determined through the STRmix™ interpretation. The comparison of the true minor contributor to the sample, who had dropped out at nearly every locus, resulted in a low LR < 1.

Despite the LRs for these presumed minor contributors, the results are not unexpected as the NOC determinations for both samples and the mixture proportion for Mock24SF_528_C are supported by the data present within the samples. Additionally, these samples were extracted using a differential extraction with the purpose of separating epithelial cells from sperm cells. For sample Mock23SF_529_P the male, major contributor represented most of the alleles observed, while the minor female contributor had dropped out at most loci. Ultimately, these results would have conformed to the analyst interpretation of the sample data.

Figure 17 shows the log(LR) for each presumed known and non-contributor profile comparison in relation to the minor contributor template RFU output as per the STRmix™ interpretation. LRs of 0 are plotted at Log(LR) = -50. Table 6 gives the count of non-contributor profiles where the LR is above a specific value.

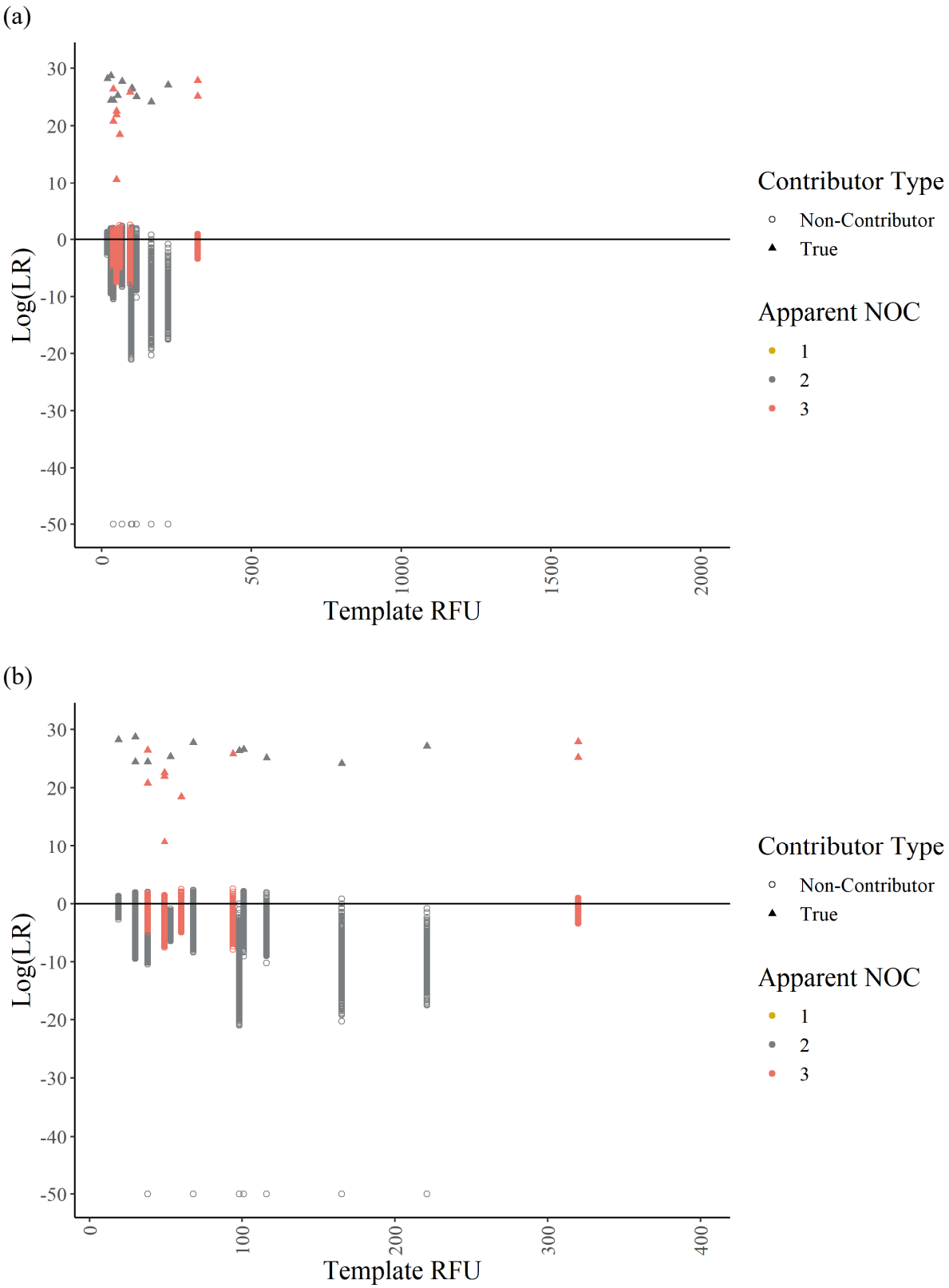


Figure 17. Log(LR) versus template RFU per contributor for mock evidence samples. (a) all data plotted; (b) zoomed in x-axis.

Table 6. Count of non-contributor database LRs above a specific value. *All other samples had 0 LRs>1 and are not listed in this table.

Sample	# of LRs >1	# of LRs >10	# of LRs >100	# of LRs >1000
Mock12_138_C 3p	113	8	2	0
Mock17_73_P 3p	576	15	0	0
Mock19_207_P 2p	100	35	1	0
Mock20_368_C 2p	170	74	5	0
Mock21_151_P 2p	17	5	0	0
Mock22EC_513_C 2p	667	77	0	0
Mock22SF_535_C 2p	773	36	0	0
Mock23EC_535_P 3p	385	20	0	0
Mock24SF_528_C 2p	747	49	0	0
Mock25_539_P 2p	1128	10	0	0
Mock2_515_P 2p	1194	3	0	0
Mock32_143_C 3p	427	30	2	0
Mock8_534_P 2p	807	5	0	0
Mock9_557_C 2p	2	0	0	0
NISTD_509_C 3p	266	0	0	0

The LR results of non-contributor profiles observed in Figure 17 were as expected with most LRs < 1. Table 6 indicates the count of LRs for the non-contributor profiles that were >1, with no LR result for any mock sample resulting in a value above 1000. The highest non contributor profile LR was 365.8 for sample Mock12_138_C 3p. This LR value is attributed to allele sharing of the non-contributor profile and the true contributors by chance. After re-running this sample to apply the HPD LR, the comparison LR for the non-contributor profile lowered to 187.

Ultimately, the LRs calculated in this experiment gave expected results with known contributor comparisons resulting in high LRs that support inclusion and non-contributor profile comparisons resulting in lower LRs with most < 1. Appendix D provides a summary of additional diagnostic values for each mock evidence sample.

Experiment 8:

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.

Nine of the known mixtures amplified at 525pg and deconvoluted by STRmix™ as a part of experiment 4 were blindly interpreted by five analysts of varying levels of experience. The analysts were asked to determine the apparent NOC for each sample and then assign alleles to each contributor. STRmix™ deconvolutions were performed using the previously determined apparent NOC (which correlated with the true NOC for the chosen samples). Based on the deconvolution by STRmix™, alleles or genotypes were assigned for all contributors where the weight for a

particular allele or genotype reached ≥99.0%, and contributor profiles were determined where at least 6 locations could be assigned a full genotype, in accordance with NYC OCME protocols (Table7). The manual deconvolutions were compared to the allele and profile designations using STRmix™.

Table 7. STRmix™ deconvolution results of samples used for manual deconvolution comparison where the determination of a profile is defined as at least 6 locations having full genotypes reaching a weight ≥99.0%.

Sample Info		**Determine Profile**		**Percentage of Mixture from STRmix™ v2.7 Deconvolution**			
NOC	Ratio	Major	Minor(s)	Contributor 1	Contributor 2	Contributor 3	Contributor 4
2	10:1	Yes	No	0.9127	0.0873		
2	4:1	Yes	Yes	0.7133	0.2867		
2	2:1	Yes	Yes	0.6566	0.3434		
2	1:1	No	No	0.5549	0.4451		
3	20:5:1	Yes	No	0.7378	0.2254	0.0367	
3	5:1:1	Yes	No	0.7867	0.1325	0.0808	
3	1:1:1	No	No	0.3837	0.3298	0.2865	
4	10:5:3:1	No	No	0.4818	0.2699	0.1838	0.0644
4	10:1:1:1	Yes	No	0.8246	0.0816	0.0594	0.0344

Figure 18 depicts an overall consistency across analysts and STRmix™. For analysts as well as for STRmix™, the ability to deconvolute contributor profiles decreases as the mixture proportion becomes less extreme; in other words, contributors are less likely to be deconvoluted as the ratio approaches 1:1 (or 1:1:1). Major contributor deconvolutions improved as the contributor percentage increased, and as the minor contributor(s) proportions decreased. Minor contributor deconvolutions were only obtained when the contributors were far enough from equal mixture proportions to distinguish the major and minor contributors from each other, as well as with a limited number of total contributors to a mixture. For more extreme mixture proportions, there was an insufficient amount of data present to get a complete minor deconvolution. Ultimately, the results show that as the ability to deconvolute a mixture increased, the ability to assign contributor alleles/genotypes by the interpreting analysts as well as STRmix™ also increased.

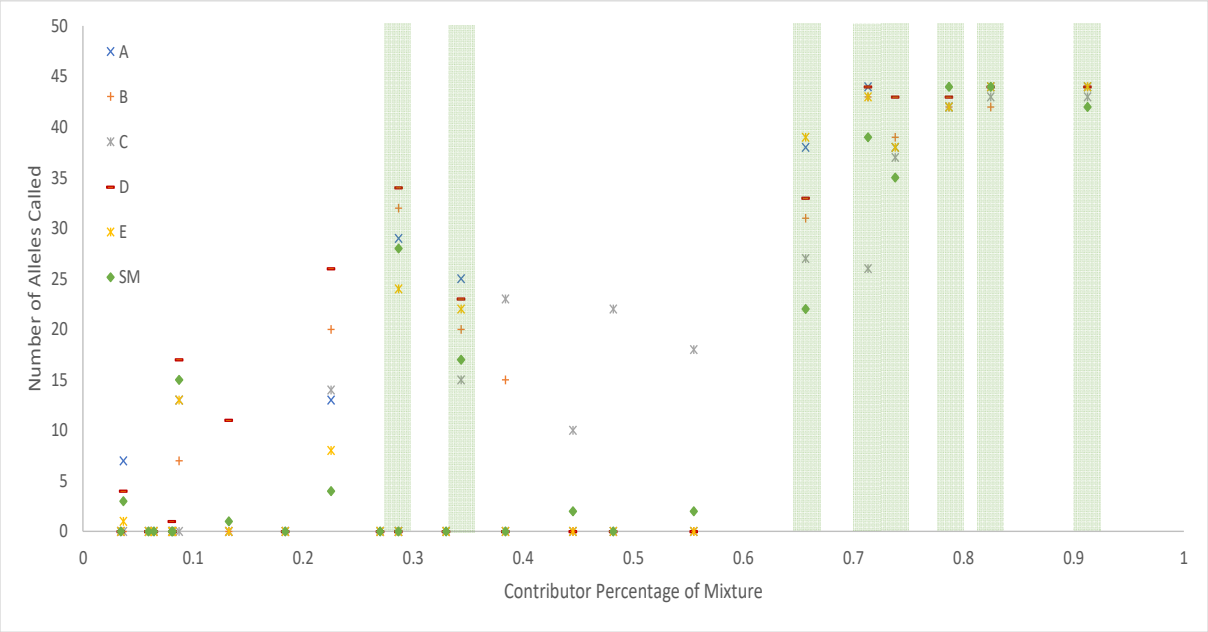


Figure 18. Contributor percentage of the total mixture versus number of alleles called by analysts A-E and STRmix™ v2.7 (SM) for all mixtures. Green bars indicate the contributors for which STRmix™ assigned full genotypes at six or more fully deconvoluted locations.

Experiment 9: Hypotheses Testing Including Different Numbers of Contributors, N, N-1, N+1, N-2, and N+2

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.

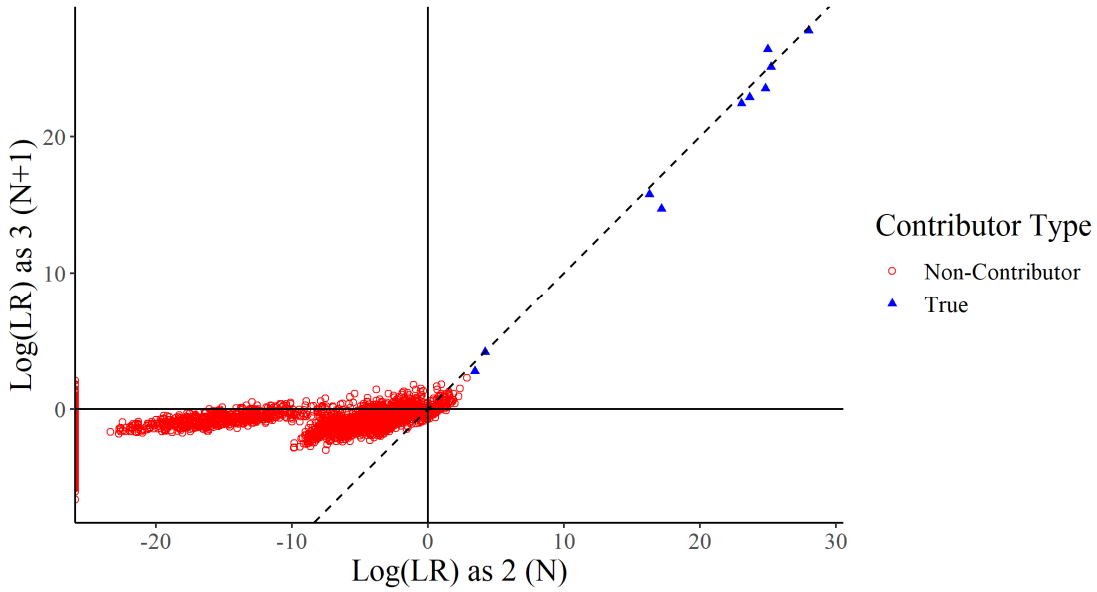
4.1.6.4. If the number of contributors is input by the analyst, both correct and incorrect values (i.e., over- and under-estimating) should be tested.

The true NOC to a crime scene profile is always unknown and the uncertainty in the NOC of a sample has been shown by testing more or fewer contributors (27-30). While analysts go through extensive training to determine the apparent NOC that most accurately represents the data, there are factors that may cause an analyst to add or subtract an additional contributor(s) to the sample. Analysts are more likely to add contributors in the presence of an increased number of artifacts or elevated stutter peaks. With the addition of a contributor(s), the LR of the trace contributor will tend to decrease and the LR of a non-contributor may increase and possibly give false support for inclusion. This is due to the genotype weights distributing more broadly to account for the

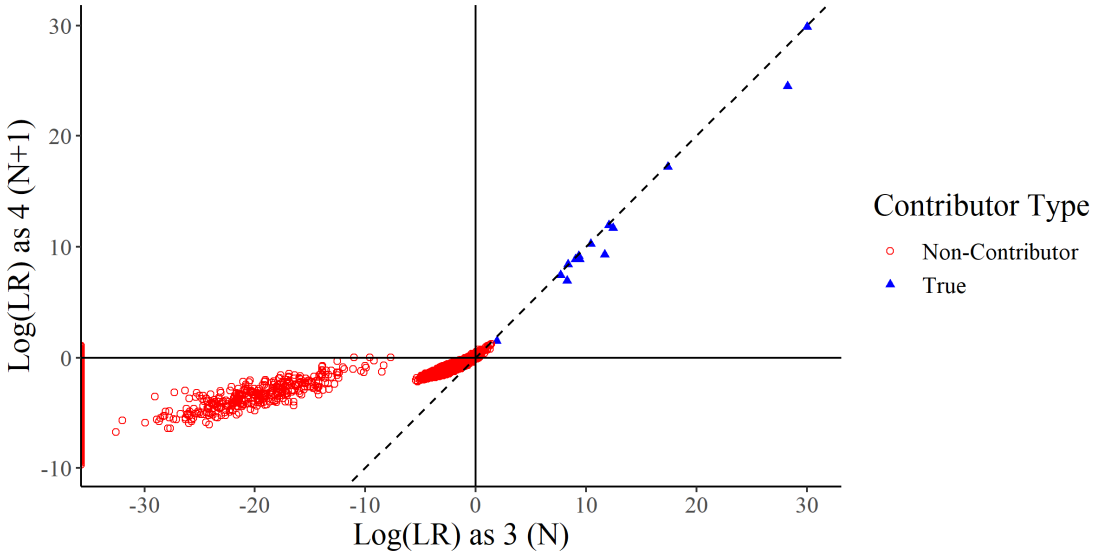
additional contributor(s). Contrarily, analysts may assume fewer contributors to a sample when there are very low, trace contributors with drop-out, samples with similar profiles at the same concentrations, and/or more sharing of alleles (family scenarios). While a major contributor or clear minor contributor may not be affected, this could lead to false exclusions of a true trace contributor (27-30).

A set of 2-, 3-, 4- and 5-person mixtures from experiment 4 were assessed to observe the effect of under- and over- estimation of the NOC on the STRmix™ interpretation. All mixture samples chosen for this experiment had an apparent NOC equal to the true NOC. For samples where additional contributors were added, N+1 or N+2, the samples were interpreted using the higher NOC and compared to their known contributors and 10,000 non-contributor profiles (the same database used in experiment 4). Figures 19 and 20 show the database LRs with an increased NOC as compared to the original database LRs calculated in experiment 4. Note that all database LRs were calculated using theta = 0.03 and the sub-source LR (factor of N!) applied. LRs = 0 are plotted directly on the x or y axis.

(a)



(b)



(c)

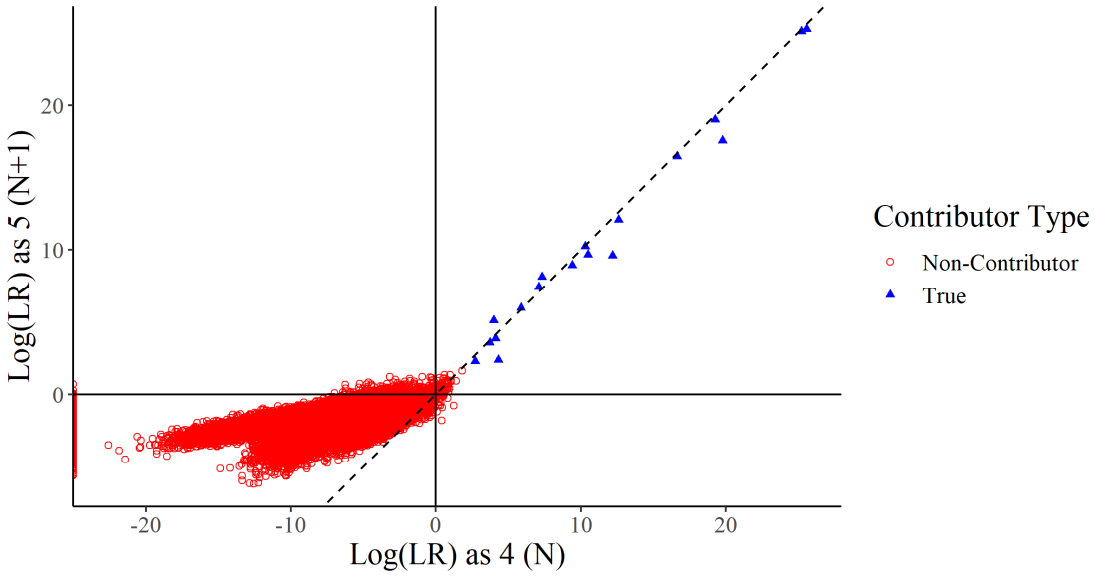
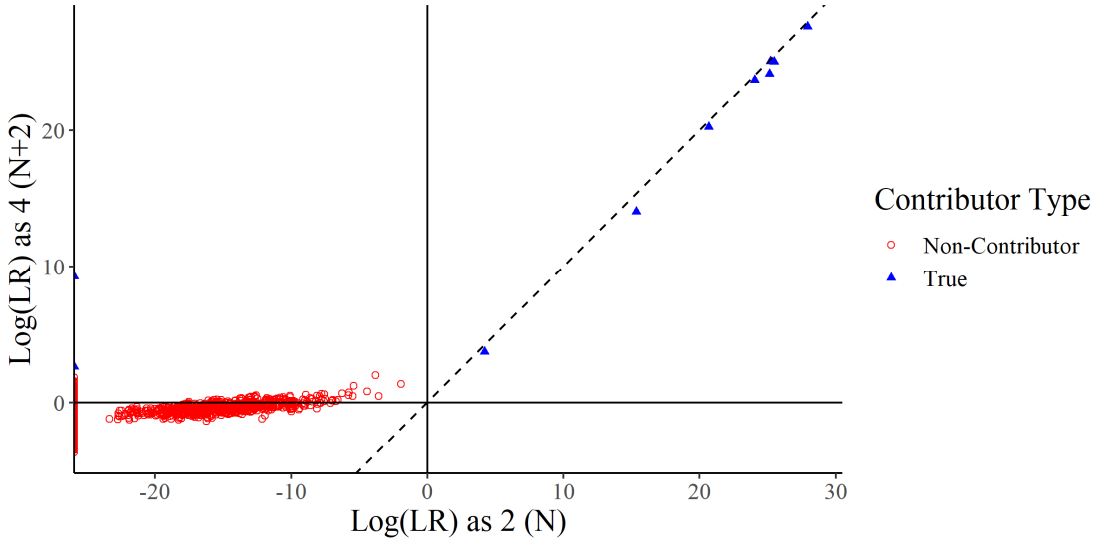


Figure 19. Comparison of log(LR) for N versus log(LR) for N+1 contributor for N = (a) 2p (5 samples), (b) 3p (5 samples), and (c) 4p (5 samples).

(a)



(b)

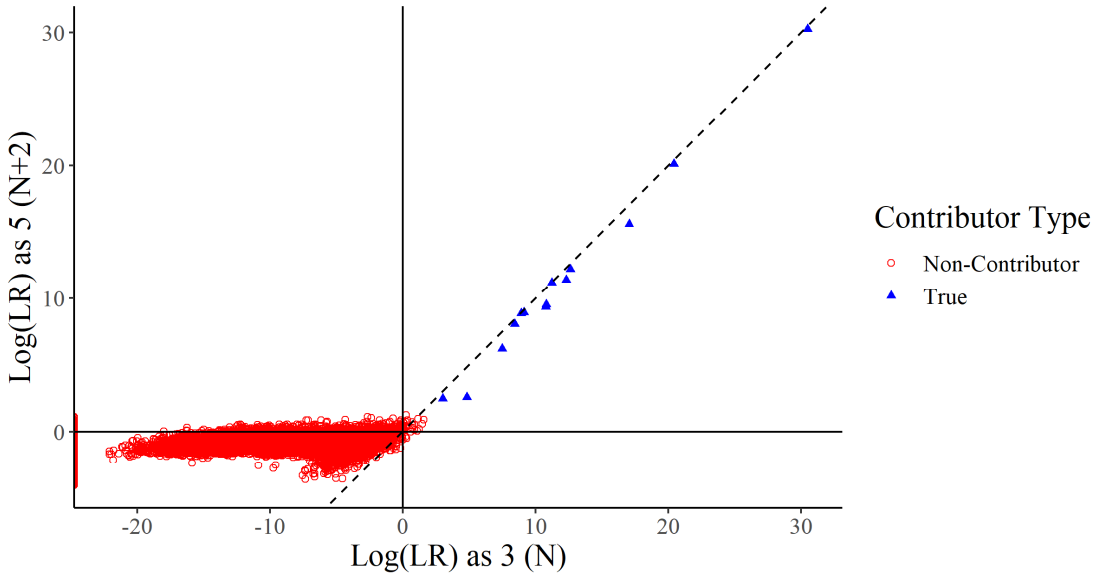
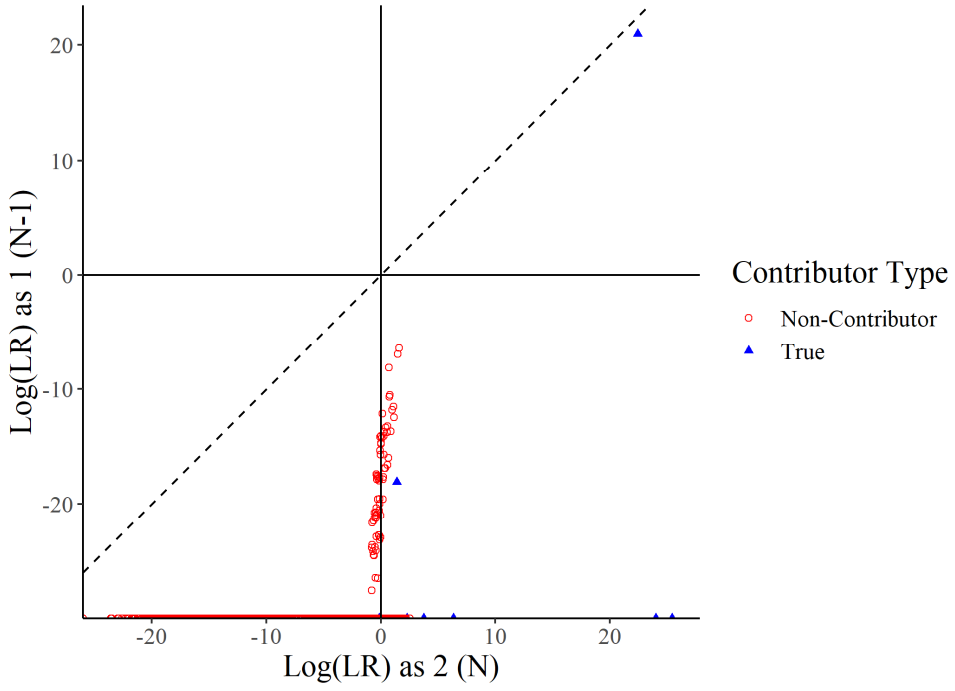


Figure 20. Comparison of log(LR) for N versus log(LR) for N+2 contributors for N = (a) 2p (5 samples) and (b) 3p (5 samples).

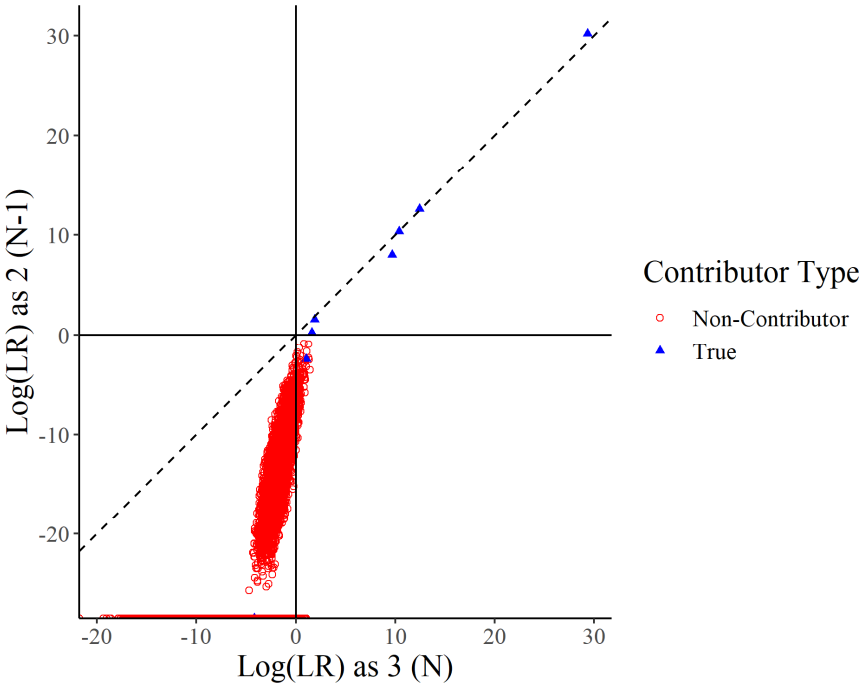
Figures 19 and 20 show that even with an increase of one to two contributors, there is no significant effect on the LR of the true major and minor contributors. There are times where the log(LR) slightly decreased when additional contributors were considered. This is an expected result as the genotypes are forced to consider other options for potential trace contributors. While the LRs for non-contributor profiles generally increased with the overestimation of a contributor(s) for the NOC, there were no high inclusionary LRs. The highest LR for a non-contributor profile was 194.7 for a 2p mixture, M2_100-1_525_C, run as N+1. Note that all 4p mixtures run as N+2 and 5p mixtures run at N+1 were attempted but did not run to completion due to the availability of computing power.

For samples where there was an assumption of fewer contributors, N-1 or N-2, the samples were interpreted using the lower NOC and compared to their known contributors and the 10,000 non-contributor profile database. Figures 21 and 22 show the database LRs with a decreased NOC as compared to the original database LRs calculated in experiment 4. Note that all database LRs were calculated using theta = 0.03 and the sub-source LR (factor of N!) applied. LRs = 0 are plotted directly on the x or y axis.

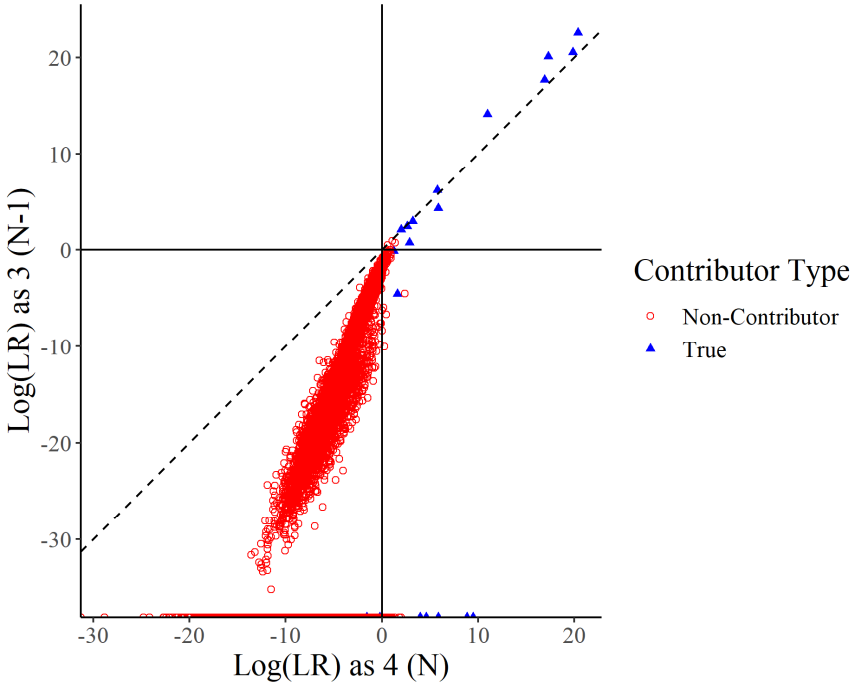
(a)



(b)


Log(LR) as 2 (N-1)
Log(LR) as 3 (N)
30
20
10
0
-10
-20
-20
-10
0
10
20
30
Contributor Type
Non-Contributor
True

(c)


Log(LR) as 3 (N-1)
Log(LR) as 4 (N)
20
10
0
-10
-20
-30
-30
-20
-10
0
10
20
Contributor Type
Non-Contributor
True

(d)

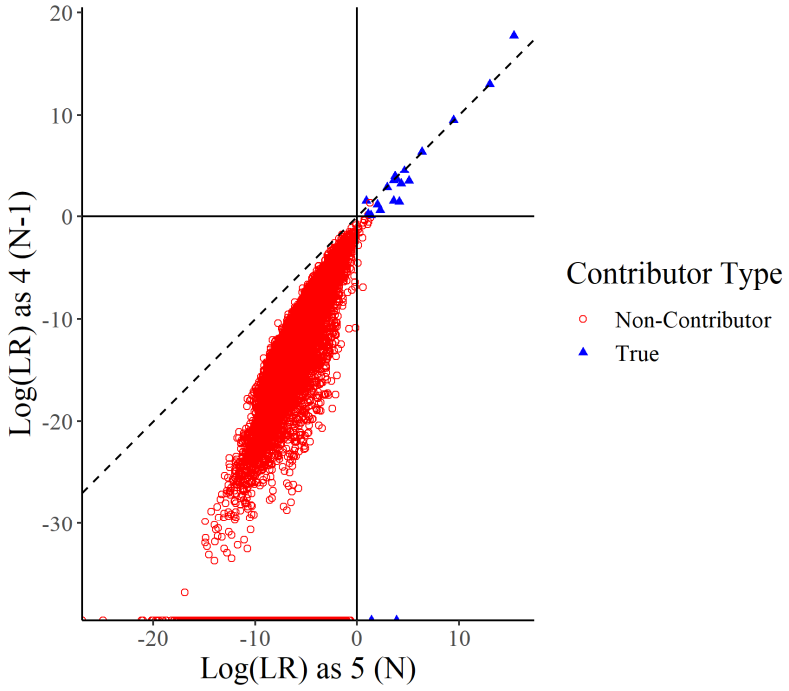
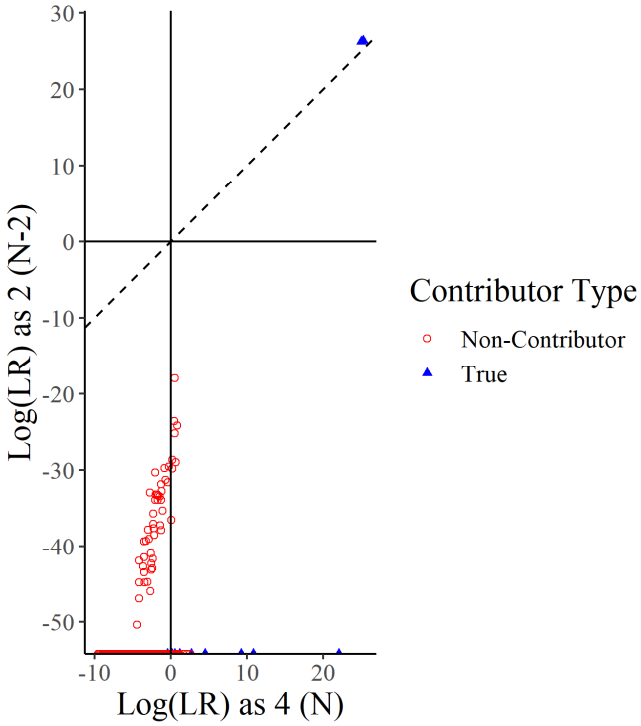


Figure 21. Comparison of log(LR) for N versus log(LR) for N-1 contributors for N = (a) 2p (4 samples), (b) 3p (3 samples), (c) 4p (5 samples), and (d) 5p (4 samples).

(a)



(b)

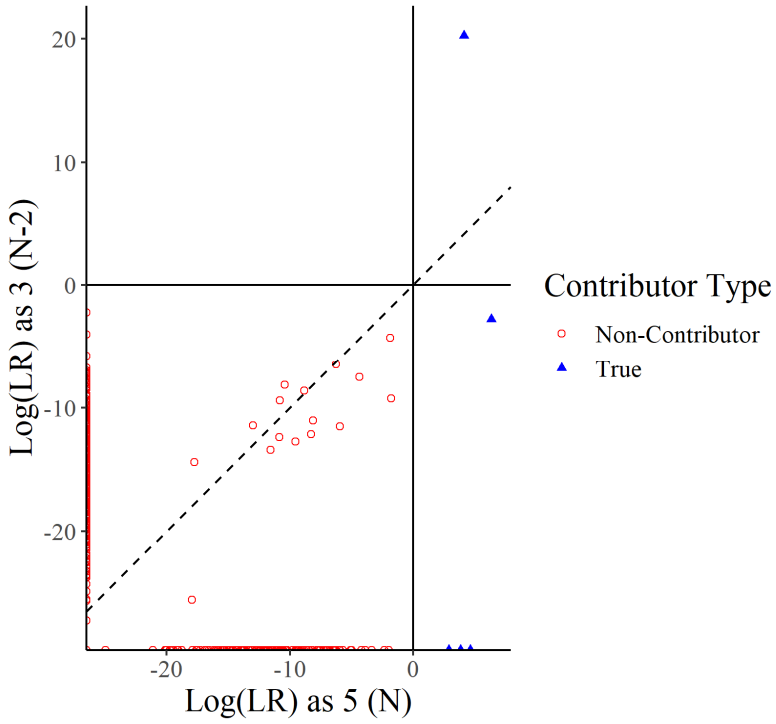


Figure 22. Comparison of log(LR) for N versus log(LR) for N-2 contributors for N = (a) 4p (3 samples) and (b) 5p (2 samples).

Figure 21 and 22 shows that underestimation of the NOC may lead to false exclusions, as noted by true contributors falling below 1 or log(0) when the NOC was underestimated. These true contributors were primarily the trace or minor contributor(s) with few alleles present, and their genotypes were no longer modeled as possible genotypes of the profile by STRmix™. The major contributors were typically not affected by underestimation of the NOC; this was increasingly true as that contributor attributed more DNA to the sample (i.e. higher-level contributors within a mixture and/or high template sample). While there was some variability in the non-contributor profile LRs across the figures, they generally decreased with the subtraction of a contributor(s) with no high inclusionary LRs.

Generally, as shown by comparison of Figures 19 and 20 to Figures 21 and 22, LRs for non-contributor profiles did increase when overestimating the NOC and decreased when underestimating the NOC. Additionally, while the results are not incorporated in the above figures, there were some 5p samples that were run as N-1 that went to completion within STRmix™ that did not complete their interpretation when initially run in experiment 4. In these instances, the LR of the major contributor was a high inclusionary LR.

As a note, there are instances in which some loci may not be accurately modeled by the NOC that is input into STRmix™. In these situations, STRmix™ is not able to describe additional peaks through stutter modeling and/or drop-in and can only explain the peaks as attributing to an additional contributor. Samples such as these will produce an error message by STRmix™ that outlines the specific location in which the profile cannot be described by the determined NOC parameter that was input into the software.

Experiment 10: The Effect of Allelic Drop-in on LR Calculations, Stutter, and Peak Height Variance

4.1.8. Allele drop-in.

In a series of experiments, drop-in peaks were artificially added *in silico* to the corresponding text files of single source samples to test the effect of allele drop-in. Results of these experiments were compared to those in experiment 1 (and experiment 4 for 10c) for the same sample and comparison LRs calculations. Observed drop-in rates for the OCME have been modelled and the appropriate parameters are incorporated within STRmix™ (13). The OCME drop-in parameters for STRmix™ 2.7 for Fusion data are as follows:

Drop-in cap	300 RFU
Drop-in frequency	0.0087
Drop-in parameters	22.31, 2.65

Experiment 10a:

For the first experiment, a drop-in peak below the drop-in cap was added *in silico* to the following high template and low template samples:

- SS_525_C_21M
    - Drop-in at D13S317 (homozygous locus), allele 15 at 200 RFU
- G07_25M_r_50pg
    - Drop-in at D22S1045 (heterozygote locus), allele 15 at 200 RFU

Both samples were interpreted with an apparent NOC of 1. As expected, the drop-in peak affected the interpretation of the low template sample and not that of the high template sample.

For the high template sample, STRmix™ modeled the artificial peak only as drop-in since it did not reasonably pair well with the true allele at a height of 11,838 RFU. The artificial peak was not given any weight as part of the genotypes in the profile, and the true genotypes of the profile were given a weight of 1 (100%). This weighting is the same as that of the interpretation of the profile without the artificial drop-in and the resulting LR for the comparison sample was the same as the previous LR calculation performed in experiment 1, with both resulting in a sub-source LR of 4.2131E+28.

For the low template sample, STRmix™ modeled the artificial peak as both drop-in and a true allelic peak. This is not unexpected as the true heterozygous peaks were below the stochastic threshold and the drop-in cap (< 300 RFU); the height of the artificial drop-in peak was similar in height to the true alleles. The resulting LR was slightly lower than the initial run of the profile in experiment 1, with a sub-source LR of 3.1577E+28 and 6.9963E+28, respectively.

Experiment 10b:

For the second experiment, an artificial drop-in peak was added just above the drop-in cap *in silico* to one high template sample and one low template sample. Each sample was interpreted using STRmix™ twice, one with the drop-in peak added at a homozygous locus and one with the drop-in peak added at a heterozygous locus. The following samples were used:

- SS_525_C_21M
    - Drop-in at D13S317 (homozygous locus), allele 15 at 400 RFU
    - Drop-in at D8S1179 (heterozygous locus), allele 9 at 400 RFU
- SS_37.5_P_25M
    - Drop-in at D10S1248 (homozygous locus), allele 18 at 310 RFU
    - Drop-in at Penta E (heterozygous locus), allele 7 at 310 RFU

As expected for the heterozygous locus of the high template single source sample, the interpretation did not proceed as the profile could no longer be explained by a NOC of 1 for this location. Figure 23 is the error message that populated in STRmix™.

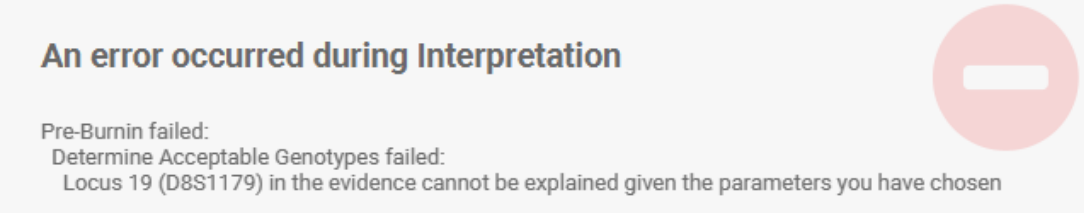


Figure 23. STRmix™ v2.7 error message when evidence cannot be explained by the given parameters for locus D8S1179.

For the homozygous locus of the high template single source sample, the artificial drop-in peak above the cap could also only be explained and modeled by STRmix™ as a true allele, despite the peak height differences. Because the drop-in peak was not in a stutter position, meaning it would not be modeled as any type of stutter, and was above the determined drop-in cap, STRmix™ could not consider this as anything other than a true allele to the profile. Subsequently, the resulting LR for the comparison was exclusionary (LR = 0). For both the homozygous locus and heterozygous locus of the low template single source sample, the artificial drop-in peak above the cap could only be explained and modeled by STRmix™ as a true allele. The resulting LRs for the comparison sample for these runs were exclusionary (LR = 0). This is not unexpected as the drop-in peaks added to this sample were close to the peak heights of the true alleles in the profile. In instances such as this, other diagnostics may be looked at to determine if the STRmix™ interpretation and comparison is nonintuitive to the data observed. This may include counter-intuitive genotype weights given to the major contributor and/or assessing the Gelman Rubin and average log(LR) to determine if the NOC needs to be adjusted.

Experiment 10c:

For the third experiment, an artificial drop-in peak below the drop-in cap was added to the following 2-person mixture with an extreme mixture ratio of 100:1:

- M1_100-1_525_P 2p
  - Drop-in at TH01, allele 6 at 150 RFU

This drop-in peak was deliberately added to potentially pair with the alleles “7, 8” of the minor contributor. This locus also had an 8.3 stutter peak that was of similar height to the minor contributor, which could also be considered as allelic and assigned to a genotype within STRmix™.

As expected, STRmix™ considered the additional peak as both a drop-in and allelic peak; as observed in Table 8, the drop-in peak was given weight within various genotype combinations for

the minor contributor. The drop-in peak had no effect on the genotype weight for the major contributor where "9.3,9.3" was given 100%. Additionally, the sub-source LRs for both the true major and true minor contributors were similar going from 2.99E+25 and 1.73E+03 without the drop-in peak to 3.20E+25 and 1.17E+03, respectively, with the drop-in peak.

Table 8. Genotype combinations and weights for minor contributor before and after the addition of a drop-in peak to a 2p mixture with an extreme mixture ratio. *True genotype of contributor. **Drop-in peak.

Genotype Combinations		**Weights Given for Minor Contributor**	
Major *(9.3,9.3)	**Minor *(7,8)**	**Before drop-in**	**After drop-in **(6)**
9.3	7, 8	0.9831	0.4270
9.3	6, 8	-	0.2901
9.3	6, 7	-	0.2685
9.3	7, 8.3	0.0029	0.0025
9.3	8, 9.3	0.0033	0.0019
9.3	8.3, 9.3	0.0000	0.0000
9.3	6, 8.3	-	0.0018
9.3	7, 9.3	0.0030	0.0017
9.3	8, 8.3	0.0020	0.0017
9.3	6, 9.3	-	0.0013
9.3	8, 8	0.0025	0.0012
9.3	7, 7	0.0016	0.0009
9.3	Q, 6	-	0.0007
9.3	6, 6	-	0.0002
9.3	Q, 7	0.0010	0.0002
9.3	Q, 8	0.0005	0.0002
9.3	9.3, 9.3	0.0000	-

4.1.9. Forward and reverse stutter.

The single source and mixture samples used throughout the experiments in this validation exhibited varying amounts of stutter for each type of stutter modeled at the designated loci (14). It has been shown throughout the validation experiments that STRmix™ can analyze data that exhibits varying amounts of stutter and ultimately return expected/reasonable results.

4.1.10 Intra-locus peak height variance.

The single source samples and mixtures used throughout this validation exhibited varying degrees of intra-locus peak height variance. Notably, Appendix C (v) presents the allele variances observed for the mixtures in experiment 4 and experiment 13 and highlights how the LSAE and APH calculated by STRmix™ are affected when there is sample degradation or inhibition. It has been shown throughout the validation experiments that STRmix™ can analyze data that exhibit intra-locus peak height variance and ultimately return expected/reasonable results.

4.1.11 Inter-locus peak height variance.

Inter-locus peak height variance is routinely encountered in forensic casework samples. Experience has found that inter-locus peak height variance increases in samples that have less than the optimal DNA target amount. The samples utilized in previous experiments contain varying amounts of inter-locus peak height variance. As expected, STRmix™ has properly deconvoluted the samples and returned appropriate LR calculations for samples at both high and low template amounts of DNA.

4.1.12 For probabilistic genotyping systems that require in-house parameters to be established, the internal validation tests should be performed using the same parameters. The data set used to establish the parameters should be different from the data set used to validate the software using those parameters.

STRmix™ requires several parameters to be established in-house before implementation on validation and casework samples. All parameters were calculated and established prior to running the internal validation studies using separate samples (11, 13-14).

Experiment 11: Additional Challenge Testing

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

The input file for a STRmix™ interpretation must be a text file generated from the analyzed data to retain only the allelic values attributed to potential DNA contributors and/or stutter peaks. It is important to remove non-numeric and non-allelic artifacts (OL - "off-ladder"/OB - "off-bin"). Failure to remove such artifacts may cause STRmix™ to fail to initiate the interpretation or yield inaccurate results. For example, STRmix™ cannot model tri-allelic loci. Additionally, the inclusion of a numeric artifact (such as a pull-up peak in an allelic position) that is of similar height to true allelic peaks may cause a false exclusion. Ultimately, these errors are prevented with appropriate review of the input file and can be identified by analyst review of STRmix™ interpretation results.

Experiment 11a:

For the first portion of this experiment, a half back stutter peak was changed from an allelic call to outside the allelic bin (i.e., an OB peak) within the text input file of a single source profile, sample SS_525_P_25M. This peak would have been manually assigned an allelic number (13.2) in accordance with the electropherogram of the corresponding allelic ladder and the parent peaks within the sample. Table 9 shows the input file for locus D1S1656, which includes the OB peak. Also shown is a screenshot, Figure 24, of the automatically populated message STRmix™ presents when attempting to deconvolute an input file that includes an OB peak. In short, STRmix™ could not initiate the run.

Table 9. Locus D1S1656 of STRmix™ input file with OB added.

Locus	Allele	Height	Size
D1S1656	13	708	172.4
D1S1656	**OB**	183	174.4
D1S1656	14	8574	176.5
D1S1656	15	740	180.5
D1S1656	15.2	125	182.5
D1S1656	16	6935	184.6

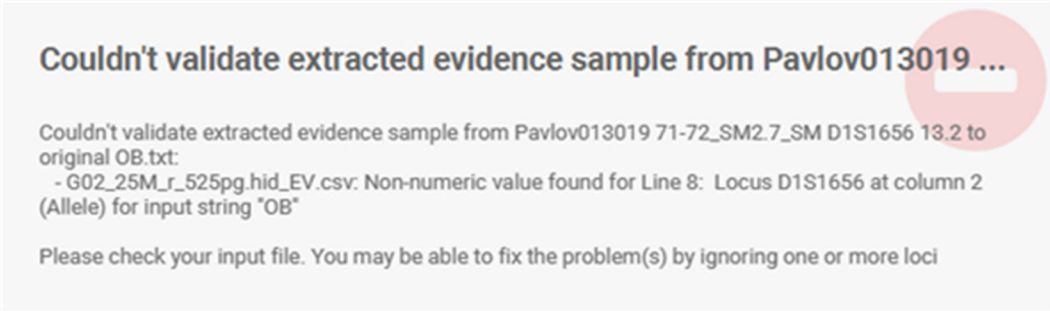


Figure 24. STRmix™ v2.7 error message when an OB peak is left in the input file.

Experiment 11b:

The second portion of this experiment involved further editing of the same single source input file to demonstrate several results that may occur with the presence of a non-allelic peak(s) or artifact(s) that have been sized within an allelic bin position. These peaks can cause an exclusionary LR, have no effect on the LR, or cause a failure to interpret.

An exclusionary LR

If the artifact is modeled as having originated from the person of interest (for example, if the peak is of a similar height to the alleles corresponding to the person of interest in a mixed DNA profile), this may result in an exclusion of a true contributor.

An allelic peak, 23, was artificially added at locus FGA to model a spike that had a similar height to the allele attributed to the true contributor "22, 22". As seen in Figure 25, this resulted in a genotype assignment including the artificially added spike peak of "22, 23", which was assigned a weight of 100%. The comparison resulted in an LR value of 0, and thus, a false exclusion of the true contributor.

Figure 25. Locus FGA genotype and genotype weight from STRmix™ v2.7 with an artificially added allele of 23 at similar height to the true profile alleles.

No effect

If drop-in is observed within the profile, the allele may be modeled as such if it is less than the drop-in cap of 300 RFU. An allelic peak, 9, was artificially added at locus TPOX below the drop-in cap that resulted in no effect on the deconvolution of the person of interest "8,8" for the locus and the subsequent likelihood ratio calculation. Table 10 and Figure 26 shows the information for the three peaks from the input file and genotype weight for TPOX. STRmix™ gave a genotype assignment excluding the artificially added peak below the drop-in cap. The genotype "8,8" was assigned a weight of 100% and the subsequent LR for the true contributor was unchanged.

Table 10. Locus TPOX of STRmix™ input file with an artificially added allele of 9 within drop-in parameters.

Locus	Allele	Height	Size
TPOX	7	264	406
TPOX	8	15408	410
TPOX	9	150	414

Figure 26. Locus TPOX STRmix™ v2.7 genotype and genotype weight with an artificially added allele of 9 within drop-in parameters.

Failure to interpret

If an artifact within an allelic bin is retained in a profile it may artificially increase the minimum number of contributors required to be interpreted in STRmix™. For example, an artifact at a heterozygous locus in a single source profile (not able to be modeled as stutter or drop-in) will increase the minimum number of contributors to two. STRmix™ will not proceed assuming only one contributor. In a case where there is an apparent tri-allelic location within a single source sample, that specific locus must be ignored during STRmix™ evaluation allowing the run to proceed with the remaining locations.

To assess this, a peak in allele bin 8 at 310 RFU was artificially added at locus TH01 to simulate an artifact above the drop-in cap. Table 11 shows the edited STRmix™ input file. Additionally, Figure 27 shows the automatically generated message indicating the failure to run the deconvolution in STRmix™.

Table 11. Locus TH01 of STRmix™ input file with artifact in allelic bin.

Locus	Allele	Height	Size
TH01	5	221	73.1
TH01	6	12019	77.5
TH01	7	13905	81.8
TH01	8	310	86.2

Figure 27. STRmix™ v2.7 error message when evidence cannot be explained by the given parameters for locus TH01.

Experiment 12: Comparisons of Manual Interpretation Guidelines with STRmix™ Analyses

4.2 Laboratories with existing interpretation procedures should compare the results of probabilistic genotyping and manual interpretation of the same data, notwithstanding the fact that probabilistic genotyping 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.

Qualitative comparisons were made to five known mixtures from experiment 4 and five mock samples from experiment 7. Analysts were given sample electropherograms with stutter filters on and off as well as a potential contributor profile for comparison. They were asked to determine a best description of the number of contributors for the sample and then indicate if the qualitative comparison of the potential contributor could be described as supporting inclusion, supporting exclusion, or uninformative (with additional options of uninformative/inclusionary and uninformative/exclusionary) based on the uninformative range defined in experiment 4. The samples were interpreted within STRmix™, LRs were calculated, and the results were compared to the manual comparisons made by the analysts.

Table 12. Summary of qualitative comparisons drawn by analysts and by sample (Uninformative is listed as Uninf, Inclusion is listed as Incl, and Exclusion is listed as Excl).

Sample Information	**Sample #**	4	6	9	5	7	3	10	2	8	1
	Amount (pg)	174	533	525	37.5	75	143	150	525	535	368
	True NOC	mock	mock	2	3	5	mock	4	4	2	mock
	Apparent NOC	2	1	2	3	4	3	4	4	3	2
	Ratio if available	-	-	1:50	10:5:1	10:5:1:1:1	-	10:5:1:1	1:3:3:1	-	-
	Comparison Position	Not a known	Not a known	Not a known	Not a known	Not a known	Not a known	A trace	A minor	Minor	Major
	Qualitative Comparison	Excl	Excl	Excl	Excl	Uninf	Uninf	Uninf /Incl	Incl	Incl	Incl
STRmix™ Results	**LR**	0.00 E+00	0.00 E+00	0.00 E+00	0.00 E+00	3.20 E-02	3.30 E-02	8.97 E+02	2.05 E+03	1.43 E+13	4.70 E+15
	Conclusion	Excl	Excl	Excl	Excl	Uninf	Uninf	Uninf	Incl	Incl	Incl
Visual Comparisons by Participant	**A**	Excl	Excl	Excl	Uninf/ Excl	Incl	Uninf/ Incl	Incl	Incl	Incl	Incl
	B	Excl	Excl	Excl	Excl	Uninf/ Incl	Excl	Incl	Incl	Incl	Incl
	C	Excl	Excl	Excl	Uninf/ Excl	Uninf/ Incl	Excl	Uninf/ Incl	Incl	Incl	Incl
	D	Excl	Excl	Excl	Uninf/ Excl	Incl	Uninf/ Incl	Incl	Incl	Incl	Incl
	E	Excl	Excl	Uninf/ Excl	Excl	Incl	Uninf/ Excl	Incl	Incl	Incl	Incl

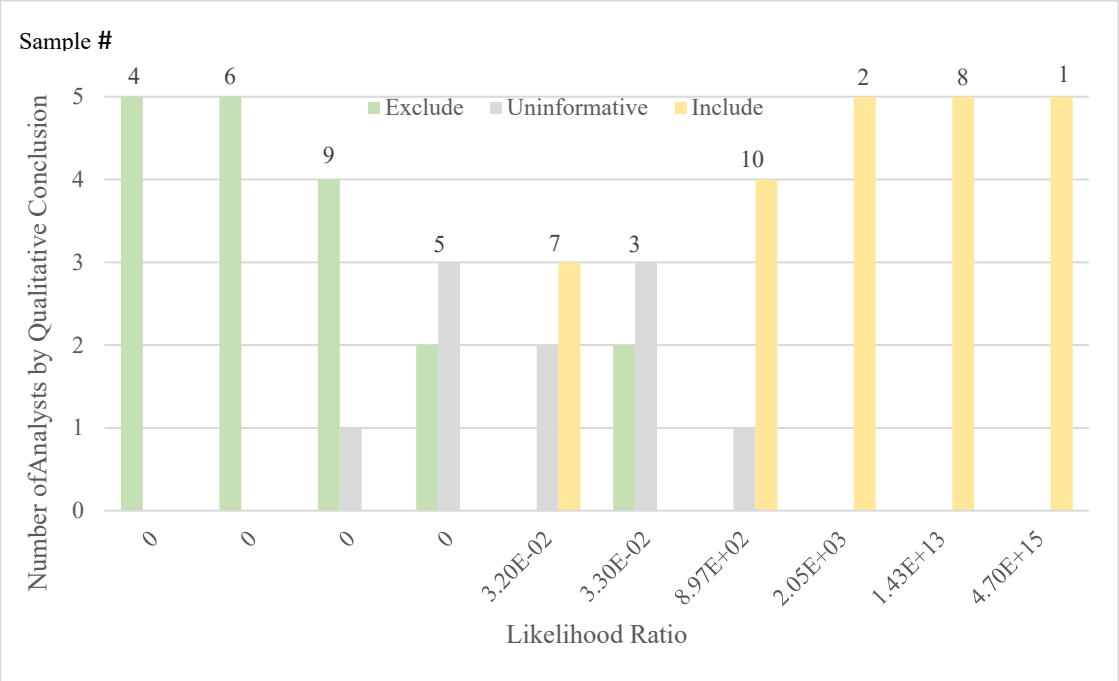


Figure 28. Number of analysts by qualitative conclusion versus the likelihood ratio calculated by STRmix™ for each sample. Sample order is consistent with Table 12. The uninformative conclusion encompasses uninformative, uninformative/inclusion, and uninformative/exclusion.

The results displayed in Figure 28 and Table 12 show that the qualitative conclusions drawn by analysts are consistent with the LRs calculated using STRmix™. For the purpose of this experiment, LRs between 0.001 and 1,000 were considered uninformative (as determined in experiment 4). The uninformative conclusions drawn by analysts show further support for using an uninformative range as it captures the challenge of comparisons that result in LRs close to 1. Despite small variations with uninformative/inclusion and uninformative/exclusion interpretations, the overall conclusions were in line with STRmix™.

Experiment 13: Partial Profile Testing

4.1.7. Partial profiles to include the following:

4.1.7.1. Allele and locus drop-out

4.1.7.2. DNA degradation

4.1.7.3. Inhibition.

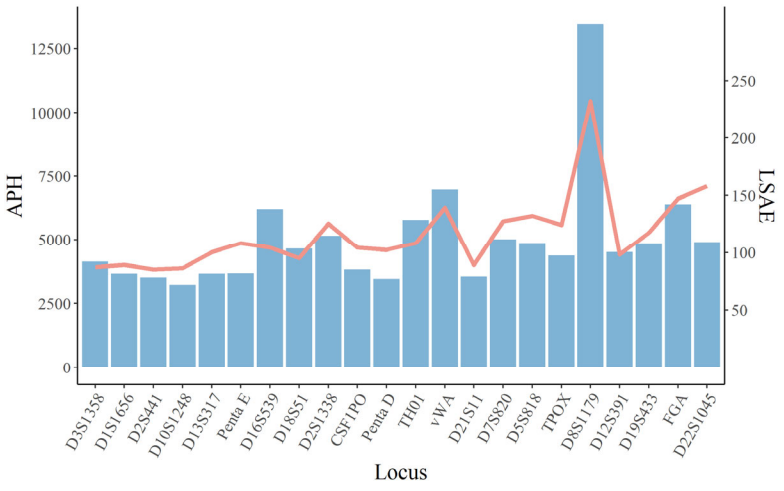
STRmix™ models inter-locus peak height variance using locus specific amplification efficiencies (LSAE). The LSAE is optimized during interpretation and is used as a diagnostic tool. Along with

the APH at the various loci, the LSAE can be helpful in the analysis of samples where there is degradation and inhibition. To demonstrate how allele/locus dropout, degradation, and inhibition within samples is interpreted through STRmix™, a series of experiments were performed.

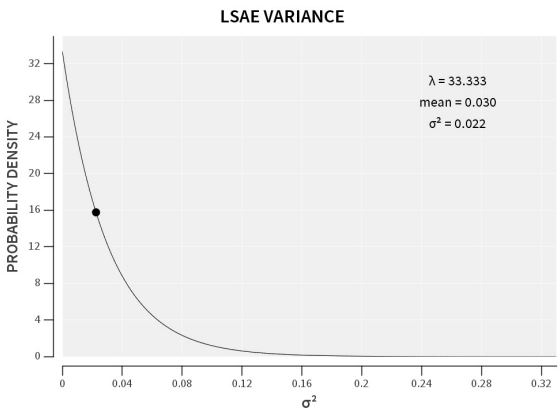
Experiment 13a:

In the first experiment, two single source samples at different amplification amounts were artificially degraded by editing the STRmix™ input text file. Figures 29 and 30 show the APH and LSAE per location, the LSAE variance graph output by STRmix™, and the degradation curve graph output by STRmix™ for the samples without any editing (originally run in experiment 1). Figures 31 and 32 show the same results for the artificially degraded samples.

(a)



(b)



(c)

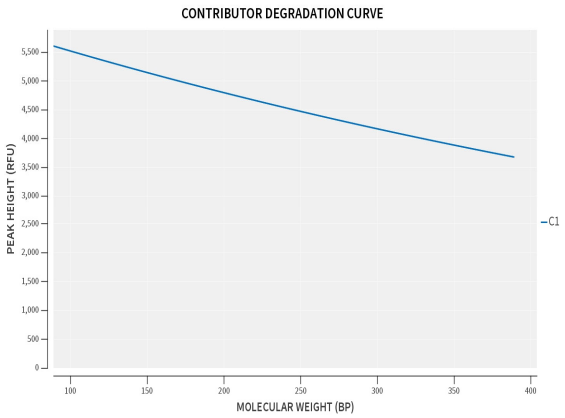
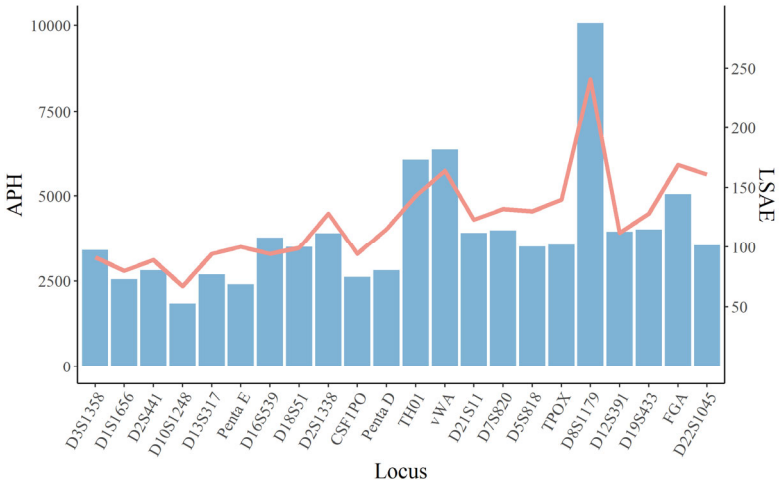
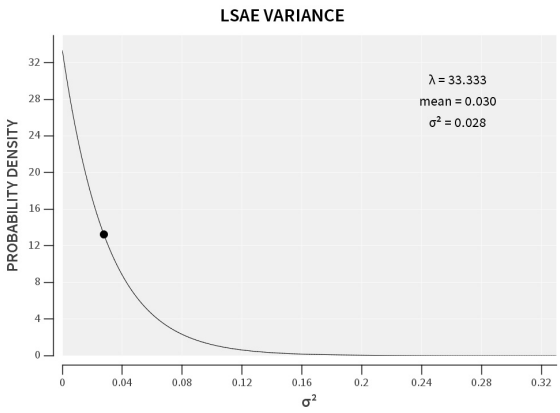


Figure 29. Plot of (a) APH (blue bars) and LSAE (orange line) per locus, (b) LSAE variance curve, and (c) degradation curve for a single source sample amped at 125 pg (Ex13_SS_125_C_12M).

(a)



(b)



(c)

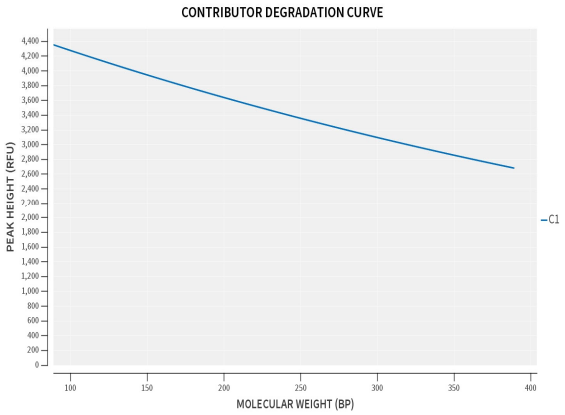
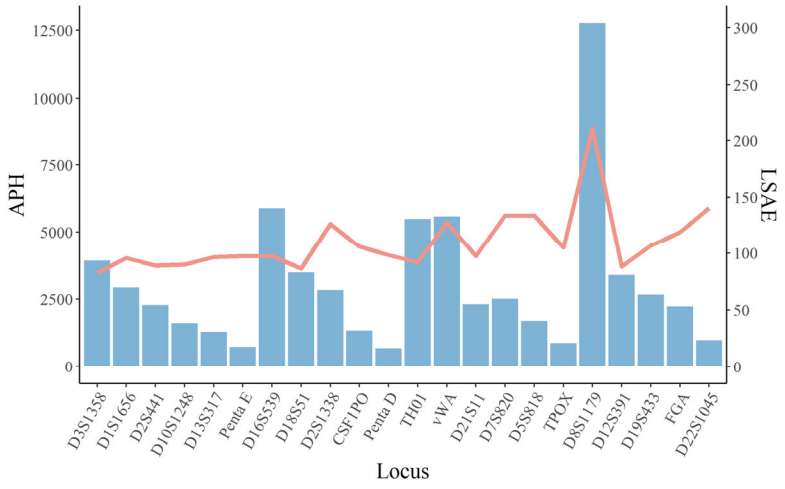
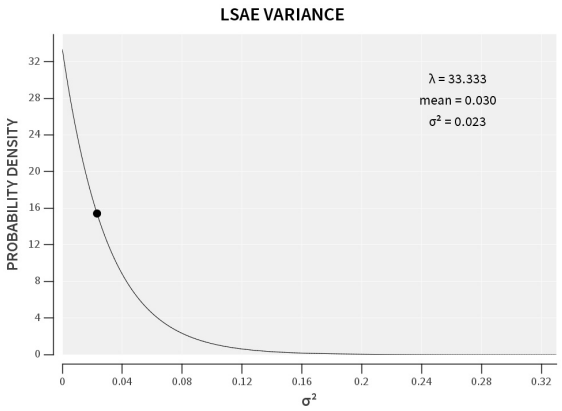


Figure 30. Plot of (a) APH (blue bars) and LSAE (orange line) per locus, (b) LSAE variance curve, and (c) degradation curve for a single source sample amped at 250 pg (Ex13_SS_250_P_12M).

(a)



(b)



(c)

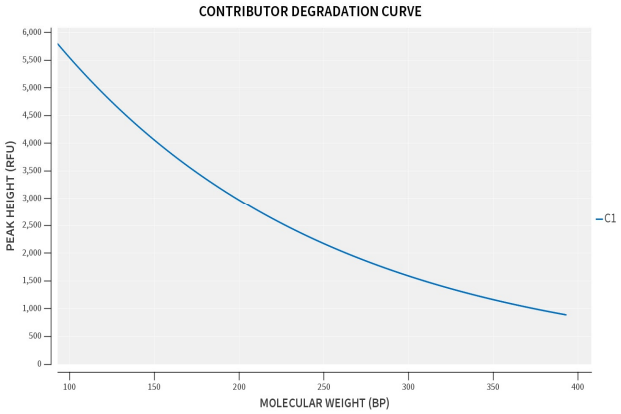


Figure 31. Plot of (a) APH (blue bars) and LSAE (orange line) per locus, (b) LSAE variance curve, and (c) degradation curve for an artificially degraded single source sample amped at 125 pg (Ex13_SS_125_C_12M).

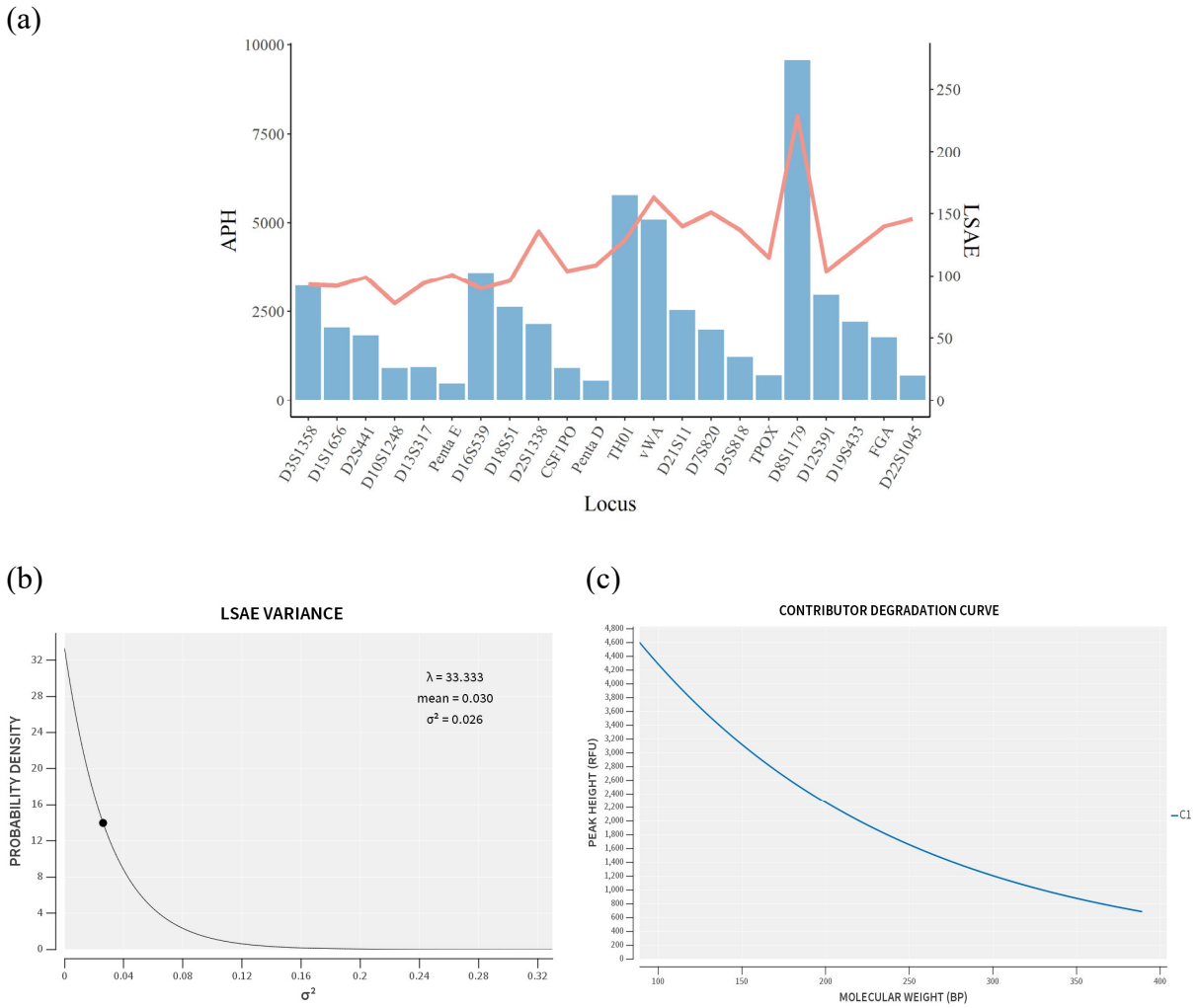


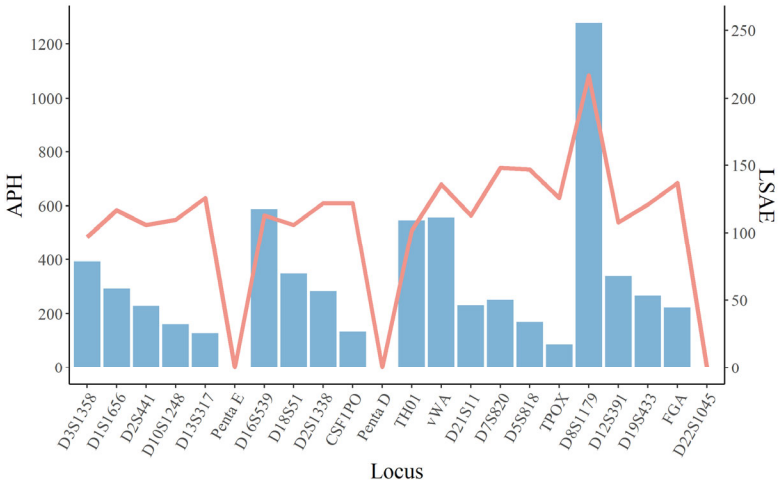
Figure 32. Plot of (a) APH (blue bars) and LSAE (orange line) per locus, (b) LSAE variance curve, and (c) degradation curve for an artificially degraded single source sample amped at 250 pg (Ex13_SS_250_P_12M).

In comparison to the original samples without artificial degradation (from experiment 1), the two single source samples have approximately the same LSAE values overall and the degradation of the artificially degraded samples are being modeled by STRmix™ as expected. The average of the accepted post burn-in values for the degradation went from 1.41E-03 (125 pg) and 1.61E-03 (250 pg) to 6.24E-03 (125 pg) and 6.35E-03 (250 pg). The degradation can be visualized in the degradation curve for each sample. All genotypes were as expected for both samples with a weighting ≥ 99% and sub-sub-source LRs were within the same order of magnitude as the unedited sample sub-sub-source LRs.

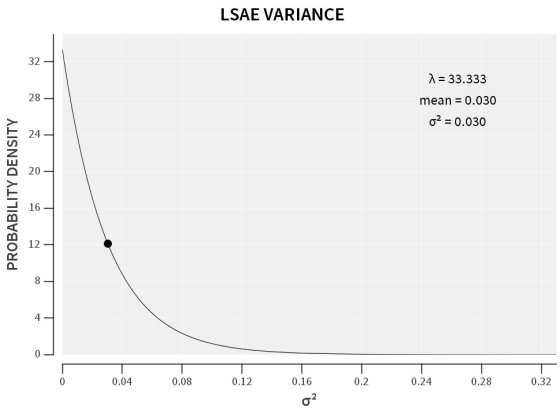
Experiment 13b.

For the second experiment, the same two single source samples were artificially degraded even further resulting in some allelic and locus drop-out. Figures 33 and 34 show the APH and LSAE per location, the LSAE variance graph output by STRmix™, and the degradation curve graph output by STRmix™ for these samples.

(a)



(b)



(c)

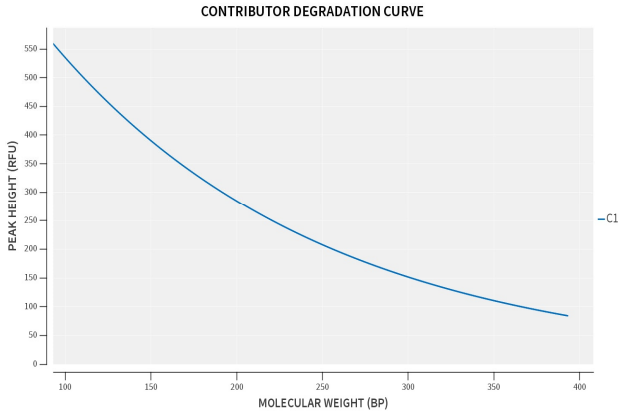


Figure 33. Plot of (a) APH (blue bars) and LSAE (orange line) per locus, (b) LSAE variance curve, and (c) degradation curve for an artificially degraded single source sample with drop-out amped at 125 pg (Ex13_SS_125_C_12M).

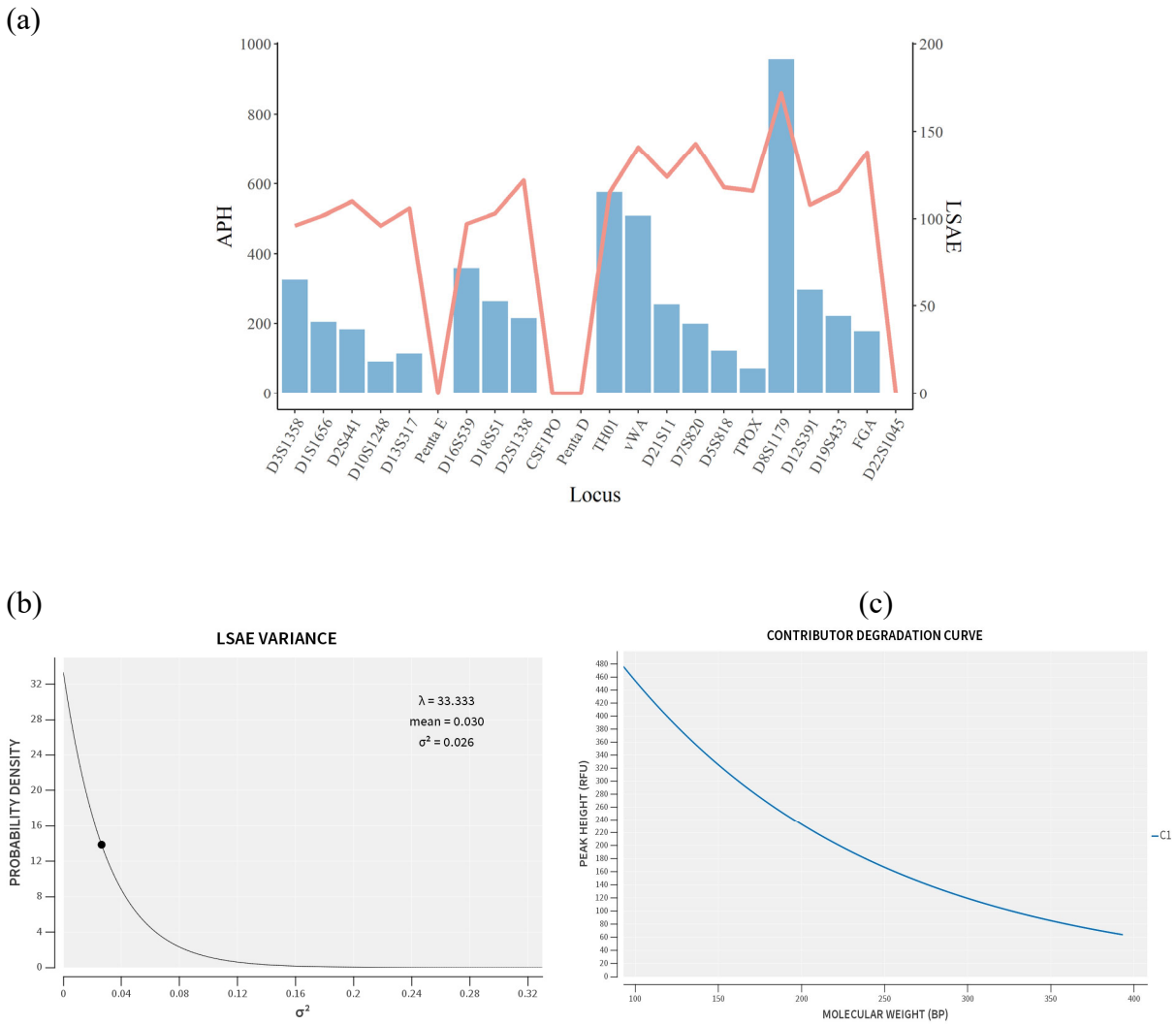


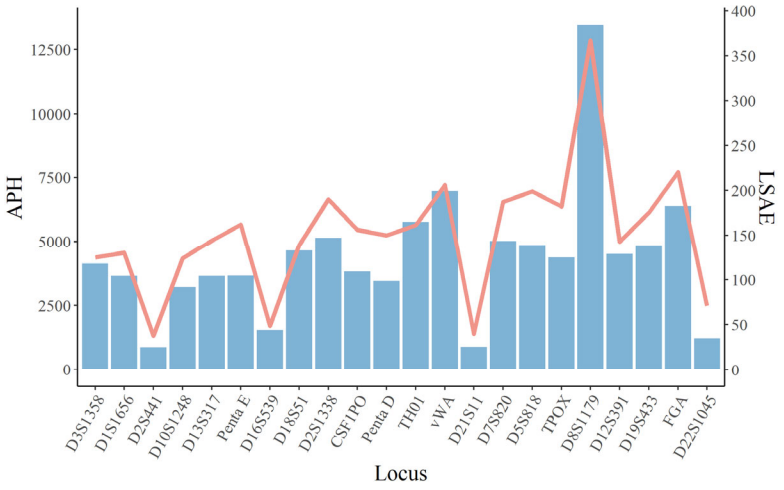
Figure 34. Plot of (a) APH (blue bars) and LSAE (orange line) per locus, (b) LSAE variance curve, and (c) degradation curve for an artificially degraded single source sample with drop-out amped at 250 pg (Ex13_SS_250_P_12M).

After further degradation to include drop-out, the LSAE values were comparable to the un-edited samples for loci in which there was no drop-out. The average of the accepted post burn-in values for the degradation increased further to 6.31E-03 (125 pg) and 6.67E-03 (250 pg) as can be seen in the degradation curves for each sample. As expected, further degradation had the effect of lowering the weights of the genotypes overall as assigned by STRmix™. This also resulted in the sub-sub-source LR of the comparison sample to lower from 6.74E+25 (un-edited sample) to 2.75E+20 for the 125 pg sample and from 6.74E+25 (un-edited sample) to 2.11E+18 for the 250 pg sample.

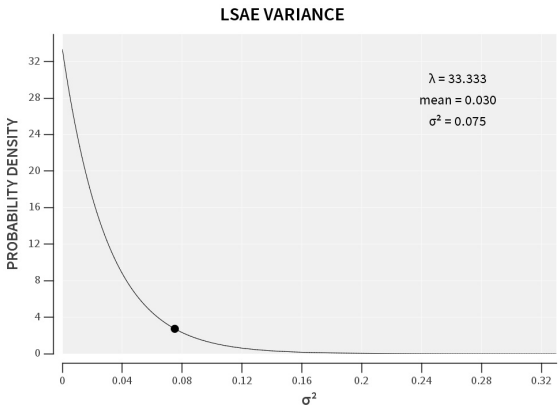
Experiment 13c.

In addition to the above, the samples were also artificially inhibited through editing of the STRmix™ input text file. Figures 35 and 36 show the APH and LSAE per location, the LSAE variance graph output by STRmix™, and the degradation curve graph output by STRmix™ for these samples. Locations of inhibition are noted in the figure descriptions.

(a)



(b)



(c)

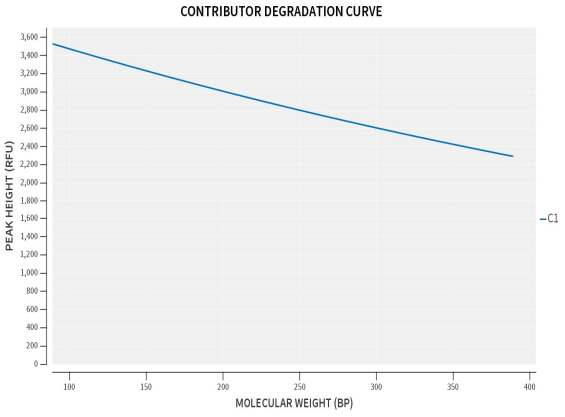
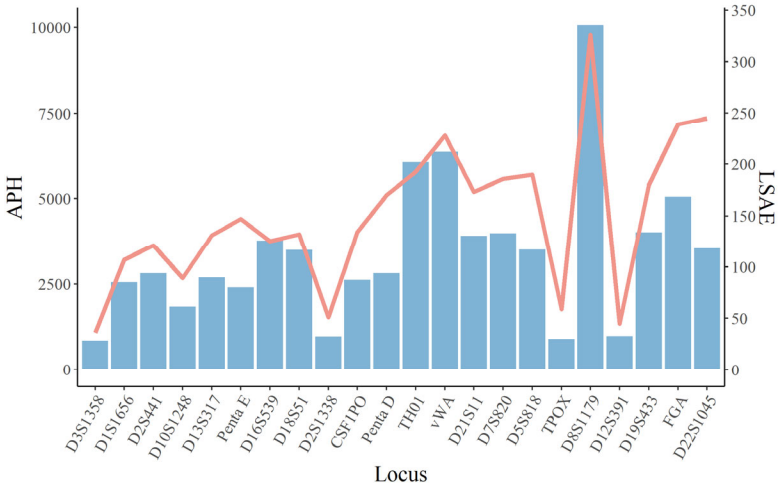
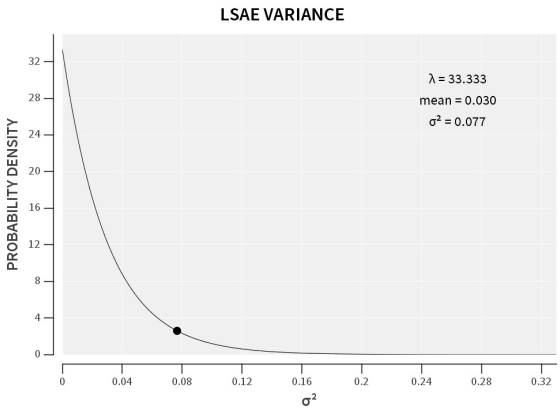


Figure 35. Plot of (a) APH (blue bars) and LSAE (orange line) per locus, (b) LSAE variance curve, and (c) degradation curve for an artificially inhibited single source sample amped at 125 pg (Ex13_SS_125_C_12M). Locations of inhibition include D2S441, D16S539, D21S11, and D22S1045.

(a)



(b)



(c)

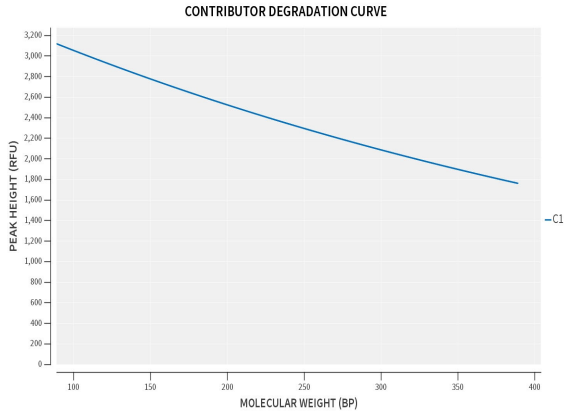


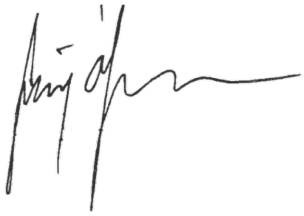
Figure 36. Plot of (a) APH (blue bars) and LSAE (orange line) per locus, (b) LSAE variance curve, and (c) degradation curve for an artificially inhibited single source sample amped at 250 pg (Ex13_SS_250_P_12M). Locations of inhibition include D3S1358, D2S1338, TPOX, and D12S391.

While the LSAE will vary slightly from run to run, the LSAE for the inhibited locations for both samples significantly decreased as is shown by the lower LSAE value for those locations where the APH is lower (i.e. the inhibited locations). This is further supported by the LSAE variance changes from 0.022 (un-edited sample) to 0.075 for the 125 pg and from 0.028 (un-edited sample) to 0.077 for the 250 pg sample. These LSAE values are as expected as inhibition of samples affects how the alleles are replicated during amplification and may sometimes only be observed for specific loci. The sub-sub-source LRs and genotype weights were the same between the un-edited samples and the inhibited samples.

Conclusion:

This document describes the NYC OCME's Department of Forensic Biology's internal validation experiments performed for samples amplified with PowerPlex® Fusion 5C (half reaction, 29 cycles) and run on 3500*xLs* genetic analyzers for STRmix™ v2.7. These experiments show that STRmix™ v2.7 functions as expected and is suitable for use on casework samples in this laboratory to aid in their interpretation and calculation of a statistic in the form of a likelihood ratio.

APPROVED
By Craig O'Connor at 10:28 pm, Sep 07, 2021



References:

1. **Institute of Environmental Science and Research Limited.** *STRmix™ V2.7 Operation manual.* 11 October 2019.

2. —. *STRmix™ V2.7 User's Manual.* 6 September 2019.

3. **Jo-Anne Bright, Duncan Taylor, Catherine McGovern, Stuart Cooper, Laura Russell, Damien Abarno, John Buckleton.** Developmental validation of STRmix™, expert software for the interpretation of forensic DNA profiles. *Forensic Science International: Genetics.* 2016, Vol. 23, pp. 226-239.

4. **NYC OCME.** *Internal Validation of STRmix™ V2.4 for Fusion NYC OCME. NYC OCME, 24 January 2017.* NYC : NYC OCME, 2017.

5. —. *STRmix™ V2.4.08 Performance Check. NYC OCME, 24 July 2018.* NYC : NYC OCME, 2018.

6. **Institute of Environmental Science and Research Limited.** *STRmix™ V2.6.0 Release and Testing Report. 24 January 2017.* s.l. : ESR, 2017.

7. —. *STRmix™ v2.7.0 Release and Testing Report.* 4 September 2019.

8. **SWGDAM.** *Guidelines for the Validation of Probabilistic Genotyping Systems.* s.l. : SWGDAM, 2015.

9. **Federal Bureau of Investigation.** *Quality Assurance Standards for Forensic DNA Testing Laboratories.* s.l. : Federal Bureau of Investigation, 2020.

10. **Institute of Environmental Science and Research Limited.** *STRmix™ V2.7 Implementation and Validation Guide.* 2 September 2019.

11. **NYC OCME.** *PowerPlex® Fusion System Amplification Kit on the Applied Biosystems® 3500xL Genetic Analyzer with GeneMarker® HID 2.9.5.* New York City : s.n., July 2019.

12. —. *NYC OCME Performance Check of GeneMarker® HID v3.0.0 with PowerPlex® Fusion 5C Data Run on 3500xL Genetic Analyzers.* 2021.

13. —. *NYC OCME Internal Validation of STRmix™ v2.7 for Fusion 5C/3500xL Data - STRmix™ Parameters .* NYC : NYC OCME, 2021.

14. —. *Stutter Study for GeneMarker® HID 3.0.0 and STRmix™ Version 2.7- PowerPlex® Fusion Data run on 3500xl Genetic Analyzers.* NYC : NYC OCME, 2021.

15. **Carolyn R Steffen, Michael D Coble, Katherine B Gettings, and Peter M Vallone.** Corrigendum to 'U.S. Population Data for 29 Autosomal STR Loci' [Forensic. *Forensic Science International: Genetics.* 2017, Vol. 31, pp. e36-e40.

16. **Nichols, David J Balding and Richard A.** DNA profile match probability calculation: how to allow for population stratification, relatedness, database selection and single bands. *Forensic Science International.* 1994, Vol. 64, pp. 125-140.

17. **Committee on DNA Forensic Science: An Update, National Research Council.** *The Evaluation of Forensic DNA Evidence.* Washington, D.C. : National Acedemy Press, 1996.

18. **NYC OCME.** *NYC OCME STR Analysis and Interpretation - PowerPlex® Fusion and STRmix™ on 3500xLs.* 2021.

19. **Promega Coorporation.** *PowerPlex® Fusion system for use on the Applied Biosystems® genetic analyzers.* Madison : Promega, 2020.

20. **Schuerman, Curt and Tim Kalafut, Clint Buchanan, Joel Sutton, and Jo-Anne Bright.** Using the nondonor distribution to improve communication and inform decision making for low LRs from minor contributors in mixed DNA profiles. July 2020, Vol. 65.

21. **Duncan Taylor, John Buckleton, Ian Evett.** Testing likelihood ratios produced from complex DNA profiles. *Forensic Science International: Genetics.* 2015, Vol. 16, pp. 165-171.

22. **Talor, Duncan.** Using continuous DNA interpretation methods to revisit likelihood ratio behavior. *Forensic Science International: Genetics.* 2014, Vol. 11, pp. 144-153.

23. **Curran, CM Triggs and JM.** The sensitivity of the Bayesian HPD method to the choise of prior. *Science and Justice.* 2006, Vol. 46, 3, pp. 169-178.

24. **Buckleton, James M Curran and John S.** An investigation into the performance of methods for adjusting for sampling uncertainty in DNA likelihood ratio calculations. *Forensic Science International: Genetics.* 2011, Vol. 5, pp. 512-516.

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

26. **Jo-Anne Bright, Kate E Stevensom, James M Curran, John S Buckleton.** The variability in likelihood ratios due to different mechanisms. *Forensic Science International: Genetics.* 2015, Vol. 14, pp. 187-190.

27. **Jo-Anne Bright, James M Curran, John S Buckleton.** The effect of the uncertainty in the number of contributors to mixed DNA profiles on profile interpretation. *Forensic Science International: Genetics.* 2014, Vol. 12, pp. 208-214.

28. **Jo-Anne Bright, Duncan Taylor, James Curran, John Buckleton.** Searching mixed DNA profiles directly against profile databases. *Forensic Science International: Genetics.* 2014, Vol. 9, pp. 102-110.

29. **Bright, Jo-Anne.** STRmix™ collaborative exercies on DNA mixture interpretation. *Forensic Science International: Genetics.* 2019, Vol. 40, pp. 1-8.

30. **Todd Bille, Steven Weitz, John S Buckleton, Jo-Anne Bright.** Interpreting a major component from a mixed DNA profile with an unknown number of minor contributors. *Forensic Science International: Genetics.* 2019, Vol. 40, pp. 150-159.

Appendix A: Experiment 2 results for single source profiles (i) 14F, (ii) 21M, (iii) 25M, (iv) 34F, and (v) 35F.

(i) Calculation of sub-sub-source LRs 'by-hand'(Microsoft® Excel) and using STRmix™ for single source profile 14F using the NIST amended Caucasian frequency population database for varying F_{ST} (θ) values.

	θ = 0		***θ = 0.03***	
Locus	**Microsoft® Excel**	**STRmix™ V2.7**	**Microsoft® Excel**	**STRmix™ V2.7**
D3S1358	1.5748E+01	1.5748E+01	1.3222E+01	1.3222E+01
D1S1656	4.8525E+01	4.8525E+01	3.2653E+01	3.2653E+01
D2S441	6.9268E+00	6.9268E+00	6.4281E+00	6.4281E+00
D10S1248	2.5883E+01	2.5883E+01	1.5519E+01	1.5519E+01
D13S317	8.7909E+01	8.7909E+01	4.9040E+01	4.9040E+01
Penta E	4.6528E+02	4.6528E+02	1.1437E+02	1.1437E+02
D16S539	5.0658E+00	5.0658E+00	4.8718E+00	4.8718E+00
D18S51	2.3845E+01	2.3845E+01	1.8716E+01	1.8716E+01
D2S1338	5.6540E+01	5.6540E+01	3.6731E+01	3.6731E+01
CSF1PO	2.7769E+01	2.7769E+01	2.0501E+01	2.0501E+01
Penta D	2.0198E+01	2.0198E+01	1.6259E+01	1.6259E+01
TH01	2.6983E+01	2.6983E+01	2.0402E+01	2.0402E+01
vWA	1.2329E+01	1.2329E+01	1.0751E+01	1.0751E+01
D21S11	6.0810E+01	6.0810E+01	3.0846E+01	3.0846E+01
D7S820	1.5330E+01	1.5330E+01	1.2941E+01	1.2941E+01
D5S818	2.3262E+01	2.3262E+01	1.6050E+01	1.6050E+01
TPOX	7.4832E+00	7.4832E+00	6.5982E+00	6.5982E+00
D8S1179	4.5548E+01	4.5548E+01	2.5062E+01	2.5062E+01
D12S391	9.5839E+01	9.5839E+01	5.2479E+01	5.2479E+01
D19S433	2.7806E+01	2.7806E+01	2.0009E+01	2.0009E+01
FGA	4.7178E+01	4.7178E+01	3.1969E+01	3.1969E+01
D22S1045	2.0803E+01	2.0803E+01	1.5581E+01	1.5581E+01
Total	1.3805E+32	1.3805E+32	3.2778E+28	3.2778E+28

(ii) Calculation of sub-sub-source LRs 'by-hand'(Microsoft® Excel) and using STRmix™ for single source profile 21M using the NIST amended Caucasian frequency population database for varying F_{ST} (θ) values.

Locus	$\theta = 0$		$\theta = 0.03$	
	Microsoft® Excel	**STRmix™ V2.7**	**Microsoft® Excel**	**STRmix™ V2.7**
D3S1358	8.7172E+00	8.7172E+00	7.9203E+00	7.9203E+00
D1S1656	2.8769E+01	2.8769E+01	2.1849E+01	2.1849E+01
D2S441	3.8976E+01	3.8976E+01	2.6162E+01	2.6162E+01
D10S1248	2.5883E+01	2.5883E+01	1.5519E+01	1.5519E+01
D13S317	6.8864E+01	6.8864E+01	2.9833E+01	2.9833E+01
Penta E	3.8897E+01	3.8897E+01	2.7128E+01	2.7128E+01
D16S539	5.0658E+00	5.0658E+00	4.8718E+00	4.8718E+00
D18S51	4.8635E+02	4.8635E+02	8.6100E+01	8.6100E+01
D2S1338	2.5620E+01	2.5620E+01	1.9692E+01	1.9692E+01
CSF1PO	7.3600E+00	7.3600E+00	6.8062E+00	6.8062E+00
Penta D	6.5107E+01	6.5107E+01	4.0241E+01	4.0241E+01
TH01	1.0964E+01	1.0964E+01	9.6969E+00	9.6969E+00
vWA	6.8240E+02	6.8240E+02	1.1332E+02	1.1332E+02
D21S11	3.4913E+01	3.4913E+01	2.5244E+01	2.5244E+01
D7S820	2.1809E+01	2.1809E+01	1.7444E+01	1.7444E+01
D5S818	3.6291E+00	3.6291E+00	3.5878E+00	3.5878E+00
TPOX	2.4017E+02	2.4017E+02	9.8799E+01	9.8799E+01
D8S1179	1.7970E+01	1.7970E+01	1.4839E+01	1.4839E+01
D12S391	1.2178E+02	1.2178E+02	5.7846E+01	5.7846E+01
D19S433	3.8407E+01	3.8407E+01	2.2096E+01	2.2096E+01
FGA	3.5481E+01	3.5481E+01	2.5213E+01	2.5213E+01
D22S1045	4.0777E+00	4.0777E+00	3.9926E+00	3.9926E+00
Total	1.2376E+33	1.2376E+33	4.2131E+28	4.2131E+28

(iii) Calculation of sub-sub-source LRs 'by-hand'(Microsoft® Excel) and using STRmix™ for single source profile 25M using the NIST amended Caucasian frequency population database for varying F_{ST} (θ) values.

	θ = 0		***θ = 0.03***	
Locus	**Microsoft® Excel**	**STRmix™ V2.7**	**Microsoft® Excel**	**STRmix™ V2.7**
D3S1358	7.7042E+00	7.7042E+00	7.1058E+00	7.1058E+00
D1S1656	4.6703E+01	4.6703E+01	3.1707E+01	3.1707E+01
D2S441	2.2598E+01	2.2598E+01	1.4062E+01	1.4062E+01
D10S1248	5.4703E+00	5.4703E+00	5.2240E+00	5.2240E+00
D13S317	6.8864E+01	6.8864E+01	2.9833E+01	2.9833E+01
Penta E	1.6550E+02	1.6550E+02	7.8358E+01	7.8358E+01
D16S539	2.8678E+01	2.8678E+01	2.1693E+01	2.1693E+01
D18S51	3.0232E+01	3.0232E+01	2.2790E+01	2.2790E+01
D2S1338	4.1347E+01	4.1347E+01	2.9217E+01	2.9217E+01
CSF1PO	2.0640E+01	2.0640E+01	1.3152E+01	1.3152E+01
Penta D	3.4543E+01	3.4543E+01	2.5357E+01	2.5357E+01
TH01	1.0964E+01	1.0964E+01	9.6969E+00	9.6969E+00
vWA	7.6375E+02	7.6375E+02	1.2345E+02	1.2345E+02
D21S11	2.4502E+01	2.4502E+01	1.4916E+01	1.4916E+01
D7S820	1.5330E+01	1.5330E+01	1.2941E+01	1.2941E+01
D5S818	4.9166E+01	4.9166E+01	2.4108E+01	2.4108E+01
TPOX	3.6367E+00	3.6367E+00	3.2072E+00	3.2072E+00
D8S1179	3.9496E+01	3.9496E+01	2.7483E+01	2.7483E+01
D12S391	8.1549E+01	8.1549E+01	3.3035E+01	3.3035E+01
D19S433	2.4416E+02	2.4416E+02	9.8736E+01	9.8736E+01
FGA	2.3842E+01	2.3842E+01	1.4623E+01	1.4623E+01
D22S1045	4.7752E+01	4.7752E+01	3.2111E+01	3.2111E+01
Total	6.5232E+33	6.5232E+33	1.4017E+29	1.4017E+29

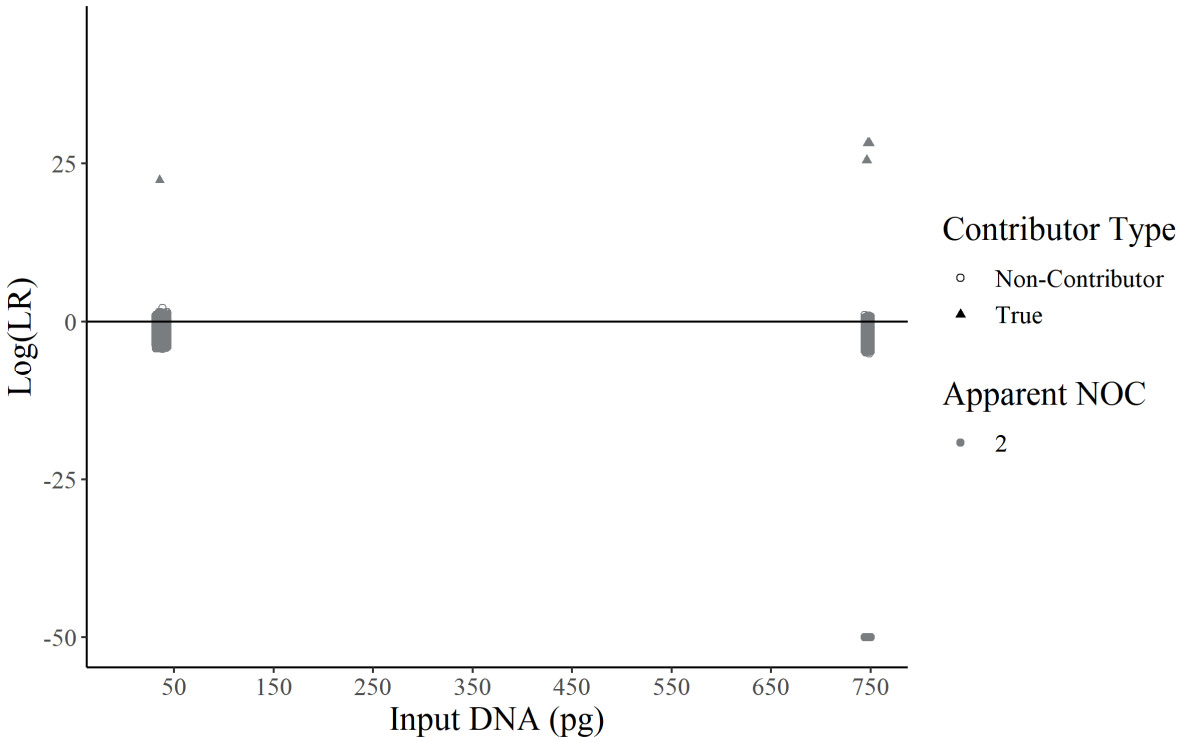
(iv) Calculation of sub-sub-source LRs 'by-hand'(Microsoft® Excel) and using STRmix™ for single source profile 34F using the NIST amended Caucasian frequency population database for varying F_{ST} (θ) values.

	θ = 0		***θ = 0.03***	
Locus	**Microsoft® Excel**	**STRmix™ V2.7**	**Microsoft® Excel**	**STRmix™ V2.7**
D3S1358	7.7042E+00	7.7042E+00	7.1058E+00	7.1058E+00
D1S1656	2.4662E+01	2.4662E+01	1.9309E+01	1.9309E+01
D2S441	2.8106E+02	2.8106E+02	1.0503E+02	1.0503E+02
D10S1248	1.1297E+01	1.1297E+01	8.2725E+00	8.2725E+00
D13S317	9.4554E+00	9.4554E+00	7.1722E+00	7.1722E+00
Penta E	1.9375E+02	1.9375E+02	6.2852E+01	6.2852E+01
D16S539	2.8678E+01	2.8678E+01	2.1693E+01	2.1693E+01
D18S51	3.0026E+01	3.0026E+01	2.2614E+01	2.2614E+01
D2S1338	3.1198E+01	3.1198E+01	2.3217E+01	2.3217E+01
CSF1PO	7.7242E+00	7.7242E+00	6.0803E+00	6.0803E+00
Penta D	1.7074E+01	1.7074E+01	1.4039E+01	1.4039E+01
TH01	1.8061E+01	1.8061E+01	1.1902E+01	1.1902E+01
vWA	1.8991E+01	1.8991E+01	1.4904E+01	1.4904E+01
D21S11	3.4913E+01	3.4913E+01	2.5244E+01	2.5244E+01
D7S820	9.5329E+00	9.5329E+00	8.5750E+00	8.5750E+00
D5S818	6.6622E+00	6.6622E+00	5.3783E+00	5.3783E+00
TPOX	3.6367E+00	3.6367E+00	3.2072E+00	3.2072E+00
D8S1179	1.4814E+01	1.4814E+01	1.2071E+01	1.2071E+01
D12S391	3.0504E+01	3.0504E+01	2.2656E+01	2.2656E+01
D19S433	9.0465E+01	9.0465E+01	3.2011E+01	3.2011E+01
FGA	4.3145E+01	4.3145E+01	2.2119E+01	2.2119E+01
D22S1045	6.8559E+00	6.8559E+00	5.5084E+00	5.5084E+00
Total	1.3095E+29	1.3095E+29	3.6913E+25	3.6913E+25

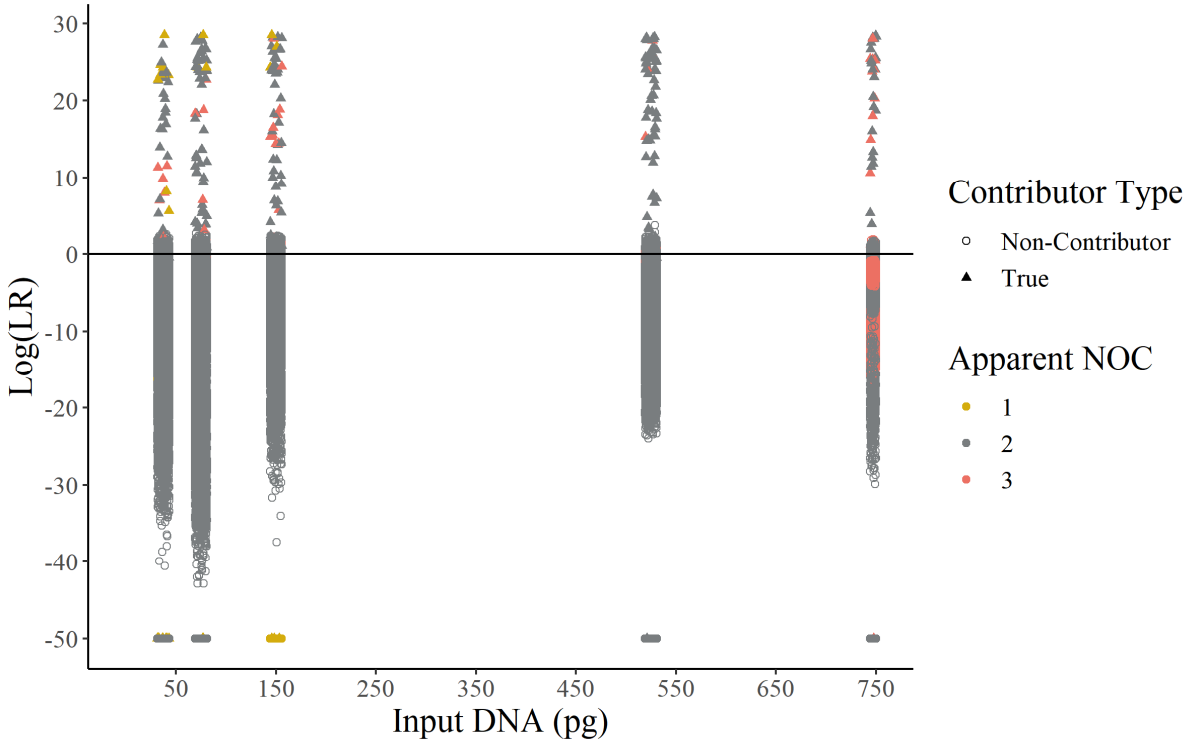
(v) Calculation of sub-sub-source LRs 'by-hand'(Microsoft® Excel) and using STRmix™ for single source profile 35F using the NIST amended Caucasian frequency population database for varying F_{ST} (θ) values.

Locus	*θ = 0*		*θ = 0.03*	
	Microsoft® Excel	**STRmix™ V2.7**	**Microsoft® Excel**	**STRmix™ V2.7**
D3S1358	9.9834E+00	9.9834E+00	8.9385E+00	8.9385E+00
D1S1656	2.7743E+01	2.7743E+01	2.1265E+01	2.1265E+01
D2S441	6.0515E+00	6.0515E+00	5.7082E+00	5.7082E+00
D10S1248	8.5498E+00	8.5498E+00	7.7650E+00	7.7650E+00
D13S317	1.2759E+01	1.2759E+01	1.0758E+01	1.0758E+01
Penta E	1.7512E+02	1.7512E+02	7.7409E+01	7.7409E+01
D16S539	2.7967E+01	2.7967E+01	1.9141E+01	1.9141E+01
D18S51	2.1223E+01	2.1223E+01	1.7034E+01	1.7034E+01
D2S1338	2.3462E+01	2.3462E+01	1.8385E+01	1.8385E+01
CSF1PO	6.1869E+02	6.1869E+02	1.2612E+02	1.2612E+02
Penta D	3.5280E+01	3.5280E+01	2.3971E+01	2.3971E+01
TH01	2.1657E+01	2.1657E+01	1.7205E+01	1.7205E+01
vWA	4.5732E+01	4.5732E+01	3.1605E+01	3.1605E+01
D21S11	1.9686E+01	1.9686E+01	1.5321E+01	1.5321E+01
D7S820	1.2266E+01	1.2266E+01	1.0633E+01	1.0633E+01
D5S818	8.4181E+01	8.4181E+01	4.6065E+01	4.6065E+01
TPOX	1.5578E+01	1.5578E+01	1.2955E+01	1.2955E+01
D8S1179	1.9922E+01	1.9922E+01	1.5032E+01	1.5032E+01
D12S391	1.2178E+02	1.2178E+02	5.7846E+01	5.7846E+01
D19S433	9.0603E+02	9.0603E+02	1.8735E+02	1.8735E+02
FGA	1.3676E+01	1.3676E+01	1.1751E+01	1.1751E+01
D22S1045	4.0777E+00	4.0777E+00	3.9926E+00	3.9926E+00
Total	3.7411E+32	3.7411E+32	5.1380E+28	5.1380E+28

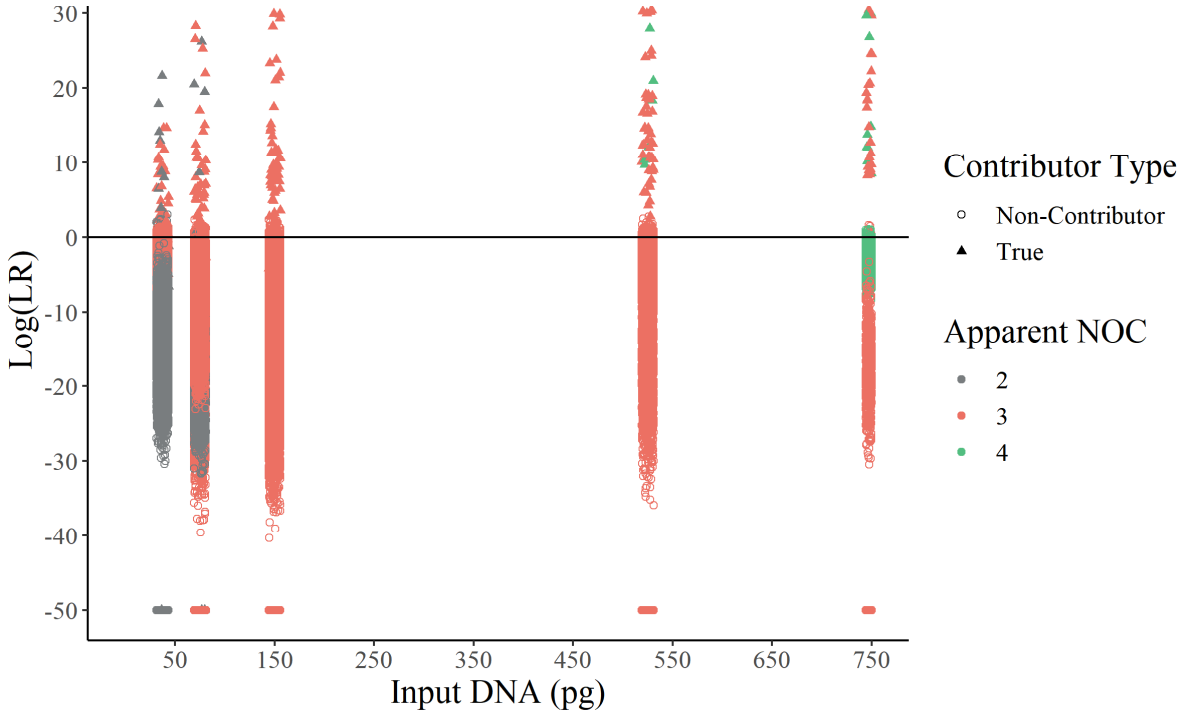
Appendix B: Log(LR) versus input DNA in pg (37.5 pg – 750 pg) per contributor for true (i) one-person samples and (ii) two-person, (iii) three-person, (iv) four-person, and (v) five-person mixtures of variable ratios from Experiment 4. Points are slightly jittered to better visualize the data.



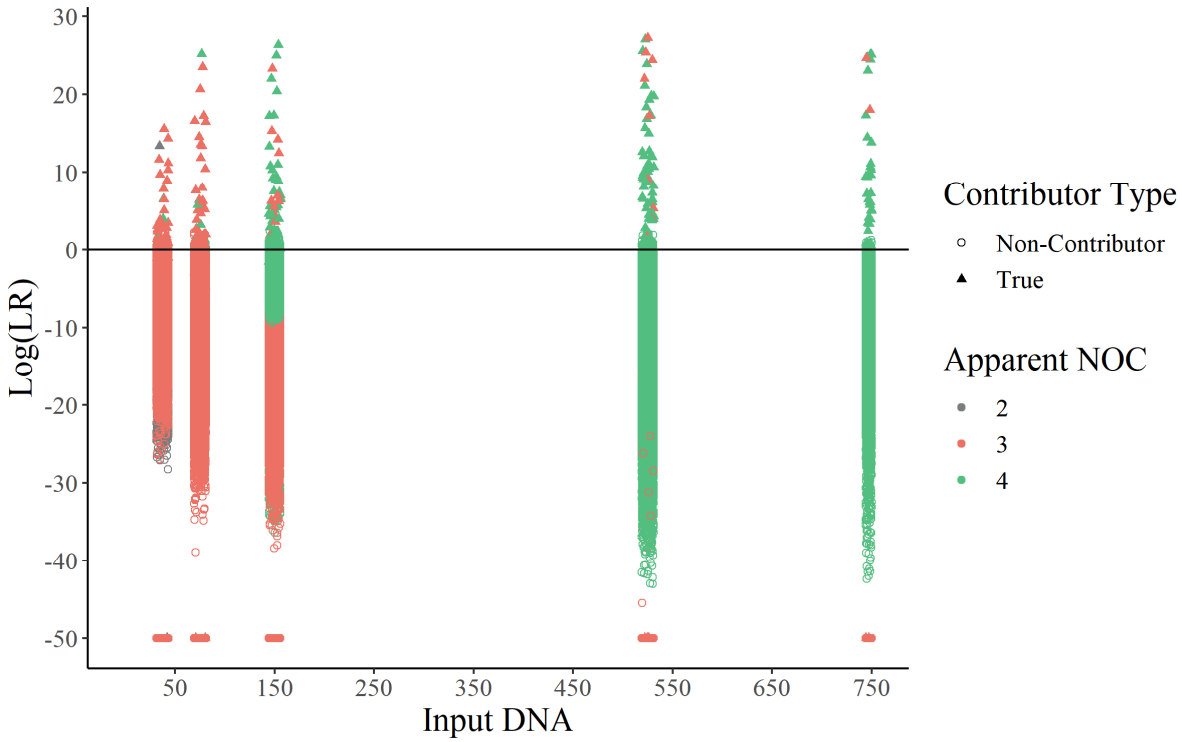
(i) Log(LR) versus input DNA amount (pg) per contributor for true one-person samples of variable ratios.



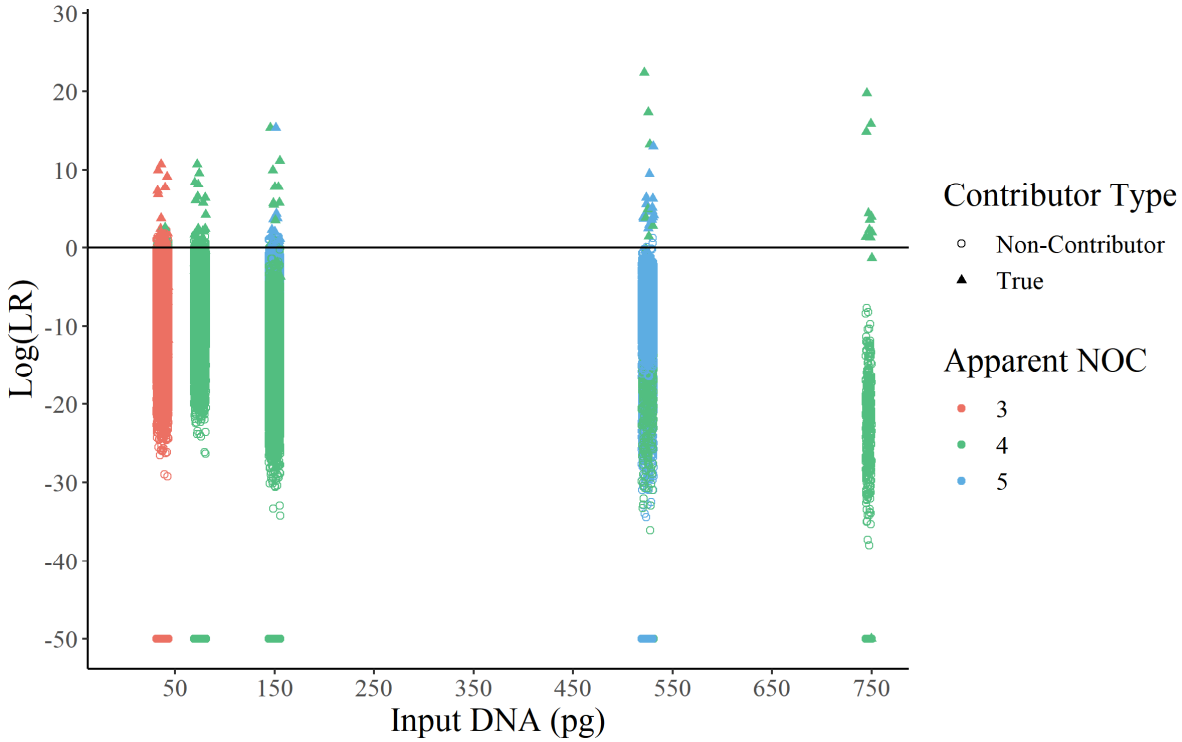
(ii) Log(LR) versus input DNA amount (pg) per contributor for true two-person samples of variable ratios.



(iii) Log(LR) versus input DNA amount (pg) per contributor for true three-person samples of variable ratios.

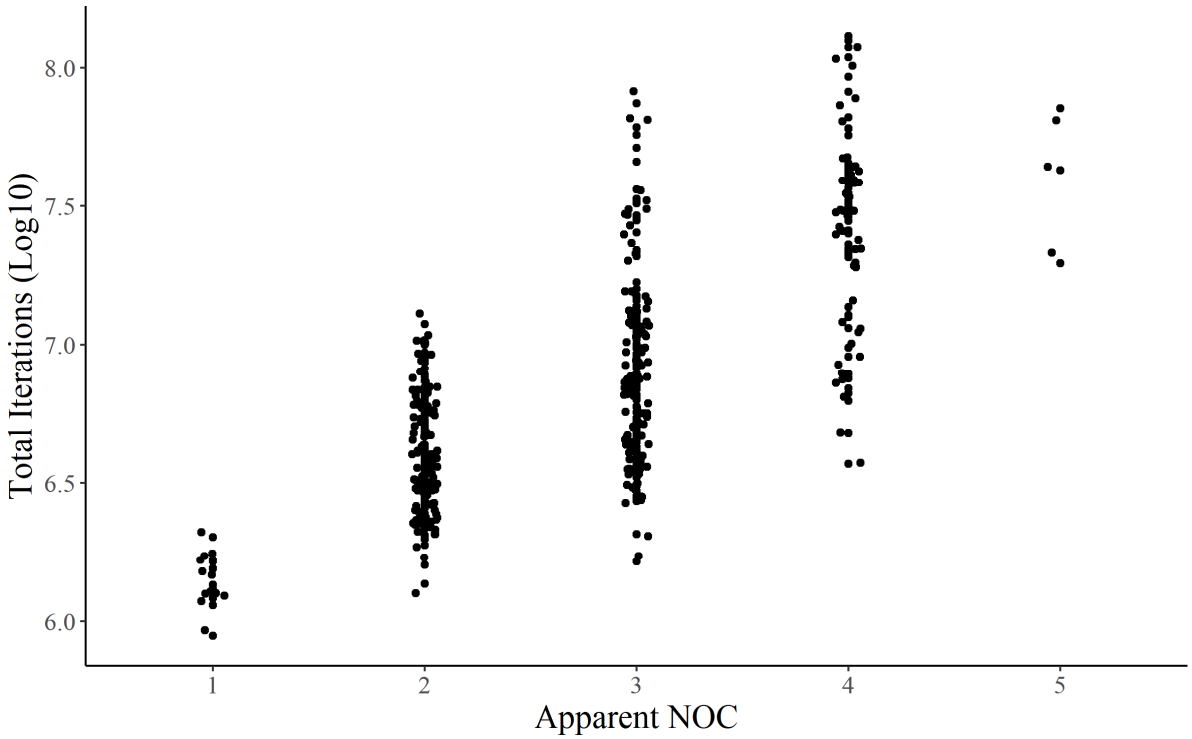


(iv) Log(LR) versus input DNA amount (pg) per contributor for true four-person samples of variable ratios.

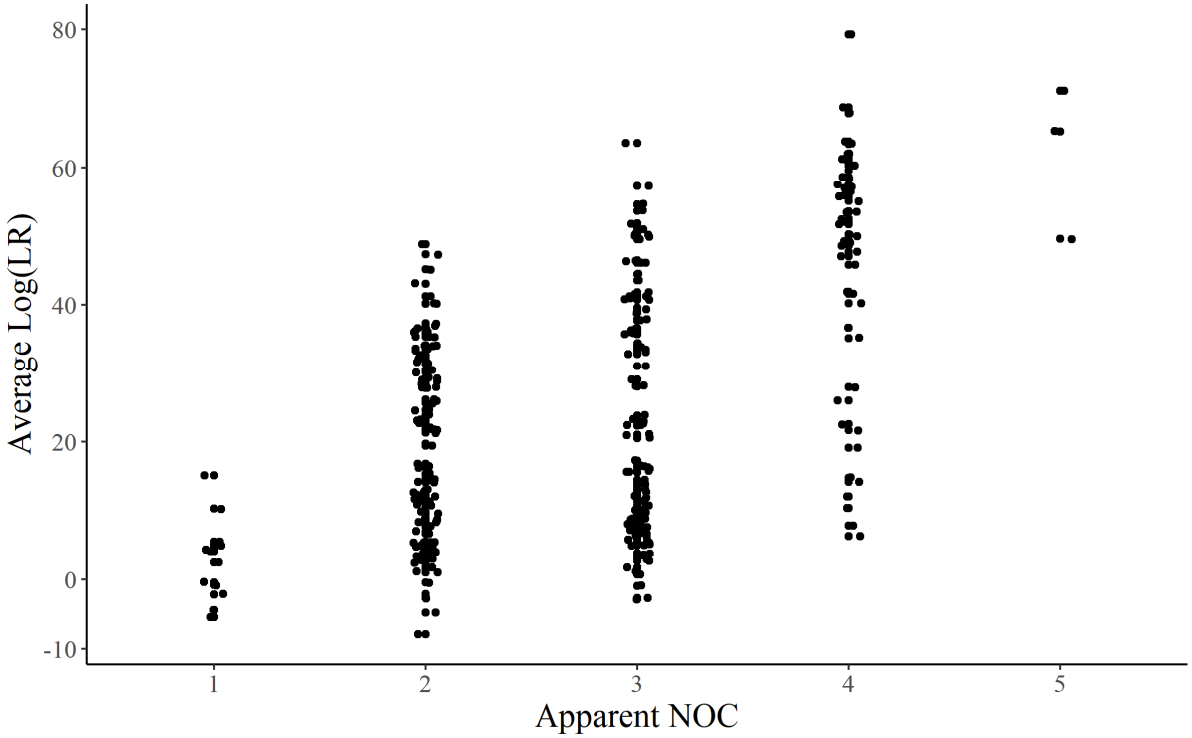


(v) Log(LR) versus input DNA amount (pg) per contributor for true five-person samples of variable ratios.

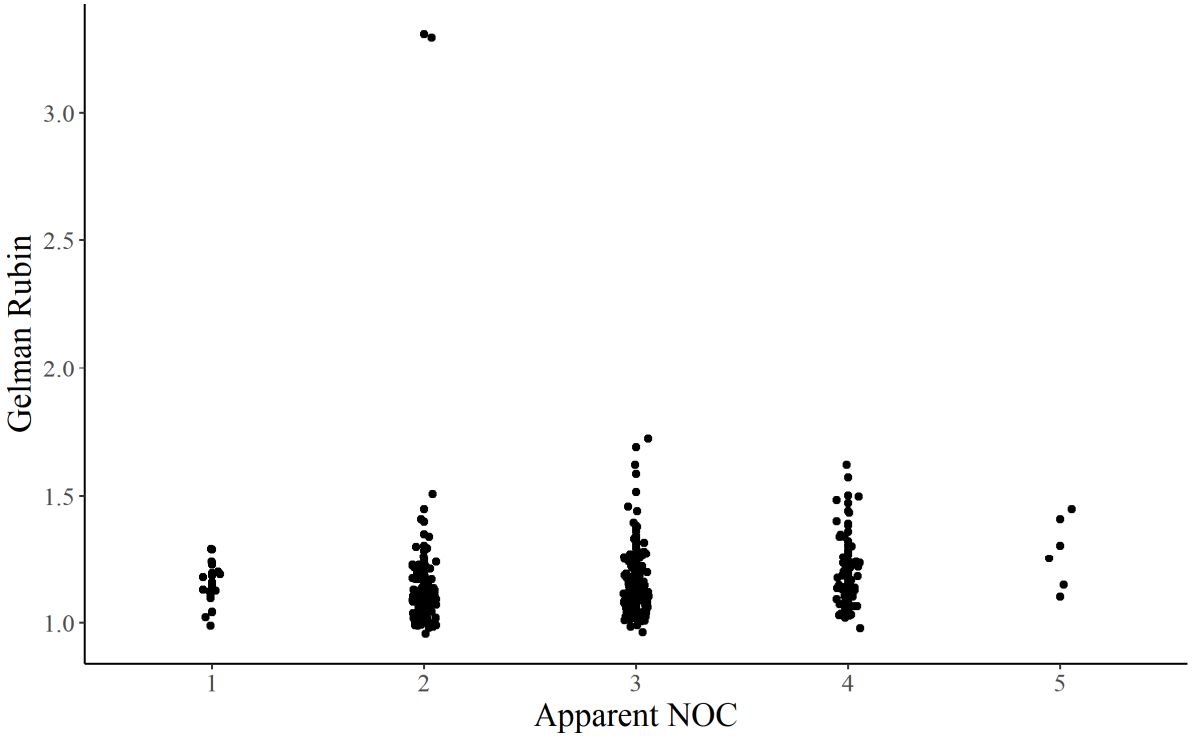
Appendix C: Summary plots of secondary diagnostics for Experiment 4 mixture samples amplified at 37.5 pg – 750 pg including (i) total iterations, (ii) average log(LR), (iii) Gelman Rubin, (iv) effective sample size, (v) allele variance, (vi) back stutter variance, (vii) forward stutter variance, (viii) half back stutter variance, (ix) double back stutter variance. Points are slightly jittered to better visualize the data.



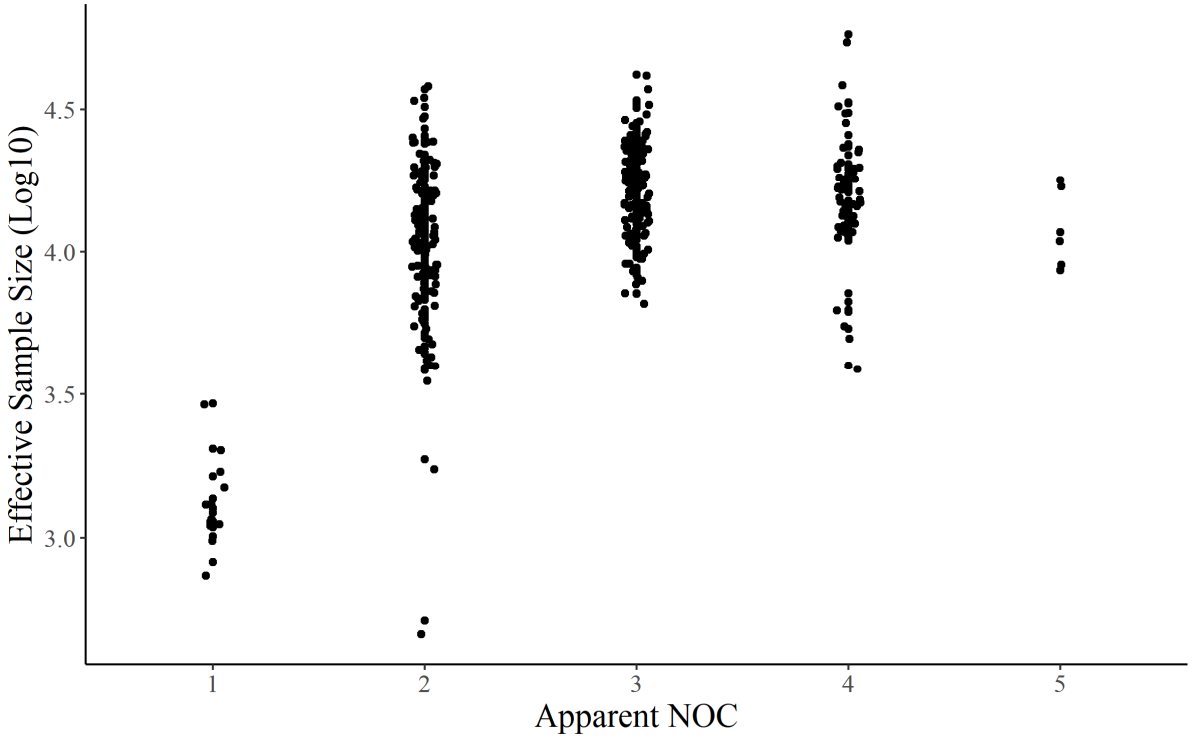
(i) Summary of the log of total iterations diagnostic for all mixtures per apparent NOC.



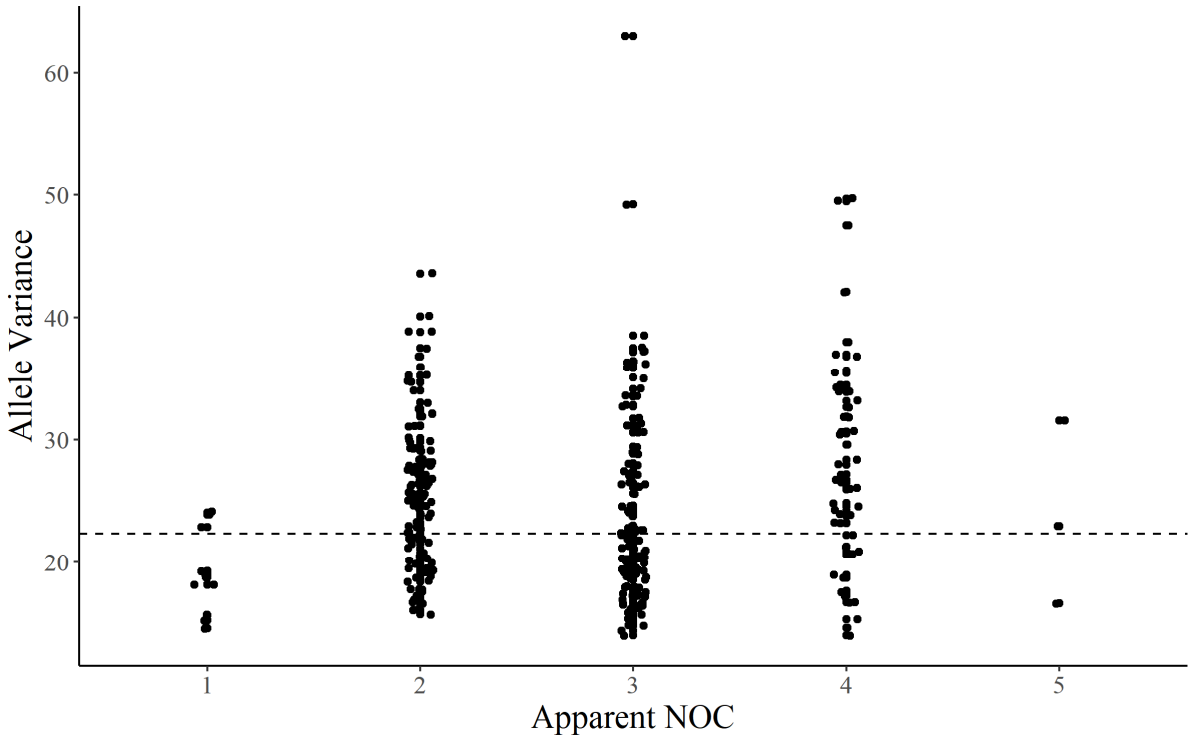
(ii) Summary of the average log(LR) diagnostic for all mixtures per apparent NOC.



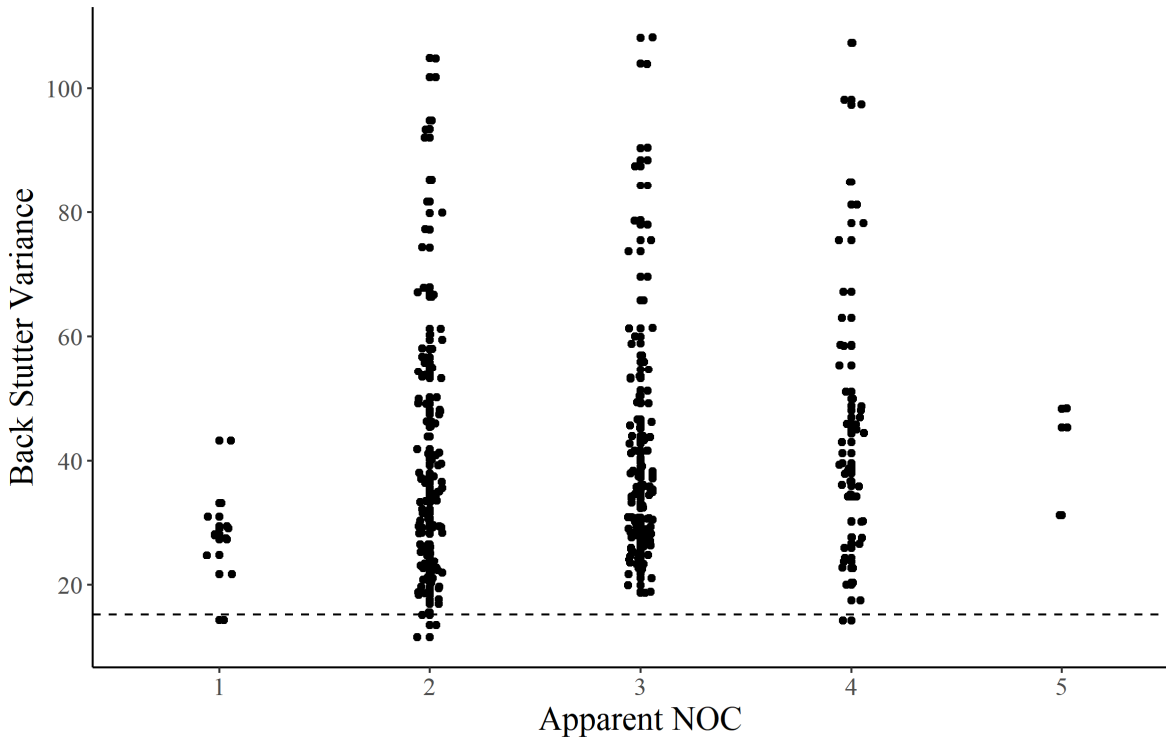
(iii) Summary of the Gelman Rubin diagnostic for all mixtures per apparent NOC.



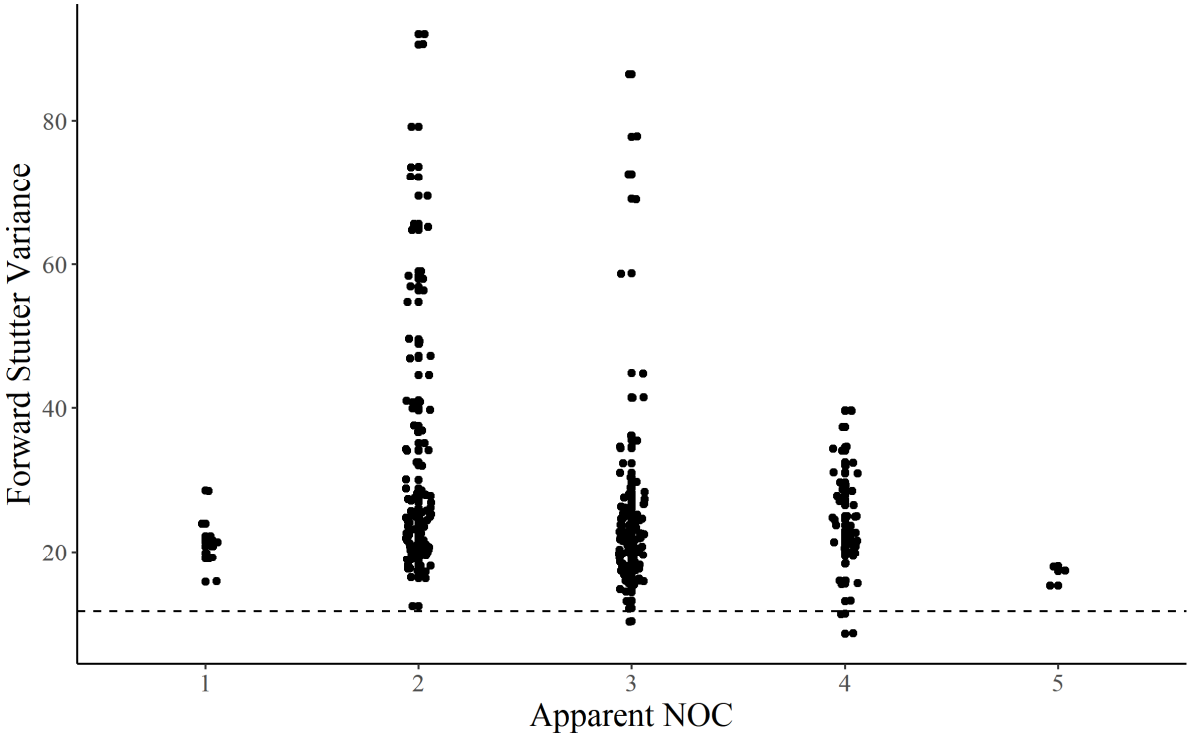
(iv) Summary of the log10 of effective sample size diagnostic for all mixtures per apparent NOC.



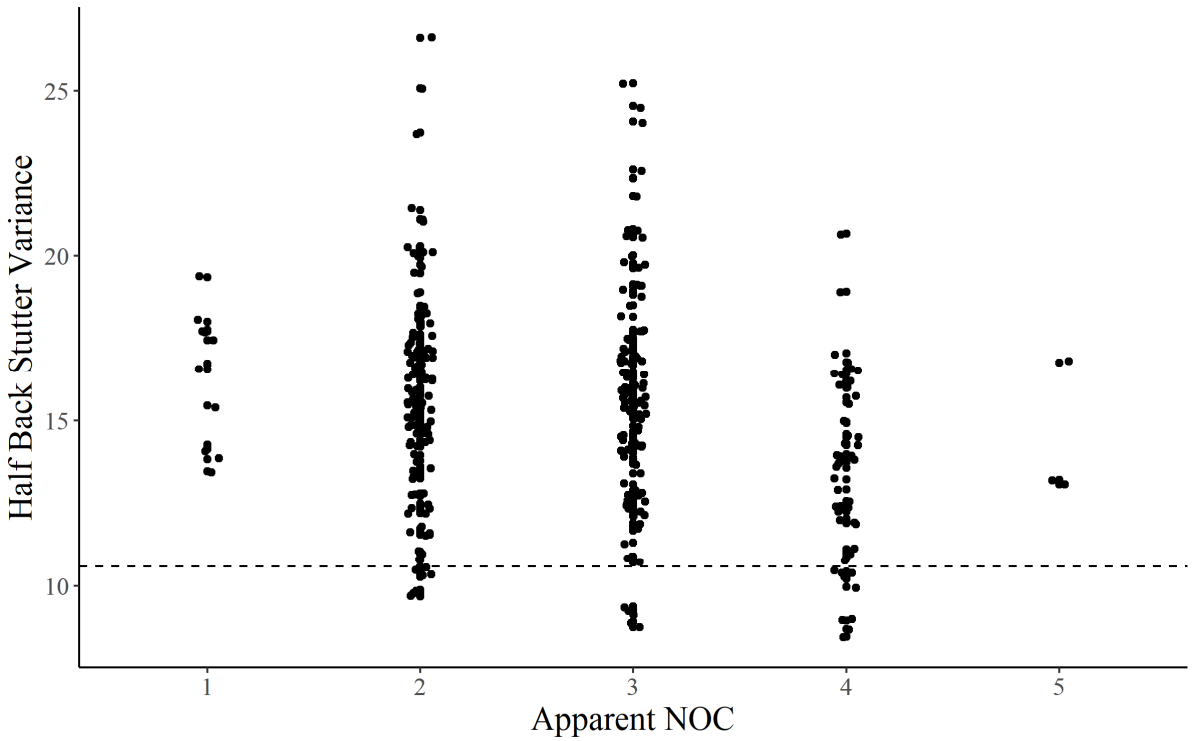
(v) Summary of the allele variance diagnostic for all mixtures per apparent NOC. *Dashed line represents the mode of the prior distribution (22.360).



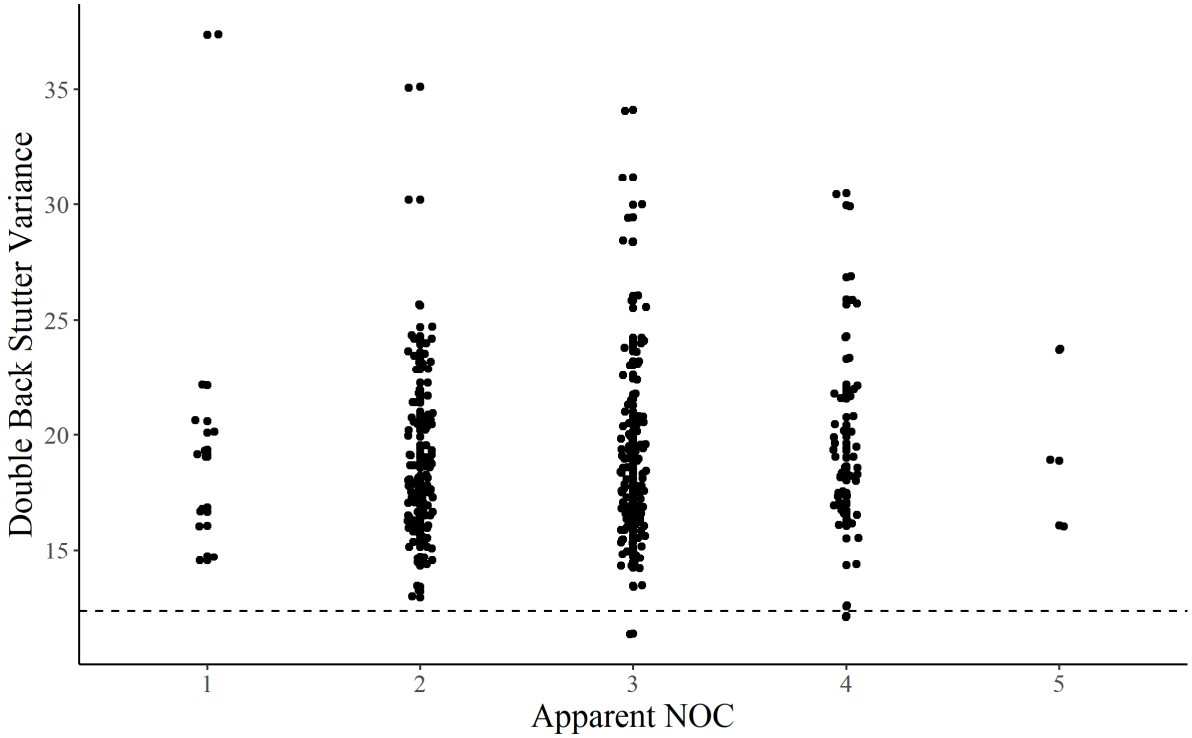
(vi) Summary of the back stutter variance diagnostic for all mixtures per apparent NOC. *Dashed line represents the mode of the prior distribution (15.223).



(vii) Summary of the forward stutter variance diagnostic for all mixtures per apparent NOC. *Dashed line represents the mode of the prior distribution (11.691).



(viii) Summary of the half back stutter variance diagnostic for all mixtures per apparent NOC. *Dashed line represents the mode of the prior distribution (10.591).



(ix) Summary of the double back stutter variance diagnostic for all mixtures per apparent NOC. *Dashed line represents the mode of the prior distribution (12.334).

Appendix D: Summary of STRmix™ results and diagnostics for all mock evidence samples from Experiment 7. BS = back stutter; FS = forward stutter; HBS = half back stutter; DBS = double back stutter.

#	Sample Name	NOC	Total Iterations	Effective Sample Size	Avg log(LR)	Gelman Rubin	Allele Variance	BS Variance	FS Variance	HBS Variance	DBS Variance	Sub-Sub Source LR	HPD LR	Unified LR
1	**Mock10_174_P 2p**	2	2707314	10505.8	14.4	1.1	22.0	44.6	18.5	20.5	20.0	2.62E+27	5.27E+26	9.89E+15
2	**Mock12_138_C 3p**	3	12443703	20849.3	22.3	1.1	22.5	15.0	21.1	12.7	19.7	1.97E+26	2.62E+25	6.11E+15
3	**Mock13_529_P 1p**	1	1569826	5307.9	38.4	1.1	22.8	31.3	23.5	13.8	13.4	7.90E+24	3.45E+24	4.36E+15
4	**Mock16_545_C 1p**	1	971533	3793.5	27.0	1.1	26.7	58.6	10.5	13.4	16.6	2.55E+25	1.04E+25	5.83E+15
5	**Mock17_73_P 3p**	3	1987401	9961.2	-1.0	1.3	19.7	30.3	19.2	17.9	18.3	1.34E+11	1.92E+10	2.06E+09
6	**Mock18_528_C 1p**	1	818104	2026.9	4.9	1.1	21.7	41.7	22.3	14.6	16.2	5.42E+25	2.07E+25	1.64E+15
7	**Mock19_207_P 2p**	2	3451799	15355.8	21.5	1.1	25.8	18.5	17.9	17.6	16.0	7.41E+26	1.52E+26	1.03E+16
8	**Mock1_521_C 1p**	1	1637505	4037.5	21.7	1.1	24.8	114.3	25.9	12.0	14.3	2.96E+30	1.14E+30	2.10E+16
9	**Mock20_368_C 2p**	2	3101239	19266.4	18.5	1.0	26.0	63.5	23.9	11.8	19.6	1.11E+28	2.07E+27	1.10E+16
11	**Mock21_151_P 2p**	2	3011882	11136.3	13.4	1.1	21.1	24.0	22.0	12.7	16.6	2.57E+25	5.13E+24	4.42E+15
12	**Mock22EC_513_C 2p**	2	8055102	17035.4	36.2	1.0	26.1	39.1	24.0	9.6	21.1	5.68E+24	1.17E+24	2.00E+15
13	**Mock22SF_535_C 2p**	2	10944175	15226.3	33.4	1.1	24.0	52.2	33.2	10.4	19.3	5.22E+24	1.10E+24	1.99E+15
14	**Mock23EC_535_P 3p**	3	19005475	38817.2	41.3	3.3	34.3	40.4	27.4	16.5	16.0	8.01E+26	1.03E+26	7.04E+15
15	**Mock23SF_529_P 1p**	1	1984937	4108.5	2.4	1.1	34.2	256.8	31.5	15.1	16.6	0	0	0
16	**Mock24EC_509_C 2p**	2	3536286	14119.9	37.7	1.1	35.3	93.4	33.8	12.1	14.6	2.45E+15	4.81E+14	2.10E+10

17	**Mock24SF_528_C 2p**	2	8706988	19350.9	35.3	1.0	22.0	47.8	43.0	11.1	19.8	0.01226	0.00335	0.00335
18	**Mock25_539_P 2p**	2	5688738	14484.7	7.7	1.4	25.1	223.3	15.2	9.7	18.1	3.17E+26	6.28E+25	6.28E+15
19	**Mock26_519_C 2p**	2	2979917	19760.7	15.2	1.1	20.2	146.4	21.9	17.0	18.8	4.83E+26	1.03E+26	6.90E+15
20	**Mock27_540_P 3p**	3	14011419	17375.7	34.1	1.6	29.6	98.2	19.9	19.6	17.7	2.57E+22	3.70E+21	1.19E+14
21	**Mock2_515_P 2p**	2	10641890	14597.5	23.0	1.1	26.8	55.5	64.1	12.7	16.1	1.00E+29	1.96E+28	1.75E+16
22	**Mock32_143_C 3p**	3	2648610	15123.4	5.5	1.2	20.7	39.5	21.9	15.7	24.2	8.63E+18	1.34E+18	1.16E+13
23	**Mock3_533_C 1p**	1	1208346	2922.3	22.8	1.1	31.0	48.6	40.7	14.3	14.8	1.14E+27	4.92E+26	7.64E+15
24	**Mock4_519_P 1p**	1	1596131	5585.9	29.5	1.1	26.6	31.5	19.2	15.8	15.7	1.32E+23	5.63E+22	1.70E+15
25	**Mock5_533_C 1p**	1	1727838	3430.4	32.9	1.1	28.1	35.1	22.8	12.2	20.9	8.37E+27	3.07E+27	9.31E+15
26	**Mock6_525_P 1p**	1	1459235	3484.1	30.8	1.2	26.3	23.9	54.0	14.9	13.1	3.20E+26	1.40E+26	7.32E+15
27	**Mock7_538_C 1p**	1	1691099	2852.7	25.1	1.1	41.1	54.0	25.1	13.2	21.6	8.21E+26	3.20E+26	1.02E+16
28	**Mock8_534_P 2p**	2	8231150	6102.4	19.0	1.0	31.9	48.7	14.5	14.7	17.9	3.32E+28	6.13E+27	1.79E+16
29	**Mock9_557_C 2p**	2	4252756	24255.9	32.2	1.0	18.7	28.2	19.8	12.4	19.1	2.98E+24	5.98E+23	3.99E+15
30	**NISTA_521_P 1p**	1	1863918	4540.6	30.9	1.1	32.1	45.8	40.1	13.0	16.0	4.35E+25	1.63E+25	4.74E+15
31	**NISTB_512_C 2p**	2	6136530	11447.8	21.2	1.1	28.8	89.4	87.2	18.2	29.0	4.01E+25	8.49E+24	1.75E+15
32	**NISTC_536_P 1p**	1	1500161	3963.8	30.5	1.1	19.5	66.8	25.1	13.0	16.7	6.50E+29	2.41E+29	2.10E+16
33	**NISTD_509_C 3p**	3	23287439	11678.0	49.9	1.2	34.3	35.2	21.0	12.7	17.0	(NISTA) 4.42E+25	5.78E+24	4.89E+15
												(NISTB) 2.31E+28	3.16E+27	8.17E+15