Figure 4.2: The effect of recombination on inferring diversifying selection. Reconstructed evolutionary history of codon 516 of the Cache Valley Fever virus glycoprotein alignment is shown according to GARD inferred segment phylogeny (left) or a single phylogeny inferred from the entire alignment (right). Ignoring the confounding effect of recombination causes the number of nonsynonymous substitutions to be overestimated. A fixed effects likelihood (FEL, Kosakovsky Pond and Frost (2005)) analysis infers codon 516 to be under diversifying selection when recombination is ignored ($p = 0.02$), but not when it is corrected for using a partitioning approach ($p = 0.28$).

Accounting for recombination

- First screen the alignment to find putative non-recombinant fragments (e.g. using GARD)

- Apply a model-based test (MEME, FUBAR) using multiple phylogenies (one per fragment), but inferring other parameters (e.g. nucleotide substitution biases and base frequencies) from the entire alignment

- This has been shown to work very well on simulated and empirical data

- This is the approach does not work for analyses assuming a single tree (BUSTED, aBSREL).