Null Model Selection, Compositional Bias, Character State Bias, and the Limits of Phylogenetic Information

James Lyons-Weiler and Guy A. Hoelzer†*

*Institute of Molecular Evolutionary Genetics, Department of Biology, Pennsylvania State University; and †Department of Biology and Department of Environmental and Resource Sciences, University of Nevada at Reno

Evolutionary trends and processes can distort phylogenetic information in sequences such that they do not reliably reflect the evolutionary processes that generate them. This fact of molecular evolution has a ubiquitous influence on the ability of researchers to adequately reconstruct genealogical relationships and histories of the processes of molecular evolution. This feature of phylogenetic inference can limit the capacity of researchers to adequately specify a relevant null hypothesis for testing hypothesis of relationships, data informativeness, and processes of molecular evolution. We show how this feature of historical inference also influences the exactness of the relative apparent synapomorphy analysis (RASA) test for phylogenetic signal and demonstrate how a permutation modification of the null hypothesis can improve the robustness of the underlying distributional assumption of the test. The RASA test (using either null model) was found not only to appropriately reject the combinability of independent lines of evidence for the relationships among the *Physalaemus pustulosus* frog species group, but also to be more appropriately sensitive to individual uninformative data sets than commonly used tree-based measures of signal, including the consistency index, the retention index, and the permutation tail probability test statistic.

Introduction

Methods for the estimation of phylogenetic relationships among organisms rely on the assumption that the genealogical relationships of those organisms are encoded in the distribution of character states among them. When this assumption is violated, the resulting inferences of genealogical relationships and character evolution are apt to be misleading (Lockhart et al. 1992). To avoid the misuse of uninformative data, this fundamental assumption can be tested using tests of matrix structure before tree topologies are derived (Lyons-Weiler, Hoelzer, and Tausch 1996; Lyons-Weiler and Hoelzer 1997). Such tests rely on the identification of an appropriate null universe of possible outcomes (Bryant 1995). The usual null model adopted is a random distribution of evidence of shared genealogy for all relationships among taxa. In this paper, we use our improved understanding of the null universe for this test to illustrate a number of generally important issues regarding the choice of null distributions in statistical molecular systematics. Specifically, we demonstrate that the tree-independent test for phylogenetic signal described by Lyons-Weiler, Hoelzer, and Tausch (1996), like most methods of phylogenetic analysis, is inappropriately sensitive to compositional biases, and we trace this sensitivity to differences between assumed and true null distributions. We further demonstrate that a permutation-based estimate of the null slope for the relative apparent synapomorphy analysis (RASA) test leads to an improvement in this aspect and discuss some of the general limitations of defining a representative null sampling distribution in specific cases. Although the limitations we discuss exist for all statistical tests of historical processes, they do not reduce the general utility of such tests.

Key words: phylogenetics, phylogenetic signal, RASA, compositional bias, permutation test.

Address for correspondence and reprints: James Lyons-Weiler, Institute of Molecular Evolutionary Genetics, 311 Mueller Laboratory, Department of Biology, The Pennsylvania State University, University Park, Pennsylvania 16802. E-mail: jfl8@psu.edu.

Mol. Biol. Evol. 16(10):1400–1405. 1999

Compositional Bias Versus Character State Bias

One potential pitfall for phylogenetic estimation from biological sequence data is compositional bias. While this problem is not unique to molecular data, it has been most clearly shown in this context. Although most genes and genomes exhibit some degree of compositional bias, it need not be accompanied by character state biases (i.e., biases in the ratios of states within characters), but it does make the latter and its concomitant problems more likely. The degree to which a compositional bias is accompanied by a character state bias is strongly dependent on the details of the distribution of character states at each site (Collins, Wimberger, and Naylor 1994). For example, it is common for a character state bias to exist at a given site without any compositional bias evident in the overall data set; indeed, all sites with configurations that do not exhibit even character state ratios are biased, but in many cases this accurately reflects phylogenetic relationships. Thus, character state biases caused by shared compositional biases can be phylogenetically informative (e.g., Powell and Moriyama 1997) without necessarily indicating the occurrence of misleading homoplasy. When this is the case, it would be counterproductive to ''correct'' for variation in nucleotide composition among taxa, making an accurate phylogenetic inference less likely. On the other hand, character state biases at individual sites are problematic for methods of phylogenetic inference when they reflect convergent trends in nucleotide composition, because they can cause unrelated organisms to cluster due to shared base composition instead of phylogenetic affinity. The challenge in dealing with variation in compositional biases is no different from that in dealing with other sorts of variation in a phylogenetic context; that is, to distinguish between synapomorphic (i.e., phylogenetically informative) and homoplastic (i.e., phylogenetically misleading) similarities. Clearly, tree-based

measures of phylogenetic signal would be hard-pressed to distinguish these cases, because the optimal tree set and characteristics of its members can be misled by variation in a substantial fraction of characters which might be mutually influenced by compositional convergence.

Compositional biases can be found throughout the living world (e.g., Hasagawa and Hashimoto 1993; Powell and Moriyama 1997; Morton and Clegg 1995) and may play an adaptive role in transcription efficiency (Eyre-Walker 1996; Morton 1996; Xia 1996), making interpretation in a phylogenetic context complex. The LogDet transformation (Steel, Lockhart, and Penny 1993; Lake 1994; Lockhart et al. 1994) and other corrections (Sidow and Wilson 1990, 1991) have been devised for dealing with compositional bias. However, it is important to consider that although these methods can be viewed as strategies for noise reduction, and this may often be the net effect of their implementation, such corrections can add new sources of phylogenetic noise by treating all individual sites, or sites within predefined classes, the same. This is inevitable, because not all sites will exhibit the influence of convergent compositional biases; thus, corrections to distances should be expected to ''overcorrect'' in many cases. All methods that treat a set of sites as a sample from a single probability density function result in a loss of information for some sites (i.e., they are ''lossy'' transforms) when the same function is not appropriate for each site in the set (e.g., when the distribution of phylogenetic signal among sites does not match designated site classes, such as sites classified according to codon position).

Excessive compositionally induced character state biases limit the phylogenetic information that can be encoded in a data matrix, and they are known to be a source of difficulty in all methods of phylogeny estimation (see review by Collins, Wimberger, and Naylor 1994). For example, it is well known that characters with the greatest possible bias in state frequencies, those with a single unique taxon and a single state shared by all other taxa, are entirely uninformative in the context of cladistic parsimony. It is also clear that the probability of inferring erroneous nodes on trees is greater for matrices with lower phylogenetic signal (Lyons-Weiler, Hoelzer, and Tausch 1996; Lyons-Weiler and Hoelzer 1997). Therefore, an abundance of characters with compositionally induced biased state frequencies is expected to obscure the hierarchical patterns in the matrix caused by phylogenetic history, making accurate phylogenetic inference less likely (Saccone, Pesole, and Preparata 1989; Saccone et al. 1990; Saccone, Lanave, and Pesole 1993; Steel, Lockhart, and Penny 1993; Lockhart et al. 1994; Steel 1994; Pesole et al. 1995). Specifically, one effect of evolving an increased compositional bias is to reduce the effective number of character states per character; thus, the increased mean character state bias that typically attends a matrixwide compositional bias can reduce the probability of accurate phylogenetic inference, even in the absence of convergent compositional biases.

The Null Distribution of t_{RASA}

We explored the influence of compositional biases on inferences drawn from one particular measure of phylogenetic signal, called RASA (Lyons-Weiler, Hoelzer, and Tausch 1996), and the role of the assumed null distribution of the RASA test statistic (t_{RASA}) on that influence. RASA is a tree-independent statistical test for phylogenetic signal that can be solved in polynomial time. The purpose of the RASA test for signal is to provide a critical test of the assumption that the combined influences of the processes of evolution and sampling of both taxa and characters have resulted in a distribution of character states among taxa that is reflective of genealogical relationships. The test statistic, t_{RASA}, provides a measure of the difference between an observed (β_{obs}) and null (β_{null}) rate of increase in apparent phylogenetic similarity (RAS) with phenetic similarity (E) for pairs of taxa in the matrix regression model

$$\widehat{\mathrm{RAS}}_m = \beta_{obs}(E_m) + \beta_0 + \epsilon_m, \qquad (1)$$

where ϵ_m is the error with which apparent cladistic similarity of the mth taxon pair is estimated by the model and β_0 is the intercept. Under the special rule of equiprobability, the null expected frequency of any event is the mean frequency of such events. This rule was invoked to achieve the analytical estimate of β_{null} (Lyons-Weiler, Hoelzer, and Tausch 1996), which, in its most elegant form, is

$$\beta_{null} = \frac{\sum \mathrm{RAS}}{\sum E} \qquad (2)$$

(D. Colless, personal communication). Lyons-Weiler, Hoelzer, and Tausch (1996) demonstrated that t_{RASA} is appropriately sensitive to several variables that influence the amount of phylogenetic signal in matrices, including the amount of mutation per internode on trees, the numbers of characters, and the number of character states. These authors also compared the effects of these factors on signal under limited null conditions. When phylogenetic information is low or absent, conflicting statements of relationships exist in the matrix, which decreases measured signal in two ways: random homoplasy tends to both increase the error term in the test statistic and reduce the difference between the observed and null slopes. Applications that use phylogenetic signal as a criterion include the identification of long edges (another source of systematic error for methods of phylogenetic inference; Lyons-Weiler and Hoelzer 1997) and outgroup selection (Lyons-Weiler, Hoelzer, and Tausch 1998). Lyons-Weiler and Hoelzer (1997) demonstrated that t_{RASA} tracks the distortion of hierarchy in matrices caused by long- edge taxa, including the attrition of signal as more characters are sampled when long-edge taxa are present. Long-edge taxa also leave a characteristic, localized footprint in the error structure of the regression model, providing a means to detect long-branch taxa. The RASA test, therefore, provides a window into the components of structure in phylogenetic matrices not available from tree-based measures of information. The combined use of the test statistic as a

Table 1
Estimated *P* Values for the Critical Value $\alpha = 0.05$ of t_{RASA} Based on Student's *t* Distribution Using Both β_{null} and β_p

Character State Bias	N_t	N_c	Empirical Alpha, β_{null}	Empirical Alpha, β_p
None (1:1:1:1) . . .	7	12	0.064	0.008
	7	40	0.022	0.055
	7	100	0.073	0.024
	9	12	0.009	0.024
	9	40	0.026	0.062
	9	100	0.027	0.067
6:1:1:1	7	12	0.175	0.043
	7	40	0.209	0.052
	7	100	0.225	0.048
	9	12	0.264	0.049
	9	40	0.329	0.053
	9	100	0.318	0.060
9:1:1:1	7	12	0.172	0.035
	7	40	0.270	0.069
	7	100	0.235	0.055
	9	12	0.283	0.046
	9	40	0.318	0.054
	9	100	0.323	0.065

NOTE.—N_t = number of taxa; N_c = number of characters; β_{null} = the analytical null slope; β_p = the permutation-based null slope.

criterion and the taxon variance comparison as a window into the phenetic-cladistic covariance structure of a matrix provides a guide for optimal taxon sampling. The test statistic is also useful outside of a strictly phylogenetic context as a measure of the degree of hierarchy in matrices of species' distribution (Lyons-Weiler and Tausch 1996), and has been used as an indicator of the effects of differential lineage sorting (Lyons-Weiler and Milinkovitch 1997). A worked example of the test is available (Lyons-Weiler, Hoelzer, and Tausch 1998).

Given the apparent general utility of t_{RASA}, its statistical properties deserve scrutiny. The sampling distribution of the test statistic (Lyons-Weiler, Hoelzer, and Tausch 1996) was originally approximated by Student's *t* distribution, as would be appropriate if the null model was both sufficiently equiprobable and independent of the observed model. However, it has since been determined that character state biases can cause the test to yield inflated empirical *P* values (and correspondingly inflated Type I error rates) under Student's *t* distribution. The reason for the undesirable sensitivity of the test to character state biases appears to lie in construction of the null model (β_{null}), which is algebraically dependent on the mean empirical values of RAS and E alone. In fact, β_{null} is defined by a line connecting the origin to the point [RAS/*E*]. However, character state biases cause the observed linear regression to pass below the origin. Constraining the null slope to pass through the origin artificially reduces β_{null} relative to β_{obs}, making the null and observed slopes more different than they ought to be when the average character state bias in the matrix is too great; hence, there is an association between increasing character state bias and inflated type I error rate (table 1).

For any statistical test, the frequency with which matrices generated under completely null conditions exceeds particular critical values of a sampling distribution (e.g., those for $\alpha = 0.05$, 0.01, etc.) reflects the fit of the test to the distributional assumption. If the appropriate null distribution is used, these empirical *P* values will approximate α. We compared the statistical properties of a revised test that uses a permutation-based estimate of the null slope with those of the original test. This null parameter, β_p, is determined by randomizing the character states within characters and then determining an observed slope for each randomized matrix. Like β_{obs}, these slopes are not constrained to pass through the origin. The mean of this distribution of observed slopes is used as an estimate of the null parameter (β_p). States are randomized without replacement within sites; thus, the randomization step is identical to that used by Archie (1989). This approach to estimating the null slope in a RASA analysis involves a modest increase in computing time, but it appears to invoke a more appropriate null distribution. Specifically, it defines a null universe in which all matrices share the same distribution of character state biases but include all possible hierarchical and nonhierarchical patterns consistent with that distribution. The result is that $\beta_p > \beta_{null}$, and t_{RASA} is usually reduced with the use of β_p.

We generated 100 random four-state matrices for a range of compositional biases (table 1) and determined t_{RASA} using β_{null} for each. We then performed 30 permutations of each matrix to obtain β_p and estimated t_{RASA} using the error term of the observed regression model. The error terms for the regression used to derive β_p were ignored because the test was designed to study the observed slope and the distribution of error around that slope. Empirical *P* values were determined for the old and new test statistics, based on both β_{null} and β_p, by determining the proportion of the randomly generated matrices that resulted in a test statistic greater than the critical value of Student's *t* distribution.

The difference between β_{obs} and β_{null} was found to be larger for matrices generated with a compositional bias. This difference was not observed when the permutation null parameter was implemented. In biological terms, character state biases can generate patterns of shared similarities and dissimilarities among taxa that mimic the pattern caused by phylogeny. In this case, the permutation estimate of the null hypothesis anticipates the steeper relationship between RAS and *E,* typical of such matrices, to provide a sufficiently exact test. Accordingly, the revised test statistic better approximates Student's *t* distribution under a wider range of conditions and produces fewer type I errors under null conditions. The results (table 1) indicate that β_p generally provides a better estimate of the null universe than does β_{null}, and its use thus improves the empirical *P* values of the RASA test. Moreover, because the original test statistic changes in tandem with the new test statistic (unpublished data), data exploration using maximum signal as a criterion remains effective and worthwhile when the more computationally efficient β_{null} is used.

Table 2
Signal is Overestimated in Data partitions of the *Physalaemus pustulosus* Frog Species Group Under the Analytical Null Model as well as Various Tree-Dependent Measures of Signal

	Tree-Dependent Measures			RASA	
Data Partition	CI[a]	RI[a]	PTP[a]	t_{RASA} (β_{null})	t_{RASA} (β_p)
Combined	0.68	0.60	0.0002	8.868	5.609
12S	0.73	0.66	0.0002	10.544	7.4006
COI	0.60	0.45	0.0002	1.848(NS)*	−1.189(NS)*
Allozymes	0.80	0.53	0.0002	4.695	4.532
Calls	0.71	0.61	0.0004	−0.7441(NS)*	−0.9718(NS)*
Morphology	1.00	1.00	0.0002	5.56	4.566

Note.—t_{RASA} (β_{null}) = the RASA test statistic derived using the analytical null model; t_{RASA} (β_p) = the RASA test statistic derived using the permutation null model described in this paper. Note that the permutation tail probability (PTP) test fails to discriminate between any of the partitions.

[a] Results from Cannatella et al. 1998.

*NS = not significant at $\alpha = 0.05$.

Limitations of Permutation Theory

The characteristics of any test are influenced by the degree and frequency with which the assumptions of the test are met in specific cases, which then largely determine the Type I and Type II error rates. One assumption is that a null distribution sufficiently represents the actual sampling distribution. The sampling distribution of any test can be thought of as the distribution of possible outcomes under the same processes minus the effect being tested for. For any test, when we use a given data set to derive a null model, that derivation may not adequately describe the range of possibilities under the correct null model. The derived null distribution might be constrained in obscure ways that invoke unwanted and implicit changes to the intended null hypothesis. In the present case, the use of β_{null} has the undesirable effect that the null expectation includes values of t_{RASA} that correspond to the degree of structure expected without considering the distribution of character state biases in the focal matrix.

An inaccurate representation of the null universe can result even when all possible permutations of observed data are considered. For example, when a sampled distribution is biased, the estimated null distribution can differ from the correct null distribution, and the result will be an inflation of Type I error rates. This type of problem is revealed for permutation tests of phylogenetic matrices when compositionally induced character state bias is considered. In the study of molecular evolution, two types of evolutionary models may be invoked: a stationary (equilibrium) model, or a dynamic model in which the model varies over time and among lineages. In the determination of the conditions that obtain for our ''random'' matrices, we were omniscient in the sense that we knew of all of the factors that influenced the character state frequencies. In empirical studies, observed (extant) compositional biases may reflect either a stationary (ergodic) process or a stage in a dynamic process with appreciable fluctuations over time, with or without an accumulating trend, and with or without reversals in trends. Historical trends can pose a problem for the accurate determination of a null sampling distribution using extant sequences. Depending on the dissimilarities in shape of the relevant distributions, a null distribution derived under an ergodic assumption may be more limited than one derived under a fluctuating model in which any and all trends may be accommodated. In such instances, the ergodic assumption will lead to overly extreme estimates of probabilities. In practice, the use of permutation to achieve any null sampling distribution imposes the assumption of stationarity of state usage. In the present case, a permutation-based point estimate of the null slope allows for improved exactness by considering current character state biases, but it cannot guarantee exactness, because by using the observed data as a guide to the null distribution of possible outcomes, permutation assumes ergodic compositional and character state biases. Tests that apply null models that implicitly or explicitly assume stationarity (as most do) should be reexamined for exactness with fixed character state biases, and perhaps for dynamic compositional biases.

It is clear that permutation-based estimates of parameters are far from assumption-free. Furthermore, evaluating the performance of tests using many permutations of only a single matrix can lead to spuriously characterized empirical *P* values. We avoided this by using permutations of 100 different matrices instead of 100 (or more) permutations of a single matrix.

Empirical Example

Table 2 provides the analytical and permutation-based measures of signal for five recently published partitions of phylogenetic data for the frogs of the *Physalaemus pustulosus* species group (Cannatella et al. 1998). Our permutation-based estimate of the null slope in each was determined using 30 randomizations for each partition. The overestimation of phylogenetic signal for most partitions (12S sequences, allozymes, and morphology) is evident from the larger test statistics generated using the analytical null model. Interestingly, for two partitions (cytochrome oxidase I [COI] sequences and call characters), the null hypothesis that the data are phylogenetically uninformative cannot be rejected. Other, tree-dependent, measures of phylogenetic informativeness, including the consistency index, the reten-

tion index, and the permutation tail probability test, all indicate that the vocalization (call) characters are at least as informative as the other partitions, despite the determination from subsequent tree congruence tests and one pairwise heterogeneity test that the call characters should not be combined with the other partitions (Cannatella et al. 1998). In addition, the COI tree was found to be topologically very distinct from the other trees; yet, again, the tree-dependent assessments of the COI data did not indicate any potential difficulties. This appears to be an example wherein tree-dependent measures of signal fail to identify individual data sets that are misleading with respect to phylogenetic relationships, while the tree-independent RASA test finds these data sets uninformative. Given the apparent poor ability of the tree-dependent tests to identify individual (as opposed to combined) problematic data sets, researchers should be very cautious in their interpretation of tree-based measures of the phylogenetic informativeness of their data and tree-dependent tests of partition heterogeneity (Lyons-Weiler, Hoelzer, and Tausch 1996).

Discussion

The fact that unknown factors may impinge on ultimate inferences is not at all restricted to measures of signal. Lyons-Weiler and Hoelzer (1997), for example, demonstrated that the topology of the true phylogeny has a major effect on the efficiency of maximum parsimony; Sullivan, Holsinger, and Simon (1996) found similar sensitivities of maximum-likelihood analysis to parameters in likelihood models of molecular evolution. Because taxon sampling largely determines the true topology, taxon sampling has a major impact on phylogenetic accuracy (Lecointre et al. 1993); however, all methods of phylogenetic estimation invoke the assumption that every topology is equivalent to others with respect to its influence on phylogenetic signal. This assumption is obviously violated in most cases.

The negative influences of compositionally induced character state biases can be detected and combated. Statistical comparisons of compositional bias within a taxon to the character state biases exhibited by the remainder of the matrix should help delineate cases in which compositional distributions and character state distributions appear to fit the ergodic assumption from cases in which localized (phylogenetic) compositional biases exist. In addition, comparisons of the RASA test statistic determined with the analytical null model to one determined via the permutation null model might provide an indication of the misleading effects of character state biases in the matrix. Such tests might also be useful in determining when to invoke the compositional corrections of Sidow and Wilson (1990, 1991) and others (such as the LogDet/paralinear distance).

Our results also suggest that researchers should consider using tree-independent measures of phylogenetic signal as criteria to examine the influences of character and taxon sampling. For example, changes in t_{RASA} can be applied as an objective (maximizing) function for testing the effects of taxon and character sampling in a particular case, even when the assumptions of the test are violated. The RASA framework fundamentally demonstrates that statistical inference, especially the type of statistical inference concerned with Type II error minimization (classical statistical inference), is especially useful in the context of phylogenetic data exploration and a priori testing of the informativeness of phylogenetic data. In this way, it is seen that statistics and cladistics are compatible and that statistical inferences of the right flavor can provide much useful information about phylogenetic data that is not always accessible when tree-dependent methods are used.

We improved the RASA test statistic for comparisons with Student's t distribution with the use of a permutation-based estimate of the null parameter in place of the original, analytical, null parameter. This improvement is accompanied by a marked reduction in the sensitivity of the test to matrixwide biases in character state usage (i.e., compositional biases) that translate into redundant patterns within characters. Developers and users of methods of phylogenetic inference should keep in mind that every method of systematic inference and every test requires assumptions and that all methods have limitations. In the long run, our comprehension of the processes of molecular evolution is aided, not hindered, by cataloging the limitations of methods of phylogenetic inference. More effort is needed to make the limitations of all of our inferential tools obvious, because only then can tests for the violations in the specific case be devised, and only then can improvements in methodology occur. Given that so much of the structure of a matrix is unveiled in the RASA regression, we expect that the test will continue to provide new insights into comparative analyses. Software available online (Lyons-Weiler 1999) now implements either null model.

Acknowledgments

We owe an immense debt to S. Sawyer for bringing to our attention the sensitivity of the analytical test statistic to character state biases, for invaluable discussions on the limits of statistical inference, and for suggesting some aspects of the regime for testing the exactness of our test. Appreciation is extended to D. Colless for a comparison of t_{RASA} with other variants of the test statistic, to G. Naylor and R. Tausch for informed discussions of null models, and to M. Nei for encouragement and support. Two anonymous reviewers provided useful comments.

LITERATURE CITED

ARCHIE, J. W. 1989. A randomization test for phylogenetic information in systematic data. Syst. Zool. **38**:219–252.

BRYANT, H. N. 1992. The role of permutation probability tests in phylogenetic systematics. Syst. Biol. **41**:258–263.

CANNATELLA, D. C., D. M. HILLIS, P. T. CHIPPENDALE, L. WEIGHT, S. RAND, and M. J. RYAN. 1998. Phylogeny of the frogs of the *Physalaemus pustulosus* species group with an examination of data incongruence. Syst. Biol. **47**:311–335.

COLLINS, T. M., P. H. WIMBERGER, and G. J. P. NAYLOR. 1994. Compositional bias, character-state bias, and character-state reconstruction using parsimony. Syst. Biol. **43**:482–496.

EYRE-WALKER, A. 1996. Synonymous codon bias is related to gene length in *Escherichia coli*: selection for translational accuracy? Mol. Biol. Evol. **13**:864–872.

HASAGAWA, M., and T. HASHIMOTO. 1993. Ribosomal RNA trees misleading? Nature **361**:23.

LAKE, J. A. 1994. Reconstructing evolutionary trees from DNA and protein sequences: paralinear distances. Proc. Natl. Acad. Sci. USA **91**:1455–1459.

LECOINTRE, G., H. PHILIPPE, H. L. VAN LÊ, and H. LE GUYADER. 1993. Species sampling has a major impact on phylogenetic inference. Mol. Phylogenet. Evol. **2**:205–224.

LOCKHART, P. J., C. J. HOWE, D. A. BRYANT, T. J. BEANLAND, and A. W. D. LARKUM. 1992. Substitutional bias confounds inference of cyanelle origins from sequence data. J. Mol. Evol. **34**:153–162.

LOCKHART, P. J., M. A. STEEL, M. D. HENDY, and D. PENNY. 1994. Recovering evolutionary trees under a more realistic model of sequence evolution. Mol. Biol. Evol. **11**:605–612.

LYONS-WEILER, J. 1999. RASA 2.3 for Macintosh and manual. http://test1.bio.psu.edu/LW/rasatext.html.

LYONS-WEILER, J., and G. A. HOELZER. 1997. Escaping from the Felsenstein Zone by detecting long branches in phylogenetic data. Mol. Phylogenet. Evol. **8**:375–384.

LYONS-WEILER, J., 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.

———. 1998. Optimal outgroup analysis. Biol. J. Linn. Soc. **64**:492–511.

LYONS-WEILER, J., and M. MILINKOVITCH. 1997. A phylogenetic approach to the problem of differential lineage sorting. Mol. Biol. Evol. **14**:968–975.

LYONS-WEILER, J., and R. J. TAUSCH. 1997. The demarcation of historical from ecological variance in species diversity patterns. Pp. 209–221 *in* T. N. KAYE, R. M. LOVE, D. L. LUOMA, R. J. MEINKE, and M. V. WILSON, eds. Conservation and management of native plants and fungi. Native Plant Society of Oregon, Corvallis, OR.

MORTON, B. R. 1996. Selection on the codon bias of *Chlamydomonas reinhardtii* chloroplast genes and the plant psbA gene. J. Mol. Evol. **43**:28–32.

MORTON, B. R., and M. T. CLEGG. 1995. Neighboring base composition is strongly correlated with base substitution bias in a region of the chloroplast genome. J. Mol. Evol. **41**:597–603.

PESOLE, G., G. DELLISANTI, G. PREPARATA, and C. SACCONE. 1995. The importance of base composition in the correct assessment of genetic distance. J. Mol. Evol. **41**:1124–1127.

POWELL, J. R. and E. N. MORIYAMA. 1997. Evolution of codon usage bias in *Drosophila*. Proc. Natl. Acad. Sci. USA **94**: 7784–7790.

SACCONE, C., C. LANAVE, and G. PESOLE. 1993. Time and biosequences. J. Mol. Evol. **37**:154–159.

SACCONE, C., C. LANAVE, G. PESOLE, and G. PREPARATA. 1990. Influence of base composition on quantitative estimates of gene evolution. Methods Enzymol. **183**:570–583.

SACCONE, C., G. PESOLE, and G. PREPARATA. 1989. DNA microenvironments and the molecular clock. J. Mol. Evol. **29**: 407–411.

SIDOW, A., and A. C. WILSON. 1990. Compositional statistics: an improvement of evolutionary parsimony and its application to deep branches in the tree of life. J. Mol. Evol. **31**: 51–68.

———. 1991. Compositional statistics evaluated by computer simulations. Pp. 129–146 *in* M. M. MIYAMOTO and J. CRACRAFT, eds. Phylogenetic analysis of DNA sequences. Oxford University Press, New York.

STEEL, M. A. 1994. Recovering a tree from the leaf colorations it generates under a Markov model. Appl. Math. Lett. **7**: 19–23.

STEEL, M. A., P. J. LOCKHART, and D. PENNY. 1993. Confidence in evolutionary trees from biological sequence data. Nature **364**:440–442.

SULLIVAN, J., K. E. HOLSINGER, and C. SIMON. 1996. The effect of topology on estimates of among site rate variation. J. Mol. Evol. **42**:308–312.

XIA, X. 1996. Maximizing transcription efficiency causes codon usage bias. Genetics **144**:1309–1320.

MICHAEL HENDY, reviewing editor

Accepted July 5, 1999