Discovered bug in EuroForMix v3.0.0 – v3.2.0.

A bug was recently discovered when all the alleles in the defined allele-frequency table are also observed in the evidence profile (for one of the markers), and at the same time a stutter model (BW or/and FW) is assumed.

This situation causes no definition of the Q-allele* which then leads to an "indexing-bug": The reason is that the Q-allele is defined to not obtain or give away any stutter, solved by applying a dummy index (-1) to this allele. However, when the Q-allele was not defined, the program still assigned the dummy index when it should not have. The implication of this was that the stutter index definition for all the remaining alleles in the profile would be shifted by one, and hence wrongly assigned. This leads to a model where the stutter model does not fit the data, and because of this the LR of a person of interest would be drastically reduced.

The conditions which trigger the bug are rarely observed since this would rely on the following (on a per-marker basis):

1) A stutter model is considered (Backward or/and Forward).
2) The evidence profile is a complex mixture (at least 4p) where stutter filter is not applied.
3) The allele outcome in the frequency table is rather small, such that it only includes all the alleles of the evidence profile.

Example: The bug was discovered in a complex 4-profile (part of ProovedIT dataset) where all stutters were kept for evaluation: *H09_RD14-0003-48_49_50_29-1;4;4;4-M2a-0.75GF-Q0.4_08.25sec.hid*

The calculation of the LR were based on the hypotheses Hp: 'K48 + 3Unknowns' vs Hd: '4 unknowns', where degradation and both stutter models were considered.

The issue was observed for marker D16S539 which consisted of the following alleles: "8/9/10/11/12/13/14". The allele frequencies also contained these alleles, but nothing more, causing the Q-allele to not be assigned. The LR values for the markers at D16S539 and "those before" were greater than one, whereas the remaining ones where close to one or lower, with a joint LR much less than one (LR=0.02). After fixing the bug issue the result was LR=6.2e14, with LR-per markers mostly being greater than one.

Numerical tests:

The example above is now included in the automated numerical tests together with a simple single-source profile example to hinder the same problem happening in the future.

Code locations (v3.3.0):

- prepareC:L244-L245; The dummy index "-1" should not be included when the last allele is not a Q-allele.

- Adding 2 blocks to src/calcloglik:(L350-357 and L777-784) to ensure that the last allele is considered when not a Q-allele.

A Q-allele represents all non-observed alleles which are defined in the allele outcome (defined from frequency database), and is considered as a "compound allele" of these.