1 Supplementary material

The linkage maps for all 21 \textit{H. erato} chromosomes are shown in Figures 1 and 2. Chromosomes 4, 5, 7, and 15 show indel artefacts where marker number alternates between two positions, whereas in chromosomes 1, 4, 5, 8, 16 and 20 there are some minor errors in the genome assembly that appear as spikes.

We also run MSTmap (Wu et al., 2008) on the \textit{erato} data. We took the genotype likelihoods outputted by ParentCall2 and took only the same markers that were assigned to chromosome 1 with Lep-MAP3. First, we picked three subsets of best quality markers, classified by the sum of genotype likelihoods for each individual. Note that the sum is 1 for a fully certain genotype and 2 for a completely missing one. We got about 45000, 15000 and 5000 markers by requiring the sum to be below 96.5, 96.25 and 96 (out of 95 individuals), respectively.

The genotypes were called with certainty 1/100 or 1%. Using parameter \text{cut\_off\_p\_value}=0.000001, the MSTmap constructed two linkage groups, one for both phases. The correlation of physical and linkage position on the larger linkage group is shown in Figure 3. The correct map length is about 59 cM, whereas the MSTmap’s result increases along the number of used markers (from about 200 cM to 1500 cM).

The results show that for very high quality data, MSTmap produces reasonable results (marker separation and genotype calling are relying on Lep-MAP3 modules, the entire dataset is too large to run on MSTmap). However, when there are multiple erroneous markers, the maps becomes too inflated to use it for example contig orientation. Moreover, certain regions (e.g. at 4Mb, 10Mb and 18Mb) show that the map order is not locally correct for the largest dataset (45000 markers).

References

Fig. 1: H. erato chromosomes 1-12
Fig. 2: H. erato chromosomes 13-21
Fig. 3: Linkage mapping results on MSTmap on H. erato chromosome 1, taking initially about 5000, 15000 and 45000 best markers. About 2/3 of the markers are included in the figure as being in the same phase. The correct map distance is about 59 cM (see Figure 1).