

However, we caution that despite obvious interest in identifying specific branch-site combinations subject to diversifying selection, such inference is based on very limited data (the evolution of one codon along one branch), and cannot be recommended for purposes other than data exploration and result visualization. This observation could be codified as the “***selection inference uncertainty principle***” — one cannot simultaneously infer both the site and the branch subject to diversifying selection. In this manuscript [MEME], we describe how to infer the location of sites, pooling information over branches; previously [aBSREL] we have outlined a complementary approach to find selected branches by pooling information over sites.

Murrell et al 2012

Purpose-built models

- It is tempting to “hack” existing tools to answer questions that they are not designed to answer
- A recent example we tackled is a rigorous test for relaxation of selection (or more generally a difference in selective regimes) in a part of the tree, relative to the rest of the tree
- Typical approaches have been to estimate dN/dS ratios from two sets of branches, and interpret an *elevation* in dN/dS as evidence of selective constraint relaxation
- Two problems with this approach
- An increase in mean dN/dS could also be caused by an **intensification** of selective forces.
- *Post-hoc* analyses (e.g., estimate branch-level dN/dS and then compare [t-test, etc] them as if they were observed quantities) discard a lot of information (e.g., variance of individual estimates), and make obviously wrong assumptions (e.g., estimates are uncorrelated).