

20 **2 Detail of simulations**

21 **2.1 Simulation plan**

22 We simulated datasets from the defaults prior-likelihood model assumed in Geneland (assuming
23 spatially and genetically structured populations, see (Guillot *et al*, 2005; Guillot, 2008; Guillot
24 *et al*, 2010) for details). Each data-set consisted of $n = 100$ individuals belonging to one of $K = 2$
25 clusters. We investigated various levels of differentiations between the two clusters reported by
26 the pairwise F_{ST} value. We considered genotypes at bi-allelic loci for dominant then codominant
27 markers. We investigated genotypes at $L = 10, 20, 50, 200, 500$ markers. Each value of L was
28 investigated through a batch of $N = 200$ independent datasets.

29 **2.2 Setting for MCMC computations**

30 For each data-set, we launched a single MCMC run of 200000 iterations with a thinning of 200 and
31 we discarded 100000 iterations for burn-in. We treated the number of clusters as unknown setting
32 the maximum number of clusters to $K_{max} = 10$. Each MCMC run was post-processed according
33 to the procedure described by Guillot (2008) to get of potential label switching issues.

34 **2.3 Detail of simulation results**

35 [Figure 1 about here.]

36 [Figure 2 about here.]

37 References

- 38 Guillot G (2008) Inference of structure in subdivided populations at low levels of genetic differentiation. The correlated allele frequencies
39 model revisited. *Bioinformatics*, **24**, 2222–2228.
- 40 Guillot G, Estoup A, Mortier F, Cosson J (2005) A spatial statistical model for landscape genetics. *Genetics*, **170**, 1261–1280.
- 41 Guillot G, Santos F, Estoup A (2010) *Geneland program documentation*. Technical University of Denmark.
42 www2.imm.dtu.dk/~gigu/Geneland/Geneland-Doc.pdf.
- 43 Wasser S, Shedlock A, Comstock K, Ostrander E, Mutayoba B, Stephens M (2004) Assigning African elephants DNA to geographic
44 region of origin: applications to the ivory trade. *Proceedings of the National Academy of Sciences*, **101**, 14847–14852.

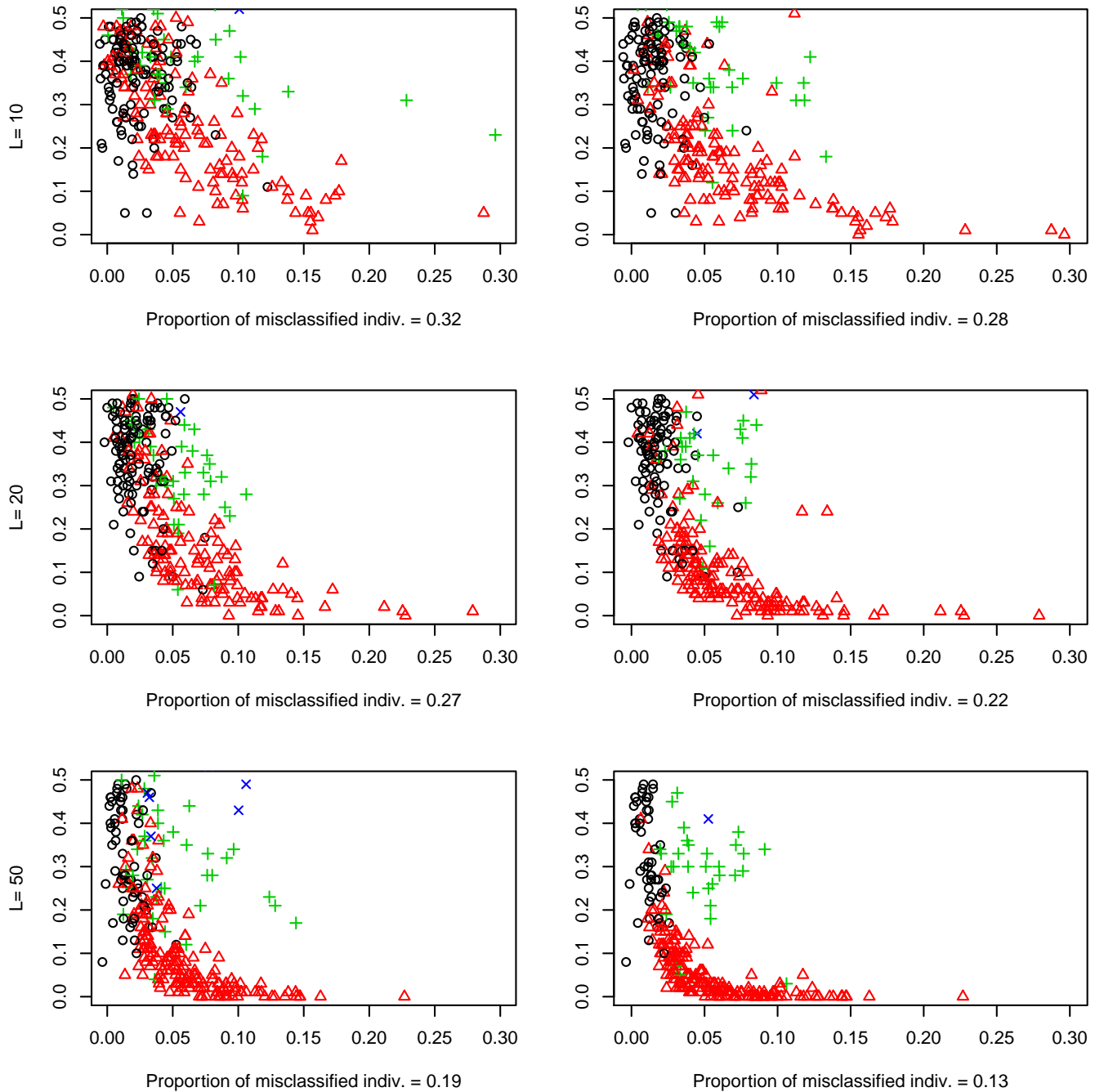


Figure 1: Clustering error (y-axis) as a function of pairwise F_{ST} (x-axis), for datasets simulated with various values of the number of loci L from the default GENELAND prior-likelihood model. Each point represents a data-set consisting of $n = 100$ individuals belonging to $K = 2$ clusters with genotypes at unlinked bi-allelic loci. The colour and shape of the symbol represent the number of clusters inferred (\circ : one cluster, \triangle : two clusters i.e. correct result, $+$: three clusters, \times : four clusters, \diamond : five clusters). Left column: dominant markers, right column: codominant markers.

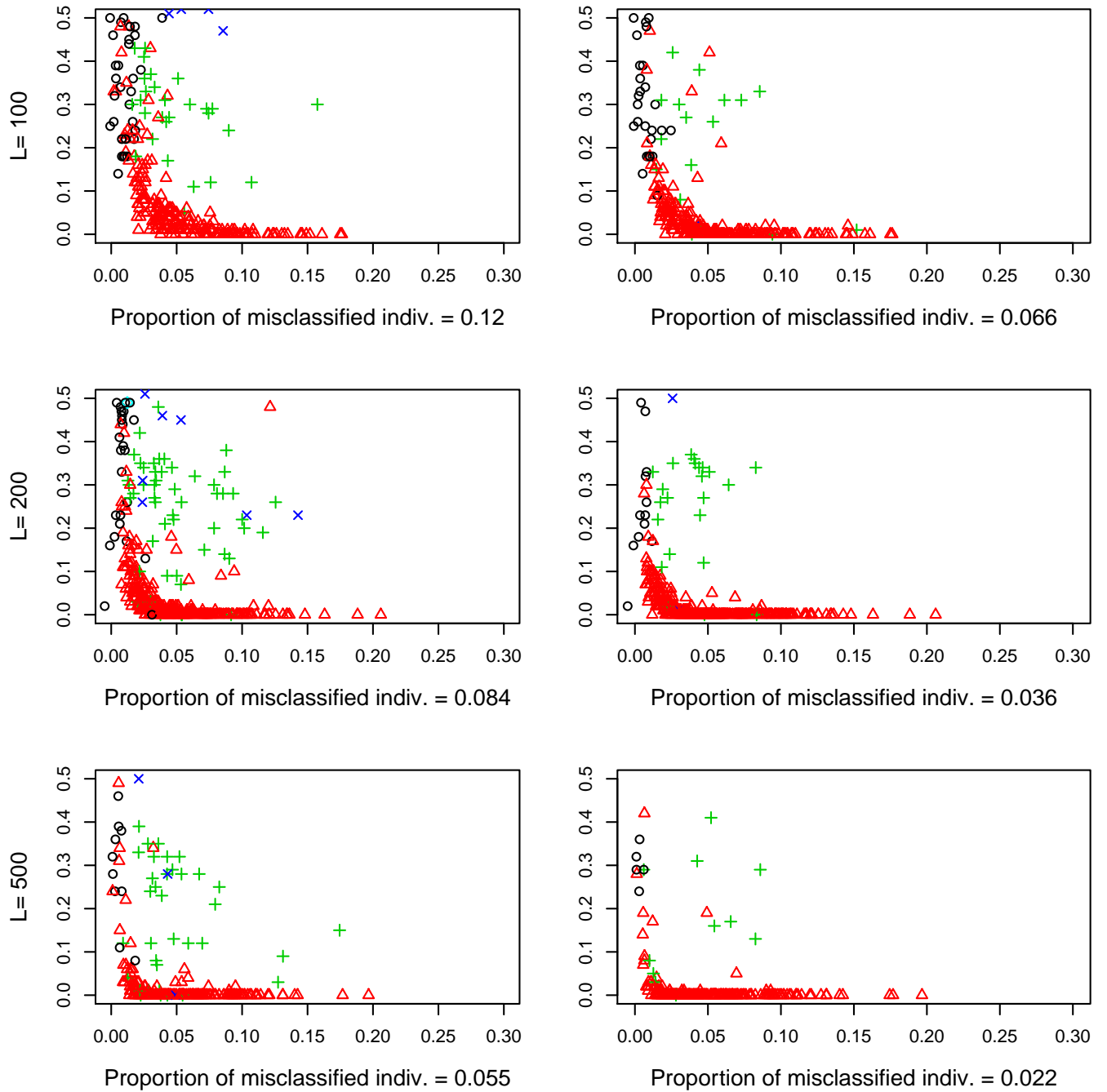


Figure 2: Clustering error (y-axis) as a function of pairwise F_{ST} (x-axis), for datasets simulated with various values of the number of loci L from the default GENELAND prior-likelihood model. Each point represents a data-set consisting of $n = 100$ individuals belonging to $K = 2$ clusters with genotypes at unlinked bi-allelic loci. The colour and shape of the symbol represent the number of clusters inferred (\circ : one cluster, \triangle : two clusters i.e. correct result, $+$: three clusters, \times : four clusters, \diamond : five clusters). Left column: dominant markers, right column: codominant markers.