

## **Supplementary material**

### **Additional power simulations**

Figures 1, 2, and 3 show the results of increasing the number of codons and the number of species on the power of the SLR and SNY tests under the conditions A, B and C described in the main text. Each graph consists of the compounded results of 100 sets of data, analyzed separately, for which a prior LRT between M8A and M8B indicates the presence of significant positive selection. Sets of data were generated until there were 100 with significant positive selection, all those not meeting this criteria being discarded.

### **Power of SLR test without prior LRT**

The ROC curves for the SLR test, shown in the main text, were produced from 100 sets of data which had significant evidence of positive selection (as indicated by a LRT between M8A and M8B). It is not necessary to filter the data in this fashion before applying the SLR test and figure 4 compares its performance on filtered and unfiltered data. The ROC curve for the SLR test after a prior LRT is identical to that in the main text; the curve for “without prior LRT” is based on one hundred sets of data generated under the appropriate conditions (see main text for details). Under all three conditions analyzed the power of the SLR test drops but not substantially so.

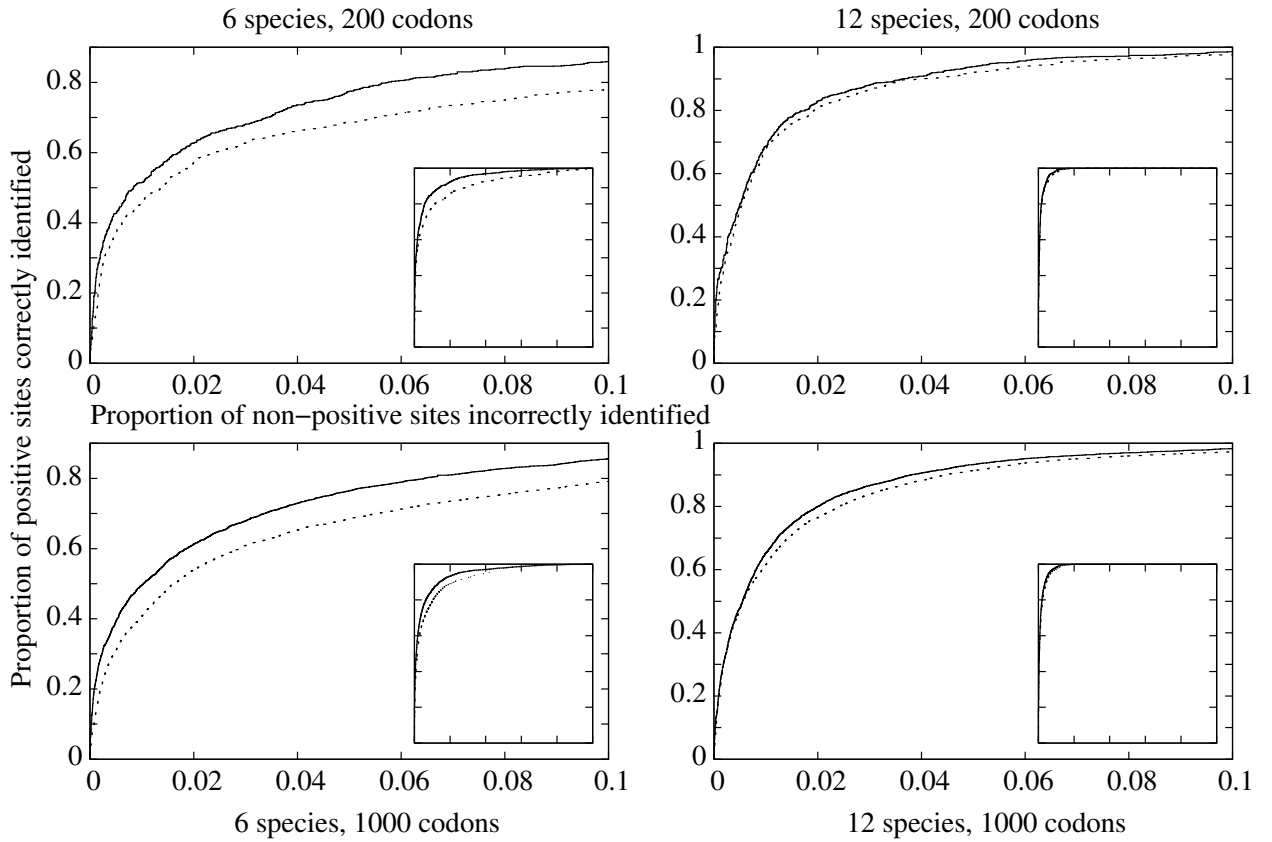


Figure 1: ROC curves for proportion of sites correctly and incorrectly identified under selection conditions (A), described in the main manuscript. The six and twelve species trees as show in the paper, figure 1. The main curves show the region with a false positive rate  $< 10\%$ ; inset are the complete curves (axes unmarked, but both between 0 and 1). Solid lines: SLR. Dotted lines: SNY.

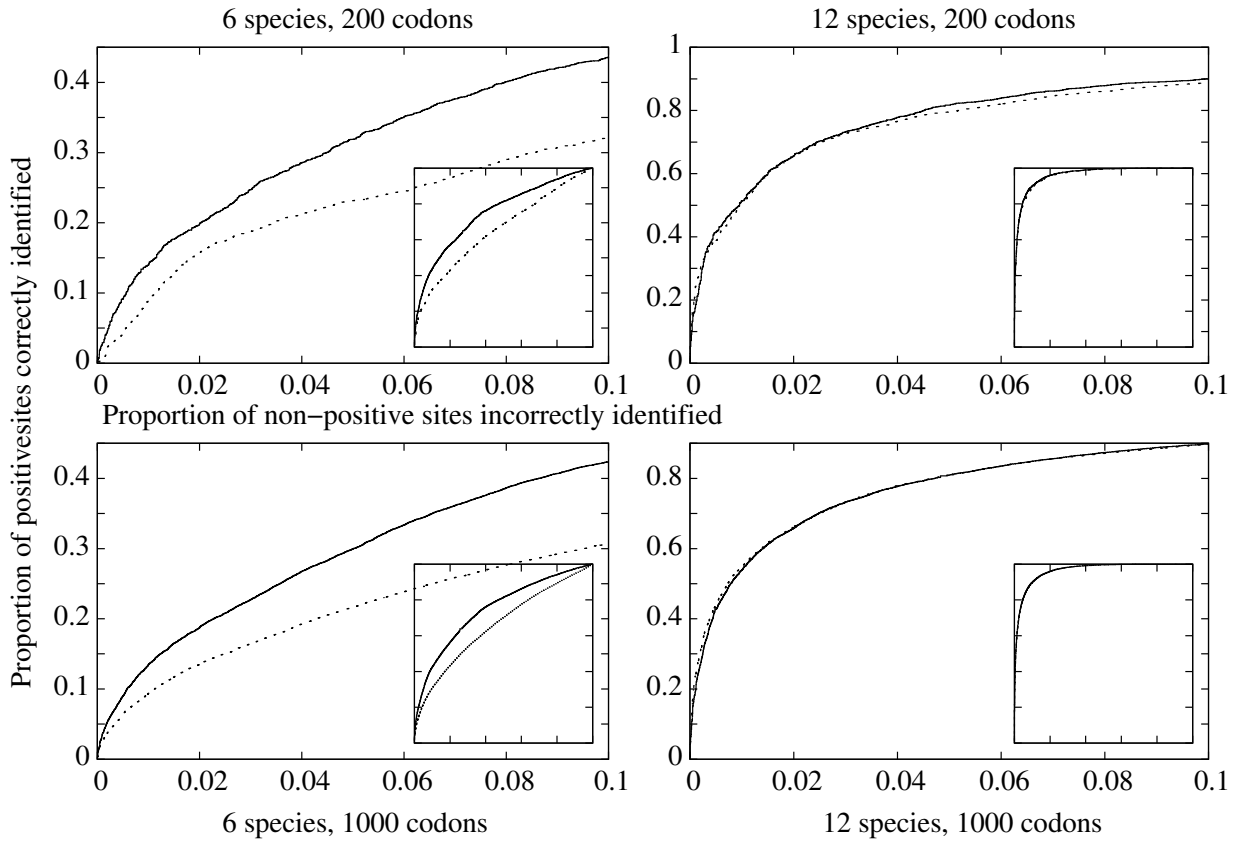


Figure 2: ROC curves for proportion of sites correctly and incorrectly identified under selection conditions (B), described in the main manuscript. The six and twelve species trees as show in the paper, figure 1. The main curves show the region with a false positive rate  $< 10\%$ ; inset are the complete curves (axes unmarked, but both between 0 and 1). Solid lines: SLR. Dotted lines: SNY.

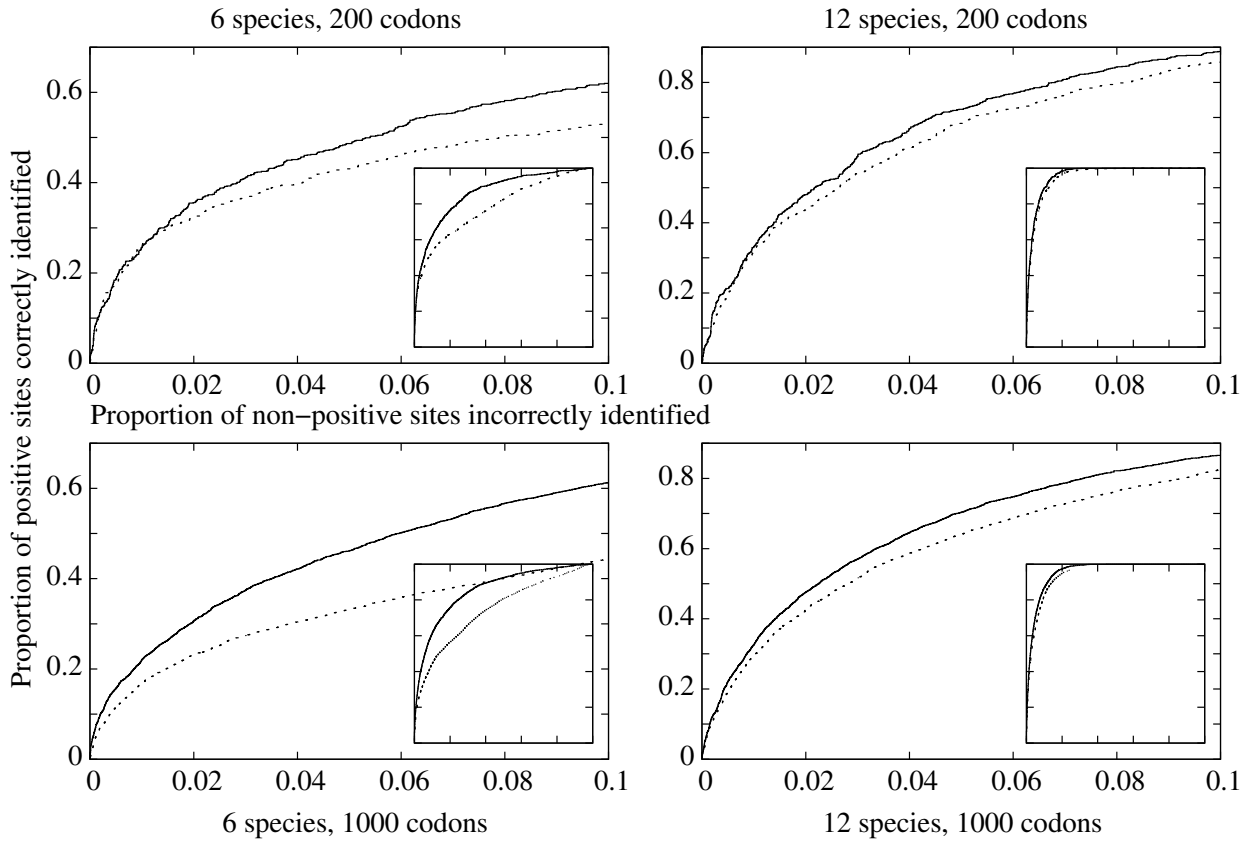


Figure 3: ROC curves for proportion of sites correctly and incorrectly identified under selection conditions (C), described in the main manuscript. The six and twelve species trees as show in the paper, figure 1. The main curves show the region with a false positive rate < 10%; inset are the complete curves (axes unmarked, but both between 0 and 1). Solid lines: SLR. Dotted lines: SNY.

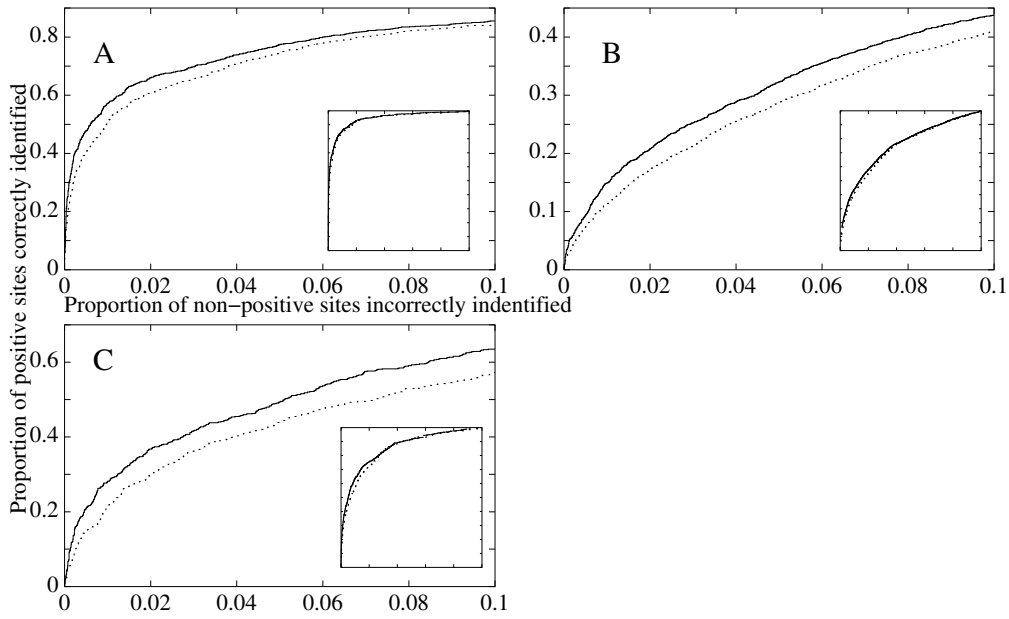


Figure 4: ROC curves for the proportion of sites correctly and incorrectly identified as positively selected in various simulation experiments. The conditions for the simulations (A, B and C) are described in the main text. The main curves show the region with a false positive rate  $< 10\%$ ; inset are the complete curves (axes unmarked, but both between 0 and 1). Solid lines: SLR with prior LRT. Dotted lines: SNY without prior LRT.