Dear STRmix Users,

This document provides further information regarding a minor miscoding of the Balding and Nichols formulae, detected in the calculation of likelihood ratios under certain circumstances, within STRmix V1.08, and all previous versions of STRmix. This affects how observed alleles are counted within the Balding and Nichols formulae. Specifically, STRmix has been 'double counting' person of interest alleles under H_1 under the following two conditions only:

1. Theta values are above 0
2. There are unknown contributors under the H_1 hypothesis.

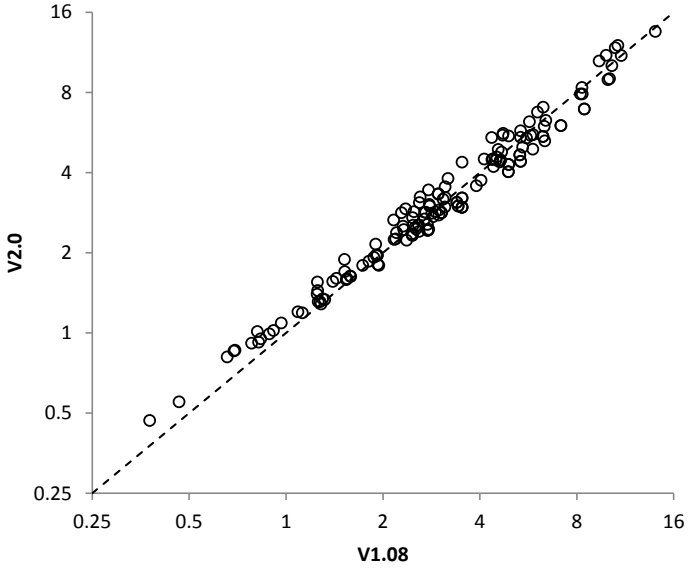
Both conditions must be met for the miscode to influence results, so does not affect database searching, or many case scenarios.

Specifically if the person of interest alleles are A_1A_2 and we seek to develop the probability of an unknown U = A_3A_4 under H_1 the sampling formula has been applied to $Pr(A_3A_4|A_1A_1A_2A_2)$ where A_1 may or may not equal A_2,A_3 and A_4 etc.

This miscode has been corrected in STRmix version 2.0, which is planned to be available very soon (end of January 2014).

A comparison of a range of profile types has been made in order to quantify the difference in LRs between V2.0 and V1.08. Generally, individual locus LRs decrease in V2.0 compared with V1.08 when there are shared alleles between the DNA contributors, and increase when there are no shared alleles.

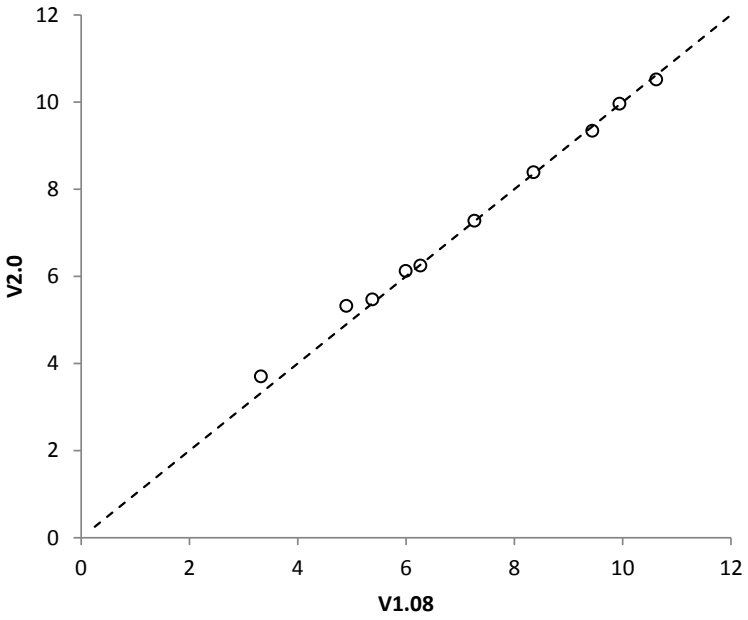
Identifiler data
Five two person and five three person Identifiler mixtures with a range of input DNA were compared. Calculations were performed using a stratified New Zealand database and a theta of 3%. A plot of the locus *LR*s for the 150 comparisons in V2.0 versus V1.08 is provided below.



The diagonal dotted line indicates x=y. Whilst there are individual loci above and below this line, this plot indicates, showing that the miscode has very minor influence on the LR in V1.08.

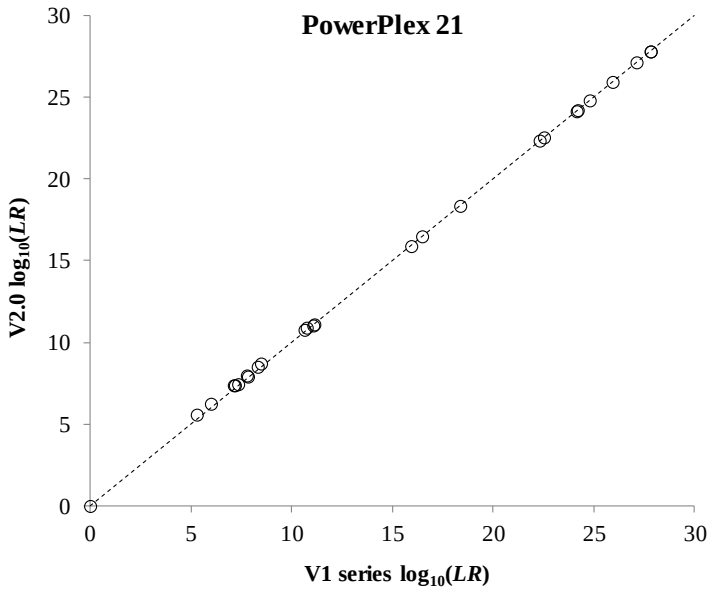
From theory we do not expect to see a systematic effect in either direction and this appears to be the empirical result.

The graph below shows the effect across the 15 loci of Identifiler (log scale).



PowerPlex 21

The graph below shows the effect across the 21 loci of PowerPlex 21 (log scale). Profiles consisted of 2, 3 and 4 person mixtures and a range of input DNA levels. LRs were calculated using a Caucasian database and a theta of 0.01.



The other improvements made from V1.08 to V2.0 include:

1) LSAE are penalised differently. They now are all individually considered as having a prior of $\log[A_i] \sim N(0, c^2)$ where c^2 is the LSAE variance determined from model maker. Each amplification efficiency is then provided a likelihood based on this prior.
2) Model Maker now optimises stutter, allele and LSAE variance for each profile individually. These values are weighted by the number of datapoints present in the profile. Any profiles with less than 10 datapoints are not included in the final curve fitting. Any profiles with peaks above the saturation level (from default settings in STRmix) are not included in the final curve fitting. Once optimised, weighted and pruned each set of variance values have a gamma curve $\Gamma(\alpha, \beta)$ fitted with the criteria:
    - $\alpha > 1.5$ (to prevent exponential curves being fit)
    - the mean of the gamma distribution must equal the mean of the weighted variances

    All fitted values and their gamma distributions are provided in the final output.
3) STRmix now defaults to view the start of the view results screen rather than the end
4) Name and identifiers for samples within a searchable database may now contain letters 'X' and 'Y', which previously would cause errors.
5) Choosing extended output will now provide that output for all chains pre and post burnin
6) Drop-in prior is now posterior allele probability and not conditional probability
7) Added 'factor of N' LRs to output
8) Updated 'About' text
9) Updated dropout generating algorithm to use normal approximation
10) Default for Model Maker are 100 000 accepts and a correlation penalty SD of 0.005

For further information please contact a member of the STRmix Development Team below.

Jo-Anne Bright, Duncan Taylor and John Buckleton