Forensic DNA Analysis

Single Source High Stutter Guidelines Assessment

Angela Adamson, Cassandra James, Emma Caunt

July 2021

Introduction

The purpose of this document is to:

A. Assess the risks involved in calling a single source profile a two contributor mixture and a two contributor mixture single source when potential high stutter peaks are present, by conducting hypothesis testing using VarNOC and DBLR™.

B. Based on the results/outcome of such testing, provide some recommendations on interpretation guidelines for DNA profiles of this type.

For a single source profile, if a high stutter is called allelic then it gives the false impression of an extra contributor. The risk of calling the profile a two contributor mixture is low since, in the absence of additional peaks, any comparison with this extra contributor is likely to provide a LR favouring non-contribution. However, if a reference sample shares a lot of information with the single source contributor, or with any stutter peaks present, then this LR could provide support for contribution when the second contributor doesn't actually exist.

For a single source profile where a peak in stutter position is actually allelic then the presence of an extra contributor can be missed. The risk of calling the profile single source is low since, in the absence of additional peaks any comparison with this extra contributor is likely to provide a LR favouring non-contribution.

It is considered that the presence of high stutters/masked alleles in a profile that otherwise appears single source has equal, but low, risk whether the profile is designated as single source or a two contributor mixture.

In order to help assess these risks we have carried out testing using real profiles with peaks in stutter position above threshold and created profiles with high stutter. These samples were tested using two systems – VarNOC and DBLR™.

Note: Stutter peaks refers to peaks found in +1 rpt and -1rpt positions.

VarNOC

VarNOC – Variable number of contributors is a feature of STRmix™ that allows the user to select a range of contributors. It will assess the profile under multiple number of contributor conditions and assign a weighting as to the most likely number of contributors. VarNOC is a system that has been validated for use by the STRmix™ Developers[1,3,4]. We utilised this tool to assess what weighting STRmix™ assigned to the number of contributors when they were analysed as one or two contributors.

DBLR™

DBLR™ (database likelihood ratios) is an application designed for the rapid calculation of likelihood ratios (LRs) using STRmix™ deconvolutions to aid forensic investigations. It can be used to calculate expected LRs for one or more components of forensic DNA profiles for true and non-contributors using randomly generated profiles. The propositions used to generate the LR for C2 are:

H1 = C2 has originated from the true contributor
H2 = C2 has originated from someone other than and unrelated to the true contributor.

If the true contributor and the true non-contributor both have the same expected LR for the second contributor (C2) in a two contributor mixture then it would not be possible to distinguish between a true contributor and a true non-contributor i.e. H1 ≈ H2

This application allowed us to assess the expected LR generated for true contributors and true non-contributors to C2 for samples that displayed high stutter but were interpreted as two contributor mixtures. The samples were deconvoluted as two contributor mixed profiles and the deconvolution imported into the DBLR™ program.

It should be noted that the Forensic DNA Analysis laboratory only had a limited time trial for use of this program and the testing was carried out within this time. DBLR™ has been validated for use by the STRmix™ Developers[2].

Results

The samples used and results generated using both VarNOC and DBLR™ can be found at I:\STRmix settings\3500 profile Issues\High stutter testing.

Table 1 – Results spreadsheet VarNOC and DBLR™

			VarNOC		DBLR™	
Barcode	Number of high stutters	% over threshold	Pr(1p)	Pr(2p)	Simulation outcome	Expected LR for true C2
[redacted]	2	53 and 16	8.05E-01	1.95E-01	H1 > H2	17
	1	15	9.99E-01	4.00E-04	H1≈ H2	1.1
	1	73	9.98E-01	1.70E-03	H1≈ H2	1.9
	1	14	9.99E-01	8.00E-04	H1≈ H2	1.1
	1	1.7	9.99E-01	1.10E-03	H1≈ H2	2.9
	1	1.70	9.94E-01	5.70E-03	H1≈ H2	1.8
	1	29	8.75E-01	1.25E-01	H1≈ H2	1.5
	1	2.3	9.99E-01	1.00E-03	H1≈ H2	2.7
	2	20 and 54	8.64E-01	1.35E-01	H1 >> H2	2200
	2	10 and 54	9.63E-01	3.70E-02	H1 >> H2	2000
	1	10	9.99E-01	1.95E-04	H1≈ H2	1.6
	2	1.3 and 25	9.99E-01	8.00E-04	H1≈ H2	2.4
	1	37	9.99E-01	2.00E-04	H1≈ H2	1.7
	1	12	9.98E-01	1.20E-03	H1≈ H2	1.3
	2	14 and 83	9.90E-01	9.81E-03	H1 > H2	54
	1	15	9.76E-01	2.42E-02	H1≈ H2	3.9
	1	9	9.58E-01	4.16E-02	H1 > H2	15
	1	30	1.00E+00	4.49E-04	H1≈ H2	2.9
	1	5	9.91E-01	8.78E-03	H1 > H2	6.3
	1	15	1.00E+00	2.23E-04	H1≈ H2	1.1
	1	3	1.00E+00	3.40E-04	H1≈ H2	1.4
	1	27	9.99E-01	1.16E-03	H1≈ H2	1.3
	1	40	9.99E-01	6.58E-04	H1 > H2	3.1
	1	2	9.99E-01	1.42E-03	H1≈ H2	2.1
	1	122	1.00E+00	3.45E-04	H1≈ H2	2.7
	1	27	9.98E-01	1.75E-03	H1≈ H2	1.2
	1	11	1.00E+00	1.75E-04	H1≈ H2	1.3
	1	1	1.00E+00	6.60E-05	H1 > H2	18
	1(2 runs)	23	9.98E-01	1.77E-03	H1 > H2	5.9
	2 (2 runs)	23 and 112	5.69E-01	4.31E-01	H1 > H2	794
	1(in 2 runs)	13 and 13	9.91E-01	8.87E-03	H1 > H2	14

VarNOC always assigned a higher probability to the DNA profile having one contributor rather than two contributors. See Table 1, column 4. In almost every case a probability over 0.95 was given to the profile consisting of one contributor. In the instances where the percentage was less than this, the sample contained multiple stutter peaks with one stutter 50-100% over the stutter threshold. There was one outlier to this, and it presented with a probability of 0.87 for the profile consisting of one contributor. This sample had a single

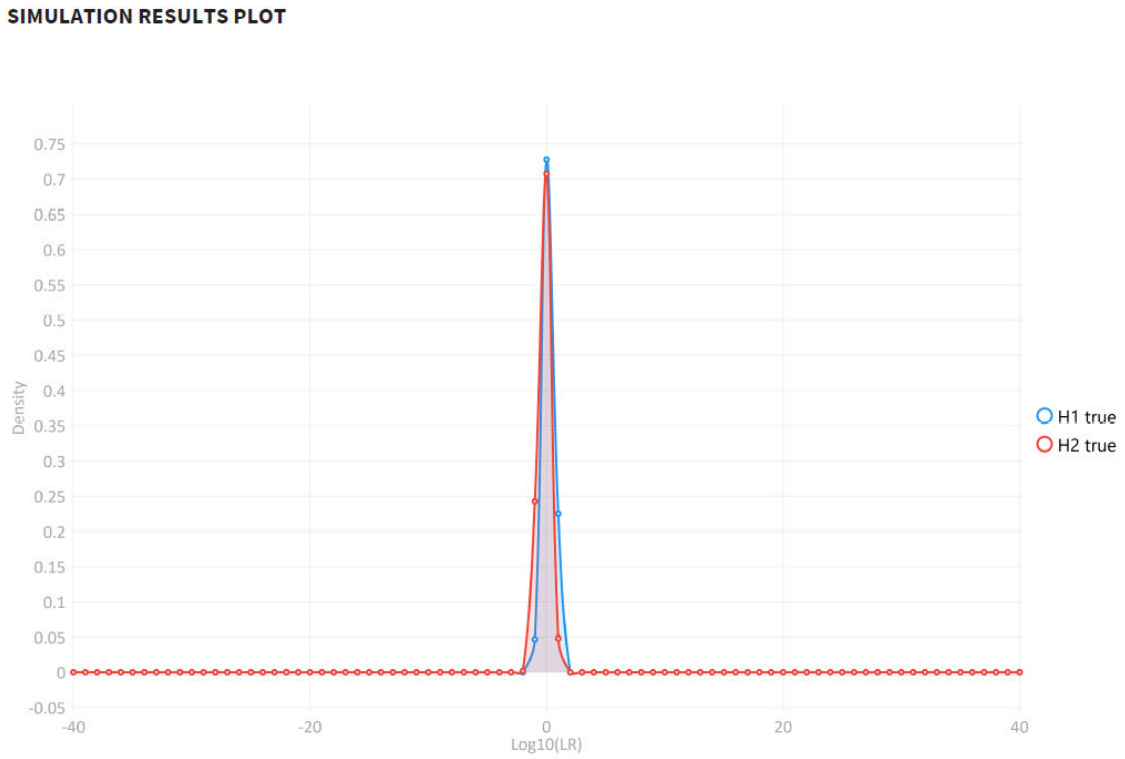
stutter value 29% above the stutter threshold. It was noted that, within the deconvolution of this sample, some stutter peaks below our laboratory stutter threshold had been modelled by STRmix™ as allelic a large proportion of the time and this may explain the probability of 0.87 for the profile having one contributor.

Using DBLR™ it was found that the results generated with almost all of the samples that displayed high stutter at one or more loci indicated that a true contributor and a true non-contributor of the minor/second contributor could not be differentiated and therefore it would be not be unexpected to obtain a similar LR if either was true i.e. H1≈ H2. The graphical representation of this can be seen in Figure 1. In a few cases it was found that they were not equal as demonstrated in Figure 2.

In most of the samples the expected LR of the true contributor to C2 was less than 20. It should also be noted that DBLR™ does not have theta or HPD considered within the generation of the LR so the calculated LR under casework conditions will be lower.

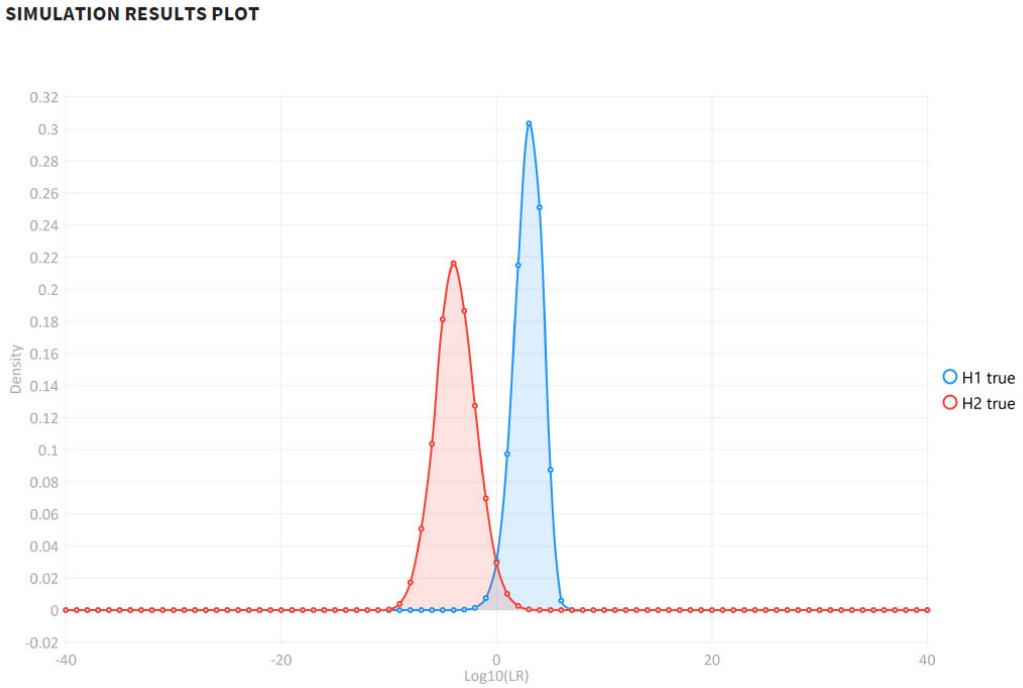
For the samples where H1 was much greater than H2 it was found that these samples had multiple high stutters with one of these stutters being extreme (up to and over 100% above threshold). One sample having the highest expected LR value of 2238.This sample showed that there is good ability to distinguish between a true contributor and true non-contributor for the minor/second contributor as seen in figure 3.

Figure 1



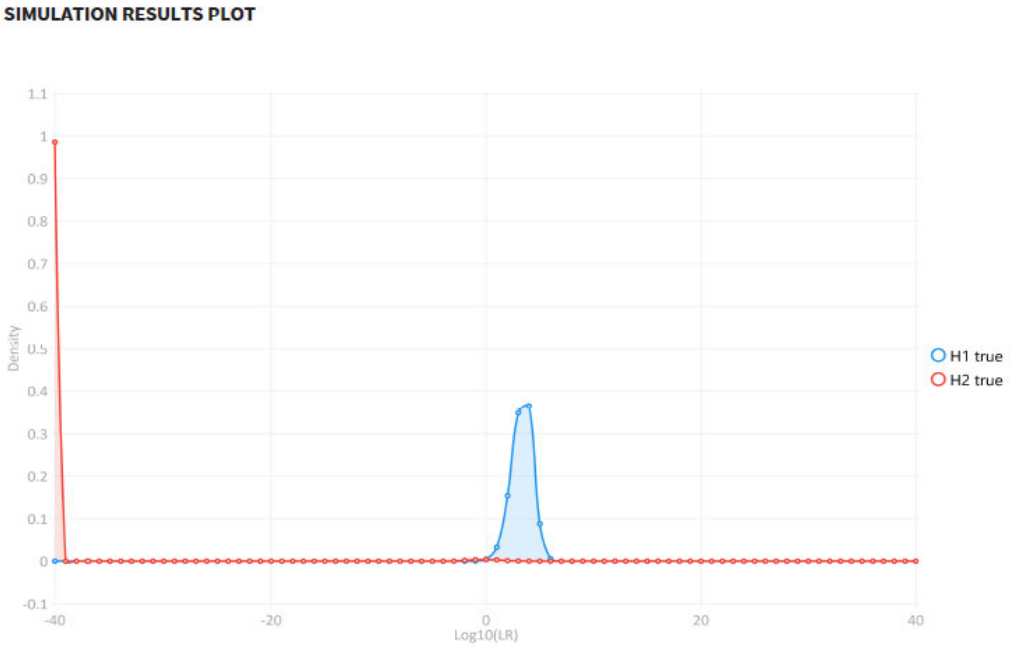
In Figure 1 the sample displayed has one high stutter peak 29% above the stutter threshold. This shows that H1 approximately equals H2.

Figure 2



In figure 2 the sample displayed has two high stutter peaks, one 23% above stutter threshold and the other 112% above stutter threshold. This shows that the H1 is greater than H2 however the overlapped area shows that it is still possible for a true contributor and true non-contributor to have the same LR.

Figure 3



In figure 3 the sample displayed has two high stutter peaks, one 20% above stutter threshold and the other 54% above stutter threshold. This shows that H1 is noticeably greater than H2.

Conclusion

The results obtained from both VarNOC and DBLR™ demonstrate that the presence of a peak in stutter position above threshold in an apparent single source profile does not suggest the addition of another contributor to the sample.

If these profiles were interpreted as two contributor mixtures, then the expected LR generated for the true second contributor would have very little evidentiary value due to the probability that the true contributor and a true non-contributor would have a similar LR.

For example, using DBLR™ the LR generated for the true second contributor in a two person mixed profile was 17, in an evidential context this value would have little meaning.

Samples tested with multiple peaks in stutter position above threshold demonstrated the same results as those samples with only one peak above threshold in stutter position and therefore can be interpreted as single source profiles. However, in samples tested with multiple peaks in stutter position above threshold and when one of these peaks are 50 to 100% over the stutter threshold it demonstrated an increase in the probability of the profile originating from two contributors rather than one and therefore should be interpreted as a mixed sample.

Recommendations

- Samples with a single peak in stutter position above threshold (labelled or unlabelled) can be interpreted as single source profiles.
- For the purposes of determining whether a peak in stutter position can be considered as high stutter, we recommend the use of the STRmix™ maximum allowable thresholds which are 30% for -1 rpt stutter and 10% for +1 rpt stutter. This means that a peak in stutter position can be considered to be high stutter up to 30% of the parent allele height for -1 rpt stutter and up to 10% of the parent allele height for +1 rpt stutter.
- Samples with multiple high stutters can be interpreted as single source, however if either of the high stutters are above the STRmix maximum allowable thresholds we recommend that theses samples are interpreted as mixed samples.
- High -2 rpt stutter is to be left labelled. STRmix™ is not modelling -2 rpt stutter but will model these peaks as drop in if they are below 250 RFU.
- These recommendations are for the determination of single source verses two contributor mixtures only. They are not intended for use for mixtures with greater than two contributors.

References

1. C. McGovern, K. Cheng, H. Kelly, A. Ciecko, D. Taylor, J.S. Buckleton, J.-A. Bright, Performance of a method for weighting a range in the number of contributors in probabilistic genotyping, Forensic Sci. Int. Genet.48 (2020)

2. DBLR V1.1 User's manual_1 Dec 2020

3. STRmix v2 6 Implementation and Validation Guide 1 August 2018 Queensland HFSS, Section 4.3 Variable Range of Contributors function (44-47)

4. STRmix™ V2.7 User's manual. 2.34 - Variable Range of Contributors function (106-115)

Additional Reading

Bright J-A, Taylor D, McGovern CE, Cooper S, Russell L, Abarno D, et al. Developmental validation of STRmix™, expert software for the interpretation of forensic DNA profiles. Forensic Science International: Genetics. 2016;23:226-39.

M. Kruijver, J.-A. Bright, H. Kelly, J. Buckleton, Exploring the probative value of mixed DNA profiles, Forensic Sci. Int. Genet. 41 (2019) 1-10.