

Matt Haber
University of Utah
Department of Philosophy

Modeling Phylogeny

abstract

Darwin discusses both the *process* of evolution by natural selection and also the *pattern* of evolution by common descent. The historical pattern of evolution, i.e., phylogeny, is not directly observable; instead, it is inferred from incomplete data. We can observe fossils, some of which may be common ancestors of extant taxa. We may also gain clues of phylogenetic relations by examining the traits and properties of extant taxa, looking for patterns that suggest common ancestry. Systematists use these, and other data, to reconstruct phylogeny (these reconstructions are models called ‘phylogenetic trees’). What is at stake over how phylogeneticists model phylogeny? What is being claimed of these models, and how is that affected by the techniques used to construct phylogenetic trees? A central challenge faced by systematists is how to justify inference of phylogeny from these models. Different phylogenetic techniques address this challenge in different ways, which has led to ongoing conceptual debates within systematics over choice of phylogenetic methodology. Some proponents of one such technique, parsimony analysis, have asserted a privileged epistemic position with regard to justifying inference of phylogeny compared to users of other, statistically based phylogenetic methods, in particular Maximum Likelihood and Bayesian methods. In this project, I am asking what phylogenetic inferences are licensed by these various techniques. Systematists are at present divided over whether the inferential claims drawn from these competing methods are equivalent. This reflects confusion over the nature of the inference licensed by these various methods, e.g., what does it mean to assign a high posterior probability to a phylogenetic tree versus assigning that tree a high Maximum Likelihood bootstrap value? It is important to comprehend in detail the nature of these claims so that they may be better evaluated and presented. Phylogenetic inference also provides a case that should be of great interest to philosophers of science. If the products of the various techniques are viewed as competing models of the same system, this raises the question of whether or not these models are inter-translatable or incommensurable. I argue that they are formally translatable, but that the inferences drawn are incommensurable. This raises the further question of what the appropriate stance ought to be when faced with multiple methods for modeling a particular system. Though I favor a particular variant of pluralism in this case, looking at other stances helps to shed light on some of the central conceptual debates in phylogenetic systematics.