Description of the likelihood ratio (LR) changes in STRmix™ v2.4.08

13 June 2017

Summary of changes:

Three changes have been made to the likelihood ratio (LR) calculations in STRmix™ resulting in v2.4.08. This updates the LR calculations in all versions preceding v2.4.08.

1. A change has been made to the LR calculation (unrelated point estimate, stratified, unified and HPD) for mixed DNA profiles when there are multiple contributors considered under H_p who are unknown under H_d. The contributors must each have dropped alleles (of either the same or different types) at the same locus. Examples of affected propositions include:

> H_{p1}: DNA has originated from POI 1 and POI 2

> H_{d1}: DNA has originated from two unknown individuals

> H_{p2}: DNA has originated from POI 1, POI 2 and POI 3

> H_{d2}: DNA has originated from three unknown individuals

> H_{p3}: DNA has originated from POI 1, POI 2 and POI 3

> H_{d3}: DNA has originated from POI 1 and two unknown individuals

This change relates to the order of the contributors within the LR calculation. The magnitude of the effect on LR calculations will depend on the proposition set used within the calculation and the number of per locus LR calculations undertaken where both contributors are dropping out of the mixture at the same locus. The effect is reduced when the number of loci of this specific configuration is reduced and when the difference in unknowns between H_p and H_d is reduced.

85 two-, three- and four -person low level mixed DNA profile interpretations were carried out in STRmix™ v2.4.06 and v2.4.08 in order to size the difference in LR. The log(LR) is compared in the plot below. The dashed line is at *x=y*. The dotted lines represented ± one order of magnitude. Different propositions were considered with varying numbers of contributors assumed under H_p (two, three or four). Forty calculations out of the 85 tested resulted in different LRs. The difference in all bar one example was within one order of magnitude. The largest difference was an LR=114 using v2.4.06 and LR=0.2 using v2.4.08. This was a known four person mixture interpreted assuming four contributors, analysed under conditions that were deliberately assigned to maximise the possible difference in LRs. There were a maximum of four peaks at any one locus and the maximum allele height was 225 rfu. The profile had previously been designated as an apparent three person mixture. Three of the four known contributors were assumed under H_p.
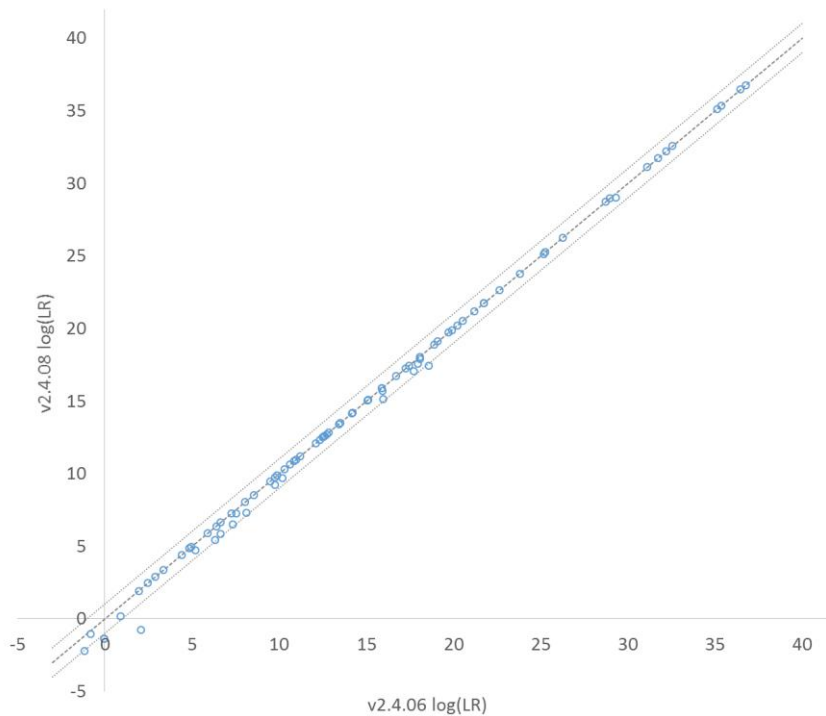
2. A change has been made to the way drop-in alleles are assigned during the determination of the genotype array within pre burn-in. This will affect all LR calculations including the database search module (standard and familial LRs), unrelated and related scenarios. Within v2.4.06 a peak in a double back stutter position (peak a in the figure to left) of an allele (peak c) that is less than the maximum back stutter parameter (percent) of the stutter peak (peak b) was assigned as 'stutter of stutter'. Within v2.4.08 this peak is now modelled as either allelic or drop-in. This change will only affect interpretations where drop-in is modelled ("turned on") within the kit.
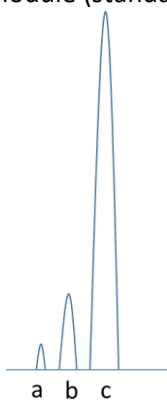
Within single source and mixed DNA profiles where the peak is solely modelled as drop-in, the drop-in penalty cancels under H_p and H_d and the LR will be numerically identical between versions. For profiles where the peak is modelled as both allelic and drop-in under various genotype combinations a difference in LR between versions is expected. A number of two-person profiles were edited by adding a peak in a double back stutter position of a major contributor's allele. The minor contributor was of a similar height to the added peak to allow consideration of the peak as both allelic and drop-in. A comparison of the log(LR) calculated using v2.4.06 and v2.4.08 is given in the plot below. The dashed line is at *x=y*. The dotted lines represented ± one order of magnitude. The comparison of LRs between versions indicates that this was predominantly within one order of magnitude.
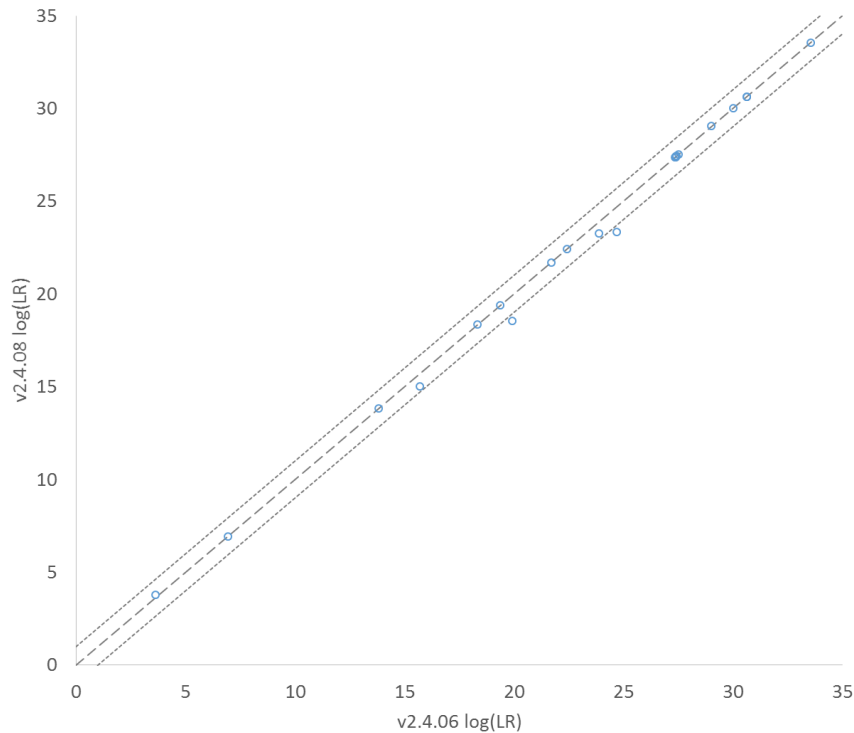
3. A minor anomaly in the familial search LR was identified. The issue affected mixed and single source profiles with dropout where on some occasions an incorrect allele frequency was assigned to the alleles. A comparison of familial LRs calculated using the database search functionality between versions is plotted below where the dashed line is at *x=y*. The dotted lines represented ± one order of magnitude. The comparison of LRs between versions indicates that this was mostly within one order of magnitude. This comparison included profiles with drop-in, dropout, and a combination of both drop-in and dropout modelled within the interpretation, with up to 4 contributors. The familial LRs were calculated for both known contributors and non-contributors to the profiles. There is more variability in the LR where more dropout is observed (approximately less than LR = 1000).