Exploring the Phylogenetic Structure of Ecological Communities: An Example for Rain Forest Trees

Campbell O. Webb[*]

Arnold Arboretum of Harvard University, Cambridge, Massachusetts 02138

abstract: Because of the correlation expected between the phylogenetic relatedness of two taxa and their net ecological similarity, a measure of the overall phylogenetic relatedness of a community of interacting organisms can be used to investigate the contemporary ecological processes that structure community composition. I describe two indices that use the number of nodes that separate taxa on a phylogeny as a measure of their phylogenetic relatedness. As an example of the use of these indices in community analysis, I compared the mean observed net relatedness of trees (≥10 cm diameter at breast height) in each of 28 plots (each 0.16 ha) in a Bornean rain forest with the net relatedness expected if species were drawn randomly from the species pool (of the 324 species in the 28 plots), using a supertree that I assembled from published sources. I found that the species in plots were more phylogenetically related than expected by chance, a result that was insensitive to various modifications to the basic methodology. I tentatively infer that variation in habitat among plots causes ecologically more similar species to co-occur within plots. Finally, I suggest a range of applications for phylogenetic relatedness measures in community analysis.

Keywords: phylogenetic conservatism, net ecological similarity, supertree, tropical rain forest, taxonomic diversity.

The search for patterns in the species composition and dynamics of local communities of interacting organisms, and for the processes that cause pattern, has seldom employed information about the phylogenetic relationships of species within those communities. Instead, species are usually treated as equivalent units, with independent functional traits (e.g., Diamond and Case 1986; Roughgarden 1989; Webb and Peart 1999; Weiher and Keddy 1999; but see Cotgreave and Harvey 1991). However, because of the conservatism of many species traits in the evolution of a

[*] E-mail: cwebb@oeb.harvard.edu.

lineage, we expect, in general, a positive relationship between a measure of the phylogenetic relatedness of two species and a measure of their overall life-history and ecological similarity (Brooks and McLennan 1991; Harvey and Pagel 1991; Silvertown et al. 1997). Hence, the analysis of the phylogenetic structure of a community should provide insight into the ecological processes that organize the community. In this article, I describe two simple, general methods for quantifying the phylogenetic structure of local communities of interacting organisms (relative to a shared species pool) and use data from a rain forest tree community in Borneo as an example of how this method can be used to address questions of community organization.

Significant advances have been made in understanding the historical and biogeographic development of communities by studying the phylogeny and taxonomy of regional species assemblages (e.g., Gorman 1992; Cadle and Greene 1993; Losos 1996; Grandcolas 1998; Qian and Ricklefs 1999). The member species of these regional assemblages (biotas) may come together to form local communities of interacting organisms. Any ecological interactions within these communities that cause the local extinction of species may in turn shape the composition of the regional assemblage (Richman and Price 1992). Several researchers have explored the taxonomic structure of regional assemblages, searching for evidence that local ecological interactions are important in this way. In particular, patterns in the "taxonomic diversity" (Simberloff 1970) of communities have been studied, especially of those on islands, comparing species-to-genera ratios to a priori expectations (Elton 1946; Moreau 1948; Williams 1964) or to explicit null models (Simberloff 1970; reviewed by Strong 1980). The expectation has generally been that competition among closely related (congeneric) species should lead to the regional extinction of all but a few species in a genus, which would be reflected in lower species-to-genera ratios in island biota than in mainland biota.

The method I describe here follows from the same ecological questions driving these analyses of species-to-genera ratios, but for the first time, considers a much

smaller spatial scale. Local communities of potentially interacting individuals (e.g., 10^1–10^3-m^2 scale for forest trees) are compared within the context of a larger species pool (e.g., 10^4–10^6-m^2 scale for trees) to assess the role of contemporary ecological interactions in determining the local community composition. This new approach also differs from earlier studies of the taxonomic diversity of assemblages in using an explicitly phylogenetic approach. The search for patterns in rank-based taxonomic structure is hampered by the partly subjective nature of traditional classification, which often represents overall similarity rather than the possession of shared derived characters and in which ranks may differ greatly in age and species richness among different lineages (Miles and Dunham 1993). Now, as well-supported, cladistic, molecular phylogenies are being published for many organisms (e.g., Sibley and Ahlquist 1990; Chase et al. 1993; Soltis et al. 1999), we can begin to examine the explicit phylogenetic structure of the species in a community.

A simple, initial way to quantify phylogenetic structure, having a goal in common with previous studies of taxonomic diversity, is to derive indices that estimate the overall phylogenetic relatedness of a community. With this approach, a community consisting of many species in the same terminal clade (e.g., genus) would have a high index of net relatedness, while one consisting of species all from different terminal clades would have a lower index. If ecological similarity is correlated with the phylogenetic relatedness of species, then this index would also reflect the "net ecological similarity" of species within a community.

The expectation that related species should be more similar than unrelated ones is the basis of the comparative method in biology (Harvey and Pagel 1991) and arises because phylogenetic constraints to evolutionary change, resulting from inherited developmental pathways, can restrict the range of available trait space on which selection can act. Selection itself may also act to prevent the niche of an organism diverging ("phylogenetic niche conservatism"; Lord et al. 1995). Under certain circumstances, where new niches are encountered (e.g., on islands) or where competition with an ecologically similar species is strong and predictable, particular ecological traits of an organism may change adaptively over time (Schluter 1994), leading to a breakdown of the correlation between phylogenetic relatedness and ecological similarity for some traits in some groups of organisms (e.g., Bohning-Gaese and Oberrath 1999). However, even in these organisms we would still expect that a measure of "overall" ecological similarity, taking many traits into account, would be correlated with phylogenetic relatedness. Indeed, speciation may even take place without any measurable change in ecological niche (Peterson et al. 1999; McPeek and Brown 2000). Ecological traits less subject to external selection

will also be more likely to be conserved in a lineage. An example of such a trait might be plant secondary chemistry; polyphagous herbivore and pathogen species are more likely to feed on more closely related host species than less related ones (Futuyma and Mitter 1996), indicating conservatism in host secondary chemistry.

Fully addressing the relationship between the phylogenetic relatedness of a community and the ecological similarity of its species will be a large, complicated research program, requiring the explicit measurement of ecological traits for individual species. Without the full study of the autecology of individual species, interpreting phylogenetic structure in terms of ecological similarity must rely on the (justified) expectation of the conservatism of autecology during lineage evolution and, thus, must be made tentatively. However, the quantification of phylogenetic structure described here is a fundamental first step in this process.

Indices of Community Phylogenetic Relatedness

The measure of the phylogenetic relatedness of two species that is most likely to be correlated with their ecological similarity is the age of the speciation event suffered by their most recent common ancestor. While this information is impossible to obtain directly, it might be approximated by the base-pair difference between two species in a neutral gene used for molecular phylogeny. However, this would require every species in a community to be sequenced in order to obtain an estimate of community phylogenetic similarity, which is currently infeasible for large communities. There is also no guarantee that the rate of molecular change in any particular gene is the same in the widely different lineages that may form an ecological community, which necessitates the sequencing of several genes for each species. Given these drawbacks to ideal methods, I propose a simple, but more limited, method that considers just the topology of a phylogeny, estimating relatedness by using the number of nodes between two species on some reference phylogenetic tree. Since only branching topology is needed, composite phylogenies ("supertrees"; Sanderson et al. 1998; Bininda-Emonds et al. 1999) can be used, which can be constructed to include all the species in a community, even if they are not part of a preexisting phylogeny. Such supertrees can be constructed from published phylogenies, attaching species by their orders and families to more well-resolved, higher-level branching structures and leaving lower nodes unresolved where information is not available. The recent publication of hypotheses of angiosperm phylogeny based on molecular data (Chase et al. 1993; Nyffeler 1999; Soltis et al. 1999) and of its order-level classification (APG 1998) now makes possible the construction of a reasonable hy-

pothesis for the phylogeny of any angiosperm plant community.

There are two main limitations of the nodal-distance method. The first is that the measures obtained are relative, not absolute and can only be used to compare communities whose species are a subset of the species in the reference phylogeny. This is because the number of nodes between two taxa is determined by the particular reference phylogeny used. The only way to avoid this relativity is to use the phylogeny of all extant plants as the reference phylogeny; I demonstrate a possible way to approximate this method (see "Sensitivity Analyses").

The second limitation is that the species richness of a clade will influence the level of relatedness of two species in that clade: unlike estimates of age of the most recent common ancestor, nodal distance is not independent of subsequent speciation events. This means that two species drawn randomly from a species-rich clade (e.g., containing 100 species) are likely to appear less related (in nodal distance) than two species from a less species-rich clade (e.g., with 10 species), even if the ages of the most basal speciation event in the two clades are the same. Using nodal distance as a proxy for ecological dissimilarity will also mean that the former two species are predicted to be less ecologically similar than the latter pair. This expectation differs from that arising out of the simplest model of ecological change, in which ecological differences accumulate as a linear function of time ("ecological Brownian motion"): we should expect similar levels of ecological differentiation in both hypothetical clades. However, while this counterintuitive tendency when using nodal distances as relatedness measures is a real problem for comparisons between pairs of species, it decreases in effect as the number of species included in a community increases. When there are more species, more widely distributed over the whole phylogeny, we expect the influence of variation in the species richness of particular terminal clades to decrease, and in general, species separated by more nodes will tend to truly share older common ancestors and be less ecological similar. Later, I will further examine the consequences of this issue for interpretation of results. Despite these drawbacks, this simple topological method is again a first step toward a more comprehensive approach that would include variation in branch length in estimates of relatedness.

In detail, then, we can use the number of nodes (x) separating two taxa in a rooted phylogeny as a simple measure of their (un)relatedness, relative to other taxa in the particular sample pool (of N_{pool} species) in which they occur. The net (un)relatedness of a particular community (of N_c species, a subset of the N_{pool} species) is then represented by the mean number of nodes (on the phylogeny of the N_{pool} species), separating all possible $[N_c(N_c -$

$1)]/2$ pairs of the N_c taxa. To standardize this mean nodal distance ($\bar{x}$) and correct its direction, we can calculate a net relatedness index (NRI) as $1 - (\bar{x}/\bar{x}_{max})$, where $\bar{x}_{max}$ is the maximum value of $\bar{x}$ possible for a given number of taxa and a given phylogeny. The $\bar{x}_{max}$ may be determined by inspection for small communities and phylogenies or by recording the maximum value of $\bar{x}_{max}$ found during many reiterations of randomized drawings of N_c taxa from N_{pool}, for large communities and phylogenies. Figure 1 illustrates the calculation of this index for two communities (A, B, C, D and A, B, E, F) drawn from a pool (A, B, C, D, E, F) for which we have a hypothetical phylogeny. This method is similar to the algorithms used to calculate "phylogenetic uniqueness" that aid in the evaluation of species and sites for conservation plans (Williams et al. 1991; Faith 1996).

This net relatedness index reflects the phylogenetic "clumpedness" of taxa over the whole pool phylogeny. We may also, however, wish to compare communities for the extent to which taxa are "locally clustered" within particular terminal clades, irrespective of the relationship among those clades. This approach is similar to using species-to-genera ratios and is less sensitive to errors in the higher-
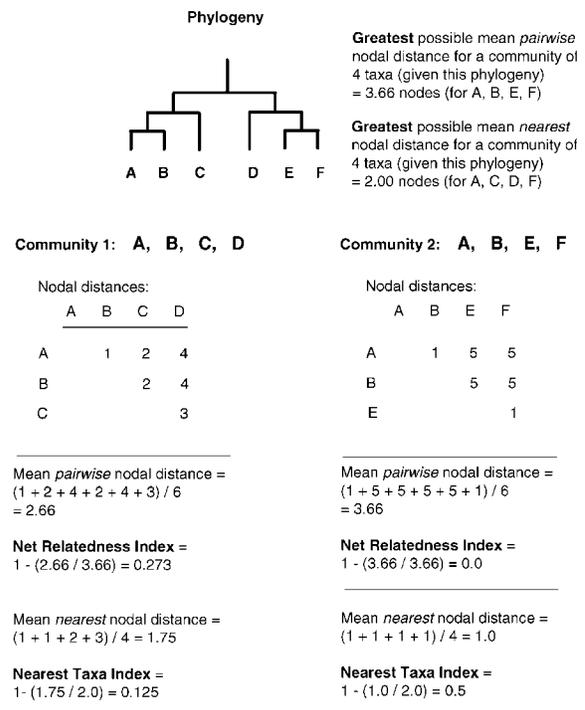
Figure 1: Example of the calculation of the net relatedness index and the phylogenetic nearest taxa index for two communities (*A, B, C, D*, and *A, B, E, F*) drawn from the pool of species included in a hypothetical phylogeny (*A, B, C, D, E, F*).

level phylogenetic structure of the species pool. We can thus construct a second index based on the mean of the nodal distance (x_{min}) to the closest relative in the community for each of the N_c species, again, using the phylogeny of the pool. The phylogenetic nearest taxa index (NTI) is then equal to $1 - [\bar{x}_{min}/(\bar{x}_{min})_{max}]$, where $(\bar{x}_{min})_{max}$ is the greatest value of $\bar{x}_{min}$ possible for a given number of taxa and a given phylogeny, determined by inspection or randomization (as for NRI).

Both indices will have a value of 0 when the community is as spread out as possible on the pool phylogeny, with both $\bar{x}$ and $\bar{x}_{min}$ reaching theoretical maxima. The indices will increase as the "clumping" of community taxa on the pool phylogeny increases and will both reach a value of 1 when the community contains only one species (by definition zero nodes from itself).

An Example: Using Rain Forest Trees

As an example of the use of these indices in the exploration of community structure, I applied them to a data set of rain forest trees at Gunung Palung (in Indonesian Borneo). Rain forest tree communities have long been a focus of attention by ecologists, primarily due to their very high species diversity (e.g., ~190 species of tree, ≥10 cm diameter at breast height [DBH], per hectare at Gunung Palung; Webb 1997). Hypotheses of the maintenance of high local species diversity can be grouped around two main concepts: the role of habitat partitioning, including partitioning of regeneration and architectural niches (Lieberman et al. 1985; Hubbell and Foster 1986; Newbery et al. 1986, 1996; Denslow 1987; Kohyama 1994; Terborgh et al. 1996; Clark et al. 1998; Webb and Peart 2000) and the role of density-dependent herbivores and pathogens (Janzen 1970; Connell 1971; Antonovics and Levin 1980; Clark and Clark 1984; Condit et al. 1996; Webb and Peart 1999), which are both set against a general null hypothesis of community drift (Hubbell 1979, 1997). In addition, competitive exclusion of inferior competitors may be prevented by disturbance and recruitment fluctuations (Chesson and Warner 1981; Chesson and Huntly 1989) and may be retarded by dispersal limitation (Hurtt and Pacala 1995; Hubbell et al. 1999; also see general review by Tilman and Pacala 1993).

The demonstration of nonrandom spatial association of species with habitats is a necessary but not sufficient criterion in demonstrating that habitat partitioning is important in enabling many species to coexist; habitat association may also be caused by dispersal limitation (Webb and Peart 2000). However, to be able to show that co-occurring species in different habitats are more ecologically similar than expected by chance would support the case that species are partitioning habitat according to their aut-

ecology. On the other hand, if species in local communities were less ecologically similar than expected by chance, this would suggest that negative neighborhood interactions were causing increased mortality among ecologically similar species, an effect that would probably be density dependent. Such negative interactions might be due to interspecific competition (e.g., Elton 1946) but might also result from attack by polyphagous herbivores and pathogens. Pathogens appear to be an important factor in the maintenance of species diversity in this Bornean forest (Webb and Peart 1999). Hence, by comparing the observed mean phylogenetic relatedness of local communities with expectations from a null model of random species co-occurrence and interpreting phylogenetic relatedness in terms of ecological relatedness, we may gain valuable insight into the community organization of rain forest trees. Tropical rain forest is a particularly suitable system with which to examine the role of ecological interactions in structuring community composition because the members of the community are sedentary and densely packed so that the potential for biotic interaction to leave a permanent "signature" in community species composition is high.

The data set consisted of trees (≥10 cm DBH) sampled in 28, 0.16-ha plots (40 m × 40 m) in lowland mixed dipterocarp forest at the Gunung Palung National Park, West Kalimantan, Indonesia (Webb 1997; Webb and Peart 1999, 2000). The plots were scattered in a stratified random sampling design over 150 ha, on homogeneous, sandy-clay soils derived from granite and contained 2,862 individuals of 325 species in 50 families. I asked the question, Are the tree species in local (0.16 ha) communities more or less related than expected if such communities were formed from a random sampling of available species in the larger area (150 ha)?

I used 324 species recorded in the total 4.48 ha as the species pool for which to assemble a supertree phylogeny (one of the 325 total taxa was omitted from analyses because it could not be reliably identified to genus). I assembled a composite tree "by hand" (Sanderson et al. 1998), using the phylogeny in the recent ordinal classification of angiosperms (APG 1998) as the backbone (fig. 2). I also used published molecular studies of relationships within orders and families where available: for the Malvales (Baum et al. 1998), Dipterocarpaceae (Kamiya and Harada 1998), and Fabaceae (Doyle et al. 1997). Species within genera and genera within families were left unresolved in the absence of sufficient molecular information. As more phylogenies of angiosperm groups are published, the resolution of a composite tree will increase, as will the power of the methods described here.

The NRI and NTI indices were calculated for the N_c species in each local community (0.16-ha plot), using the
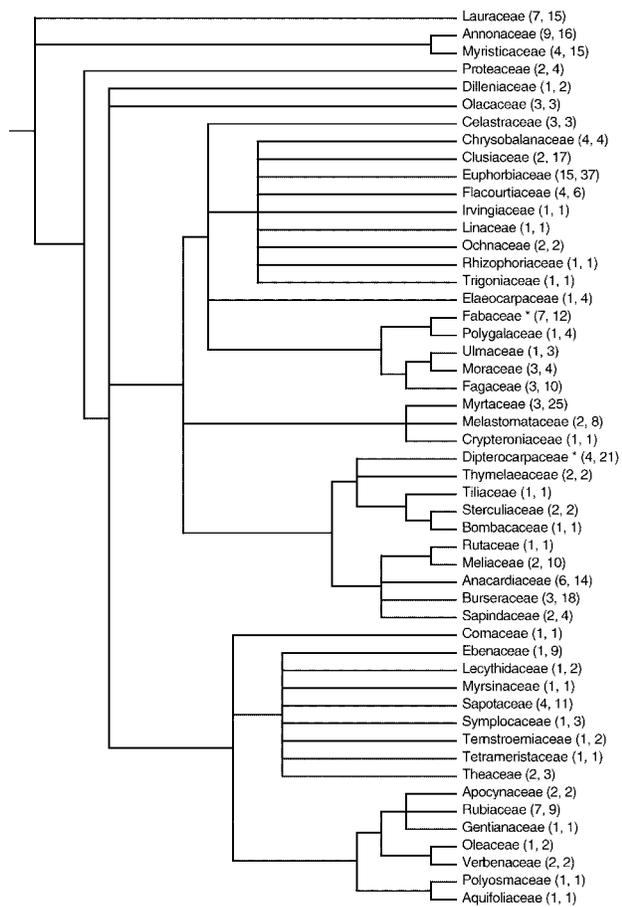
Figure 2: Summary of the composite phylogenetic tree assembled for all the tree taxa in 4.48 ha of rain forest at Gunung Palung. Numbers after family names indicate the number genera and species within that family. An asterisk indicates that published sources (see text) were used to resolve relationships among some genera or species within the family; all other families were modeled as an unresolved polytomy of monophyletic genera, containing unresolved species.

supertree phylogeny of the 324 species (N_{pool}) as the pool. The mean of the index values of the 28 plots was compared to a distribution of similarly calculated mean values from 1,000 runs using random plot species lists created under the following rules: species were shuffled randomly among plots (using only trees in the appropriate size class to form the species pool) while maintaining the same total number of plots in which a species occurred and the same total number of species within each particular plot as in the observed set. The main assumption implicit in this null model is that species have been able to disperse (possibly over many generations) anywhere within the 150 ha. This assumption is likely to be correct in this relatively small area without major barriers to dispersal, especially when the majority of species are animal dispersed (Webb 1997).

An observed value more extreme than 975 of the randomized values was used as an indicator of significance, equivalent to a two-tailed *P* value of 0.05. I implemented all algorithms in the C programming language, compiled with GNU *gcc* under the Linux operating system (Free Software Foundation 1996).

Results

I found that the mean NRI of all plots did not differ significantly from the expectation of the null hypothesis, whether whole plots were used (NRI = 0.091, quantile = 0.836) or when smaller subplots were used (112, 20 × 20-m plots or 448, 10 × 10-m plots, within the 40 × 40-m plots). The NTI (=0.231), however, was significantly greater than expected by the H_0 (quantile = 0.991), indicating that, overall, species in the 0.16-ha plots were more likely to be found with species separated by fewer nodes than expected by chance (e.g., congeners). This pattern in the two indices differed when only the small, understory trees (10–16 cm DBH) were considered (a class containing ∼50% of all individuals): the NRI was significantly more than expected by the null model (NRI = 0.137, quantile = 0.989; fig. 3*A*, 3*B*), while the NTI was not significantly different than expected (NTI = 0.278, quantile = 0.940).

We can therefore reject our H_0 that species are assembled into local communities at random; there is evidence that species occur with closely related species more than we expect by chance. Insofar as our assumption of a correlation between phylogenetic relatedness and ecological similarity is correct, this finding supports the general hypothesis that habitat partitioning is important in the maintenance of rain forest tree species diversity. Higher seedling and sapling mortality in suboptimal habitats in this forest, with those suboptimal habitats being more similar for more closely related taxa, may have created a spatial association among trees of closely related taxa. Species-level physiographic habitat association (with ridges, plateaus, or gullies) has been found in this forest, with seedlings being less associated with habitat than adults (Webb and Peart 2000). When the NTI result is considered in the light of a field botanist's experience, it is perhaps not surprising to find phylogenetic "clumping" in different habitats: one notices that some genera tend to be associated with particular habitats. However, this technique now provides rigorous support for this subjective perception.

Sensitivity Analyses

To increase confidence in the main results presented above, I assessed the effect of changing the assumptions of the basic analysis. I tested two alternate ways to deal with
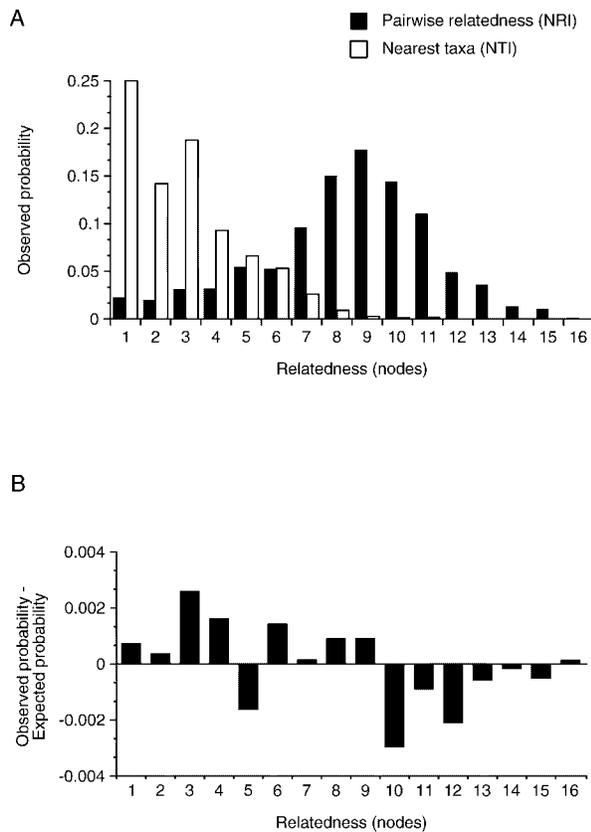
Figure 3: *A*, Probability distribution of mean nodal distances between all possible pairs of species of tree (10–16 cm DBH) in each of 28 (0.16-ha) plots (used to calculate *NRI*; see text) and the most closely related taxon for each taxon in the same plots (used to calculate *NTI*), both averaged over the 28 plots. *B*, Difference between the observed distribution of mean nodal distances between all possible pairs of species (*A*, *NRI* method) and that generated by 1,000 randomized sets of plots with species drawn without replacement from the same pool. The *NRI* of the observed plots was greater than that of 987 of the *NRI* of the randomizations.

polytomies in the supertree, used a different species pool for the supertree, and used an alternate model for the creation of random lists of species per plot.

Species in more well-resolved clades will have more effect on relatedness than species in clades that are large polytomies (an issue closely related to that arising from variation in a clade's true species richness; see the introduction to this article). I used two alternate ways to assess the effect of polytomies in the supertree on the main results. First, and most simply, I randomly resolved all polytomies in the basic supertree of 324 species, using MacClade (Maddison and Maddison 1992) and reran the analyses. Out of five iterations of this procedure, I found, for all trees, NRI was significantly greater than expected

once, while NTI was significantly greater than expected five times, and for small trees (10–16 cm DBH), NRI was significantly greater than expected twice, while NTI was never significantly greater than expected. Hence, we can be quite confident that the lack of resolution in the original supertree does not influence the basic result for the NTI measure but must be less sure that the significance of NRI for small trees is not an artifact of the lack of resolution. A similar robustness to variation in phylogeny was also found by Ackerly and Reich (1999).

This conclusion of robustness was also supported by a second method of assessing the effect of polytomies. When a polytomy of *t* taxa (or internal nodes) is fully, randomly resolved many times, the average nodal distance between any taxon (or node) and the clade root will converge on a value D that depends only on t; D is approximated by $2 \times \ln(t) - 0.85$ (Aldous 1996; M. Steel, personal communication). Hence, by weighting each polytomous node with D, we can estimate what NRI and NTI would be for a fully resolved version of the supertree. I modified the index algorithms to weight polytomous nodes by D and reran the basic analyses. For all trees, NRI (=0.096, quantile = 0.800) was not significantly greater than expected, but NTI was (NTI = 0.204, quantile = 0.981), and for small trees (10–16 cm DBH), neither NRI (=0.129, quantile = 0.933) nor NTI (=0.249, quantile = 0.802) differed from expected.

The choice of the appropriate sample of taxa to use as the basis for phylogenetic comparisons is a perennial problem (e.g., Ackerly and Donoghue 1998). The current choice of the total of all species in the plots as the pool will tend to underestimate the effect of species from poorly sampled clades on measures of relatedness; that is, the inclusion in a plot of a species belonging to a high-level clade (e.g., order) that is represented in the species pool by that species alone will decrease the net relatedness less than the inclusion of species from a clade that has several species in the pool (again, a sampling problem closely related to the effect of variation in clade species richness; see the introduction to this article). While this should not influence the significance level when communities are compared within a species pool, or against a null model, it does contribute to the inability to compare communities from different species pools.

An alternative species pool for this example, and one that would enable community indices to be compared outside this particular rain forest region, is the entire set of extant angiosperms. Obviously, creating a supertree for all angiosperms is impossible. It is, however, possible to modify any particular community supertree, adding additional branches at each node so that within the "phylogenetic locality" of the tree's topology the full angiosperm tree structure is reconstructed. For example, using the recently

published phylogenetic classification of the angiosperms (APG 1998), the Malpighiales can be treated as a polytomy of 31 unresolved families of which nine are included in the forest pool considered in this study. By weighting the nodes of these polytomies by the expected number of nodes in a randomly resolved clade, as in the methodology described above (e.g., 6.02 nodes for the 31 Malpighiales taxa), the calculated nodal distance between any two taxa in a local community will now estimate the true nodal distance between those taxa on the fully resolved phylogeny of all the angiosperms.

I constructed a weighted tree of this kind, using Mabberley (1997) as the source for the total number of species within genera, and genera within families, and APG (1998) as the tree backbone and as the source for the number of unresolved families within an order. For all trees (≥ 10 cm DBH), I found that both the NRI and NTI were significantly greater than expected (NRI = 0.052, quantile = 0.980; NTI = 0.205, quantile = 1.000), and for small trees (10–16 cm DBH), the NRI was almost significantly greater than expected (NRI = 0.071, quantile = 0.972; NTI = 0.245, quantile = 0.958). These results, using an estimate of the full angiosperm phylogeny as the species pool, are similar to those from the basic method, further supporting the conclusion that local tree communities in this forest are not assembled at random.

The third way that I assessed the sensitivity of the results to changes in the methods was by using an alternate randomization model. The choice of the appropriate null model against which observed communities should be compared, in order to infer the presence of structuring ecological forces, has always been contentious (Connor and Simberloff 1979; Strong 1980). The randomization model used in all the above tests creates random communities very similar in structure to the observed communities, by preserving the abundances of the various species: an abundant species that was observed in many of the plots in nature would also occur in the same number of plots in the randomization. The simplest change in the randomization method is to relax this requirement and allow all species to be either abundant or rare (or absent) in the randomized plots, drawing species from the pool with replacement. I modified the randomization algorithms in this way and reran the basic tests (i.e., species not resolved, nodes not weighted, supertree of 324 species). For all trees (≥ 10 cm DBH), I found that only the NTI was significantly greater than expected (NRI = 0.095, quantile = 0.063; NTI = 0.229, quantile = 0.983), and for small trees (10–16 cm DBH), neither index was significantly different than expected (NRI = 0.143, quantile = 0.854; NTI = 0.279, quantile = 0.788).

Discussion

The consistency of results within these various modifications to the basic model increases confidence that the NTI of locally occurring trees is truly greater than expected by chance (although the result of greater than expected NRI in small trees is more equivocal). Hence, despite the limitations of using nodal distance as a measure of phylogenetic relatedness (discussed above), this example demonstrates the potential of this simple method to extract a meaningful phylogenetic signal from community structure data.

When the taxonomic structure of other natural communities has been examined, a similar result has been found: taxonomic diversity has been lower than expected (i.e., relatedness has been higher than expected, measured by higher than expected species-to-genera ratios) and has been attributed to the effect of the similar habitat requirements and/or dispersal abilities in closely related species (Williams 1964; Simberloff 1970).

The differences in the significance of the NRI and NTI may reflect habitat-related associations of taxa at different phylogenetic levels (i.e., taxonomic ranks). The higher than expected NTI in the full set of trees indicates the "clumping" of a number of taxa within clades that are themselves well spread out on the phylogeny of the species pool and, thus, giving a nonsignificant NRI. Among the small trees though, there appears to be less "lower-level" clumping (the NTI is nonsignificant), while higher clades are more clumped than expected (the NRI is significantly greater than expected).

In interpreting the results, we must be aware of the potential effect of variation in clade species richness on node-based relatedness measures and therefore on our inferences about ecological similarity (see the introduction to this article). A subcommunity that has many species from a particularly species-rich terminal clade may appear less related overall than one with species mainly from a less species-rich clade. However, while this may be problematic for interpreting the relatedness of pairs of species, or the net relatedness of single, small communities, the current analysis, comparing average relatedness levels of species-rich communities to a null model, will be less affected. First, the more species there are in a community, the more scattered they will tend to be over the reference phylogeny and the less affected net relatedness will be by variation in clade species richness. Second, because the randomized communities are drawn from the same phylogeny as the observed communities, the influence of species-rich clades will thus be shared, and the relatedness of the observed set relative to the randomized set should not be affected.

For a bias to be introduced in the randomization anal-

ysis, the average proportion of species from a clade represented by many species in observed plots must, for some reason, be quite different from the overall proportion of those species in the species pool. For example, in the species pool used here, the most resolved clade represented by many species is that of the dipterocarp family. The overall observed result of higher than expected relatedness in local plots could have been generated spuriously if fewer than expected dipterocarp species occurred in observed plots. In fact, however, the average proportion of dipterocarp species in each plot was actually slightly higher than the overall proportion of dipterocarps in the species pool: for all trees, dipterocarps comprise 6.46% (21 out of 325 species) of the pool but an average of 10.7% in observed plots, and for 10–16 cm DBH trees, dipterocarps comprise 6.08% (16 out of 263 species) of the pool but an average of 8.4% in observed plots. While addressing only one of many clades, this simple check of the data structure does not support the idea that the overall results are spuriously generated. That these results were also robust to major changes in the resolution of the reference tree (see "Sensitivity Analyses") also indicates that they may be quite robust to variation in the species richness of different clades. Further work is planned to thoroughly assess this issue, using artificial phylogenies constructed with different, controlled, branching patterns and communities assembled under differing, specified rules. Additionally, an important goal of this research program will be to actually include branch length information into the analysis of community structure, thus voiding the problem discussed here.

More generally, these methods have many other applications in community analysis. First, the process of creating supertrees for local communities is not difficult (especially using the TreeBASE database; Sanderson et al. 1993; Piel et al. 1999), and many plant communities have been sampled in a manner similar to that which I used. Hence, there are many data sets, in forest and in other vegetation types (e.g., fynbos, grassland), that can easily be analyzed in a similar fashion, allowing comparisons to be made along diversity gradients and among vegetation types.

Second, the relatedness indices for separate plots within a site can be compared with other measured plot factors, particularly moisture, soil nutrients, and light (acknowledging the potential problems arising from variation in clade species richness; see above). For example, at Gunung Palung, per plot NTI in the basic analysis above was positively correlated with the clay fraction (vs. sand and silt; Webb and Peart 2000) of soil cores (10–15 cm depth) taken at the center of each plot ($r = 0.431$, Spearman's rank correlation $P = .019$), indicating that drought-stressed, sandy plots may have communities that are less closely related than those on less drought-prone soils. Further investigation of this result is intended.

Third, the incorporation of phylogenetic information may bring new power to the study of neighborhood interactions among sessile organisms. For example, neighborhood competition in plants has traditionally been framed in terms of intra- and interspecific interactions. However, classing neighbors in this way hides great ecological difference among species, and the effect of a neighbor on a focal individual may be correlated with the phylogenetic distance between them. Hence, using nodal distance between a focal plant and its neighbor as a weighting factor, in addition to spatial distance and size difference, may help explain much of the residual variation in focal performance (e.g., Pacala and Silander 1985).

Finally, estimates of phylogenetic relatedness might be incorporated into standard community ordination and classification methods, by weighting the interspecies correlation coefficients by the phylogenetic relatedness of the two species involved. This would have the effect of increasing the clustering of sample units in eigenvector space where those sample units contained phylogenetically related species, and reducing the clustering among sample units where those sample units contained phylogenetically unrelated species, thus causing the resulting ordination or clustering to reflect the "phylogenetic affinity" of sample units, as well as their simple species-level similarity.

To conclude, most studies that have incorporated evolutionary relationships into ecology have primarily used phylogenies to control for phylogenetic constraints in analyses of trait evolution (the comparative method; Felsenstein 1985; Harvey and Pagel 1991; Silvertown et al. 1997) or to deduce the historical context for the assembly of communities (Brooks and McLennan 1993; Cadle and Greene 1993; Losos 1996; McPeek and Miller 1996; Grandcolas 1998). One of the most powerful applications of better phylogenies will be to add a new dimension to the study of how contemporary ecological interactions influence the species composition of communities. The methods and findings presented here illustrate the great potential of this approach.

Literature Cited

Ackerly, D. D., and M. Donoghue. 1998. Leaf size, sapling allometry, and Corner's rules: phylogeny and correlated evolution in maples (*Acer*). American Naturalist 152:767–791.

Ackerly, D. D., and P. B. Reich. 1999. Convergence and correlations among leaf size and function in seed plants: a comparative test using independent contrasts. American Journal of Botany 86:1272–1281.

Aldous, D. 1996. Probability distributions on cladograms. Pages 1–18 *in* D. Aldous and R. Pemantle, eds. Random discrete structures. Springer, New York.

Antonovics, J., and D. A. Levin. 1980. The ecological and genetic consequences of density-dependent regulation in plants. Annual Review of Ecology and Systematics 11:411–452.

APG (Angiosperm Phylogeny Group). 1998. An ordinal classification for the families of flowering plants. Annals of the Missouri Botanical Garden 85:531–553.

Baum, D. A., W. S. Alverson, and R. Nyffeler. 1998. A durian by any other name: taxonomy and nomenclature of the core Malvales. Harvard Papers in Botany 3:313–330.

Bininda-Emonds, O. R. P., J. L. Gittleman, and A. Purvis. 1999. Building large trees by combining phylogenetic information: a complete phylogeny of the extant Carnivora (Mammalia). Biological Reviews of the Cambridge Philosophical Society 74:143–175.

Bohning-Gaese, K., and R. Oberrath. 1999. Phylogenetic effects on morphological, life-history, behavioural and ecological traits of birds. Evolutionary Ecology Research 1:347–364.

Brooks, D. R., and D. A. McLennan. 1991. Phylogeny, ecology, and behavior: a research program in comparative biology. University of Chicago Press, Chicago.

———. 1993. Historical ecology: examining phylogenetic components of community evolution. Pages 267–280 *in* R. E. Ricklefs and D. Schluter, eds. Species diversity in ecological communities: historical and geographical perspectives. University of Chicago Press, Chicago.

Cadle, J. E., and H. W. Greene. 1993. Phylogenetic patterns, biogeography, and the ecological structure of Neotropical snake assemblages. Pages 281–293 *in* R. E. Ricklefs and D. Schluter, eds. Species diversity in ecological communities: historical and geographical perspectives. University of Chicago Press, Chicago.

Chase, M. W., D. E. Soltis, R. G. Olmstead, D. Morgan, D. H. Les, B. D. Mishler, M. R. Duvall, et al. 1993. Phylogenetics of seed plants: an analysis of nucleotide sequences from the plasmid gene *rbc*L. Annals of the Missouri Botanical Garden 80:528–580.

Chesson, P. L., and N. Huntly. 1989. Short-term instabilities and long-term community dynamics. Trends in Ecology & Evolution 4:293–298.

Chesson, P. L., and R. R. Warner. 1981. Environmental variability promotes coexistence in lottery competitive systems. American Naturalist 117:923–943.

Clark, D. A., and D. B. Clark. 1984. Spacing dynamics of a tropical rain forest tree: evaluation of the Janzen-Connell model. American Naturalist 124:769–788.

Clark, D. B., D. A. Clark, and J. M. Read. 1998. Edaphic variation and the mesoscale distribution of tree species in a neotropical rain forest. Journal of Ecology 86:101–112.

Condit, R., S. P. Hubbell, and R. B. Foster. 1994. Density dependence in two understory tree species in a neotropical forest. Ecology 75:671–680.

Connell, J. H. 1971. On the role of natural enemies in preventing competitive exclusion in some marine animals and in rain forest trees. Pages 298–312 *in* P. J. Den Boer and G. R. Gradwell, eds. Advanced Study Institute Symposium on Dynamics of Numbers in Populations, Oosterbeek. Pudoc, Wageningen.

Connor, E. F., and D. S. Simberloff. 1979. The assembly of species communities: chance or competition? Ecology 60:1132–1140.

Cotgreave, P., and P. H. Harvey. 1991. Bird community structure. Nature (London) 353:123.

Denslow, J. S. 1987. Tropical rainforest gaps and tree species diversity. Annual Review of Ecology and Systematics 18:431–451.

Diamond, J., and T. J. Case, eds. 1986. Community ecology. Harper & Row, New York.

Doyle, J. J., J. L. Doyle, J. A. Ballenger, E. E. Dickson, T. Kajita, and H. Ohashi. 1997. A phylogeny of the chloroplast gene *rbc*L in the Leguminosae: taxonomic correlations and insights into the evolution of nodulation. American Journal of Botany 84:541–554.

Elton, C. 1946. Competition and the structure of ecological communities. Journal of Animal Ecology 15:54–68.

Faith, D. P. 1996. Conservation priorities and phylogenetic pattern. Conservation Biology 10:1286–1289.

Felsenstein, J. 1985. Phylogenies and the comparative method. American Naturalist 125:1–15.

Free Software Foundation. 1996. GNU *gcc.* http://www.gnu.org.

Futuyma, D. J., and C. Mitter. 1996. Insect-plant interactions: the evolution of component communities. Philosophical Transactions of the Royal Society of London B, Biological Sciences 351:1361–1366.

Gorman, O. T. 1992. Evolutionary ecology vs. historical ecology: assembly, structure, and organization of stream fish communities. Pages 659–690 *in* R. L. Mayden, ed. Systematics, historical ecology and North American freshwater fishes. Stanford University Press, Stanford, Calif.

Grandcolas, P. 1998. Phylogenetic analysis and the study of community structure. Oikos 82:397–400.

Harvey, P. H., and M. D. Pagel. 1991. The comparative method in evolutionary biology. Oxford University Press, Oxford.

Hubbell, S. P. 1979. Tree dispersion, abundance and diversity in a tropical dry forest. Science (Washington, D.C.) 203:1299–1309.

———. 1997. A unified theory of biogeography and relative species abundance and its application to tropical rain forest and coral reefs. Coral Reefs 16(suppl.): S9–S21.

Hubbell, S. P., and R. B. Foster. 1986. Commonness and rarity in a neotropical forest: implications for tropical tree conservation. Pages 205–231 *in* M. Soulé, ed. Conservation biology: the science of scarcity and diversity. Sinauer, Sunderland, Mass.

Hubbell, S. P., R. B. Foster, S. T. O'Brien, K. E. Harms, R. Condit, B. Wechsler, S. J. Wright, and S. Loo de Lao. 1999. Light-gap disturbances, recruitment limitation, and tree diversity in a neotropical forest. Science (Washington, D.C.) 283:554–557.

Hurtt, G. C., and S. W. Pacala. 1995. The consequences of recruitment limitation: reconciling chance, history and competitive differences between plants. Journal of Theoretical Biology 176:1–12.

Janzen, D. H. 1970. Herbivores and the number of tree species in tropical forests. American Naturalist 104: 501–528.

Kamiya, K., and K. Harada. 1998. Molecular phylogeny of dipterocarp species using nucleotide sequences of two non-coding regions in chloroplast DNA. Tropics 7: 195–207.

Kohyama, T. 1994. Size-structure-based models of forest dynamics to interpret population-level and community-level mechanisms. Journal of Plant Research 107: 107–116.

Lieberman, M., D. Lieberman, G. S. Hartshorn, and R. Peralta. 1985. Small-scale altitudinal variation in lowland wet tropical forest vegetation. Journal of Ecology 73:505–516.

Lord, J., M. Westoby, and M. Leishman. 1995. Seed size and phylogeny in six temperate floras: constraints, niche conservatism, and adaptation. American Naturalist 146: 349–364.

Losos, J. B. 1996. Phylogenetic perspectives on community ecology. Ecology 77:1344–1354.

Mabberley, D. J. 1997. The plant-book: a portable dictionary of the vascular plants. 2d ed. Cambridge University Press, Cambridge.

Maddison, W. P., and D. R. Maddison. 1992. MacClade: analysis of phylogeny and character evolution. Version 3. Sinauer, Sunderland, Mass.

McPeek, M. A., and J. M. Brown. 2000. Building a regional species pool: diversification of the *Enallagma* damselflies in eastern North American water. Ecology 81:904–920.

McPeek, M. A., and T. E. Miller. 1996. Evolutionary biology and community ecology. Ecology 77:1319–1320.

Miles, D. B., and A. E. Dunham. 1993. Historical perspectives in ecology and evolutionary biology—the use of phylogenetic comparative analyses. Annual Review of Ecology and Systematics 25:587–619.

Moreau, R. E. 1948. Ecological isolation in a rich tropical avifauna. Journal of Animal Ecology 17:113–126.

Newbery, D. M., J. S. Gartlan, D. B. McKey, and P. G. Waterman. 1986. The influence of drainage and soil phosphorus on the vegetation of Douala-Edea Forest Reserve, Cameroun. Vegetatio 65:149–162.

Newbery, D. M., E. J. F. Campbell, J. Proctor, and M. J. Still. 1996. Primary lowland dipterocarp forest at Danum Valley, Sabah, Malaysia—species composition and patterns in the understorey. Vegetatio 122:193–220.

Nyffeler, R. 1999. A new ordinal classification of the flowering plants. Trends in Ecology & Evolution 14:168–170.

Pacala, S. W., and J. A. Silander. 1985. Neighborhood models of plant population dynamics. I. Single-species models of annuals. American Naturalist 125:385–411.

Peterson, A. T., J. Soberon, and V. Sanchez-Cordero. 1999. Conservatism of ecological niches in evolutionary time. Science (Washington, D.C.) 285:1265–1267.

Piel, W. H., M. J. Donoghue, and M. J. Sanderson. 1999. TreeBASE: a database of phylogenetic information. Abstracts of the XVI International Botanical Congress, St. Louis.

Qian, H., and R. E. Ricklefs. 1999. A comparison of the taxonomic richness of vascular plants in China and the United States. American Naturalist 154:160–181.

Richman, A. D., and T. Price. 1992. Evolution of ecological differences in the Old World leaf warblers. Nature (London) 355:817–821.

Roughgarden, J. 1989. The structure and assembly of communities. Pages 203–226 *in* J. Roughgarden, R. M. May, and S. A. Levin, eds. Perspectives in ecological theory. Princeton University Press, Princeton, N.J.

Sanderson, M. J., B. G. Baldwin, G. Bharathan, C. S. Campbell, D. Ferguson, J. M. Porter, C. Von Dohlen, M. F. Wojciechowski, and M. J. Donoghue. 1993. The growth of phylogenetic information and the need for a phylogenetic database. Systematic Biology 42:562–568.

Sanderson, M. J., A. Purvis, and C. Henze. 1998. Phylo-

genetic supertrees: assembling the trees of life. Trends in Ecology & Evolution 13:105–109.

Schluter, D. 1994. Experimental evidence that competition promotes divergence in adaptive radiation. Science (Washington, D.C.) 266:798–801.

Sibley, C. G., and J. E. Ahlquist. 1990. Phylogeny and classification of birds: a study in molecular evolution. Yale University Press, New Haven, Conn.

Silvertown, J., M. Franco, and J. L. Harper, eds. 1997. Plant life histories: ecology, phylogeny and evolution. Cambridge University Press, New York.

Simberloff, D. S. 1970. Taxonomic diversity of island biotas. Evolution 24:23–47.

Soltis, P. S., D. E. Soltis, and M. W. Chase. 1999. Angiosperm phylogeny inferred from multiple genes as a tool for comparative biology. Nature (London) 402:402–404.

Strong, D. R. 1980. Null hypotheses in ecology. Synthese 43:271–285.

Terborgh, J., R. B. Foster, and P. Nunez. 1996. Tropical tree communities: a test of the nonequilibrium hypothesis. Ecology 77:561–567.

Tilman, D., and S. W. Pacala. 1993. The maintenance of species richness in plant communities. Pages 13–25 *in* R. E. Ricklefs, and D. Schluter, eds. Species diversity in ecological communities: historical and geographical perspectives. University of Chicago Press, Chicago.

Webb, C. O. 1997. Seedling ecology and tree diversity in a Bornean rain forest. Ph.D. diss. Dartmouth College, Hanover, N.H.

Webb, C. O., and D. R. Peart. 1999. Seedling density dependence promotes coexistence of Bornean rain forest trees. Ecology 80:2006–2017.

———. 2000. Habitat associations of trees and seedlings in a Bornean rain forest. Journal of Ecology 88:464–478.

Weiher, E., and P. Keddy. 1999. Ecological assembly rules: perspectives, advances, retreats. Cambridge University Press, Cambridge.

Williams, C. B. 1964. Patterns in the balance of nature. Academic Press, New York.

Williams, P. H., C. J. Humphries, and R. I. Vane-Wright. 1991. Measuring biodiversity: taxonomic relatedness for conservation priorities. Australian Systematic Botany 4: 665–679.