Syst. Biol. 42(1):90-92, 1993

# Experimental Approaches to Phylogenetic Analysis

DAVID M. HILLIS, J. J. BULL, MARY E. WHITE,
MARTY R. BADGETT, AND IAN J. MOLINEUX

*Departments of Zoology and Microbiology, The University of Texas, Austin, Texas 78712, USA*

Our paper on an experimental approach to phylogenetics (Hillis et al., 1992) had two general goals: (1) to establish the feasibility of generating phylogenies of organisms that exhibit considerable molecular variation and (2) to use these organismal phylogenies to test methods of phylogenetic estimation. We argued that the performance of a reconstruction method on the experimental phylogeny gives a measure of its likely performance on other (natural) phylogenies. The advantage of the experimental approach is that because we know a priori the relevant details of the evolutionary history in the lab, we can determine with precision whether and when a reconstruction is accurate. Natural phylogenies are not known a priori, so one never knows with certainty whether their reconstruction is correct.

Methods of phylogenetic analysis should be general for biological phylogenies, and simulations have shown that the methods are effective over a wide range of model conditions. However, it is not known if the necessary simplifications of numerical models reduce the relevance of simulations. Worse, simulations may make explicit assumptions that exactly match the assumptions of a method of analysis preferred by the investigator. For example, Jin and Nei (1990) found that the neighbor-joining algorithm using Kimura's (1980) two-parameter estimate of distance outperformed competing algorithms when applied to simulations of sequence evolution that were also based on Kimura's two-parameter model. The point of the experimental approach is to avoid approximating biological evolution by examining actual cases of biological evolution. Sober (1993) has challenged our illustration of experimental phylogenetics on the grounds that the bacteriophage system we used may not represent "nature," and thus a method's ability to reconstruct the phage phylogeny could be irrelevant to its ability to reconstruct other phylogenies. Here we attempt to clarify the foundations of his argument and suggest that, despite his objections, there is ample reason to pursue the approach we developed.

A general model of molecular evolution within populations entails two components: the origin of mutations and their subsequent fate (whether they persist or are lost). Evolution of multiple lineages is in turn connected through phylogenies. Sober argues that we must demonstrate an adequate parallel between these processes in our phage system and in nature. How then does our system fare?

Consider first the particulars of the phylogenetic tree. Our system allows the topology to be determined by the investigator, and this flexibility is one of the system's main advantages. No single topology can be regarded as representative of nature. Rather, exploration of a range of topologies is required, and our study illustrated the feasibility of this goal by providing a first example. Our conclusion that the results supported the legitimacy of phylogenetic estimation is derived from that fact that for the first time, given a phylogeny completely known in advance, the methods correctly reconstructed the known history. This does not mean that the methods will always work; we explicitly chose a "best case" tree, because if the methods failed for this tree then there would be little point in attempting to reconstruct more difficult topologies. We did not consider, as suggested by Sober, that the methods would "retrieve true phylogenies in nature when the methods all agree." In

fact, there are topologies suggested by simulations (e.g., those with alternating long and short terminal branches for four taxa separated by a short internal branch; Felsenstein, 1978) that are predicted to result in congruent, but wrong, results from all of the methods we tested (and we are in the process of testing actual trees of this type).

A second concern is whether our phage cultures undergo molecular evolution in a manner adequately similar to that of natural systems. Of course, there is no single model of molecular evolution that represents nature: selective constraints vary with nucleotide site, gene, and species, and the spectrum of mutations that arise also varies from taxon to taxon. Our system elevated the mutation rate by using a chemical mutagen, but this mutagen's spectrum of base-pair changes is within the bounds observed in natural systems (Gojobori et al., 1982; Li et al., 1984; Moriyama et al., 1991). Although our system cannot be imagined to represent all of nature, it clearly falls within the realm of known natural processes, and it does not appear to be less representative than would any other single taxon chosen for study. Reconstruction methods purport to be robust to variations in mutation and tree topology. Our experimental system provides an absolute test of their fallibility.

Sober incorrectly claims that we designed and tested a constant-rate model of evolution (a model that has been simulated in detail). However, we did not build a phylogeny with equal rates of change; the phage cultures were not forced to evolve at equal rates, as could have been the case in a simulation. As noted by Sober, systematists have been reticent to regard uniform rates as a model with any validity; in the experimental system, the rates of evolution are actual rates, controlled only by the constraints of real biological organisms. What we controlled were the times between branch points, so that the time from the roots to the tips (as measured in number of lytic cycles) was held constant in each of the lineages. The numbers of changes among the branches did vary substantially (greater than a three-fold range), even though the heterogeneity is not significantly greater than that expected for a Poisson distribution. Thus, the experiment differs from a simulation under the equal rates model in that the rates were determined by the constraints of biological organisms, rather than by constraints imposed a priori by a simulator. As we noted in the introduction to our paper (Hillis et al., 1992), the primary reason to produce experimental phylogenies is to avoid (as much as possible) making untested assumptions about evolutionary processes (many of these assumptions are necessary in numerical simulations).

We evaluated three areas of performance for common phylogenetic methods: ability to reconstruct branching order, branch lengths, and ancestral genotypes (the last applicable only to parsimony). Sober concentrated on the first of these performance criteria and found it unusual that all the methods found the same (correct) topology. However, as we originally stated, we intentionally chose an initial topology that simulations had suggested would be amenable to accurate reconstruction so that we could evaluate the feasibility and usefulness of the approach. Now that the feasibility has been demonstrated, we are examining other topologies that will be more difficult for some methods to reconstruct. The relevant differences among these experiments are in tree topology, which can be controlled by the investigator. We maintain that the other relevant parameters fall within the range seen in nature (or can also be studied in a controlled fashion, e.g., population size). For instance, our study demonstrated that the ancestral genotypes reconstructed by the parsimony method matched the genotypes of the actual ancestors with >98% accuracy. Although the accuracy of such reconstructions undoubtedly varies with the level of homoplasy (convergence and reversal), the amount of homoplasy in our study (consistency index of 0.75) was almost exactly the mean level of homoplasy in existing empirical studies involving nine taxa (Sanderson and Donoghue, 1989).

Saying that the processes of molecular evolution are constrained by the biology of the organisms rather than by the investigator is not the same as saying that the experiments are being interpreted atheoretically, as suggested by Sober. Rather, the difference between experimental and simulated phylogenies is like the difference between experimental and simulated bombs: the explosion of an experimental bomb does not indicate what will happen every time a bomb explodes, but it does provide information on one actual explosion. Simulations, on the other hand, provide exhaustive information on an idealized set of conditions that never actually exist in nature. Combining simulations with experiments is likely to result in the refinement of both, and the theory of phylogenetic estimation can thereby advance. Definitive tests of any theory rely on experimentation, whether this means building real bombs or building real phylogenies. This endeavor is especially important in phylogenetics because we know so little about the relevant processes of molecular evolution.

We see no reason to deemphasize any of the claims or implications of our paper. Phylogenetic systematics is largely a field of reconstruction methods and unknowable (but nonetheless estimable) phylogenies. The system we offered is a first step in providing an experimental, empirical foundation for this discipline in which methods of reconstruction can be tested unambiguously. Our study does not indicate what should happen with real organisms, but rather what did happen with real organisms.

## REFERENCES

FELSENSTEIN, J. 1978. Cases in which parsimony and compatibility methods will be positively misleading. Syst. Zool. 27:401–410.

GOJOBORI, T., W.-H. LI, AND D. GRAUR. 1982. Patterns of nucleotide substitution in pseudogenes and functional genes. J. Mol. Evol. 18:360–369.

HILLIS, D. M., J. J. BULL, M. E. WHITE, M. R. BADGETT, AND I. J. MOLINEUX. 1992. Experimental phylogenetics: Generation of a known phylogeny. Science 255:589–592.

JIN, L., AND M. NEI. 1990. Limitations of the evolutionary parsimony method of phylogenetic analysis. Mol. Biol. Evol. 7:82–102.

KIMURA, M. 1980. A simple method for estimating evolutionary rate of base substitutions through comparative studies of nucleotide sequences. J. Mol. Evol. 16:111–120.

LI, W.-H., C.-I. WU, AND C.-C. LUO. 1984. Nonrandomness of point mutation as reflected in nucleotide substitutions in pseudogenes and its evolutionary implications. J. Mol. Evol. 21:58–71.

MORIYAMA, E. N., Y. INA, K. IKEO, N. SHIMIZU, AND T. GOJOBORI. 1991. Mutation pattern of human immunodeficiency virus genes. J. Mol. Evol. 32:360–363.

SANDERSON, M. J., AND M. J. DONOGHUE. 1989. Patterns of variation in levels of homoplasy. Evolution 43:1781–1795.

SOBER, E. 1993. Experimental tests of phylogenetic inference methods. Syst. Biol. 42:85–89.

*Received 12 May 1992; accepted 27 August 1992*

*Syst. Biol.* 42(1):92–102, 1993

# When Theories and Methodologies Clash: A Phylogenetic Reanalysis of the North American Ambystomatid Salamanders (Caudata: Ambystomatidae)

THOMAS R. JONES, ARNOLD G. KLUGE, AND ALAN J. WOLF

*Museum of Zoology, University of Michigan, Ann Arbor, Michigan 48109, USA*

Shaffer et al. (1991; hereafter referred to as SCK) took positions on several theoretical and methodological issues of interest to historical biologists and in so doing provided the primary focus for the following discussion. We consider their use of the generally acknowledged controversial subjects of taxonomic congruence, char-