NINE EASY STEPS FOR CONSTRUCTING RELIABLE TREES FROM PUBLISHED PHYLOGENETIC ANALYSES[1]

Richard H. Zander[2]

Abstract

Classification and biogeography should be informed by reliable phylogenetic reconstruction. This may be recovered from the mainly speculative literature by a nine-step protocol of critical methods, the "operative transform." Piecemeal total evidence is attained by combining separately published support values for the same branch arrangement with Bayes' Formula. Joint probability correction is advanced for multiple test problems of conclusions of monophyly necessarily regarded as sets. Tables are given for conversion of nonparametric bootstrap values and decay indexes into probabilities, based on four-taxon simulations. Uncontested morphology-based branch arrangements are assigned high priors. Probabilities associated with unaccounted assumptions are addressed by a general correction factor. A probabilistic calculation allows the inference of two or more reliable internodes from a series of less well-supported internodes. A patristic distance of at least two internodes is used to support inference from molecular data alone of an intermediate lineage for recognition of supraspecific segregate taxonomic entities. Recognizing paraphyletic taxa when appropriate preserves attributes of inferred functional evolution providing taxic unity and value. Evolutionary Lazarus taxa contravening Dollo's Rule against re-evolution of complex traits justify interpretations of reticulation. This method of evaluation of the decisiveness of published molecular studies is exemplified with previously published cladograms focused on Malpighiales and the moss family Pottiaceae. Additional consideration of evolutionary process not reflected in phylogenetic analysis leads to reliable hypotheses of evolutionary relationship.

Key words: Bayesian, binomial confidence interval, bootstrap, bryophyta, cladograms, decay index, Malpighiales, multiple tests, operative transform, paraphyly, reliability.

It is clear from a perusal of published papers on phylogenetics that the past two or three decades of excitement in clarifying ancestral relationships, first with morphological traits and now with molecular or genomic data, have fueled a largely speculative literature. Whether the authors are theorists or working systematists, an emphasis has been laid on statistical "power" in distinguishing lineages rather than vigorous avoidance of type 1 error that the null hypothesis of no or at least no distinguishable phylogenetic signal is false when in fact it is true. Confirmation of conclusions of monophyly is commonly left to others. Although nowadays, with ever-increasing data, support values for branch arrangements (BAs) can be very high or apparently statistically certain, many assumptions (e.g., regarding alternative sequence alignments) remain unaddressed or are only now being investigated, and the contribution of this additional uncertainty impacts the reliability of published results. For any hypothesis test, a third alternative is needed: the metahypothesis that there is insufficient information from all relevant factors to tell if the null is rejected or not.

This paper is written solely for the second-level users of published cladograms, such as systematists, biogeographers, ecologists, and population biologists, who require well-grounded theories of ancestral relationship on which further, contingent study may be based. Nine simple steps are presented here as a method of identifying reliable conclusions of monophyly (resolved BAs) and, beyond this, hypotheses of descent with modification from the published literature. Frequentist statistics are avoided, and somewhat easier probabilistic or Bayesian concepts are emphasized. The nine steps are for situations in which published papers are the only source of data, and we must then necessarily resort to some shortcuts and ad hoc solutions. It is out of the question, however, that the second-level user, commonly not a phylogeneticist, might simply redo the phylogenetic study and investigate any additional possible sources of uncertainty.

"Reliability" is here defined in the Bayesian context as that level of confidence or credibility needed to confidently place a bet on being correct in light of what might be lost if wrong. Given that what is risked is not only phylogenetic science in advancing a possibly wrong conclusion of monophyly, but necessarily also the dependent science of second-level users, then high confidence is required. In this paper, a probability of 0.95 is accepted as minimum reliability, meaning that no more than one out of every 20 BAs is wrong (and which one is unknown). This level is standard for fields, such as experimental psychology, physiology, and ecology, that commonly involve statistical evaluation of test results, and is not arbitrary in that it establishes the expected across-the-board level of reliability of research in the whole field.

MATERIALS AND METHODS

Software used in this study was MrBayes (Huelsenbeck & Ronquist, 2001) for Monte Carlo Markov chain (MCMC) Bayesian analysis; PAUP* (Swofford, 1998) for maximum parsimony analysis, nonparametric bootstrapping, and jackknifing; RANDSET (Zander, 1999) for generating totally contrived random sets of either binary digital or DNA data; and VassarStats (Lowry, 2005) for binomial probabilistic calculations.

Two published treatments are here used as examples for operative transformation into cladograms with all branch arrangements supported at 0.95 or greater, or collapsed. The Savolainen et al. (2000) maximum parsimony cladogram of the Malpighiales based on *rbcL* gene sequences, among other results, split the Euphorbiaceae sensu Webster into five separate families, these accompanied by certain morphological differences. Of these, Chase et al. (2000: 689) wrote: "Nearly all families are well supported as monophyletic..." The present work asks: "Are these families supported by molecular data alone as distinct from the traditional Euphorbiaceae?" A second study, by Werner et al. (2004), focused on the moss family Pottiaceae using *rps4* chloroplast gene sequences and Bayesian analysis. In this study of 73 species, 51 of Pottiaceae and 22 of species from related families, the Pottiaceae group (with *Scopelophila* (Mitt.) Lindb. as basal) was almost entirely terminal. One genus, *Timmiella* (De Not.) Limpr. in Rabenh., was segregated toward the base of the cladogram, while exemplars of two genera of other families, Ephemeraceae and Cinclidotaceae, were embedded in the Pottiaceae. The present study investigates the questions: "Is the Pottiaceae to lose *Timmiella*? Should Ephemeraceae and Cinclidotaceae be synonymized with Pottiaceae to avoid paraphyly of the Pottiaceae?" More generally, the present paper asks: "What reliable hypotheses of evolutionary relationship can be extracted from published cladograms?"

In short, continuing refinement of a method (Zander, 2001, 2004, 2006) resulted in an optimal nine-step procedure for evaluating the reliability of cladograms published in the literature. As this paper is methodological, the Results summarize the technique.

RESULTS

This paper introduces several new or possibly unfamiliar concepts and acronyms, summarized here: **AB** represents a posteriori evolutionary support indicating shared ancestry of taxa A and B as measured by branch length (steps or state changes in maximum parsimony); (**AB**) is a lineage of two taxa of shared ancestry, in which A and B in this paper are always the terminal sister taxa or lineages of the optimal arrangement among three taxa A, B, and C; **BPP** is Bayesian posterior probability, calculated via the well-known Bayes' Formula or as given in the publication being analyzed; **PBA** is the probability of branch arrangement, and is here used as equivalent to the Bayesian credible interval and is a general term for both the BCI and BPP; **BP** is nonparametric bootstrap proportion; **JP** is jackknife proportion; **DI** is decay index or Bremer support; **BCI** is binomial credible interval, essentially one minus the chance by chance alone of support of AB steps or better out of all support for A or B in all NNI combinations of branches in a four-taxon data set; **NNI** is nearest neighbor interchange, or switching of each of two terminal lineages with the next lower branch, e.g., (AB)C, (AC)B, and (BC)A; **IRI** is implied reliable internode, a single reliable internode representing the pooled less-than-reliable support of two or more chained internodal branch arrangements later collapsed into one reliable internode; **FIRI** is the formula for calculating the probability of an implied reliable internode; α (alpha) in hypothesis testing is probability of a type I error or, in a nonparametric context (Kolaczkowski & Thornton, 2004), the chance of obtaining these data by chance alone, where BCI is the complement or one minus α.

Efforts at eliminating intolerable uncertainty long embedded in standard methodologies resulted in a nine-step "operative transformation" of published cladograms into trees with resolution restricted to conclusions of monophyly at 0.95 or greater probability of being correct. The chance of a BA being correct is represented here as the probability of branch arrangement (PBA), and refers to either the Bayesian posterior probability (BPP) or a nonparamet-

Table 1.   Approximate maximum assignable fail-safe four-taxon binomial credible intervals (BCIs) calculated from the nonparametric bootstrap proportion (BP) for various branch lengths as the BCI. The probability of branch arrangements being correct (PBA) is approximated by the chance of at least one trait shared by the terminal pair representing joint ancestry, being the chance of n steps (the branch length) minus one step by chance alone out of the total pairs of shared traits among three non-root branches (i.e., AB + AC + BC), using exact binomial calculation at one third probability. Interpolate intermediate values as necessary.

BP	Approximate PBA (as BCI)				
	5 steps	10 steps	15 steps	20 steps	25 steps
1.00	0.99	0.99	0.99	0.99	0.99
0.95	0.95	0.97	0.98	0.98	0.98
0.90	0.93	0.95	0.95	0.95	0.95
0.85	0.90	0.90	0.92	0.92	0.93
0.80	0.85	0.88	0.90	0.90	0.90
0.75	0.83	0.83	0.85	0.86	0.87
0.70	0.79	0.78	0.81	0.80	0.82
0.65	0.70	0.75	0.78	0.77	0.78
0.60	0.66	0.70	0.74	0.75	0.76
0.55	0.62	0.64	0.65	0.71	0.72
0.50	0.57	0.60	0.61	0.64	0.65

ric bootstrap value matched with a probability of the branch arrangement being correct with four-taxon simulations and previously dubbed (Zander, 2001, 2004) the binomial confidence interval (BCI) or here renamed binomial credible interval to avoid mixing frequentist and Bayesian processes of evaluation.

STEP ONE

Change support measures to PBAs.   Nonparametric bootstrap values are commonly given in published cladograms, but these are not generally taken as direct measures of probability. If the published study were redone using one of the newer bootstrap corrections, such as double-bootstrap or similar methods (e.g., Rodrigo, 1993; Zharkikh & Li, 1995; Efron et al., 1996; Sanderson & Wojciechowski, 2000; J. Farris in Salamin et al., 2002; Aris-Brosou, 2003; Susko, 2006), then this could be used as PBA. In lieu of redoing the study, however, an approximation table (Table 1) is given here for conversion of BP to BCI values (Zander, 2004). While the relationship between the BP and BCI is not linear, interpolation of intermediate BPs is not difficult.

Four-taxon data simulations at various branch lengths were contrived to give particular BPs. Given that convergence affects all possible pairings of the three non-root branches (a.k.a. nearest neighbor interchange), the chance of at least one trait from shared ancestry belonging to the terminal pair is the chance of n steps (the branch length of the terminal pair) minus one step occurring by chance alone out of the total pairs of shared traits among three non-root branches, using exact binomial calculation at one-third probability. Alternatively, it is the same as one minus the chance that all steps of the terminal pair occurred by chance alone among the total pairs of shared traits among the three non-root branches.

We here define A, B, C, etc., as representing lineages: (AB), (AC), and (BC), etc., as lineages hypothesized to have shared ancestry; and AB, AC, BC, etc., as the branch length (support in numbers of shared advanced traits) leading to those lineages. As example of conversion of BP to BCI, a BP of 0.85 is produced by lineages A and B sharing 10 traits, when AB = 10, AC = 5, and BC = 5. So of the 10 shared traits of (AB), about five are expected to be convergent, matching the five of (AC) and (BC), but this expected number has random variation and there is a certain chance that all 10 traits of (AB) are randomly generated (by convergence). With a null hypothesis of all data randomly generated, the chance of (AB) having nine or fewer out of 20 total synapomorphies (paired traits) at one third by chance alone is 0.91, which is the same as the chance of one true synapomorphy for (AB), the minimum number of synapomorphies needed for support of a hypothesis of shared ancestry for (AB)C.

Four taxa are sufficient for determining an effective approximation of the true probability of shared ancestry from the BP given in the literature even if the original BP was generated on the basis of the whole data set, not just that of the three terminal taxa. One does not know if the original BP was calculated largely on the basis of NNI or involves branch exchanges elsewhere in the cladogram, but NNI may

be assumed for purposes of approximation. "Privileged" branch arrangements are those segregated by the software as having much higher PBAs than expected from chance alone, and which are thus non-null cases from a null hypothesis of a star at the base of the cladogram and all variation random, similar to the "thinned counts" of Efron (2005). In fact, the BP is rarely over 0.50 from random data alone.

Both maximum parsimony and Bayesian procedures, when used with several small data sets generated by RANDSET (Zander, 1999) of ca. 20 taxa and 20 to 200 characters (no parsimony non-informative characters) of either 2- or 4-state random data, usually but not always resulted in fully resolved cladograms. Reliability measures of 0.01–0.50 BPP were obtained for the Bayesian analysis (with datatype = standard, 500,000 generations, burnin = 1000), and < 0.05–0.26 BP (heuristic search, with 2000 replications) for the parsimony analysis (BP values less than 0.50 may not be fully compatible with the tree of maximum parsimony). With larger random data sets, e.g., 20 or 73 taxa and 588 characters, BPs were commonly smaller, < 0.05–0.12, while BPPs were more variable, averaging about 0.20 BPP but rarely individually as high as 0.82. With large random data sets, branch lengths are generally between 100 and 200 steps, indicating that length of branches is no more indicative of accuracy than is degree of cladogram resolution.

One might then assume that in practice, all branch arrangements with BPPs and BPs less than 0.50, figures that are usually not given anyway, are indistinguishable from those generated by totally random data. On the other hand, all BPPs and BPs larger than 0.50 are amenable to probabilistic generation of single composite, reliable internodes via the IRI formula (Step Five) because they represent the equivalent of tosses of phylogenetically loaded dice.

Such BAs may be viewed as sufficient statistics (including all relevant information and, thus, independent) in the same way that uncontested groups based on descriptive data are considered phylogenetically isolated except for convergence with immediately alternate branch arrangements (a local star). Uncertainty is probably local and well approximated by the four-taxon tables for BAs with high BPs, which are the most critical branch arrangements. Given that the BCI approximates the likelihood of (AB)C,D being correct, it is also the local BPP, given uninformative priors and a normalizing constant of unity. Additional information and discussion are detailed in the Web Supplement (Zander, 2007). A similar local measure of branch arrangement reliability, the approximate

likelihood-ratio test (Anisimova & Gascuel, 2006) is also based locally (the branch of interest and four adjacent branches) and apparently performs well in simulations with larger trees; it is, however, implemented by specialized software using the original data and cannot be used to easily interpret published cladograms.

In Figure 1, we examine a much modified version of the Malpighiales *rbcL* maximum parsimony cladogram of Savolainen et al. (2000) for the level of molecular support for splitting the Euphorbiaceae into five families. The bold lines show the tree supporting the split. From the original cladogram, branch lengths are given above the lines and BP below. With branch lengths and BP, the BCI can be estimated from Table 1, and because the BPs were all above 0.50 and thus not expected by chance alone, the BCI can be taken as PBA. The five-part partition comprises the pre-cladogram concept of Euphorbiaceae.

The four-taxon BCI equivalent of Bremer support or the decay index (DI) is easily calculated, but a table (Table 2) is given here for convenience. This is the "local DI" (Zander, 2004), with no provision for collapse of a clade due to influence of a distant lineage, but justifications given above for approximations using four-taxon simulations also apply. Like the BCI for the BP, it is equal to AB minus one step out of AB + AC + BC trials where AC and BC are equal (as possible), and AC and BC are each set equal to AB minus the DI in number of steps. Alternatively, it is equal to one minus the probability of AB out of all shared pairs of traits. Thus, for a branch length of 10 steps and a DI of 6, AB = 10 and AC and BC are assumed both equal to four as a maximum to be failsafe by maximizing the possible contrary data, and the BCI is the chance of nine steps out of 18 at probability one third by chance alone.

Thus, when branch length and either BP or DI are given, the probability that the branch arrangement is correct may be approximated by reference to Table 1 or 2. To repeat, the BCI is an acceptable approximation of the BP or DI of the true PBA, in lieu of fully redoing the analysis with maximum parsimony and a technique like double bootstrapping, or with Bayesian MCMC.

When Bayesian posterior probabilities are given, they may be at least provisionally accepted directly as a support measure (PBA). There are certain papers (Suzuki et al., 2002; Zander, 2004; Lewis et al., 2005) that indicate that these values are too low, at least with short branches (but see Kolaczkowski & Thornton, 2006). Table 3 demonstrates with artificial four-taxon data sets that BPP is too low with short branches. Further study with longer branches indicates that BPP and BCI begin to approximate at about
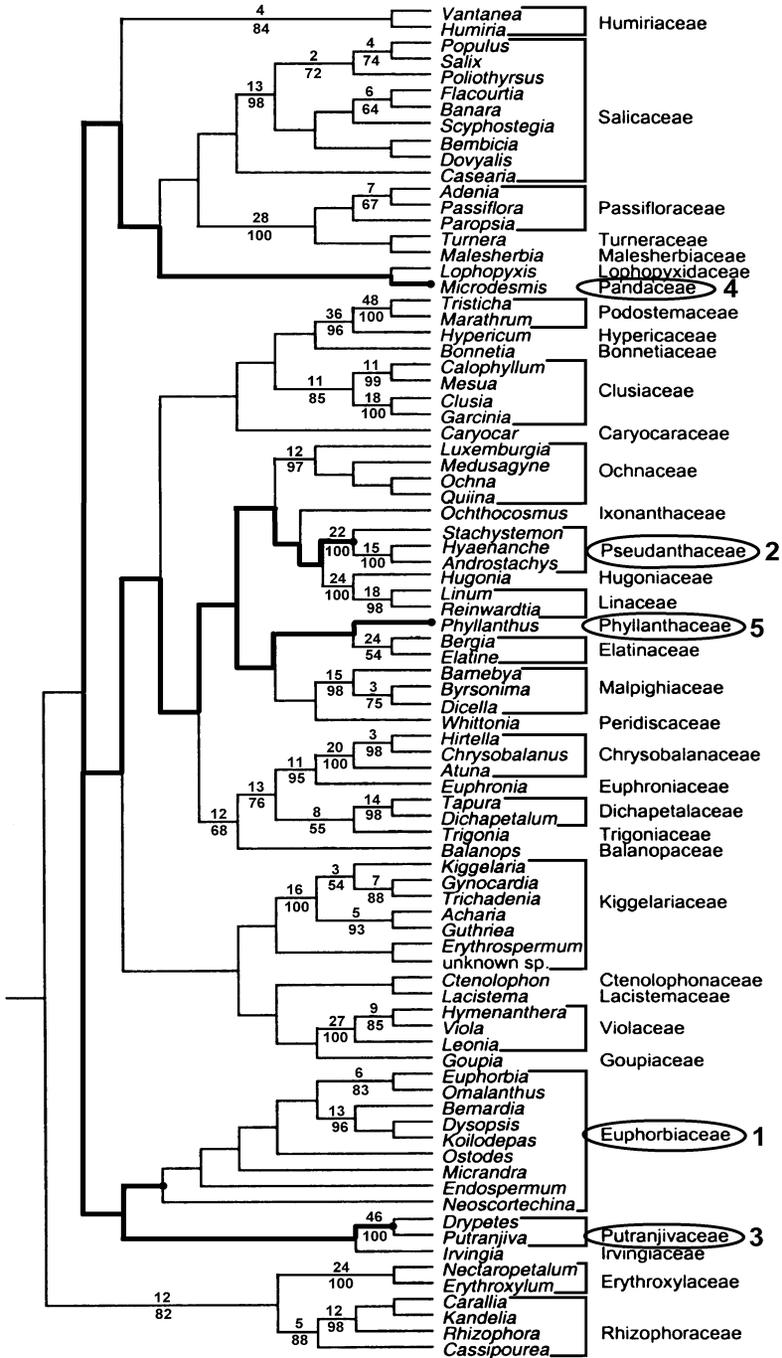
Figure 1. A single *rbcL* tree of Malpighiales from Savolainen et al. (2000) as presented by Chase et al. (2000). Euphorbiaceae sensu Webster is split (bold lines) into five families. These are numbered in order of importance for evaluation of multiple test problems. Branch length is above lines, nonparametric bootstrap proportion below. Reprinted with permission.

Table 2. Approximate maximum assignable fail-safe four-taxon PBAs calculated from the DI for various branch lengths. Interpolate intermediate values when necessary.

DI	Approximate PBA (as BCI)						
	3 steps	4 steps	5 steps	10 steps	15 steps	20 steps	25 steps
1	0.57	0.56	0.52	0.54	0.53	0.53	0.52
2	0.79	0.74	0.71	0.64	0.62	0.60	0.59
3	0.96	0.95	0.85	0.75	0.70	0.67	0.65
4	–	0.99	0.95	0.84	0.78	0.74	0.72
5	–	–	0.99	0.91	0.85	0.80	0.78
6	–	–	–	0.96	0.90	0.86	0.83
7	–	–	–	0.98	0.94	0.90	0.87
8	–	–	–	0.99	0.97	0.94	0.91
9	–	–	–	0.99	0.99	0.96	0.94
10	–	–	–	0.99	0.99	0.98	0.96

30 steps. Because the window of reliability is only 0.05, a slight difference matters greatly.

STEP TWO

Begin evaluation of published conclusions of monophyly. Partition the cladogram into phylogenetically isolated groups, using pre-cladogram concepts of classification or relationship, as best possible. Multiple test (multiple comparison) problems are a major fault with phylogenetic analysis as presented in the literature. Any conclusion of monophyly is a test. One problem, for example, is when one of two or more results is selected as "better resolution" when, in fact, the result with higher support is simply due to chance variation. Most researchers, in addition, will acknowledge that of 20 conclusions of monophyly each at 0.95 chance of being correct, one of them is on the average wrong, and one does not know which one.

A third problem is generally unrecognized though pervasive. Although it is perfectly understandable that multiple test problems are ignored by theoreticians requiring statistically powerful discrimination of possible clades (Miller, 1981), uncertainty is contributed when a set of two or more well-supported conclusions of monophyly must be correct at the same time, such as a statement that a family is demonstrably split into two or more monophyletic families. Given that a set of two or more conclusions of monophyly must be correct at 0.95, and the probability of the set being correct is the product of the individual probabilities of being correct, then the joint probability of two conclusions of monophyly at 0.95 being correct is only 0.90 (or about one set in 10 incorrect). Thus, individual PBAs of 0.975 (or 0.97 and 0.98) are needed for a set of two to be taken as reliable, and 0.99 each for a set of five to be acceptable as correct. Partitioning the cladogram is an

Table 3. Comparison of BP, JP, and BPP at various short branch lengths using contrived four-taxon data sets with PAUP* (BP and JP with fast heuristic, 2000 replicates) and MrBayes (datatype = standard, 300,000 generations, burnin = 1000). AB:AC:BC are ratios of synapomorphic traits unique to pairs of three terminal branches A, B, and C. BPP is clearly too high at low branch lengths in this simulation.

AB:AC:BC	BCI	BP	JP	BPP
1:0:0	0.67	1	0	0.79
2:0:0	0.89	1	1	0.96
2:1:0	0.74	0.75	0.66	0.83
2:1:1	0.59	0.59	0.50	0.69
3:0:0	0.96	1	1	0.99
3:1:0	0.89	0.85	0.75	0.97
3:1:1	0.79	0.77	0.61	0.91
3:2:0	0.79	0.69	0.60	0.84
3:2:1	0.68	0.63	0.58	0.77
3:2:2	0.57	0.52	0.50	0.64
4:0:0	0.99	1	1	1
4:1:0	0.95	0.95	0.79	0.99

initial means of eliminating lineages that need not be considered as part of a set.

In statistics, likelihoods of raw data are particular manifestations of multiple events, e.g., in 100 tosses, a coin came up heads 45 times, and thus the likelihood of heads is 45% for that coin, while probabilities are ideal generalizations of expectations, e.g., a fair coin has a 50% probability of coming up heads. Frequentist methods of dealing with multiple tests, such as Bonferroni or sequential Bonferroni correction (Holm, 1979; Hochberg, 1988), control of False Discovery Rate (Benjamini & Hochberg, 1995), or new "sharpened" methods (Hochberg, 1988; Verhoeven et al., 2005) that take into account the nonuniform distributions of confidence intervals generated by phylogenetic signal, convert likelihoods to P values and examine how close they come to a separately determined (for instance by maximum likelihood) probability required for acceptance as correct. The present method follows the Bayesian convention of converting likelihoods to probabilities whenever they may be taken to well approximate the ideal, generalized expectation, such when information is deemed sufficient or, as in this case, likelihoods are demonstrably non-null, e.g., BPs and BPPs greater than 0.50 are seldom obtained by random data under parametric bootstrap (Sanderson & Wojciechowski, 2000; but see Dolphin et al., 2000).

Two branches of a tree are, of course, dependent on each other to some extent through interpretation of the data. But, after the optimum tree is settled on, any observation that the uncertainty attributed to one branch arrangement (as contrary data as opposed to lack of data) affecting the other BA may be viewed as an artifact of small taxon sampling. The data are independent (orthogonal), even when BAs are occasionally not. The joint probability reflects this, and, like Bonferroni correction (Rice, 1989), the present method does not require independence of branch arrangements because the data are independent and may be extended to data sets with more OTUs (operational taxonomic units) as a kind of pseudocount or reverse bootstrap.

As to conflict of branch arrangements, if NNI were the only results of alternatives to optimal BAs, then no lineage would change in composition, i.e., of ((AB)C)D, then C would not embed itself inside lineage A or B or D but would only switch places with A or B. If all nests in the cladogram, however, were strongly supported, e.g., by moderate to high PBAs, then C would have to crack two or more nested BAs to jump across the cladogram and modify another lineage. The chances of that are powers of 0.05 for 0.95 PBAs. Thus, a cladogram of only BAs with 0.95 and above support is sufficiently nested for probabi-

listic calculations by this argument and by independence of data. Using IRI with BAs of less than 0.95 is justified in the logic that (1) such BAs are probably due to signal because they are above 0.50 and thus privileged (Efron, 2005), and (2) reliability of monophyly should increase with phylogenetic distance.

The single *rbcL* tree of Malpighiales from Savolainen et al. (2000) as presented by Chase et al. (2000) showed the traditional Euphorbiaceae split into five families (indicated in Fig. 1). These and only these are the focus of this study, and these are subject to multiple test strictures. The only exception is the possible splitting of two family lineages by a third intermediate lineage (e.g., sister to one of two families) to ensure molecular support as separate families, as discussed in Step Seven.

STEP THREE

Preselect conclusions of monophyly to test, prioritized to break ties. Choose those lineages that must be viewed as a set for the conclusion of joint monophyly. In this case, we choose the five families of Step Two: (1) Euphorbiaceae (s. str.), (2) Pseudanthaceae, (3) Putranjivaceae, (4) Pandanaceae, and (5) Phyllanthaceae (see Fig. 1). Here, the most taxon-rich (in the cladogram) lineages are selected as most important and are prioritized as above. Thus, if 0.95 was the highest level of support among them, and there were two or more, then Euphorbiaceae s. str. would be the one supported at that level, and all other branch arrangements would be considered speculative. If Euphorbiaceae s. str. was not supported at 0.95, then the next in priority would be the one considered reliable. One may wonder that only one of several conclusions of monophyly would be acceptable when all have the same support, but over the course of several publications, the sets accepted would be wrong much more often than one out of 20 sets. Preselection and prioritization are thus important.

STEP FOUR

Impose a penalty for unaccounted assumptions. It is clear that there remains considerable uncertainty not reflected in BA support values given in the literature. I can only explain this as a function of three preconceptions on the part of phylogeneticists: (1) Speculation is presently considered more important than consolidation of well-grounded theory, and such investigation necessarily involves possibly critical assumptions and choice of statistical discriminative power over reliability. (2) The likelihood principle indicates that the probability of the data given the

hypothesis is the same as the probability of the hypothesis given the data, and therefore one may assume (wrongly) that the data as presented in the aligned sequence contain all the information needed to evaluate the hypothesis (given a model). (3) The probability of a given branch arrangement being correct is simply presented as the posterior probability given by Bayes' Formula. The last two are simplistic in the extreme, ignoring the necessity that any relevant factors not included in the data or model must be attended to as a joint probability (the product) of the chance all assumptions are correct. Thus, the final product of a Bayesian analysis is properly not the posterior probability but the Bayes' Solution (Kendall & Buckland, 1971), which takes into account and minimizes risk, such as uncertainty contributed by assumptions not dealt with in the method, model, or data.

In the late 1970s and early 1980s, cladistics initially presented "optimality-alone" as a justification for choosing a particular tree, in that scientists supposedly always choose the simplest answer, the one that "converges on the truth," or that which is least falsifiable. The obvious objection has now generally been accepted, namely that a branch arrangement of interest in the best tree may be contradicted in many respects by the next several best trees, which may be only slightly less optimal. With molecular data, abundant statistical support is available through bootstrapping and Bayesian posterior probabilities. Such support is sometimes quite strong. Unfortunately, the old criterion of optimality alone has been shifted from the optimal tree to other assumptions, such as the sequence alignment, for which there may be other slightly less optimal alignments, each of which may produce a different tree (Giribet et al., 2002). A recently devised method of estimating reliability of both alignment and phylogenetic trees (Redelings & Suchard, 2005) has much promise, but the literature lacks a joint probability assessment of branch arrangement and alignment reliability. In addition, using a different, less "optimal" outgroup may produce a different branch arrangement or different support measure for an arrangement of interest. In the case that several outgroups are used, the branch support for both the immediate outgroup and the tree as sister lineages could be used in a joint probability measure for a branch arrangement when the next nearest outgroup affects the result, but this kind of modified branch support value is not to be expected from the literature, and in this respect, too, the tree, however well supported prima facie, is to some extent merely "optimal."

I have examined a selection of recent phylogenetic papers and found (Zander, 2005) more than 100

biological, methodological, and computational assumptions (or contingencies) cited by authors as problematic. These assumptions are not commonly included as sources of uncertainty reflected in support values in the final cladogram as published. Examples, in addition to the egregious alignment and outgroup problems above, include: insufficient or polyphyletic exemplars, wrong gap costs, differential lineage sorting, hybridy or polyploidy, lack of clocklike behavior, rates other than gamma-distributed, differences between results of "total evidence" and evaluations based on separate gene studies, too few generations in MCMC analysis or of bootstrap replicates, possible strong selection pressure on strongly conserved non-coding sequences and persistent pseudogenes, too few exemplars, inclusion or exclusion of fossil evidence, endogenous retroviruses, gene conversion, high evolution rates, self-correction of flawed DNA, paralogy, codon bias, chloroplast capture and other horizontal gene flow, novel clades, among-lineage and among-site heterogeneities, saturation, differences when including or excluding third nucleotide position, lack of vouchers, choice of ACCTRAN or DELTRAN in PAUP*, long-branch attractions, unusual noise, differences between results from different analytic programs, sample error, and model insufficiency. Most of the 100, doubtless, are inapplicable or too improbable in any one particular instance to affect the PBA. The uncertainty may be considerable from all remaining assumptions, and some correction is necessary, even ad hoc. Consider the following: if only 10 of 100+ assumptions apply to one branch arrangement of interest and each of these assumptions were wrong and changed the branch arrangement only one out of 1000 times, then the contributed uncertainty is $10 \times 1/1000 = 1/100$, or 1% of the PBA. At this point, a penalty of 0.01, or 20% of the window of reliability, can be levied without qualm. If there are additional sources of uncertainty relevant to only the branch arrangement of interest, such as suspicion of misidentification, an additional judicious penalty can be set. A penalty of 0.01 on six chained internodes, each of 1.00 BPP, means that the chance of all branch arrangements being correct at once is reduced to 0.94. Thus, if corrected for unaccounted assumptions, no cladogram of seven or more terminal taxa can be entirely reliable. Figure 2 shows branch arrangements with BPs converted to BCIs and each reduced by 0.01 for unaccounted assumptions. This is then taken as the working PBA.

This penalty is certainly a ballpark estimate that may be changed by the user, and it has been accused of being ad hoc by reviewers. There is, however, no alternative correction for second-level users of
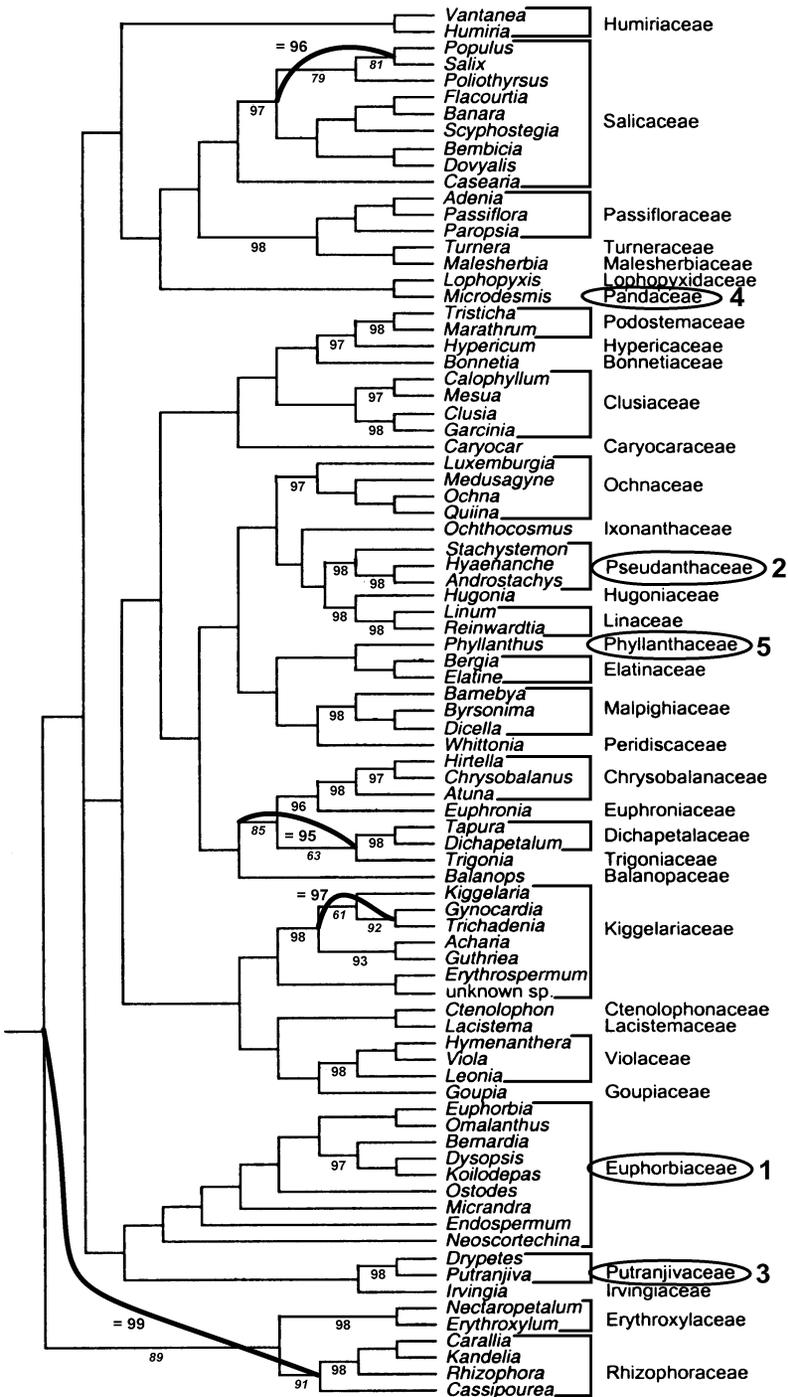
Figure 2.   Malpighiales tree after conversion of bootstrap proportions into binomial credible intervals and subtraction of 0.01 penalty for unaccounted assumptions, these taken as equivalent to probability of branch arrangements being correct. All possible contiguous internodes with less than 0.95 PBA are fused into one at 0.95 or above with formula for implied reliability (curved lines). All remaining values less than 0.95 are deleted. (Modified from Chase et al., 2000; reprinted with permission.)

published analyses. Instead of the penalty for unaccounted assumptions, one can either reject all even slightly unreasonable molecular results as too speculative, or redo molecular analyses offering contrary solutions to accepted monophyly, and test whether the branch arrangements of interest are robust to reasonable variation in all relevant assumptions. One might note that accepting many assumptions, because if false they would not significantly affect any branch arrangements, was also ad hoc when statistical phylogenetic analysis began; however, the practice has been embedded in the analytic method for the last 30 years.

STEP FIVE

Combine two or more contiguous branch arrangements of PBAs of 0.50 to 0.94 into one of 0.95 or higher using the Formula for Implied Reliable Internodes (FIRI). Basically, the simplest probabilistic calculation of the chance that at least one of two or more events will happen is one minus the product of the chances they will not happen.

$$\mathrm{PBA_{implied}} = 1 - ((1 - \mathrm{PBA_{internode\ 1}})$$
$$\times\ (1 - \mathrm{PBA_{internode\ 2}})\dots$$
$$\times\ (1 - \mathrm{PBA_{internode\ n}}))$$

In phylogenetics, this is equivalent to the chance that one (or more) of two chained branch arrangements is correct, or (more easily calculable) one minus the product of chances that each is incorrect, where the chance of being incorrect is one minus the PBA for the branch arrangement. For example, the chance of at least one of two chained internodes at 0.81 and 0.79 PBA being correct is $1 - (0.19 \times 0.21)$, or 0.96. Compare this to the chance both are correct at the same time (as a set), which is the product, or 0.64. There is clearly a large difference when depending on the resolved branch arrangement for support of a biogeographic hypothesis. Do not use this formula with branch arrangements of < 0.50 BP or BPP or BCI, because support in those cases may be due to chance alone and the additional apparent certainty though small is inapplicable. If all branch arrangements that reach 0.95 or greater using this formula are dealt with and those that reach less than 0.95 are collapsed, a cladogram (or partition) with only branch arrangements of 0.95 or more is obtained (Figs. 3 and 4).

STEP SIX

Make corrections for multiple tests. Figure 3 shows no support for Euphorbiaceae s. str., but PBAs of 0.98

for both Pseudanthaceae and Putranjivaceae, and tentative 1.00 for Pandanaceae and Phyllanthaceae of one exemplar each. The product of the support values for the last four is 0.96, so we have acceptable monophyly of each lineage in the set of four lineages, which, however, are all basal branches when considering the molecular data set alone. There is no reliable phylogenetic structure.

Figure 4 of the operatively transformed Werner et al. (2004) cladogram shows many branch arrangements at 0.99 probability. For any particular previous hypothesis of relationship that is to be tested, up to five of these BAs can be true as a set. These should be selected from only those with 0.50 (or dubious) morphological priors, and not those with 0.01 priors because these are clearly contrary to accepted use and without alternative morphological traits supportive of the molecular BA. One can involve as a set as many as needed from those that are morphologically uncontested (therefore assigned 0.99+ PBA when using Bayes' Formula).

STEP SEVEN

With molecular data alone, reliable demonstration of paraphyly will split a taxonomic group, i.e., distinguish two groups as warranting separate names. In the case of several exemplars, such as splitting a genus or higher taxonomic category, a distance of two internodes reliable as a set with an intermediate lineage not related to either (morphologically or molecular with a distinction greater than mere monophyly) is needed. The first distance distinguishes the terminal isolate of the two as distinct from the disparate sister lineage; the second establishes the second isolate that is lowermost in the tree as not being merely a basal branch (and therefore potentially monophyletic). Conditions for distinguishing taxonomic groups based on molecular data alone, without reference to morphology, are not established. There is progress being made in species delimitation using multiple molecular samples (Pons et al., 2006), but molecular data as used in the literature are generally sparse, represented by one or a very few exemplars, and are without the extensive analysis of variation represented by morphology-based taxonomy. Two lineages, separated by a third reliable lineage, must both be true at the same time (as a set), and so the product of their support values should reach 0.95. Two lineages among the split portions of the Euphorbiaceae in Figure 3 reach 0.98 and are reliable as a set. None of the five lineages, however, are reliably other than a basal branch of the Euphorbiaceae. Their status as separate families remains entirely morphological (based on this one publication). The high (0.97) PBA support for mono-
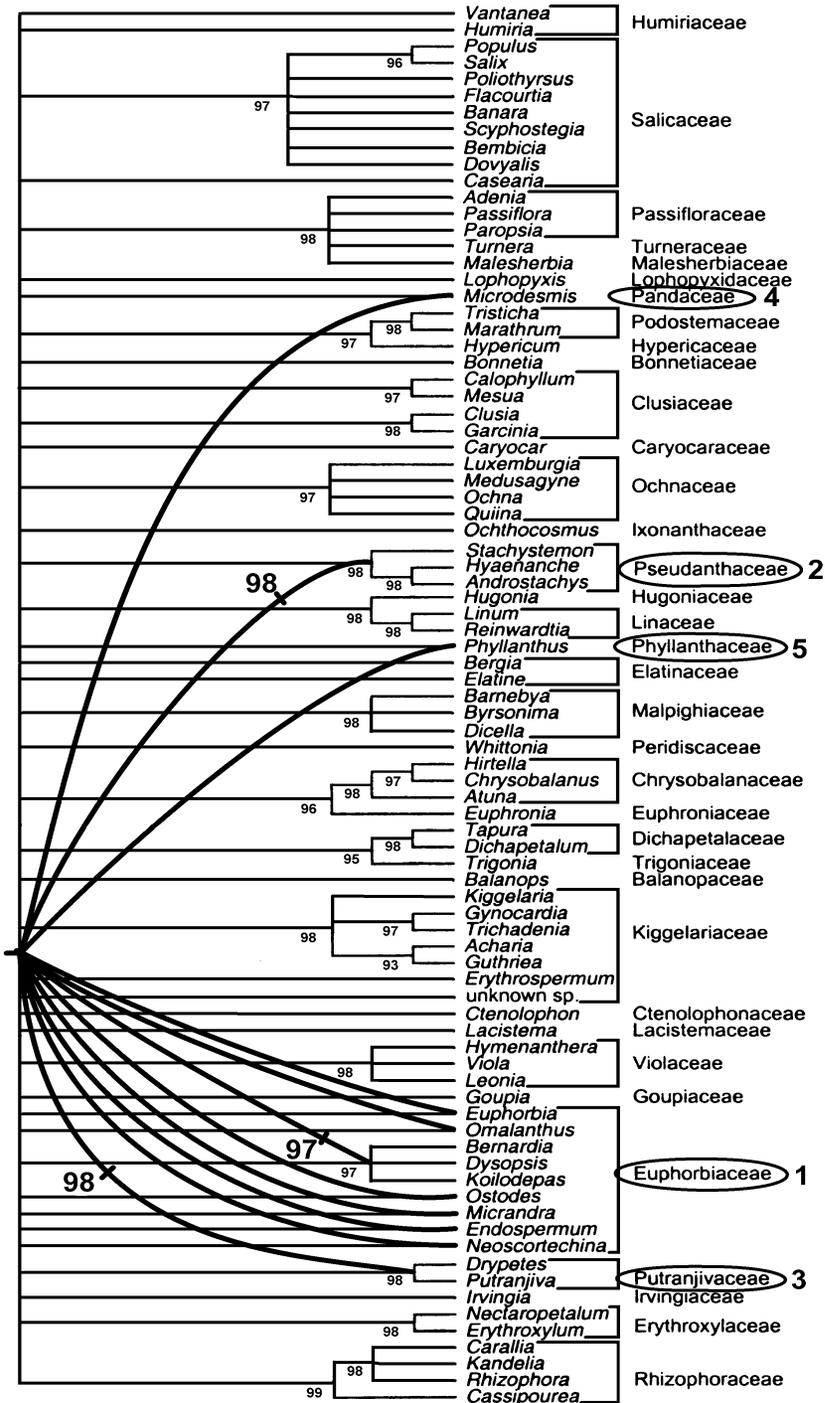
Figure 3. Operative transform of Malpighiales tree with all branch arrangements less than 0.95 PBA collapsed. The five families of Euphorbiaceae sensu Webster are found to be arranged without reliable phylogenetic structure (as multifurcative sister groups). No family is isolated from the traditional Euphorbiaceae by a disparate intermediate lineage, and thus only non-molecular traits sustain them as distinct. Two families are each monophyletic as a set (joint probability = 0.96), while the three exemplars in Euphorbiaceae s. str. united at 0.97 must be viewed as speculative (joint probability with other two conclusions of monophyly is 0.93). (Modified from Chase et al., 2000; reprinted with permission.)

Figure 4.   Operative transform of the *rps4* tree of Pottiaceae and associated families of Werner et al. (2004). *Timmiella* is retained in the Pottiaceae by recourse to Dollo's Rule. *Ephemerum* (Ephemeraceae) and *Cinclidotus* (Cinclidotaceae) are excluded from the Pottiaceae, which is retained in spite of being paraphyletic, justification being similar phenotype and survival strategy throughout the family leading to a single type of functional evolution.

phyly of a portion of Euphorbiaceae s. str. must be regarded as speculation, as 0.98, 0.98, and 0.97 are not reliable as a set (product = 0.93).

The same publication (Savolainen et al., 2000) also split the Flacourtiaceae into Salicaceae and Kiggelar-

iaceae, which is an entirely different multiple test problem. Examination of Figure 3 indicates that the Salicaceae and Kiggelariaceae, though widely separated in the original cladogram (Fig. 1), are doubtfully more than sister groups. Molecular coherence of large

internal portions of these families plus phytochemical distinctions combine, however, in rather reliably distinguishing them.

The stricture of Step Seven is true only when recognizing paraphyletic taxa proves no purpose. There may well be situations in which two or more patristically distant lineages are not functionally evolutionarily divergent (as inferred from phenome) though divergent in time (as tracked by neutral DNA base changes). In that case, such two or more lineages may be considered taxonomically one, with a functionally divergent separate taxon emerging from anywhere within that single ancestral group (phylogenetically well-structured species, genus, or family). An example would be an ancestral population isolated into two populations, then both continue to accumulate neutral mutations but retain the same phenome, a second population is isolated from the terminal population and diverges in phenome–environmental interaction through selection while the ancestral population is unchanged except for continued accumulation of neutral mutations. All three groups are different if viewed from a neutralist standpoint, yet the ancestral population and the first isolate are phenomically identical though paraphyletic. There may even be several phenomically evolutionarily divergent populations separating a surviving ancestral population and a daughter group with identical evolutionary strategies, which would argue against distinction of the two at higher taxonomic levels. A two-internode distance based on genes responding to recombination is only equivalent to a minimum distance of two genetically isolated panmictic or tokogenetic populations, while with matrilineally inherited (chloroplast or mitochondrial) DNA it is equivalent to a pedigree of a minimum of two ancestral individuals.

STEP EIGHT

Use Bayes' Formula to combine separately published branch arrangements based on different data. When two or more published trees are relevant to a BA of interest, select a branch arrangement of three taxa or lineages of interest, commonly from the most well-supported and complex cladogram. This one is then the "nominal branch arrangement." We ask, now, is this particular arrangement, say (AB)C, on the whole supported or falsified by previous studies? Bayes' Formula may be used. Likelihoods of model conditions have been previously addressed in computerized analysis of the published study, so the Formula as applied here is now a simple proportion. All probabilities must be conditional on being expressed as support for (AB)C, and this must be estimated from the different published studies, even

when lineages A, B, and C are presented as a bush, i.e., as (ABC). Any introduction of uncertainty must be addressed, and support for (AC)B of 0.70 can be represented as support of 0.30 for (AB)C, though this is, of course, generous. If (AB)C is supported by a value given as merely less than 0.50, 0.42 (midway between 0.33 and 0.50) is an appropriate assignment, though less than generous.

In cladograms with some branch arrangements collapsed because the support is less than 0.50, the multifurcation (ABC) should be interpreted as 0.33 support for (AB)C because, if the same data set does well support many branch arrangements in other areas of the cladogram, then (ABC) is evidence of a hard polytomy or resolved branch arrangement indistinguishable from such and cannot be ignored. In cases when support values may have been generated by randomness in the data set, such low support values can be ignored only if they tend to increase certainty (see Step Six); any decrease of certainty must be addressed in Step Eight.

The Bayes' Formula expresses a percentage when the joint probability (chance of the set of prior and nominal probability being both correct) is divided by a normalizing constant (chance of the set being correct plus the chance that the set is incorrect, and therefore summing all possible cases). Because of the normalizing constant, the chance of a hypothesis being correct increases with additional probabilities greater than 0.50. The Bayes' Formula is commonly used when the prior is a subjective measure of prior knowledge, but here the prior is entirely empirical, being the result of a previous analysis. Given that both prior and nominal branch arrangements must evaluate probabilistically (AB)C because the Formula is conditional on (AB)C being correct at some level of certainty, the posterior probability for (AB)C is:

$$\text{(Prior correct} \times \text{Nominal correct)} /$$
$$\text{((Prior correct} \times \text{Nominal correct)}$$
$$+ \text{(Prior incorrect} \times \text{Nominal incorrect))}$$

or

$$(\text{PBA}_{\text{paper 1}} \times \text{PBA}_{\text{paper 2}}) /$$
$$((\text{PBA}_{\text{paper 1}} \times \text{PBA}_{\text{paper 2}}) +$$
$$((1 - \text{PBA}_{\text{paper 1}}) \times (1 - \text{PBA}_{\text{paper 2}})))$$

This may be continued with several papers, with the posterior probability of one calculation serving as the prior (or "paper 1") of the next.

Inclusion of evidence from morphology (or chemistry, ultrastructure, etc.) cannot be ignored, although, with addition of data from several molecular data sets,

it justifiably may be overwhelmed by successive Bayesian analyses if contrary. When exact likelihoods are unavailable, Bayesian subjective priors may be assigned. If the morphological branch arrangement is "uncontested" or "accepted," and it agrees with the nominal (molecular) branch arrangement, assign it a prior of 0.99. If it disagrees, assign it 0.01. If the morphological branch arrangement is equivocal, assign a 0.50 prior; with Bayes' Formula, a 0.50 value does not change the posterior probability of the nominal branch arrangement.

Imaginary examples for piecemeal "total evidence" from separately published studies are here presented using Bayes' Formula with each posterior probability treated as the prior for the next calculation:

(1) The nominal branch arrangement of a molecularly derived conclusion of monophyly (AB)C supported by probability 0.80 is combined with a second result for (AB)C of 0.70: $(0.80 \times 0.70)/((0.80 \times 0.70) + (0.20 \times 0.30)) = \mathbf{0.90}$ as BPP.

(2) A third, morphological data set is equivocal between two branch arrangements, (AB)C and (AC)B. Using 0.90 as a prior, we combine it with 0.50, which is the probability from the equivocal data that either (AB)C or (AC)B is correct: $(0.90 \times 0.50)/((0.90 \times 0.50) + (0.10 \times 0.50)) = \mathbf{0.90}$, and the posterior probability stays the same.

(3) A fourth molecular data set supports the contrary tree (BC)A at 0.80, and therefore (AB)C is at most supported by 0.20: $(0.90 \times 0.20)/((0.90 \times 0.20) + (0.10 \times 0.80)) = \mathbf{0.70}$ as BPP that (AB)C is correct.

(4) A fifth data set with many other branch arrangements well resolved presents a bush (ABC) for the three taxa of interest, and (AB)C can be assigned a fail-safe 0.33: $(0.70 \times 0.33)/((0.70 \times 0.33) + (0.30 \times 0.67)) = \mathbf{0.54}$ BPP.

(5) A sixth data set with (AB)C at 0.80: $(0.54 \times 0.80)/((0.54 \times 0.80) + (0.46 \times 0.20)) = \mathbf{0.82}$ BPP.

(6) A seventh data set with (AB)C at 0.80 increases the BPP to $\mathbf{0.95}$, which indicates a reliable conclusion of monophyly based on all the data. Bayesians may or may not select beforehand a stopping rule for sequential procedures, but in any case many tests are needed to decide if the data are adequate for reliable resolution. One should not, however, expect the great variation expressed in this imaginary set of examples. Modern molecular analyses with large data sets from multiple loci, using robust technique, often produce very well-supported molecular trees.

It appears in Figure 4 that species of *Barbula* Hedw. are strongly supported as split into two groups: *B. bolleana* (Müll. Hal.) Broth. and *B. indica* (Hook.) Spreng. in Steudel associated with mostly species of subfamily Trichostomoideae (lineage "A" in Fig. 4),

and *B. unguiculata* Hedw. with mostly species of subfamily Barbuloideae (basal branches of "B"). The Trichostomoideae is held together at 0.96 probability, and the other BA at 0.99; thus a joint probability of 0.95 supports a hypothesis that the two groups of *Barbula* species are indeed separated by two internodes. Because the Werner et al. (2004) study used chloroplast DNA, which is matrilineally inherited as a haplotype, the exemplars represent individuals, not species. Genetic isolation begins with the individual with chloroplast and mitochondrial DNA. There are thus potentially as few as two ancestral individuals in the pedigree separating the two groups. On the other hand, the *Barbula* species in the Trichostomoideae have the plane leaves characteristic of the Trichostomoideae, which provides a 0.95 morphological prior yielding a 0.99 probability by Bayes' Formula that they do belong in the Trichostomoideae, while the other species of *Barbula* has the recurved margins of the Barbuloideae. This evaluation, thus, supports reclassification of plane-leaved species of *Barbula* with Trichostomoideae, although corroboration with additional study would be helpful.

STEP NINE

Examine the potential value of recognizing paraphyletic taxa by tempering strict monophyly with evidence for taxic unity based on inferred functional evolution and non-dichotomous evolutionary processes. Avoiding paraphyletic classificatory elements in favor of strict monophyly has been recently identified as a major problem with modern systematics (e.g., Brummitt, 2003, 2006; Nordal & Stedje, 2005; Hörlandl, 2006). Of three taxa or exemplar specimens, modern molecular phylogenetic analysis asks: "Which two share a recent event of genetic isolation, and how probable is this event?" After analysis, morphological traits are now commonly mapped on a molecular tree of such inferred events and conclusions made about morphological trait evolution, often involving discovery of abundant morphological homoplasy and convergence. The problem involves neutralist evolutionary philosophies coupled with cladistic insistence on strict monophyly.

Any long period of isolation of two or more populations results in accumulation of theoretically neutral DNA base changes in each population that clearly matches traditional cladistic speciation (two new taxa from one ancestral taxon, which then disappears). Requirement of strict monophyly assuming the biological species concept (involving genetic isolation) implies that any phylogenetic structure may be documented as a dichotomous tree down to basic tokogenetic or panmictic units. Many species, espe-

cially plants, have molecular phylogenetic structure associated with full or partial genetic isolation of populations, but such multiplex species can retain an essential morphological and ecological unity, for instance through processes like stabilizing selection. If paraphyletic groups may be described as having a unity in descent with modification by showing a distinctive morphology with reproductive and survival strategies, then they should be recognized in classification as single taxa. That such hypotheses of adaptation and selection sometimes have alternative explanations (Gould & Lewontin, 1979), albeit less persuading, is part of the incremental nature of scientific analysis.

Classification based solely on monophyletic groups (Ebach et al., 2006) is then crippled in explanatory power when a portion of such a group is nothing more than an isolated element of a greater assemblage that is biologically better recognized as a unit, and has practical value (Raven, 1974) across science. By extension, taxonomic splitting of genera and families to fit a relentless pursuit of molecular pedigree cannot produce concepts of taxa that document facts or hypothetical explanations of evolutionary process that involve multiple populations, but instead produce just a chart of inferred molecular ancestral relationships. Strict monophyly splits evolutionary units of multiple populations into cryptic species of no particular value except possession of a line of neutral mutations. Such finely dissected lineages are of dubious value in biodiversity triage and mask truly important differences among species needing conservation protection.

Computer-generated schemes of monophyly, either as reflected in morphological cladograms sorting all traits, or morphological traits or taxa mapped to molecular cladograms based on nuclear or organellar genes, cannot address evolutionary change adequately. When classifying evolutionarily divergent taxa, it is best to continue to recognize paraphyletic groups when their functional differences are not involved. For example, the Bayesian *rps4* cladogram of genera of the Pottiaceae (Bryopsida) and related genera of Werner et al. (2004) showed the family Ephemeraceae embedded deeply (two internodes, 0.95 joint probability) in the Pottiaceae (Fig. 4). Functional evolution is rather different for the two somewhat morphologically different families, with Ephemeraceae adapted to avoidance of harsh conditions while the Pottiaceae generally tolerates them. Though apparently molecularly embedded in the Pottiaceae, the Ephemeraceae should be recognized as a family with a separate evolutionary unity, leaving the Pottiaceae paraphyletic. Likewise the Cinclidotaceae, adapted to aquatic habitats, rates family status distinct from the paraphyletic Pottiaceae though well embedded (three

internodes, joint probability 0.97). The presence of *Goniomitrium* J. D. Hook. & Wils. (Funariaceae) was apparently based on a misidentification of an *Acaulon* Müll. Hal. species (Werner et al., 2007) and is here discounted. That systematists should accept at least occasional non-dichotomous evolution in classification has been well discussed by Hörandl (2006) in the context of recognizing paraphyly when appropriate.

In addition, one should check for contravention of Dollo's Rule against re-evolution of complex traits requiring consideration of reticulation and possible evolutionary Lazarus taxa. There are many examples of complex deep homologies re-expressed higher in a phylogenetic tree. These include: apparent re-evolution of wings in wingless walking sticks (mantids), shell coiling in snails, single eyes (a pleisiomorphic trait) in brine shrimp, eyes in eyeless copepods, modes of vulva formation in nematodes, ancestral features of the lateral lines, muscles, and gill rakers of cichlid fishes, teeth in chickens (laboratory), occurrence of a second molar in lynx (reviewed by Zander, 2006), and, recently, patterns of water storage tissue in *Adenia* Forrsk. (Passifloraceae; Hearn, 2006). These phenomena have been explained by gene cluster silencing (through, for example, cytidine methylation or histone deacetylation), surviving millions of years with little degradation (Fryer, 1999; Lynch & Force, 2000; Collin & Cipriani, 2003), and subsequent reactivation. Convergent evolution of complex organs and other massive homoplasy at a particular point become extremely improbable and may be taken as evidence of evolutionary Lazarus taxa (as opposed to geological Lazarus taxa that reappear after long periods in more recent strata). If identified correctly, reticulate evolution is the proper explanation, and taxonomic splitting is unwarranted.

As example, the same Bayesian *rps4* cladogram of genera of the Pottiaceae and related genera (Werner et al., 2004) showed *Timmiella* to occur toward the base of the tree far from the Pottiaceae proper. *Timmiella* is morphologically pleisiomorphic to the Pottiaceae, and it is found at the base of the morphological tree generated by Zander's (1993) cladistic analysis. This genus is rich in the characters of the family, possesses the distinctive twisted peristome of the family, and is not morphologically similar to any family nearby in the Werner et al. (2004) cladogram. A transformed version (Fig. 4) of the Werner et al. (2004) cladogram with all branch arrangements at 0.95 PBA or higher definitely shows *Timmiella* distinct from the Pottiaceae by three branch arrangements that are reliable as a set, with intervening lineages of other families. Although one might consider making *Timmiella* representative of a new family, when morphology is considered the cladogram contravenes Dollo's Rule,

and reticulation through gene cluster silencing and reactivation is a better and more parsimonious explanation.

An alternative explanation for the above (B. Goffinet, pers. comm.) is that lineages (other families) between *Timmiella* and Pottiaceae s. str. simply involved deactivation of the Pottiaceae traits, and there was no reactivation. In this case, however, populations comprising the sister groups of the intermediate lineages (the main line to Pottiaceae) would clearly be Pottiaceae (though gradually undergoing anagenetic changes in non-coding or evolutionarily neutral sequences), and the acceptance of intermediate lineages from a paraphyletic basal group is justified. Continued identity as one family through concatenated sister groups of other families, though more dubious than the former explanation, is equivalent to reticulation in the cladogram and can be interpreted as Lazarus-like survival. *Timmiella* is in any case justifiably retained in the Pottiaceae. The phylogenetic distance between *Timmiella* and the Pottiaceae may be considered merely three inferred events of genetic isolation following the Biological Species Concept, and it is less than the five events inferred between the base of Pottiaceae s. str. and the terminal group involving *Syntrichia* Brid. Given that tracking events of neutral evolution here involved matrilineally inherited chloroplast traits, as few as three ancestral individuals may separate *Timmiella* from the Pottiaceae s. str.; this admittedly extreme suggestion involves assumption of great clock heterogeneity.

DISCUSSION

Optimality alone in phylogenetic analysis as justification for public presentation is rarely seen today. Philosophical justifications for optimal or shortest trees, such as "simplicity," "converging on the truth," "most parsimonious," or "least falsified," as sufficient criteria of satisfactory results in phylogenetic analysis have been replaced with various methods of gauging statistical support (with reference to the Central Limit Theorem, which basically comes from physics) for trees or branches. In addition, as noted above, resolution alone is insufficient to demonstrate reliability because random data usually produce resolved trees; length of branches alone is insufficient because large random data sets generate long branches; and agreement alone between two cladograms is insufficient because if one or both arrangements are at less than 0.50 probability (given sufficient data), then, by Bayes' Formula, the BPP must be reduced, not increased. Phylogeneticists' use of Occam's Razor (Posada & Buckley, 2004) continues

to be abused, however, in that the past difficulty of evaluating the relative importance of suboptimal solutions has been repositioned to acceptance of optimal sequence alignments (analysis of Redelings & Suchard, 2005) and model selection (e.g., critique of Alfaro & Huelsenbeck, 2006), and simply assuming that all the other many assumptions associated with methods of phylogenetic reconstruction are correct or correct enough. Any use of "best" solutions when "next best" is large enough to crowd the reliability window of 0.05 is inappropriate. Although phylogenetic analysis of bryophytes using molecular data largely agrees with traditional concepts, major revisions in taxonomy have been urged for vascular plants on its basis, and clearly a revisitation of relevant molecular studies (e.g., APG II, 2003) is in order.

Huelsenbeck and Ronquist (2001) and Huelsenbeck and Immenov (2002) have described making the posterior probability of one analysis the prior for the next (in empirical Bayesian analysis) by substituting the posterior probability of the previous study prior for a non-informative prior during the next analysis. The present simple method using Bayes' Formula advocated here is, however, nonparametric, and can be used without complex analysis.

Information from both molecular and morphological studies must be dealt with, but just how is the puzzle. Morphological characters are good clues for inference of organismal evolutionary strategies for fitness and adaptation. Recognizing species and higher categories that document or hypothesize functional evolution (e.g., Robinson, 1990) or otherwise hypothetically explain the phenome and environment interaction are important in an age of rapid environmental change associated with human activity. Zander (1993) proposed a reclassification of the moss family Pottiaceae emphasizing traits of the gametophyte over those of the sporophyte, with the justification that parallel and largely identical transformation series of reduction in the sporophyte in many lineages well explains the similarities of species apportioned among traditional genera. Some but not all large genera in the family demonstrated a vast range of sporophyte reduction from peristomate, cylindrical capsules dehiscent by a lid and situated on elongate seta to nearly sessile, spherical capsules rupturing irregularly. Selective pressures on the phenotype resulting in atelochory (precinctiveness) in often patchy and harsh environments is an appropriate hypothesis accounting for this reduction. Molecular results to date commonly support to some extent Zander's new classification, but one might question whether probabilistic inferences of evolutionarily neutral ancestral traits really test these new hypotheses when neutral mutations provide only indirect and fragmented evidence of directional and

disruptive selection acting as a whole on several genetically isolated populations of a species, insistence on strict monophyly further scrambles interpretations of evolutionary process, and consideration of adaptation to particular environmental niches is generally given short shrift in modern systematics beyond mapping morphological traits on a molecular tree.

Step Nine may be the most controversial of the here recommended elements of reanalyzing published cladograms. In systematics, the major issues of our times are the phylogenetic or biological species concept versus the ecological species concept (e.g., Anderson, 1990; Grant, 1992; Pianka, 2000), the neutralist versus the selectionist positions on evolution (e.g., Ohta & Gillespie, 1996), and strict monophyly versus paraphyly (e.g., Brummitt, 2003; Ebach et al., 2006; Hörlandl, 2006). Within the scientific culture of phylogenetics, there are three commonly presented alternatives to traditional evolutionary hypotheses of adaptation and fitness involving phenome and environment. (1) Morphological phylogenetics produces a parsimonious tree or trees involving all presumably independent morphological traits, yet not all traits are evolutionarily functional. (2) With nuclear genes, molecular tracking goes right down to the nut of a tokogenetic or panmictic population. Yet above the panmictic population, there may be at times some or abundant gene flow between populations, of importance in evolution. Molecular phylogenetics is based on the idea that neutral base changes track events of genetic isolation, in practice equivalent to pan-neutralism. The phylogenetic species concept actually assumes the biological species concept, namely, it asks of three exemplar OTUs, which two probabilistically share a recent event of genetic isolation, then extends the inference to whatever the exemplar is supposed to represent. Molecular analysis of this sort should not work well with groups that are not particularly well described by the biological species concept (e.g., plants). There can be no "surviving ancestors" because as soon as a population is genetically isolated, neutral traits begin to accumulate divergently in both the surviving ancestral population and all divergent lineages. All new lines, given time, can be identified by accumulations of neutral traits whether or not their essential biorole or evolutionary strategy changes. (3) With matrilineally inherited molecular traits, such as mitochondrial and chloroplast loci, neutral base changes track the individual (e.g., "African Eve"); this is true also for patrilineal traits on Y-chromosomes, though these are seldom used in systematics. But above the individual, populations may be genetically isolated but retain phenomic and re-

productive similarity and an essentially identical interaction with the environment. Some genetically isolated populations are essentially identical to each other due to, for instance, stabilizing selection on exons. Evolution in the ecological context is not directly tracked in monophyletic classifications in modern work, though some studies may also treat it as "mapped." Thus, the fan-shaped pedigrees of matrilineally inherited neutral traits will support any exemplar associated with taxonomic splitting, because the molecular analysis is biased toward excessive taxonomic atomization. The consilience of molecular and morphological systematics may be more a function of similar bias in splitting than true support for the taxa supposedly represented by the exemplars.

Exemplars often represent quite different concepts: with matrilineal inheritance of molecular traits, it is the individual; with nuclear genes, the panmictic population; and with cladograms of morphological traits, a series of all traits whether involved in functional evolution or not. In effect, an exemplar enters phylogenetic analysis as a traditional concept including inferred evolutionary process but emerges as simply a molecular genealogy.

Any family newly discovered to be deeply embedded in a lineage of another family can and should be continued to be recognized at the family level if it is considered to represent a new evolutionary direction with associated changes in morphology, reproductive strategies, and ecological niche. Strict monophyly is a classificatory shortcut having the attraction of simplicity. Inferring a process of functional evolution eventually involves a large amount of work in population and reproductive biology, biophysics (for exactly what reason does this plant have this structure?), and biosystematics. We made a good start with the Modern Synthesis, but 30 years ago got sidetracked into a simplistic methodology focused on reconstruction of genealogy. Computerized parsimony and Bayesian analyses of morphology and molecular data are, of course, helpful in classification, but an evolutionarily-based classification requires evaluation, however difficult, of descent *with* modification, not just genealogies, however detailed.

In summary, the Nine Easy Steps provide a first pass at assessing uncertainty in published phylogenetic analyses with a view to providing well-supported theory on which to base new research. Those with interest in the fundamentals involved in the question of descent alone with traits mapped on dichotomous trees (phylogenetic analysis) versus descent with modification with trees or retes reflecting hypotheses of trait change evaluated in the context of reproductive and environmental strategies and the concept of niche (evolutionary analysis) need to closely examine

the conflict between purveyors of the biological species concept (enforced by standard phylogenetic analysis) and the ecological species concept, between neutralist (including mutationist) and selectionist evolutionary philosophies, and between those requiring strict monophyly and those allowing paraphyly.

Literature Cited

Alfaro, M. E. & J. P. Huelsenbeck. 2006. Comparative performance of Bayesian and AIC-based measures of phylogenetic model uncertainty. Syst. Biol. 55: 89–96.

Anderson, L. 1990. The driving force: Species concepts and ecology. Taxon 39: 375–382.

Anisimova, M. & O. Gascuel. 2006. Approximate likelihood-ratio test for branches: A fast, accurate, and powerful alternative. Syst. Biol. 55: 539–552.

APG II. 2003. An update of the Angiosperm Phylogeny Group classification for the orders and families of flowering plants: APG II. Bot. J. Linn. Soc. 141: 399–436.

Aris-Brosou, S. 2003. Least and most powerful tests to elucidate the origin of the seed plants in the presence of conflicting signals under misspecified models. Syst. Biol. 52: 781–793.

Benjamini, Y. & Y. Hochberg. 1995. Controlling the false discovery rate: A practical and powerful approach to multiple testing. J. Roy. Statistical Soc. B 57: 289–300.

Brummitt, R. K. 2003. Further dogged defense of paraphyletic taxa. Taxon 52: 803–804.

——. 2006. Am I a bony fish? Taxon 55: 268–269.

Chase, M. W., M. F. Fay & V. Savolainen. 2000. Higher-level classification in the angiosperms: New insights from the perspective of DNA sequence data. Taxon 49: 685–704.

Collin, R. & R. Cipriani. 2003. Dollo's law and the re-evolution of shell coiling. Proc. Roy. Soc. London, Ser. B, Biol. Sci. 270: 2551–2555.

Dolphin, K., R. Belshaw, C. D. Orme & D. L. Quicke. 2000. Noise and incongruence: Interpreting results of the incongruence length difference test. Molec. Phylogen. Evol. 17: 401–406.

Ebach, M. C., D. M. Williams & J. J. Morrone. 2006. Paraphyly is bad taxonomy. Taxon 55: 831–832.

Efron, B. 2005. Local false discovery rates. Dept. of Statistics, Stanford University, Technical Report 2005-20B/234.

——, E. Halloran & S. Holmes. 1996. Bootstrap confidence intervals for phylogenetic trees. Proc. Natl. Acad. Sci. U.S.A. 93: 7085–7090.

Fryer, G. 1999. Opinion: The case of the one-eyed brine shrimp: are ancient atavisms possible? J. Nat. Hist. 33: 791–798.

Giribet, G., W. C. Wheeler & J. Muona. 2002. DNA multiple sequence alignments. Pp. 107–114 in R. DeSalle, G. Giribet & W. Wheeler (editors), Molecular Systematics and Evolution: Theory and Practice. Birkhäuser Verlag, Basel, Switzerland.

Gould, S. J. & R. C. Lewontin. 1979. The spandrels of San Marco and the Panglossian paradigm: A critique of the adaptationist programme. Proc. Roy. Soc., London, Ser. B, Biol. Sci. 205: 581–598.

Grant, V. 1992. Comments on the ecological species concept. Taxon 41: 310–312.

Hearn, D. J. 2006. Adenia (Passifloraceae) and its adaptive radiation: Phylogeny and growth form diversification. Syst. Biol. 31: 805–821.

Hochberg, Y. 1988. A sharper Bonferroni procedure for multiple tests of significance. Biometrika 75: 800–802.

Holm, S. 1979. A simple sequentially rejective multiple test procedure. Scand. J. Stat. 5: 65–70.

Hörandl, E. 2006. Paraphyletic versus monophyletic taxa—evolutionary versus cladistic classifications. Taxon 55: 564–570.

Huelsenbeck, J. P. & N. S. Imennov. 2002. Geographic origin of human mitochondrial DNA: Accommodating phylogenetic uncertainty and model comparison. Syst. Biol. 51: 155–165.

—— & F. Ronquist. 2001. MrBayes: v30B4. Bayesian Analysis of Phylogeny. University of California, San Diego, and Dept. of Systematic Zoology, Uppsala University.

Kendall, M. G. & W. R. Buckland. 1971. A Dictionary of Statistical Terms, 3rd ed. Oliver & Boyd, Edinburgh.

Kolaczkowski, B. & J. W. Thornton. 2004. Performance of maximum parsimony and likelihood phylogenetics when evolution is heterogeneous. Nature 431: 980–984.

—— & ——. 2006. Is there a star tree paradox? Molec. Biol. Evol. 23: 1819–1823.

Lewis, P. O., M. T. Holder & K. E. Holsinger. 2005. Polytomies and Bayesian phylogenetic inference. Syst. Biol. 54: 241–253.

Lowry, R. 2005. VassarStats: Web Site for Statistical Computation. Department of Psychology, Vassar College, Poughkeepsie, New York, <http://faculty.vassar.edu/~lowry/VassarStats.html>, accessed 16 March 2005.

Lynch, M. & A. Force. 2000. The probability of duplicate gene preservation by subfunctionalization Genetics 154: 459–473.

Miller, Jr., R. G. 1981. Simultaneous Statistical Inference, 2nd ed. Springer, New York.

Nordal, I. & B. Stedje. 2005. Paraphyletic taxa should be accepted. Taxon 54: 5–6.

Ohta, T. & J. H. Gillespie. 1996. Development of neutral and nearly neutral theories. Theor. Populat. Biol. 49: 128–142.

Pianka, E. R. 2000. Evolutionary Ecology, 6th ed. Benjamin Cummings, San Francisco.

Pons, J., T. Barraclough, J. Gomez-Zurita, A. Cardoso, D. Duran, S. Hazell, S. Kamoun, W. Sumlin & A. Vogler. 2006. Sequence-based species delimitation for the DNA taxonomy of undescribed insects. Syst. Biol. 55: 595–609.

Posada, D. & T. R. Buckley. 2004. Model selection and model averaging in phylogenetics: Advantages of Akaike Information Criterion and Bayesian approaches over likelihood ratio tests. Syst. Biol. 53: 793–808.

Raven, P. H. 1974. Plant systematics 1947–1972. Ann. Missouri Bot. Gard. 61: 166–178.

Redelings, B. D. & M. A. Suchard. 2005. Joint Bayesian estimation of alignment and phylogeny. Syst. Biol. 54: 401–418.

Rice, W. R. 1989. Analyzing tables of statistical tests. Evolution 43: 223–225.

Robinson, H. 1990. A functional evolution of the Leucobryaceae. Trop. Bryol. 2: 223–237.

Rodrigo, A. G. 1993. Calibrating the bootstrap test of monophyly. Int. J. Parasitol. 23: 507–514.

Salamin, N., T. R. Hodkinson & V. Savolainen. 2002. Building supertrees: An empirical assessment using the grass family (Poaceae). Syst. Biol. 51: 136–150.

Sanderson, M. J. & M. F. Wojciechowski. 2000. Improved bootstrap confidence limits in large-scale phylogenies, with an example from neo-Astragalus (Leguminosae). Syst. Biol. 49: 671–685.

Savolainen, V., M. F. Fay, D. C. Albach, A. Backlund, M. van der Bank, K. M. Cameron, S. A. Johnson, M. D. Lledo, J. C. Pintaud, M. Powel, M. C. Sheahan, D. E. Soltis, P. S. Soltis, P. Weston, W. M. Whitten, K. J. Wurdack & M. W. Chase. 2000. Phylogeny of the eudicots: A nearly complete familial analysis based on *rbcL* gene sequences. Kew Bull. 55: 257–309.

Susko, E. 2006. Using minimum bootstrap support for splits to construct confidence regions for trees. Evol. Bioinf. Online 2: 137–151.

Suzuki, Y., G. V. Glazko & M. Nei. 2002. Overcredibility of molecular phylogenies obtained by Bayesian phylogenetics. Proc. Natl. Acad. Sci. U.S.A. 99: 16138–16143.

Swofford, D. L. 1998. PAUP*. Phylogenetic Analysis Using Parsimony (and Other Methods). Ver. 4. Sinauer Associates, Sunderland, Massachusetts.

Verhoeven, K. J. F., K. L. Simonsen & L. M. McIntyre. 2005. Implementing false discovery rate control: Increasing your power. Oikos 108: 643–647.

Werner, O., R. M. Ros, M. J. Cano & J. Guerra. 2004. Molecular phylogeny of Pottiaceae (Musci) based on chloroplast *rps4* sequence data. Pl. Syst. Evol. 243: 147–164.

———, ——— & B. Goffinet. 2007. A reconsideration of the systematic position of *Goniomitrium* (Funariaceae) based on chloroplast sequence markers. Bryologist 110: 108–114.

Zander, R. H. 1993. Genera of the Pottiaceae: Mosses of harsh environments. Bull. Buffalo Soc. Nat. Sci. 32: 1–378.

———. 1999. RANDSET generates random data files for phylogenetic analysis. Res Botanica Web site, Missouri Botanical Garden. <http://www.mobot.org/plantscience/ResBot/Phyl/Rand/Randset.htm>, accessed 28 June 2007.

———. 2001. A conditional probability of reconstruction measure for internal cladogram branches. Syst. Biol. 50: 425–437.

———. 2004. Minimal values for reliability of bootstrap and jackknife proportions, decay index, and Bayesian posterior probability. Phyloinformatics 2: 1–13.

———. 2005. Unaccounted Assumptions. Res Botanica Web Site, Missouri Botanical Garden. <http://www.mobot.org/plantscience/ResBot/phyl/unaccounted.htm>, accessed 28 June 2007.

———. 2006. The Pottiaceae s. str. as an evolutionary Lazarus taxon. J. Hattori Bot. Lab. 100: 581–602.

———. 2007. Web supplement: nine easy steps for constructing reliable trees from published phylogenetic analyses. Res Botanica, a Missouri Botanical Garden Web Site. <http://www.mobot.org/plantscience/resbot/Phyl/9Easy/WebSuppl.htm>, accessed 28 June 2007.

Zharkikh, A. & W.-H. Li. 1995. Estimation of confidence in phylogeny: Complete-and-partial bootstrap technique. Molec. Phylogen. Evol. 4: 44–63.

```
Note in PDF reprint: The last paragraph of page 693 says "these values are too low."
This is a mistake. The Bayesian posterior probabilities of short branches are actually too high.
See Table 3.
```