Supporting Text

Matrix of synonymous substitutions. For each gene, the matrix of the probabilities of synonymous substitutions for all pairs of nucleotides was constructed using only the edges of the tree with less than 0.05 synonymous substitutions per codon site. For each such edge, we compared the sequences at the beginning and at the end of the edge at fourfold degenerate synonymous sites. The matrix was then constructed from the numbers of the observed synonymous substitutions at such sites, divided by the lengths of the corresponding edges.

Codon-specific opportunity for substitution. The opportunity for nonsynonymous substitution *o(c)* for a codon *c* is the number *J(c)* of one-step neighbors of *c* that encode a different amino acid, each taken with the weight *m*, which is the probability of the corresponding nucleotide substitution taken from the matrix of synonymous substitutions:

$$o(c) = \sum_{j=1}^{J(c)} m(j)$$

The opportunity for synonymous substitution was defined analogously. When multiple synonymous or nonsynonymous substitutions occurred between two successive nodes, these substitutions generally had different opportunities. In such cases, the order of the substitutions was reconstructed, and the opportunity for each substitution was inferred as described. If the order of substitutions could not be reconstructed unambiguously, the opportunity for each of the substitutions was estimated by averaging over the possible orders of substitutions.

Estimating site-specific d_N/d_S values. For each codon site, the total opportunity of nonsynonymous mutation *O* was defined as

$$O = \sum_{k=1}^{N} \frac{\sum_{c=1}^{C_k} o(c)/C_k}{a_k},$$

where *N* is the total number of edges in the tree, a_k is the length of the *k*th edge, and C_k is the inferred number of different codons that existed in the *k*th edge. The effective number of nonsynonymous substitutions was then defined as the inferred number of nonsynonymous substitutions at this site, divided by the total opportunity for nonsynonymous mutation. The effective number of synonymous substitutions was determined analogously. The d_N/d_S value at a site was then calculated as the ratio of the effective numbers of nonsynonymous substitutions at this site to the average effective number of synonymous substitutions per site at this gene.