TOBY KENNEY, Dalhousie University

A Generalized Codon-based Model of Nucleotide Substitution for Protein-coding DNA Sequences

A generalized codon-based model for protein-coding DNA sequences is proposed for phylogenetic analysis. This model framework provides a unified framework for existing codon (or DNA or amino acid) models. Furthermore, it offers greater flexibility in the choice of rate matrix, allowing existing models to be easily extended to incorporate more of the possible driving forces in molecular evolution, such as structure information and amino acid properties. We provide a software package called Codon Optimal Likelihood Discoverer (COLD) to implement these proposed generalized codon models. We demonstrate how our model framework allows model selection based on standard likelihood theory.