

A Conditional Probability of Reconstruction Measure for Internal Cladogram Branches

RICHARD H. ZANDER

Buffalo Museum of Science, 1020 Humboldt Parkway, Buffalo, New York 14211-1293, USA;
E-mail: rzander@sciencebuff.org

Abstract.—The conditional probability of reconstruction is a measure of the robustness of cladogram internodes and, unlike Bremer support and bootstrapping values, directly gauges probability. The new method compares the three putative branch lengths (the optimal and two alternatives) obtained through branch recalculation after nearest neighbor interchange and recalculation under constraint. With rooted trees, one switches the two free lineages attached at the distal end of an internal branch with the basal lineage. Probabilistic reconstruction of a branch for small data sets (e.g., morphological) is defined as having no contrary support for the two alternative branches and, when sufficient data are available (e.g., molecular studies), as meeting a selected confidence limit in chi-squared analysis. The exact probability that the internal branch is reconstructed is the same as the preselected confidence level met with chi-squared analysis; alternatively, it is a simple calculation of the length of the optimal branch divided by the sum of the lengths of all three putative branches. This new measure of robustness allows calculation of summary probabilities of subclade and tree reconstruction. The measure is conditional on a particular data set and optimization method but may also compare support from conflicting gene trees. Examples are provided by a morphological data set (the bryophyte *Didymodon*) and a molecular data set (primates). [Branch length; chi-squared; *Didymodon*; multiple gene trees; primates; probabilistic reconstruction; support]

Although optimality methods maximize phylogenetic information, such information may be conflicting or (in the case of morphological data) paltry, and relative support for the optimal versus alternative but somewhat longer trees is not easily measured. Parsimony methods clearly eliminate grossly unreasonable trees in terms of total length, but generally an additional support measure is necessary to distinguish the optimal tree from the many reasonable but less optimal solutions if the result is to be useful and reliable (Zander, 1998a). The methodology of parsimony analysis allows measurement, in terms of length, of alternative internal branches that are suboptimal to those in the shortest tree—the shorter the alternative branches (the fewer the putative synapomorphies), the less the contradiction to the optimal branch. In recent literature, measures of support for internal branches are generally limited to bootstrapping (BS) and Bremer support (BR; the decay index). These, however, are not the familiar confidence values or probability measures of other biological disciplines and are difficult to interpret as measures of reliability of the reconstruction.

The number of alternative branches to be assayed for conflict may be reduced by decomposing a cladogram into a set of triplets of alternative positions of subclades arising

from any one internal branch (internode), the unrooted four-taxon tree being the simplest, most general case of this. In particular, in a rooted tree, the three free subclades of any internal branch may be interchanged under constraint, switching either of the two that terminate the internode with that at the base of the internode. To recalculate branch lengths by this nearest neighbor interchange, one may either use constraint trees with a parsimony program such as PAUP (Swofford, 1998) or simply move branches in MacClade (Maddison and Maddison, 1999) with the same optimization settings. Recalculation thus results in three branch length values that may be compared: the length of the optimal internode in the shortest tree and the lengths of the two alternative internodes. Here I advance the rationale that if the two closest alternatives (i.e., the most reasonable for that optimization) to the optimal internode are found to have no statistically significant conflict with the optimal branch, the optimal branch may be considered a probabilistic reconstruction—given that particular data set, optimality method, weighting, and other assumptions—and is thus a “conditional” reconstruction. This also avoids multiple-test problems, in that randomly distributed “better” alternatives to the optimal branch can

often be found if enough searching is done. In effect, this method evaluates the relative robustness of a number of different trees equal to one more than twice the number of internal branches in the optimal tree and is a methodological alternative to evaluating either a set of trees some n steps longer than the optimal tree or subsampling iterations the whole data set.

Although in the calculation of optimality the characters and portions of trees are not independent (Pamilo and Nei, 1988; Lyons-Weiler and Takahasi, 1999), when using constraint trees to confine analysis to one internal branch, the variables, even when somewhat stochastically dependent, may be considered functionally independent (see Mikhail, 1976:15). For each internode, the three branch lengths obtained from reanalysis after nearest neighbor interchange may be evaluated as evidence of phylogenetic relationship, each step (reflecting any prior weighting) being of equal evidence. If the branch lengths are identical, the probability of correct choice of a reconstruction is one-third, given all other assumptions. If one branch is somewhat longer than the two alternatives (in terms of numbers of putative synapomorphies), a simple proportion of the branch length in the optimal tree divided by the sum of all three branch lengths (i.e., Bayes' formula with uniform priors) provides a probability measure. If the data are many, the standard chi-squared test may be used instead. The chi-squared statistic can be used to test goodness of fit, that is, to determine whether the observed frequencies fit with what was expected in the null hypothesis. The chi-squared test is nonparametric, in that it does not assume any particular data distribution, but is here used to test for the absence of a random generation of data at some level of confidence. Descriptions and justifications for this well-known test are found in most statistics manuals (e.g., D'Agostino and Stephens, 1987; Kendall et al., 1987).

By chi-squared analysis, the calculated numbers of traits shared by the optimal branch and the two alternative constraint-generated branches may be determined as being significantly different or not from a random distribution, in which the null hypothesis (H_0) is that all three lengths are identical, based on a random probability distribution, whereas the alternative hypothesis (H_a) is that one of the lengths is longer, gov-

erned by a distribution affected by additional shared ancestral characters.

If chi-squared analysis of the three values demonstrates that the three branch lengths are significantly different from a random distribution, and a second chi-squared analysis gives directionality to the method by demonstrating that the lengths of the two shorter, alternative branches are indistinguishable from those in a random distribution (at some previously selected level of confidence), then the internal branch is probabilistically reconstructed at the significance level attained by the chi-squared test of the three alternative branch lengths.

In all other cases, for example, relatively large values for one or both of the shorter alternative branch lengths or too few data for chi-squared analysis, the probability of branch reconstruction is calculated simply as the length of the optimal internal branch divided by the sum of the lengths of all three branches, the basic Bayes formula (with uniform priors) providing a genuinely directional measure of reliability. Given functional independence, the optimal and two constraint branch lengths can be viewed as mutually exclusive and exhaustive (Maurer and Ralston, 1991:457) evidence for phylogenetic relationships, and the lengths are thus directly proportional to signal.

This same simple proportional method of determining probability can also be used in cases where the morphological data generate short branches in the optimal tree. There, probabilistic reconstruction of an internal branch requires zero support for both of the two alternative constrained branches, given the paucity of data. The exact probabilities calculated for any internal branch by either chi-squared or simple proportion are true for that particular cladogram and that particular data set alone, because different branch lengths will generally be obtained when using different trees, data, and optimality assumptions (e.g., analysis under DELTRAN or ACCTRAN) about the same set of terminal taxa, hence the name conditional probability of reconstruction (CPR). Also, the measure does not (at present) enhance finding the most probable tree or best optimization settings but merely gauges the reliability of the tree at hand under the assumptions and data involved.

A very short branch, or various non-random processes, including convergence,

long-branch attraction, parallelism, and introgression (Templeton, 1986; Doyle, 1992; Avise, 1994), would confound the phylogenetic signal, whereas lineage sorting (Pamilo and Nei, 1988; Doyle, 1992; Hudson, 1992; Moore, 1995; Maddison, 1996; Lyons-Weiler and Milinkovitch, 1997) can produce false-positive results for particular data sets for the same taxa.

The CPR measure is modified and extended from the method of comparing numbers of synapomorphies of different arrangements of tree branches introduced by Gouy and Li (1989), who compared chi-squared values of branch lengths of 5.8S rRNA sequences of *Homo sapiens*, *Physarum polycephalum*, *Thermus thermophilus*, and *Escherichia coli* in an effort to demonstrate differences between various estimation methods. Though less similar in method, comparison of alternative branch lengths has been used previously; it has since been limited to four-taxon (three and an outgroup) studies or applied to whole trees but not presented as a general method of evaluating branch robustness. The nonparametric Templeton test (Templeton, 1987) involves differences in total tree length between constraint trees and the optimal tree, assuming that more recently evolved taxa have smaller distances (between character states) than do more anciently evolved taxa. Pamilo and Nei (1988) decomposed trees into three-component trees to simplify comparison of gene trees and species trees. Futuyma (1998) pointed out that "fourteen synapomorphies (including both deletions and base pair substitutions) unite chimpanzee and human as sister groups... whereas only three support the hypothesis that chimpanzee and gorilla are closest relatives...." Patterson et al. (1993) also compared mitochondrial DNA (mtDNA) molecular characters of alternative branches of *Homo*, *Pan*, and *Gorilla* "by presenting the putative synapomorphies... supporting each of the possible groupings among hominids." They used a binomial test that supported the *Homo*-*Pan* clade against *Pan*-*Gorilla* at a significance of $P = 0.0011$. Felsenstein (1988) computed the significance of hominid trio branch combinations and found significance at $P = 0.05$ if a clock was assumed but not otherwise. Templeton (1986) summarized signed ranks tests on mtDNA data for chimp-human-gorilla relationships to arrive at no signifi-

cant difference in the question of hominoid relationships. In the "triples distance" of Critchlow et al. (1996), the number of subclades of three taxa that are different between two trees are used as a measure of the distance between them, an idea rather different from that presented here. Other papers dealing with alternative branches may be cited, some using modified BS methods, but none is a direct probabilistic measure of internal branch robustness suitable for general use.

To simplify analysis (although the method doubtless can be extended to polytomies), one should select a single fully resolved tree from any set of equally parsimonious trees by considering additional criteria such as functional ingroup analysis (Watrous and Wheeler, 1981), maximum likelihood, or cluster analysis (e.g., UPGMA). The product of the probabilities of internal branches of subclades or trees provides a summary probability of reconstruction for lineages for that tree.

METHODOLOGY

To provide examples of CPR analysis of support, a morphological data set (Zander, 1998b) and a molecular data set (Hayasaka et al., 1988; included in PAUP* by Swofford, 1998) were selected. The primate molecular data set, consisting of 12 taxa and 898 characters, was analyzed with PAUP* (Swofford, 1998) with the following conditions: a branch and bound search, gaps treated as missing, MulTrees in effect, unordered, under ACCTRAN. The morphological data set of 23 taxa (22 of *Didymodon* plus *Barbula unguiculata* as outgroup), 20 characters, and the methods for the original analysis of the *Didymodon* study were as detailed by Zander (1998b); and the work was replicated with PAUP* with heuristic search, the tree-bisection-reconnection (TBR) algorithm, unordered, steepest descent not in effect, and under ACCTRAN and MulTrees.

Two equally parsimonious trees were obtained with the molecular data set, and the greater likelihood was used as a basis for choosing one (Fig. 1) of these two trees for CPR analysis. Two trees differing in placement of *D. nicholsonii* and *D. sinuosus* were obtained with the *Didymodon* morphological data set, and an optimal tree (Fig. 2) was selected for analysis because it was closest to the UPGMA result obtained by Zander (1998b). For each data set, the internal

TABLE 1. Analysis of mitochondrial data set for primates. AB:AC:BC are optimal and two alternative interior branch lengths, exact values in proportion. Significance of AB,AC,BC is from chi-squared tests; significance of AB,BC is from tests against random distribution (tests for directionality). For CPR, 0.95 or better is a probabilistic reconstruction.

Internode	AB:AC:BC lengths	Significance (P)		Calculation	CPR	BS	JK	BR
		AB,AC,BC method	AC,BC					
22–21	90:90:75	>0.05	>0.05	AB/ \sum all values	0.353	1	1	11
21–20	60:37:30	0.005–0.01	>0.05	AB	0.995	0.83	0.81	4
20–19	79:20:20	<0.001	>0.05	AB	0.999	1	1	42
20–16	48:41:24	0.025–0.01	0.05–0.025	AB/ \sum all values	0.425	0.99	0.99	12
19–18	32:11:11	<0.001	>0.05	AB	0.999	1	0.99	8
18–17	27:7:1	<0.001	>0.05	AB	0.999	1	1	12
16–15	37:24:18	0.025–0.01	>0.05	AB	0.972	0.88	0.87	4
15–14	51:16:12	<0.001	>0.05	AB	0.999	1	1	19
14–13	19:11:14	>0.05	>0.05	AB/ \sum all values	0.432	0.51	0.53	0

BS, bootstrap proportion; JK, jackknife proportion; BR, Bremer Support (decay index).

swapping) to establish the tree lengths at which branches collapse; the difference in length between these trees and that of the optimal tree is the amount of BR for each branch.

In both molecular and morphological analyses, constraint trees were used to exchange each of the sister branches terminating an internal branch of the fully resolved tree with the next more-basal branch. The lengths of these two constrained internal branches were obtained with PAUP*, using heuristic search with the morphological data set and branch and bound searching with the molecular set. In this way three branch length values were determined, one for the optimal branch and two corresponding to the lengths of the two alternative constraint-generated branches (Tables 1 and 2). This was done for each internal branch on the molecular and morphological trees.

The two terminating sister lineages ending an optimal internal branch are here conventionally named A and B (reading from top to bottom on Figs. 1 and 2), and the sister group to A + B is identified as C. The interior branch lengths were labeled AB (for the optimal internode terminated by the sister groups A and B), and AC and BC for the two constrained branches. The three branch lengths, AB, AC, and BC, were analyzed through a chi-squared (Lowry, 2000) test of the null hypothesis that AB, AC, and BC are equal, generated randomly, with 2 df. A second test using 1 df and a confidence limit of 0.95 compared each combination of two constraint-generated lengths to ensure directionality of the test. CPR analysis of the

most basal interior branches used multiple outgroups.

To clarify the constraint operation in detail with the molecular example, two alternative constraint-generated branch lengths were determined for the optimal internal branch AB, the terminal sister groups for

TABLE 2. Analysis of morphological data set for the moss *Didymodon*. AB:AC:BC are optimal and alternative interior branch lengths, exact values in proportion. CPR is calculated as AB divided by the sum of all three putative branch lengths; a CPR value of 1 indicates no conflicting support from the two nearest (most reasonable) alternative branches.

Internode	AB:AC:BC lengths	CPR	Significance (P)			BR
			BS	JK	BR	
44–34	1:0:0	1	<0.5	<0.5	0	
44–43	1:1:1	0.33	<0.5	<0.5	0	
43–42	2:1:0	0.67	<0.5	<0.5	0	
43–41	1:1:0	0.50	<0.5	<0.5	0	
41–40	1:0:0	1	<0.5	<0.5	0	
40–39	1:0:1	0.50	<0.5	<0.5	0	
39–38	2:0:0	1	<0.5	<0.5	0	
39–35	1:0:1	0.50	<0.5	<0.5	0	
38–37	3:1:0	0.75	<0.5	<0.5	0	
37–36	4:1:0	0.80	0.81	0.73	4	
34–33	1:0:0	1	<0.5	<0.5	0	
33–32	1:1:1	0.33	<0.5	<0.5	0	
32–31	2:0:0	1	<0.5	<0.5	0	
32–25	1:0:1	0.5	<0.5	<0.5	0	
31–30	1:0:0	1	<0.5	<0.5	0	
30–29	3:0:0	1	<0.5	<0.5	0	
29–27	2:1:1	0.50	<0.5	<0.5	0	
29–28	2:0:2	0.50	<0.5	<0.5	0	
27–26	3:0:1	0.75	0.61	0.58	3	
25–24	1:0:0	1	0.53	<0.5	0	

BS, bootstrap proportion; JK, jackknife proportion; BR, Bremer support (decay index).

which are (*Homo Pan*) in the optimal tree, which includes ((*Homo Pan*) *Gorilla*). The two alternative internal branches were AC, terminated by (*Homo Gorilla*) in an alternative tree modified to ((*Homo Gorilla*) *Pongo*), and BC, terminated by (*Pan Gorilla*) in another alternative tree modified to ((*Pan Gorilla*) *Homo*). The operation is continued with the next more proximal interior branch, and A is re-assigned to subclade (*Homo Pan*), B becomes *Gorilla*, and C becomes *Pongo*. Length was determined (1) for optimal internal branch AB terminating with the subclade ((*Homo Pan*) *Gorilla*) in the optimal tree consisting of (((*Homo Pan*) *Gorilla*) *Pongo*) and its alternative internal branch AC being the basal branch of the subclade terminated by ((*Homo Pan*) *Pongo*) in an alternative tree with (((*Homo Pan*) *Pongo*) *Gorilla*) and (2) for branch BC terminating with the subclade (*Pongo Gorilla*) in an alternative tree modified to ((*Pongo Gorilla*) (*Homo Pan*)). Thus, through these nearest neighbor interchanges, two alternative lengths were determined for each internode in a selected fully resolved tree.

Chi-squared analysis was done only for the molecular data set because internal branch lengths were too short for such analysis with the morphological data set. Involvement of a subjective element means the lowest acceptable numbers of expected values are prescribed differently by different authors; for example, by Cochran's Rule, no predicted values (branch lengths if actually random) can be <1 and not more than 20% can be <5 (Sellers and Vardeman, 1982) or, more commonly but more conservatively, no predicted values can be <5 (Newmark, 1977). Probabilistic branch reconstructions were identified when the following two criteria were met: Optimal branch lengths (AB) were statistically significant at the 0.95 level of confidence against both constraint-generated branch lengths (AC and BC), and the two constraint-generated branch lengths (the shorter lengths AC and BC) were demonstrably indistinguishable from a random distribution at the same level of confidence.

The second chi-squared test determines directionality of the chi-squared test of all values and evaluates the chance that the conflicting evidence of the constraint-generated branch lengths may be discounted as theoretically resulting from stochastic processes. If so, the CPR is then the same as the ac-

tual confidence level attained by chi-squared analysis of all three lengths. Yates' correction for continuity (Spiegel, 1988) was not used in this analysis, although it is often recommended for chi-squared calculations with 1 df (here, the second chi-squared test) when the estimated frequencies are low (5–10). Not using Yates' correction where it would otherwise be appropriate slightly relaxes the criteria for probabilistic reconstruction but only by affecting the non-critical directionality determination. For even lower estimated values of the two constraint-generated branch lengths, an exact binomial probability calculation (Lowry, 2000) may be used to determine directionality: that is, what is the chance of AC or greater out of AC + BC trials?

In all other cases (very short lengths, or two long lengths, or all three lengths similar and not distinguishable from a random distribution at the 0.95 confidence level), no theory excludes at any confidence level either the shortest alternative branch or both shorter alternative branches from probability calculations, and the CPR is necessarily the simple proportion $AB/(AB + AC + BC)$. Although no positive identification criteria for a reconstruction are used, if one assumes that evolution in fact occurred, then probabilistic reconstruction may be evaluated in this manner.

RESULTS

Molecular Analysis

From anatomical studies Futuyama (1998) has postulated that the hominoids *Homo* (man), *Gorilla* (gorilla), *Hylobates* (gibbon), *Pan* (chimpanzee), and *Pongo* (orangutan) were related as ((((*Gorilla Pan*) *Pongo*) *Homo*) *Hylobates*). Molecular studies, however, have demonstrated a very close relationship between *Gorilla*, *Pan*, and *Homo* (e.g., Goodman, 1963; Kishino and Hasegawa, 1989), especially between *Homo* and *Pan* (King and Wilson, 1975). Various cladograms, most agreeing in a terminal pair of *Homo* and *Pan*, have been published (e.g., Miyamoto et al., 1987, 1988; Goodman et al., 1989; Kishino and Hasegawa, 1989; Yang, 1994, 1996; Yang et al., 1995; Yang and Rannala, 1997) based on new methods or additional molecular data, but support for the alternative optimal solutions presented are difficult to compare against closely suboptimal solutions. Many studies (summarized by Ruvolo et al., 1993) have

produced evidence that *Homo* and *Pan* are more closely related to each other than either is to *Gorilla*.

The example data set analyzed here demonstrates that for this data set, the relative phylogenetic positions of *Homo*, *Gorilla*, and *Pan* are clearly doubtful because of conflicting evidence, though the three are almost certainly reconstructed as terminal to the hominoid clade. This also agrees with Felsenstein's (1988) judgment that the nearly three-way split of the Africa ape (*Gorilla Homo Pan*) issue is far from resolved. The exact phylogenetic positions of *Pongo* and *Hylobates* are well reconstructed.

Because of abundant data, the molecular data set allowed chi-squared analysis for each branch. Six of nine internal branches (66%) were probabilistically reconstructed for this cladogram. By the multiplication rule, the chance of all internal branches of the entire tree (Fig. 1) being correctly chosen from the three reasonable alternatives at each branch is 0.063 under CPR. The chance of *Homo* and *Pan* being correct is 0.432; the chance of *Homo*, *Pan*, and *Gorilla* being a clade is 0.999. The chance of hominoids (through *Hylobates*) being a subclade, given this particular data and optimality method, is 0.178. Templeton (1986), on the other hand, found ((*Homo Pan*) *Gorilla*) significant at the 0.05 level in signed rank tests of mtDNA restriction site data, but Kishino and Hasegawa (1989), analyzing the mitochondrial data of Brown et al. (1982), found no such significance when using maximum likelihood and no assumption of uniformity of evolutionary rate. For additional likelihood and Bayesian studies of this problem, see also Yang (1994, 1996) and Yang and Rannala (1997).

For the chi-squared level of significance values for the primate mitochondrial data set (Table 1), lower values indicate less chance of being a random distribution, where >0.05 means a <0.95 level of confidence that distribution is nonrandom; all other values are a >0.95 level of confidence. Thus, probabilistic reconstructions are assured for six internal branches calculated at the 0.99 confidence level, but three other internal branches are doubtful.

None of the internal branches that are considered probabilistic reconstructions at the 0.95 level are contiguous. Because switching sister branches with a second more basal internal branch presumably must be even

less probable than any of the three values used in CPR estimation, given optimality, then only when two less-than-probabilistic branches are contiguous are more than two alternatives for a branch in the optimal tree possibly worth considering. Given the low probability of reconstruction of these or any poorly supported two or more contiguous internodes, the problem can be ignored because reliably reconstructed internodes are not affected, given the present definition of probabilistic reconstruction, and multiplication of probability values provides an appropriate measure of doubt for the contiguous internodes as a unit.

Morphological Analysis

The lengths of all branches in Figure 2 are proportional to steps. Eight of 20 internal branches (40%) are probabilistically reconstructed in this cladogram. The chance of all internal branches of the morphological tree being correctly chosen is 0.000169. The chance of the *Didymodon* sect. *Fallaces* subclade (with a branch leading to *Didymodon leskeoides* and *D. tophaceus* at its base) being entirely correct at every node is 0.017; however, all species are present in the subclade and two internal branches with CPR of 1 support the degree of nesting of this section of the genus. The chance of the other major subclade (with *D. asperifolius* at its base) being entirely correct is 0.031. Species from sect. *Didymodon* (*D. anserinocapitatus*, *D. johansenii*, and *D. rigidulus*) form a subclade but have only half a chance of correct reconstruction. *Didymodon vinealis* could join other species of sect. *Vineales* if its branch were switched with that leading to sect. *Didymodon*, easily imagined given the CPR of 0.33. Section *Asteriscium* (*D. revolutus*, *D. australasiae*, and *D. umbrosus*) is rendered terminal to species of sect. *Vineales*; this may be due to long-branch attraction, possibly reflected in the poor (0.50) CPR value of the connecting internal branch. In sum, details of the cladogram of inferred evolutionary relationships in *Didymodon* are adequate for classification purposes but are generally insufficiently reliable to guide evolutionary and biogeographic study.

DISCUSSION

Several methods have been proposed to evaluate support for individual branches

of a parsimony tree—including the clade stability index (Davis, 1993), random resampling (nonparametric and parametric BS and jackknifing) (Felsenstein, 1985; Sanderson, 1989, 1995; Zharkikh and Li, 1992, 1995; Huelsenbeck and Hillis, 1996; Huelsenbeck and Rannala, 1997; Flook et al., 1999), signed rank test of branch order (Templeton, 1986), the distance Hadamard analysis (Penny et al., 1999); hierarchic signal (Faith, 1991; Naylor, 1992; Lyons-Weiler et al., 1996), and BR or the decay index (Bremer, 1988, 1994). Complete trees have been evaluated by the permutation tail probability test (Faith, 1991; Faith and Cranston, 1991), total support tests (Källersjö et al., 1992), total support index (Bremer, 1994), full Bayesian Markov chain Monte Carlo analyses (Mau et al., 1997; Yang and Rannala, 1997), and, of course, the consistency index (Farris, 1989). Additional discussion and citations are given by Huelsenbeck and Hillis (1996). Beyond intuitive evaluations reflected in “accepted” or “uncontested” clades, which are commonly used to select taxa to be included in a data set or to test analytic methods (Miyamoto et al., 1994; Miyamoto and Fitch, 1995; Milinkovitch et al., 1996; Rzhetsky and Sitnikova, 1996), two measures of internode reliability or relative certainty are commonly used in the literature: (1) support from subsampling, including BS and jackknifing, and (2) BR.

According to Felsenstein (1985:786), BS is “a confidence interval within which is contained not the true phylogeny, but the phylogeny that would be estimated on repeated sampling of many characters from the underlying pool of characters.” This is clearly begging the question. He suggested BS levels of 0.95 or greater as necessary for significance, though this is a rule of thumb and not statistically based (according to Felsenstein [1985], but may be very conservative according to Zharkikh and Li [1995]). BS and the similar jackknifing (e.g., as implemented in PAUP*) are especially problematic because the entire pool of characters may be coarsely heterogeneous, the “multiple tests” problem discussed by Felsenstein (1985) may be operative, and values commonly are not given for branches of interest not on the shortest tree. The universe of trees subsampled is only that of shortest trees (which may or may not include branches of trees somewhat longer than the optimal), not the universe of the

shortest trees plus all trees some n steps of interest longer.

BR has a similar problem. For morphological data sets that involve many taxa and few data, BR for internal branches can be impressive. This breaks down with molecular data, however, because internal branches may be much longer than their assigned BR and length alone is not a measure of confidence or robustness (Rice et al., 1997). For instance, a branch terminating a tree may have BR = 10 but be 30 steps in length. This, therefore, allows for the possibility that an alternative internal branch up to 20 steps in length can position attached lineages differently. Oxelman et al. (1999) address additional problems with BS and BR.

Maximum likelihood (Yang, 1994, 1996; Huelsenbeck and Crandall, 1997; DeBry, 1999; Pagel, 1999a) involves probabilities but has the drawback that likelihood values generated for trees are already optimizations (Nei, 1987; Yang, 1996). Thus, the likelihood ratio statistic usually used to measure support in maximum likelihood studies applies to nested models (DeBry, 1999), for example, variants of the same trees such as those comparing models involving molecular clocks and theoretically approximating a chi-squared distribution. For instance, likelihood values obtained from constraint trees in the present study with the molecular data for just *Homo*, *Pan*, *Gorilla*, and *Pongo* were -2216.89 for (((*Homo Pan*) *Gorilla*) *Pongo*), -2222.51 for (((*Pan Gorilla*) *Homo*) *Pongo*), and -2225.22 for (((*Homo Gorilla*) *Pan*) *Pongo*), results that cannot be interpreted statistically in terms of support for the trees of maximum likelihood. Also, the ability to fit the model apparently differs between various methods of calculation (Pagel, 1999b). Eernisse and Kluge (1993) and Schultz and Churchill (1999), among others, point out that the final precision of maximum likelihood estimation may be strongly affected by guesswork associated with initial selection of parameters, models, and alignments.

Markov chain Monte Carlo analysis (Mau et al., 1997) provides posterior probabilities for a set of trees, but these are merely relative probabilities; though they add to “probability one,” all trees with likelihoods too small to be calculated have been eliminated and their posterior probabilities may sum to a value that places the tree of maximum posterior probability (Yang and Rannala, 1997)

below an acceptable level of confidence. Both maximum likelihood and Markov chain Monte Carlo studies produce a tree that best explains the data, but even if that tree happens to be demonstrably much better than the second best tree, the data may poorly support certain branches.

Cladistic literature commonly describes the results of parsimony or statistical analysis as “approximating” or “converging on” the true tree (e.g., Huelsenbeck and Hillis, 1996: 38). With CPR analysis, this may mean either that many or most branches are probabilistic reconstructions or that the calculated probability of all branches being correctly selected is very low, depending on whether the bottle is presented as half empty or half full. Prediction in systematics is generally a case of betting on correct interpretations of often poor and conflicting information (Avice, 1994: 314; Philippe et al., 1996; Sites et al., 1996; Seberg et al., 1997) about a single historical event. Harper (1979) demanded that for scientifically plausible models in phylogenetic analysis, the probability of two taxa sharing closer ancestry to one another than to others in the group (given that the two taxa share one or more uniquely derived advanced character states and the others do not) be >0.5 . This minimal definition of “probable,” however, applies only to long-run frequency distributions in which prediction, or at least winning a bet more often than losing it, can be almost sure. In systematics, a good bet on correctly reconstructing a single historical happening is largely psychological and must depend on whatever a scientist is willing to lose when wrong. In the present paper, the level of confidence is set such that an average of >19 of 20 probabilistic reconstructions are estimated to be correct interpretations of the data, thus requiring far more evidence than merely a greater-than-even bet for any one reconstruction.

When the probability of correctly distinguishing evolutionary synapomorphy from apparent synapomorphy (Wiley et al., 1991; Lyons-Weiler et al., 1996) is low in subclades or trees, maximally predictive classification methods (e.g., cluster analysis) may be valid alternatives to parsimony analysis. Maximal predictivity (Gower, 1974; Sneath, 1989) is something of a misnomer, meaning more directly maximal descriptiveness.

The CPR probabilities found in this study often parallel BS (and jackknifing) values. In

Table 1, six of nine internal branches (given the method, there is no calculable most basal branch leading to the outgroup, which is why trees are conventionally output by PAUP* with a trichotomy at the base) of the molecular tree have CPR and BS values that are relatively proportional, either relatively high or low, but conflict is apparent for the remaining three nodes. For node 22–21, CPR = 0.353 versus BS = 1; for node 21–20, CPR = 0.995 versus BS = 0.83; and for node 20–16, CPR = 0.425 versus BS = 0.99. In Table 2, for morphological data, node 25–24 has CPR = 1 but BS = 0.53; however, comparatively few data are available per taxon and subsampling is therefore not very informative. A similar general matching of values is found when using BR. CPR values for the molecular results (Table 1) are high for branches with BR = 4, 8, 12, 19, or 42, but low for other branches with BR = 0 or 12. For the morphological results (Table 2), however, the two high BR values of 3 and 4 are paired with relatively low CPR values (0.75 and 0.80, respectively).

This general similarity with familiar support measures, at least with molecular data, may promote acceptance of a probabilistic basis for branch support. Unlike other measures of branch support, CPR values can be multiplied to arrive at a summary estimate for probabilistic reconstruction of subclades or the entire tree. CPR also may be used with morphological data sets that are commonly short on data and for which BS and BR analysis may return support measures for very few branches. In the case of two data sets for the same organisms, if data sets that produce congruent optimal clades are alone not probabilistically reconstructed, the additive property of chi-squared analysis (Hayes and Winkler, 1971; Spiegel, 1988) may be applied: Chi-squared values and degrees of freedom may each be summed and the significance levels redetermined with the possibility of attaining the desired confidence level.

Internal branches whose optimal length and two alternate lengths are together considered “indistinguishable from a random distribution” may be considered as not actually conflicting with a contrary probabilistically reconstructed branch generated from a different data set (because they may be treated as lacking information). The degree of conflict can be clearly appraised in other

cases of incongruence. The CPR method may be extended to trees generated by methods other than parsimony, e.g., neighbor joining or cluster analysis, by mapping state changes under parsimony and constraint, then calculating CPR, or it can be generalized onto any method capable of calculating alternative branch lengths.

A recent meta-analysis by Satta et al. (2000) surveyed data for 45 loci consisting of 46,855 bp. The data sets were much in conflict with one another—23 loci supporting the ((*Homo Pan*) *Gorilla*) gene tree, 8 supporting ((*Homo Gorilla*) *Pan*), 8 supporting ((*Gorilla Pan*) *Homo*), and 6 supporting a (*Homo Gorilla Pan*) trichotomy. This incongruence was attributed in part to differential sorting of gene and species phylogenies. Recombination and genetic drift were described as the cause of between- and within-loci conflict, and the authors found that a preponderance of data supported ((*Homo Pan*) *Gorilla*) but gave no exact measure of the reliability of this tree. A CPR analysis was therefore done for this study under a null hypothesis of equal numbers of confounding incongruent gene trees for each of the three tree hypotheses, tentative elimination of the six loci that support a polytomy and presumably demonstrate a lack of signal rather than a true trichotomy, and a rationale that each gene tree is a single character providing equal evidence of species phylogeny (Doyle, 1992; Slowinski and Page, 1999), such that any significant difference is due to the true species phylogeny. A confidence level of 0.997 was found for ((*Homo Pan*) *Gorilla*) for the proportion of 23:8:8 conflicting gene trees (values serving as an optimal and two alternative branch lengths). Reducing the data to those loci with BS > 0.80 gave a proportion of 14:2:3 conflicting gene trees, which resulted in a confidence level of 0.999 for the ((*Homo Pan*) *Gorilla*) species tree, which is thus again probabilistically reconstructed from the data given in the above cited meta-analysis. A chi-squared test with Yates' correction (for the proportion 8:8), and in the second case (2:3) an exact binomial probability calculation, demonstrated that the number of gene trees of (*Gorilla Homo*) and (*Gorilla Pan*) are indistinguishable from a random distribution at the standard 0.95 confidence level. Because the null hypothesis involving the three val-

ues was always rejected, the six gene trees reported in the original meta-analysis that supported a polytomy are correctly disregarded as phylogenetically uninformative—even though they were each supported at BS = 1.00 according to the original meta-analysis. These six gene trees also are not included in the chi-squared analysis because they do not constitute an exclusive statistical category (or "bin"). These above conclusions also imply that, short of this kind of meta-analysis, the prior probability that any internal branch in a molecularly based cladogram actually tracks species evolution rather than a contrary gene lineage may be on the order of 0.60—a disconcerting notion; much more study is needed.

In sum, chi-squared analysis or, in other cases, simple proportional probabilities of correct selection of an optimal internal branch against the two most reasonable alternative branches provides a generally applicable support measure that allows identification of probabilistic reconstructions. The CPR measure requires that probabilistic reconstruction be defined as the optimal branch—being either significantly longer than the immediate constraint-generated alternatives whereas the alternative branches are not themselves significantly different from a random distribution, or simply not contradicted by immediate alternatives. It is conditional on the data set and on all optimality criteria. The CPR may be better than commonly used measures of branch robustness because it involves a familiar measure of expectation, uses standard confidence levels, is understandable to most scientists with a basic grasp of statistics, can be multiplied to get a probabilistic measure of reconstruction of absolute branch order in a lineage, and can deal with results from multiple data sets in meta-analyses. Possibly CPR analysis may provide a means of identifying a single best tree of many equally parsimonious trees, based on searching among these for the tree of maximum probability.

ACKNOWLEDGMENTS

I appreciate the comments of J. Lyons-Weiler and two reviewers on drafts of this paper, and reassuring discussions with J. Felsenstein and others on the basic idea behind this paper.

REFERENCES

- AVISE, J. C. 1994. Molecular markers, natural history and evolution. Chapman and Hall, New York.
- BREMER, K. 1988. The limits of amino acid sequence data in angiosperm phylogenetic reconstruction. *Evolution* 42:796–803.
- BREMER, K. 1994. Branch support and tree stability. *Cladistics* 10:295–304.
- BROWN, W. M., E. M. PRAGER, A. WANG, AND A. C. WILSON. 1982. Mitochondrial DNA sequence of primates: Tempo and mode of evolution. *J. Mol. Evol.* 18:225–239.
- CRITCHLOW, D. E., D. K. PEARL, AND CHUNLIN QIAN. 1996. The triples distance for rooted bifurcating phylogenetic trees. *Syst. Biol.* 45:323–334.
- D'AGOSTINO, R., AND M. STEPHENS. 1987. Goodness of fit techniques. Dekker, New York.
- DAVIS, J. I. 1993. Character removal as a means for assessing stability of clades. *Cladistics* 9:201–210.
- DEBRY, R. W. 1999. Maximum likelihood analysis of gene-based and structure-based process partitions, using mammalian mitochondrial genomes. *Syst. Biol.* 48:286–299.
- DOYLE, J. J. 1992. Gene trees and species trees: Molecular systematics as one-character taxonomy. *Syst. Bot.* 17:144–163.
- EERNISSE, D. J., AND A. G. KLUGE. 1993. Taxonomic congruence versus total evidence, and amniote phylogeny inferred from fossils, molecules, and morphology. *Mol. Biol. Evol.* 10:1170–1079.
- FAITH, D. P. 1991. Cladistic permutation tests for monophyly and nonmonophyly. *Syst. Zool.* 40:366–375.
- FAITH, D. P., AND P. S. CRANSTON. 1991. Could a cladogram this short have arisen by chance alone? On permutation tests for cladistic structure. *Cladistics* 7:1–28.
- FARRIS, J. S. 1989. The retention index and rescaled consistency index. *Cladistics* 5:417–419.
- FELSENSTEIN, J. 1985. Confidence limits on phylogenies: An approach using the bootstrap. *Evolution* 39:783–791.
- FELSENSTEIN, J. 1988. Phylogenies from molecular sequences: Inference and reliability. *Annu. Rev. Genet.* 22:521–565.
- FLOOK, P. K., S. KLEE, AND C. H. F. ROWELL. 1999. Combined molecular phylogenetic analysis of the Orthoptera (Arthropoda, Insecta) and implications for their higher systematics. *Syst. Biol.* 48:233–253.
- FUTUYMA, D. J. 1998. *Evolutionary biology*, 3rd edition. Sinauer Associates, Sunderland, Massachusetts.
- GOODMAN, M. 1963. Man's place in the phylogeny of primates as reflected in serum proteins. Pages 204–234 in *Classification and human evolution* (S. L. Washburn, ed.). Aldine, Chicago, Illinois.
- GOODMAN, M., B. F. KOOP, J. CZELUSNIAK, D. H. A. FITCH., D. A. TAGLE, AND J. L. SLIGHTOM. 1989. Molecular phylogeny of the family of apes and humans. *Genome* 31:316–335.
- GOUY, M., AND W. H. LI. 1989. Phylogenetic analysis based on rRNA sequences supports the archaeobacterial rather than the eocyte tree. *Nature* 339:145–147.
- GOWER, J. C. 1974. Maximal predictive classification. *Biometrics* 30:643–654.
- HARPER, C. W., JR. 1979. A Bayesian probability view of phylogenetic systematics. *Syst. Zool.* 28:547–553.
- HAYASAKA, K., T. GOJOBORI, AND S. HORAI. 1988. Molecular phylogeny and evolution of primate mitochondrial DNA. *Mol. Biol. Evol.* 5:626–644.
- HAYES, W. L., AND R. L. WINKLER. 1971. *Statistics: Probability, inference, and decision*. Holt, Rinehart and Winston, New York.
- HUDSON, R. R. 1992. Gene trees, species trees and the segregation of ancestral alleles. *Genetica* 131:509–512.
- HUELSENBECK, J. P., AND K. A. CRANDALL. 1997. Phylogeny estimation and hypothesis testing using maximum likelihood. *Annu. Rev. Ecol. Syst.* 28:437–466.
- HUELSENBECK, J. P., AND D. M. HILLIS. 1996. Parametric bootstrapping in molecular phylogenetics: Applications and performance. Pages 19–45 in *Molecular zoology: Advances, strategies, and protocols* (J. D. Ferraris and S. R. Palumbi, eds.). Wiley-Liss, New York.
- HUELSENBECK, J. P., AND B. RANNALA. 1997. Phylogenetic methods come of age: Testing hypotheses in an evolutionary context. *Science* 276:227–232.
- KÄLLERSJÖ, M., J. S. FARRIS., A. G. KLUGE, AND C. BULT. 1992. Skewness and permutation. *Cladistics* 8:275–287.
- KENDALL, M., J. K. ORD, A. STUART, AND A. O'HAGAN. 1987. *Kendall's advanced theory of statistics*. Edward Arnold, London.
- KING, M.-C., AND A. WILSON. 1975. Evolution at two levels in humans and chimpanzees. *Science* 188:107–116.
- KISHINO, H., AND M. HASEGAWA. 1989. Evolution of the maximum likelihood estimate of the evolutionary tree topologies from DNA sequence data, and the branching order in Hominoidea. *J. Mol. Evol.* 29:170–179.
- LOWRY, R. 2000. VassarStats: Web site for statistical computation. Department of Psychology, Vassar College, Poughkeepsie, New York. Jan. 25, 2000. <http://faculty.vassar.edu/~lowry/VassarStats.html>
- LYONS-WEILER, J., AND G. A. HOELZER, AND R. J. TAUSCH. 1996. Relative apparent synapomorphy analysis (RASA) I: The statistical measurement of phylogenetic signal. *Mol. Biol. Evol.* 13:749–757.
- LYONS-WEILER, J., AND M. C. MILINKOVITCH. 1997. A phylogenetic approach to the problem of differential lineage sorting. *Mol. Biol. Evol.* 14:968–975.
- LYONS-WEILER, J., AND K. TAKAHASHI. 1999. Branch length heterogeneity leads to non-independent branch length estimates and can decrease the efficiency of methods of phylogenetic inference. *J. Mol. Evol.* 49:392–405.
- MADDISON, W. P. 1996. Molecular approaches and the growth of phylogenetic theory. Pages 47–63 in *Molecular zoology: Advances, strategies, and protocols* (J. D. Ferraris and S. R. Palumbi, eds.). Wiley-Liss, New York.
- MADDISON, W. P., AND D. R. MADDISON. 1999. *MacClade: Analysis of phylogeny and character evolution*. Version 3.08. Sinauer Associates, Sunderland, Massachusetts.
- MAU, B., M. A. NEWTON, AND B. LARGET. 1997. Bayesian phylogenetic inference via Markov chain Monte Carlo methods. *Mol. Biol. Evol.* 14:717–724.
- MAURER, S. B., AND A. RALSTON. 1991. *Discrete algorithmic mathematics*. Addison-Wesley Publ., Reading, Massachusetts.
- MILINKOVITCH, M. C., R. G. LEDUC, J. ADACHI, F. FARNIR, M. GEORGES AND M. HASEGAWA. 1996.

- Effects of character weighting and species sampling on phylogeny reconstruction: A case study based on DNA sequence data in cetaceans. *Genetics* 144:1817-1833.
- MIYAMOTO, M. M., M. W. ALLARD, R. M. ADKINS, L. L. JANECEK AND R. L. HONEYCUTT. 1994. A congruence test of reliability using linked mitochondrial DNA sequences. *Syst. Biol.* 43:236-249.
- MIYAMOTO, M. M. AND W. M. FITCH. 1995. Testing species phylogenies and phylogenetic methods with congruence. *Syst. Biol.* 44:64-76.
- MIYAMOTO, M. M., B. F. KOOP, J. L. SLIGHTOM, M. GOODMAN, AND M. R. TENNANT. 1988. Molecular systematics of higher primates: Genealogical relations and classification. *Proc. Natl. Acad. Sci. USA* 85:7627-7631.
- MIYAMOTO, M. M., J. L. SLIGHTOM, AND M. GOODMAN. 1987. Phylogenetic relations of humans and African apes from DNA sequences in the psi eta-globin region. *Science* 238:369-373.
- MIKHAIL, E. 1976. Observations and least squares. IEP, Dun-Donnelley, New York.
- MOORE, W. S. 1995. Inferring phylogenies from mtDNA variation: Mitochondrial-gene trees versus nuclear-gene trees. *Evolution* 49:718-726.
- MORT, M. E., P. S. SOLTIS, D. E. SOLTIS, AND M. L. MABRY. 2000. Comparison of three methods for estimating internal support of phylogenetic trees. *Syst. Biol.* 49:160-171.
- NAYLOR, G. J. P. 1992. Plotting frequency distributions of phylogenetic groupings found among sets of most parsimonious trees. *Cladistics* 8:161-164.
- NEI, M. 1987. Molecular evolutionary genetics. Columbia Univ. Press, New York.
- NEWMARK, J. 1977. Statistics and probability in modern life, 2nd edition. Holt, Rinehart and Winston, New York.
- OXELMAN, B., M. BACKLUND, AND B. BREMER. 1999. Relationships of the Buddlejaceae s. l. investigated using parsimony, jackknife and branch support analysis of chloroplast *ndhF* and *rbcL* sequence data. *Syst. Bot.* 24:164-182.
- PAGEL, M. 1999a. Inferring the historical patterns of biological evolution. *Nature* 401:877-884.
- PAGEL, M. 1999b. The maximum likelihood approach to reconstructing ancestral character states of discrete characters on phylogenies. *Syst. Biol.* 48:612-622.
- PAMILO, P., AND M. NEI. 1988. Relationships between gene trees and species trees. *Mol. Biol. Evol.* 5:568-583.
- PATTERSON, C., D. M. WILLIAMS, AND C. J. HUMPHRIES. 1993. Congruence between molecular and morphological phylogenies. *Annu. Rev. Ecol. Syst.* 24:153-188.
- PENNY, D., M. HASEGAWA, P. J. WADDELL, AND M. D. HENDY. 1999. Mammalian evolution: Timing and implications from using the LogDeterminant Transform for proteins of differing amino acid composition. *Syst. Biol.* 48:76-93.
- PHILIPPE, H., G. LECOINTRE, HOC LANH VAN LÊ, AND H. LE GUYADER. 1996. A critical study of homoplasy in molecular data with the use of a morphologically based cladogram, and its consequences for character weighting. *Mol. Biol. Evol.* 13:1174-1186.
- RICE, K. A., M. J. DONOGHUE, AND R. G. OLMSTEAD. 1997. Analyzing large data sets: *rbcL* 500 revisited. *Syst. Biol.* 46:554-563.
- RUVOLO, M., S. ZEHR, M. VON DORNUM, D. PAN, B. CHANG, AND J. LIN. 1993. Mitochondrial COII sequences and modern human origins. *Mol. Biol. Evol.* 14:248-265.
- RZHETSKY, A., AND T. SITNIKOVA. 1996. When is it safe to use an oversimplified substitution model in tree-making? *Mol. Biol. Evol.* 13:1255-1265.
- SANDERSON, M. J. 1989. Confidence limits on phylogenies: The bootstrap revisited. *Cladistics* 5:113-129.
- SANDERSON, M. J. 1995. Objections to bootstrapping phylogenies: A critique. *Syst. Biol.* 44:299-320.
- SATTA, Y., J. KLEIN, AND N. TAKAHATA. 2000. DNA archives and our nearest relative: The trichotomy problem revisited. *Mol. Phylogenet. Evol.* 14:259-275.
- SCHULTZ, T. R., AND G. A. CHURCHILL. 1999. The role of subjectivity in reconstructing ancestral character states: A Bayesian approach to unknown rates, states, and transformation asymmetries. *Syst. Biol.* 48:651-664.
- SEBERG, O., G. PETERSEN, AND C. BADEN. 1997. Taxonomic incongruence—a case in point from plants. [Abstract.] *Cladistics* 13:180.
- SELLERS, G. R., AND S. A. VARDEMAN. 1982. Elementary statistics. Saunders College Publishing, New York.
- SITES, J. W., JR., S. K. DAVIS, T. GUERRA, J. B. IVERSON, AND H. L. SNELL. 1996. Character congruence and phylogenetic signal in molecular and morphological data sets: A case study in the living iguanas (Squamata, Iguanidae). *Mol. Biol. Evol.* 13:1087-1105.
- SLOWINSKI, J. B., AND R. D. M. PAGE. 1999. How should species phylogenies be inferred from sequence data? *Syst. Biol.* 48:814-825.
- SNEATH, P. H. A. 1989. Predictivity in taxonomy and the probability of a tree. *Plant Syst. Evol.* 167:43-57.
- SPIEGEL, M. R. 1988. Schaum's outline of theory and problems of statistics, 2nd edition. McGraw-Hill, New York.
- SWOFFORD, D. L. 1998. PAUP*. Phylogenetic Analysis Using Parsimony (*and other methods). Ver. 4. Sinauer Associates, Sunderland, Massachusetts.
- TEMPLETON, A. 1986. Relation of humans to African apes: A statistical appraisal of diverse types of data. Pages 365-388 in *Evolutionary processes and theory* (S. Karlin and E. Nevo, eds.). Academic Press, New York.
- TEMPLETON, A. 1987. Nonparametric inference from restriction cleavage sites. *Mol. Biol. Evol.* 4:315-319.
- WATROUS, L. E., AND Q. D. WHEELER. 1981. The out-group comparison method of character analysis. *Syst. Zool.* 30:1-11.
- WILEY, E. O., D. SIEGEL-CAUSEY, D. R. BROOKS, AND V. A. FUNK. 1991. The compleat cladist: A primer of phylogenetic procedures. University of Kansas Museum of Natural History, Special Publication 19.
- YANG, Z. 1994. Statistical properties of the maximum likelihood method of phylogenetic estimation and comparison with distance matrix methods. *Syst. Biol.* 43:329-342.
- YANG Z. 1996. Maximum-likelihood models for combined analyses of multiple sequence data. *J. Mol. Evol.* 42:587-596.
- YANG, Z., N. GOLDMAN, AND A. FRIDAY. 1995. Maximum likelihood trees from DNA sequences: A peculiar statistical estimation problem. *Syst. Biol.* 44:384-399.
- YANG, Z., AND B. RANNALA. 1997. Bayesian phylogenetic inference using DNA sequences: A Markov chain Monte Carlo method. *Mol. Biol. Evol.* 14:717-724.
- ZANDER, R. H. 1998a. Phylogenetic reconstruction, a critique. *Taxon* 47:681-693.

- ZANDER, R. H. 1998b. A phylogrammatic evolutionary analysis of the moss genus *Didymodon* in North America north of Mexico. *Bull. Buffalo Soc. Nat. Sci.* 36:81–115.
- ZHARKIKH, A., AND W.-H. LI. 1992. Statistical properties of bootstrap estimation of phylogenetic variability from nucleotide sequences. I. Four taxa with a molecular clock. *Mol. Biol. Evol.* 9:1119–1147.
- ZHARKIKH, A., AND W.-H. LI. 1995. Estimation of confidence in phylogeny: Complete-and-partial bootstrap technique. *Mol. Phylogenet. Evol.* 4:44–63.

Received 8 March 2000; accepted 19 July 2000

Associate Editor: A. Smith