Segmentation of maize actin genes

July 15, 2008

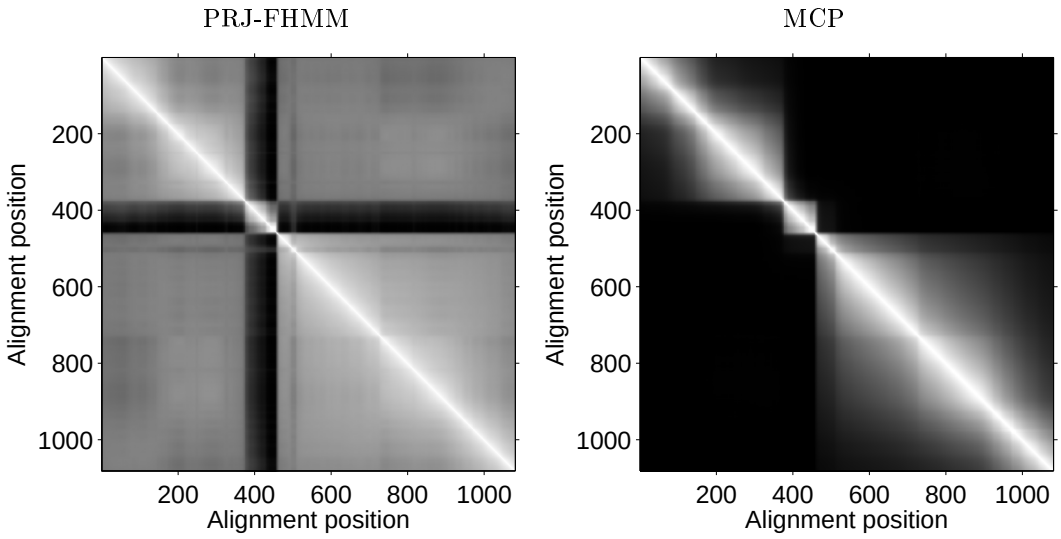

Figure 1: The posterior probability that two alignment positions in the maize actin gene sequence alignment are in the same rate state. The layout is identical to Figure 6 in the paper.

We have applied our method to four maize actin gene sequences with the following GenBank/EMBL accession numbers: U60514, U60513, U60508, and U60507. We used the same alignment as in (Husmeier, 2005), and detected the same gene conversion event as found in this earlier study, in confirmation of the finding reported in Moniz de Sa and Drouin (1996). The new result we have obtained is shown in the left panel of Figure 1, which shows the posterior probability that two alignment positions are in the same rate state.

Our method clearly detects an outlying region between sites 370 and 475. As it turns out (Lehrach, 2007), this region constitutes an inadvertently included intron, which owing to the lack of selective pressure is significantly more diverged than the coding regions by which it is flanked. Note that the corresponding plot obtained for the MCP model, shown in the right panel of Figure 1, cannot detect

this outlier, as it cannot be distinguished from an alignment of, say, three genes or three differently diverged coding regions. For further details, see Lehrach (2007).

References

Husmeier, D. (2005). Discriminating between rate heterogeneity and interspecific recombination in DNA sequence alignments with phylogenetic factorial hidden Markov models. *Bioinformatics 21*, ii166–ii172.

Lehrach, W. P. (2007). *Predicting protein-protein interactions and characterising rate hetrogeneity along DNA sequence alignments.* Ph. D. thesis, University of Edinburgh, School of Informatics.

Moniz de Sa, M. and G. Drouin (1996). Phylogeny and substitution rates of angiosperm actin genes. *Molecular Biology and Evolution 13*, 1198–1212.