

Effect of Ignoring Linkage Information in Likelihood Computation

Table 1 shows the pairwise prediction accuracy for Simulation A, where the likelihoods were computed without taking into account the linkage information between the markers [1]. For sibling relationships, the accuracy rates remained as high as when linkage information was used in the likelihood computation. For first cousins, however, the accuracy rates decreased significantly. For example, about a quarter of full cousin pairs were classified as half cousins and about 50 percent of half cousins were classified as unrelated. Furthermore, N_e was overestimated (Figure 1), which is consistent with the fewer number of cousin pairs that were estimated in the pedigrees.

Table 1: **Pairwise Prediction Accuracy**

		Predicted				
		FS	HS	UR	FC	HC
True	FS	106	0	0	0	0
	HS	0	136	0	1	0
	UR	0	1	600000	0	0
	FC	1	0	0	374	103
	HC	0	0	246	3	281

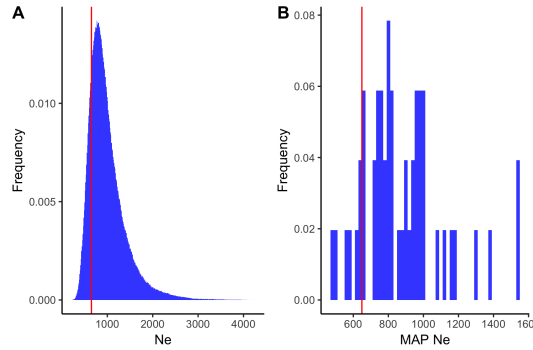


Figure 1: (A) Estimated posterior distribution of N_e from MCMC samples aggregated over 50 datasets in Simulation A, where likelihoods were computed without linkage information between markers. (B) Distribution of MAP N_e for the 50 datasets.

References

- [1] Bruce S Weir, Amy D Anderson, and Amanda B Hepler. Genetic relatedness analysis: modern data and new challenges. *Nature Reviews Genetics*,

7(10):771–780, 2006.