

# 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).

Reference Branches

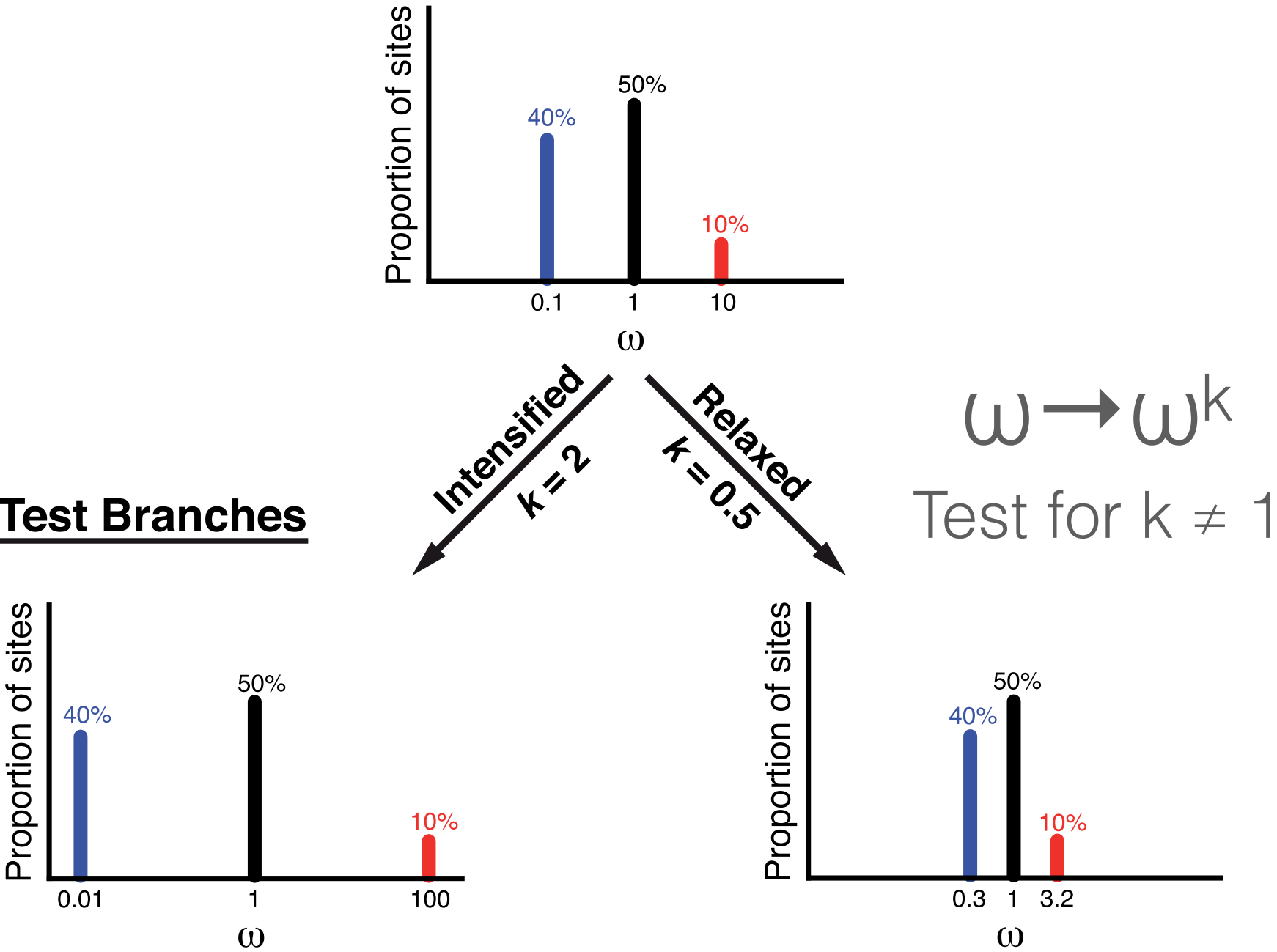


Table 1. Test for Relaxed Selection Using RELAX in Various Taxonomic Groups.

Taxa	Gene/Genes	Test Branches	Reference Branches	<i>k</i> <sup>a</sup>	<i>P</i> -Value
<i>γ</i> -proteobacteria	Single-copy orthologs	Primary/secondary endosymbionts	Free-living <i>γ</i> -proteobacteria	0.30	< 0.0001
		Primary endosymbionts	Free-living <i>γ</i> -proteobacteria	0.28	< 0.0001
		Secondary endosymbionts	Free-living <i>γ</i> -proteobacteria	0.61	< 0.0001
		Primary endosymbionts	Secondary endosymbionts	0.56	< 0.0001
Bats	SWS1	HDC echolocating and cave roosting (pseudogenes)	LDC echolocating and tree roosting (functional genes)	0.16	< 0.0001
		LDC echolocating	Tree roosting	1.07	0.577
		HDC echolocating and cave roosting	LDC echolocating and tree roosting	0.70	0.495
	M/LWS1	Echolocating species	Tree- and cave-roosting species	0.21	0.0005
		HDC echolocating	LDC echolocating	0.84	0.427
		Endogenous viral elements	Exogenous virus	0.02	< 0.0001
Bornavirus	Nucleoprotein	Asexual	Sexual	0.63	< 0.0001
<i>Daphnia pulex</i>	Mitochondrial protein-coding genes				

<sup>a</sup>Estimated selection intensity.