

Codon substitution model:

Goldman-Yang (1994) → Muse-Gaut 1994 very similar

Characters are now codons. There are 64 (or 61) of these

$$P_{xy} = \begin{cases} 0 & \text{if } x \text{ and } y \text{ differ by more than one nucleotide} \\ \pi_y & \text{if } x \text{ and } y \text{ differ by synonymous transversion} \\ \kappa \pi_y & \text{if } x \text{ and } y \text{ differ by synonymous transition} \\ \omega \cdot \pi_y & \text{if } x \text{ and } y \text{ differ by nonsynonymous transversion} \\ \omega \cdot \kappa \cdot \pi_y & \text{if } x \text{ and } y \text{ differ by nonsynonymous transition} \end{cases}$$

Where do π_y come from? F1X4, F3X4, empirical vs. estimated

How do we interpret ω ? As dN/dS.

If ω is > 1 , gene is under positive selection.

Do we expect $\omega > 1$ for the average of all sites?
Not usually.

Testing for specific sites under positive selection.

I will describe one method (FEL).

Fit the entire tree and all model parameters using Goldman-Yang model. Then fixing everything else, test $\omega_r = 1$ versus $\omega_r \neq 1$ for each

r. How do we test if $\omega_r \neq 1$?

Likelihood ratio test. Multiple hypothesis testing.

Try this with DataMonkey.

