

# **“Quantifying Optimal Level of Sequence Divergence for Phylogenetic Reconstruction”**

**Robert Makowsky, Ph.D**

University of Texas- Arlington,  
Quantitative Biology

## **ABSTRACT**

Determining the appropriate gene for phylogeny reconstruction can be a difficult process. Rapidly evolving genes tend to resolve recent relationships, but suffer from alignment issues and increased homoplasy between distantly related species. Conversely, slowly evolving genes work well for deeper relationships, but lack sufficient variation to resolve recent relationships. We determine the relationship between sequence divergence and Bayesian phylogenetic reconstruction ability using a known phylogeny and simulated datasets. Across all genes we found that an optimal divergence for resolving the correct relationships does exist. However, gene type (mitochondrial protein, mitochondrial ribosomal, nuclear protein, and nuclear ribosomal) was found to be an important factor, so the type of gene being considered is an important consideration during phylogenetic analyses. Simulations show that the general pattern exists regardless of how the sequences evolved with respect to parameters or topology. We determine that a simple to measure property of genetic sequences (genetic distance) is highly correlated with phylogenetic reconstruction ability in Bayesian analyses. This information should be useful for selecting the right gene to clarify intractable relationships, as well as minimizing both cost and confounding information during project design.