Chapter 6

ECOLOGICAL ORGANIZATION, BIOGEOGRAPHY, AND THE PHYLOGENETIC STRUCTURE OF TROPICAL FOREST TREE COMMUNITIES

Campbell O. Webb, Charles H. Cannon, and Stuart J. Davies

OVERVIEW

The assembly of local tropical forest tree communities is influenced by abiotic filters from a larger regional species pool (e.g., habitat differentiation, mass effects, dispersal limitation) and local biotic interactions (e.g., density dependence, resource competition; summarized in Figure 6.1). These assembly processes are mediated by the phenotypic similarities or differences of individuals, which are the outcome of evolutionary change in historical communities, and ultimately the composition of taxa in a regional species pool is the outcome of biogeographic processes. Given the great diversity of tropical tree species, we are unlikely ever to know enough about the ecologically important phenotypes or precise spatial ranges of species to be able to predict local community species composition based on detailed attributes of every species. However, we suggest in this chapter that because species similarity and difference are strongly influenced by common ancestry, as is the presence or absence of a taxonomic clade in a geographic area, a phylogenetic approach may be most effective for understanding and predicting local community composition.

In this chapter, we briefly review current understanding of abiotic and biotic controls of local species composition, and of evolutionary patterns in ecological characters. We then describe phylogenetic analyses that explore the outcomes of neighborhood interactions, habitat filtering, climatic gradients, and biogeographic history by analyzing the phylogenetic patterns of species composition at nested spatial scales (Figure 6.1). We test these methods with data from forests in Southeast Asia. Finally, we discuss the association of ecological and biogeographic characteristics with internal nodes of plant phylogenies and the creation of predictive models for the general taxonomic and ecological composition of communities.
and Leighton 2004), elevation (van Steenis 1972, Lieberman et al. 1985, Ashton 2003a), rainfall (Gentry 1982, Schnitzer 2005), understory light (Swaine and Whitmore 1988, Clark and Clark 1992, Davies et al. 1998), and architectural position (Kohyama 1993). However, many species at a local site also appear to share the same realized niche (Potts et al. 2004, Valencia et al. 2004). While species may sort into their appropriate habitats at a local scale, it is unlikely that they occupy all sites on a landscape where they might grow, because of the continual perturbation of climate oscillations and temporal variation in biotic interactions. Over short time scales (100–1000 years), the geographic distributions of some taxa will be expanding, and those of others will be shrinking (Bennett 1997). Over long time scales (10,000–1,000,000 years), biogeographic connections (e.g., land-bridges) and barriers (e.g., mountain ranges) change, and at even longer time scales (10–100 My [million years]), land areas and geologies will be appearing and disappearing, again changing the potential geographic distribution of taxa. Hence, the geographic distribution of most taxa will not be in equilibrium with the contemporary abiotic environment, but will represent a dynamic balance of large-scale climatic oscillations and gradients, location of species origin, rate of dispersal, and availability of dispersal routes. This disequilibrium is vital to keep in mind when fitting environmental niche envelopes (on axes of rainfall, elevation, temperature, etc.) using geographical information systems (GIS)-based interpolation (e.g., Austin 2002, Graham et al. 2004).

While species differ in their local realized distribution, it is less clear to what extent local biotic interactions, such as ubiquitous competition for light and physical space (Hubbell 2001, Kitajima and Poorter Chapter 10, this volume) or for pollinators and dispersers, modify growth and survival under these different abiotic conditions. For example, is the absence of “poor-soil” species on rich-soil sites due to some fundamental cost associated with ecological specialization that restricts their fundamental niche, or to their local exclusion from rich sites (included in their fundamental niche) by faster-growing but less stress-tolerant “rich-soil” species (Fine et al. 2004)? Some seedling growth experiments (e.g., Hall et al. 2003, 2004, Palmiotto et al. 2004) suggest that optimal performance in the absence of competition is achieved in the soils on which a species is most abundant, implying that competition may not greatly shift the position of the peak of the realized niche away from that of the fundamental niche. However, the ability of many species to prosper under conditions in which they are not normally found, when potential competition is reduced (e.g., in botanical gardens), suggests that generalized competition may also play a large role in compressing the boundaries of species’ fundamental niches.

Even if competition for space and/or light is experienced by all forest plants, does the negative effect of neighboring plants vary with the neighbor’s identity, that is, whether a neighbor is a conspecific, a phylogenetically closely related species or a distantly related one? In a temperate forest, Canham et al. (2004, 2006) detected different effects on focal species of different neighbor species. These effects might also be mediated by competition for “mobile links,” the pollinators and dispersers plants depend upon (Vamosi et al. 2006), or experienced as indirect competition resulting from pathogen or herbivore population dynamics. There is abundant evidence that plants do respond more negatively to increasing density of conspecifics than heterospecifics (“negative density dependence”; reviewed by Wright 2002). However, ecological exclusion and eventual character displacement in sibling species, and ecological speciation (Schluter 2001), depend upon the most closely related taxa experiencing the strongest negative interactions. The limited data for tropical trees support this relationship at some spatial scales (Uriarte et al. 2004, Webb et al. 2006). However, the ultimate outcome of this process, resulting in “checkerboard” patterns where certain combinations of species in the same habitat are never found due to strong competitive effects among species (e.g., Graves and Gotelli 1993), has not yet been reported. The only demonstrated example of over-dispersion of a character in tropical forest trees, of the kind generally thought to indicate biotic, competitive structuring of a community (Bowers and Brown 1982, Wilson 1999), is the segregation of
flowering and fruiting times (Miconia, Snow 1966; Shorea section Mutica, Ashton et al. 1988).

On the other hand, the frequently reported association of species of particular taxonomic groups with different habitats (e.g., Gentry 1988, Webb 2000) suggests that while generalized competition for shared resources in forests is ubiquitous, and biotic interactions may negatively influence the performance of similar or phylogenetically close neighboring plants, species distributions at habitat-wide scales generally result from the “attractive” effect of abiotic conditions.

We note that positive advantages of having related taxa nearby (e.g., higher pollinator and seed disperser availability, sharable ectomycorrhizae) may reinforce similarity in the distributions of related taxa (see Momose et al. 1998).

**EVOLUTION OF ECOLOGICAL CHARACTERS IN TROPICAL FOREST TREES**

If ecological processes affecting today’s tropical forest trees are the same as in the past (i.e., ecological uniformitarianism on a time scale of millions of years), we expect the evolution of ecological characters to have been shaped by the ecological conditions discussed above, that is, competition among large numbers of similar taxa filtered into particular forest habitats (ridge-tops, gullies, swamps, gaps). The diversity of such communities (itself probably ancient, e.g., 64 My; Johnson and Ellis 2002) means that neighborhood encounters between any specified pair of species will be relatively infrequent, and divergent selection among phenotypically similar taxa may be very weak (Connell 1980, Stevens 1980, Ashton 1988). Additionally, the long generation time of many tropical trees (100 + years), in combination with climate changes over 1000–10,000 years, will further weaken such divergent selection. This is a very different community scenario to the species-poor, small island systems which have provided much of our empirical knowledge of ecological divergence (Darwin 1859). The animal and plant species on islands may frequently experience sustained pairwise competition (causing trait “push”; Silvertown 2004) and the frequent opportunity to fill empty niches (trait “pull”), leading to rapid selective divergence in ecological characters (Givnish 1998, Schluter 2000, Silvertown 2004). Species of tropical forest trees clearly differ in autecology in many ways, but these differences are likely to be the result of drift within a species or subset of metapopulations, or weak selection on peripheral founder populations in slightly different conditions. If the accumulation of ecological changes is essentially random and relatively slow compared with speciation, then closely related taxa will generally share the same character (“symplesiomorphy”), showing an overall conservative pattern of character evolution (Harvey and Pagel 1991, Ackerly and Donoghue 1998, Webb et al. 2002).

What evidence is there that a random accumulation of character changes best represents ecological evolution in tropical trees? The most rigorous assessments of the pattern of ecological character evolution use standard methods of phylogenetic ancestral state reconstruction (or divergent tendency; see Moles et al. 2005) for niche-related characters (see Linder and Hardy 2005 for a good example). There are now a handful of phylogenetically based studies of ecological character evolution in tropical and subtropical forest plant genera (Davies 1996, Givnish et al. 2000, Dubuisson et al. 2003, Bramley et al. 2004, Cavender-Bares et al. 2004, Fine et al. 2004, Fine et al. 2005, Plana et al. 2004), but none that we know of that has attempted to sample all the extant species in a lineage. There is no consensus view arising from these studies. Some studies report that closely related species are ecologically similar, others that related species differ in a way that may indicate adaptive divergence, but since the taxon sampling is usually sparse, and the characters examined so different, it is impossible yet to generalize. No study we are aware of has rigorously tested for significant conservatism in ecological characters in a densely sampled group of closely related tree species, as has been done for animals (e.g., Losos et al. 2003, Stephens and Wiens 2004).

Because standardized data on niche parameters for tropical trees are hard to collect, an alternative approach is to estimate niche parameters by modeling the distribution of species in
multidimensional environment space, which has been successful for other taxa (e.g., Peterson and Holt 2003, Raxworthy et al. 2003). Tools such as GARP (Scachetti-Pereira 2002), BIOCLIM (http://biogeo.berkeley.edu/worldclim) and WhyWhere (http://biodi.sdsc.edu) use the values of various spatially modeled factors (rainfall, temperature, elevation, etc.) at geographical points of known species occurrence to create a niche envelope in multidimensional factor space. This envelope can be re-projected onto the GIS landscape to predict the potential range of the species. Niche parameters derived in this way can be reconstructed on a phylogeny, as with other characters (Graham et al. 2004).

Supplemental information on ecological evolution comes from comparing the ecological character of congeners that systematists consider to be sister species (although usually no molecular evidence exists to confirm this impression). These supposed close relatives are often ecologically similar (Forman 1966), but have non-overlapping (allopatric) distributions (e.g., Stevens 1980 table 4, Prance and White 1988). This suggests that sibling taxa cannot co-occur until they have diverged sufficiently to overcome local competitive exclusion, and that rapid ecological divergence is common. A similar interpretation might apply to studies that find substantial variation in ecological character among locally co-occurring congeners (Grubb and Metcalfe 1996, Osunkoya 1996, Thomas 1996, Davies 1998, Smith-Ramirez et al. 1998), and cases of congeners occupying markedly different habitats within a site (Valencia et al. 2004).

But local niche variation more than expected from models of slow character evolution? We suggest not. Because much speciation in tropical trees appears to be allopatric (e.g., Stevens 1980, Ashton 1988, Gentry 1989, although again this has seldom been confirmed with phylogenetic or population genetic data), the spatial segregation of ecologically similar, new sibling species is expected to be observed frequently. Over time niche parameters may then drift randomly and, independently, species ranges will shift such that sibling species become sympatric (Barraclough and Vogler 2000). Despite the great noise added to a system due to climate oscillations and temporal variation in biotic interactions, we expect in a system dominated by allopatric speciation that degree of range overlap will be positively associated with time since divergence (but see critique by Losos and Glor 2003). If variance in trait characters is also correlated with time since divergence, then we expect that the greater the range overlap of sibling taxa the more likely they are to differ in autecology. An appropriate null model is required to test whether niche segregation in congeners, implying either contemporary competitive exclusion or historical selection for niche divergence, is greater than expected by chance.

We must also note that in addition to cases of striking habitat variation there are also many cases of congeners sharing habitats (Webb and Peart 2000), and of general association of large taxonomic groups (e.g., families) with particular habitats (Ashton 1988, Gentry 1988, Davis et al. 2005). The existence of ecological conservatism (Wiens 2004) in plants has also been demonstrated by the ability to predict a species’ distribution based on observations of related species on a different continent (Huntley et al. 1989, Ricklefs and Latham 1992). Morphological characters with clear ecological significance (e.g., pollen ultrastructure, pollination syndrome, seed dispersal mode, and tree architecture) are also usually strongly conserved. Overall, phylogenetically based information on the evolution of autecology is scanty, and we have no basis to reject the “null” hypothesis that ecological character changes have generally accumulated at random in tropical forest tree species, and thus display a pattern of significant phylogenetic conservatism in most species (accepting the existence of great variation in evolutionary history).

A final question, fundamental to interpreting patterns of species association at geographic scales is: How consistent in space are species ecological characters and species boundaries? Some species are morphologically and ecologically constant, and studies of their breeding systems show them to be strongly resistant to hybridization (e.g., cerrado species studied by Barros 1989). Other species may not be ecologically constant across their range, and attempts to detect community assembly rules (Gotelli and McCabe 2002) over large areas may fail because species behave
Phylogenetic Community Structure and Biogeography 83
differently in different places. Where species exhibit large genotypic variation, subspecies have often been recognized (van Steenis 1948), usually in disjunct populations – clinal (continuous) variation in morphology is extremely rare in tropical tree species. Variation is not problematic in a phylogenetic context if different morphs are monophyletic (i.e., all share the same most recent common ancestor). However, we are becoming increasingly aware of complex gene-flow reticulations in some tropical tree clades within and even among species (Cannon and Manos 2003; see also Mallet 2005). More than in perhaps any other functional group of organisms, individuals of tropical forest trees are long lived and rare, which has profound consequences for the evolutionary coherence of these species. The relatively few generations between major climatic events, and the danger of erosion of necessary genetic variability when rare, may select for individual-level characters that maintain openness to gene flow (van Valen 1976, Grant and Grant 1996). The existence of “swarms” of related species or subspecies with near-complete combinations of morphological traits (“ochlospecies”; White 1962, Cronk 1999) may indicate that the terminal phylogenetic structure of some tropical tree lineages may actually be highly reticulate (Veron 1995, Funk and Omland 2003). Analyzing variation in single genes with methods that always produce a bifurcating phylogeny will seriously mislead us about the true history of evolution in these groups. If genetic information controlling the ecological niche of a taxon can be easily exchanged, then we should really treat either the inclusive clade as the effective ecological entity, or just the single individual. Incorporating a more dynamic species concept into community ecology will be a major enterprise over the coming years.

THE PHYLOGENETIC STRUCTURE OF SPECIES ASSEMBLAGES

While clade-based studies of ecological character evolution in tropical forest trees are few, the number of studies of community composition for tropical forests is far greater. A phylogenetic approach to examining assemblages can reveal patterns of non-random species composition which are otherwise hidden, patterns resulting from contemporary ecological processes (Simberloff 1970, Enquist et al. 2002), biogeographic history, and the evolutionary history of ecological characters. The important components for analyzing the composition of assemblages are: (1) lists of species in a sample at some explicit scale, and an estimate of the pool of species from which sampled taxa are drawn; (2) a phylogenetic hypothesis for the species in the pool; and (3) means of quantifying phylogenetic structure. In this section, we provide an overview of each of these three components separately, and then bring them together to discover and interpret the phylogenetic structure of tree communities at three different spatial scales.

Species composition, pools, and samples. Defining meaningful spatial scales in ecology has always been a problem, partly because the scale of sampling must relate to the scale of phenomena that we want to measure (but commonly do not know), and partly because the scales of ecological and evolutionary processes can merge continuously (Chave Chapter 2, this volume). That said, we can still define and refer to explicit scales, and we can attempt to pick natural breaks in the continuum that correspond to the scale of biogeographic, climatic, lithologic, and topographic transitions. We will refer to six potential levels: global (the total extent of tropical forests), continental (1000–10,000 km, in which climatic and biogeographic gradients may occur, e.g., Borneo), regional (10–1000 km, fairly homogeneous in climate and biogeographic history, e.g., northwest Borneo), community or “local” (1–10 km, a scale of mixing of tree seeds in one or a few tree generations), habitat (10–1000 m, a lithologically or topographically defined patch, e.g., a sandstone ridge-top), and neighborhood (0–10 m, the scale of direct inter-plant interactions). The meaningful definition of these scales will vary by taxon and geography, but the key characteristics of nested scale from an evolutionary ecology point of view are (1) a pool of species influenced by biogeographic history (including climatic barriers), and (2) a sample influenced by contemporary ecological interactions (Figure 6.1).
Figure 6.1  Illustrative example of the sampling of lineages through different scales, depending on ecological characters and biogeography. Negative (or positive) interactions may occur among individual neighbors (1a). Habitat filtering for traits that permit survival in abiotic habitat H may occur among taxa in a community pool, modified by competitive interactions (1b). The community pool is in turn an environment- and dispersal-dependent sampling of a regional pool (2), which is structured by the geographic history of lineage diversification (3). The particular physical size of each sample/pool pair, and the number of levels of nestedness, depend on biological and physical circumstances (see text). Note the intentional similarity to the “life-cycle filtering” model in Harper (1977).

*Phylogenies.* The great expansion in numbers of plant species sequenced and included in phylogenetic analyses offers an increasingly resolved picture of the relationships among angiosperms (Stevens 2001, Angiosperm Phylogeny Group 2003); for example, as of mid-2005, some 6500 taxa in GenBank had been sequenced for *rbcL* (A. Driskell personal communication). Although relatively few tropical forest tree species are included in these analyses, enough exist to draw a reasonable picture of relationships at the “generic” level, and it has become increasingly possible for ecologists themselves to undertake the molecular work required to produce a phylogeny. There are even online tools that permit the rapid retrieval of a coarse phylogenetic hypothesis for any list of angiosperm species (Webb and Donoghue 2005).

A number of serious caveats are necessary concerning phylogenies for tropical trees. First, these phylogenies are hypotheses, and some areas of the angiosperm tree may still undergo serious re-organization, especially with the increasing evidence that “deep” hybridization events may have been frequent (Davis and Wurdack 2004,
Soltis et al. 2004, Mallet 2005). Second, the lack of resolution and comprehensive sampling among congenerics restricts our understanding of the vital recent evolution of species’ autecology (Malcomber 2002); data on the rates of divergence in ecological character among sibling species are perhaps the key missing information in our models of tropical tree speciation and ecological evolution (see character evolution section above). Third, even if resolution among species existed, we may need to use individuals as the terminal units for ecological analysis, given the great intraspecific variation that has been noted for some tropical forest trees (Cannon and Manos 2003, Dick et al. 2003). Despite these issues, when forest communities comprise many different genera, we do have sufficient power to detect non-random patterns in phylogenetic structure.

**Metrics of phylogenetic structure.** The distribution of a subset of taxa on a phylogeny can be summarized with single indices (Phylogenetic Diversity, Faith 1992; Net Relatedness Index [NRI], Nearest Taxon Index [NTI], Webb 2000; Webb et al. 2002, Cavender-Bares et al. 2004). The primary characteristic of the distribution captured by these indices is how phylogenetically concentrated or clustered the subset is. In this, they are functionally equivalent to metrics of trait conservatism (Consistency Index; QVI, Ackerly and Donoghue 1998): community structure can also be analyzed by treating the presence of a taxon in a sample as a binary trait (Chazdon et al. 2003). More detailed aspects of phylogenetic structure can be assessed by observing the ratio of whole-tree clustering (NRI) to “tip-clustering” (NTI), which indicates whether samples occur in a single cluster on a phylogeny (high NRI/NTI) or in several (low NRI/NTI). By comparing the number of sampled taxa subtending each node in a phylogeny with the number expected at each node under an appropriate null model of random phylogenetic structure, the over- and under-representation of sampled taxa in each clade can be determined (the “Nodesig” algorithm in Phylocom, Webb et al. 2004; Figure 6.2). This “node-loading” result permits overall measures of phylogenetic distribution to be interpreted in terms of bias in individual clades. Finally, the phylogenetic similarity of multiple samples can be compared by using mean phylogenetic distances between all pairs of taxa in each of every pair of samples to build a pairwise phylogenetic distance matrix for samples (the “Comdist” algorithm in Phylocom, Webb et al. 2004). This distance matrix can be visualized using standard clustering or ordination techniques, and rather than reflecting the shared presence or absence of taxa, it represents the “shared evolutionary heritage” among samples, and can show similarity in deep phylogenetic structure even when no taxa are shared (see below and Figure 6.3).

Bringing the above components together, we can think of the members of any sample as being distributed on the phylogeny of the larger pool of species. At each scale transition, different ecological and biogeographic processes determine which species will be “filtered” into the sample (Figure 6.1), influencing the phylogenetic distribution of the sample members, or their “phylogenetic structure.” The ecological interpretation of phylogenetic structure is therefore wholly dependent on the particular scales involved. The following analyses move up in increasing geographic scale, from a single habitat with different topographic features, to a single watershed containing a number of habitats, and finally to the continental scale between landmasses.

**Local processes: plant–plant interactions and habitat filtering from the community pool**

The evolutionary distribution of ecological characters, either conserved or convergent, interacts with the ecological organizing processes in communities, which either draw phenotypically similar taxa together in habitats (“phenotypic attraction”) or force them apart via local exclusion of ecologically similar individuals (“phenotypic repulsion”; Webb 2000, Webb et al. 2002, Cavender-Bares et al. 2004, Cavender-Bares et al. 2006). The balance between the abiotic filter (or “funnel”) and biotic “spreader” will determine the gross phylogenetic structure of the assemblages on habitats, that is, whether closely related taxa co-occur, or whether the taxa in a sample are
Figure 6.2 Results of significance tests on the association of clades of the angiosperms with samples of rainforest at different spatial scales. Black symbols indicate significantly more taxa in the subtended clade are in the “sample” than expected by chance on the larger-scale “pool,” while white symbols indicate fewer. “Ridge in granite” refers to the distribution of species sampled in ridge-top 0.16 ha plots on the pool of all species in the lowland granite zone at Gunung Palung (Webb and Peart 2000). “Granite in G. Palung” refers to the distribution of species sampled in the granite zone (all habitats) within the list of all species at Gunung Palung (Cannon and Leighton 2004). Both of these represent “habitat within community” sampling processes (see text). “Lambir in SE Asia” refers to the distribution of species sampled in the Lambir 50 ha CTFS plot within the full species list of Lambir plus Pasoh plus Huai Kha Khaeng, and represents “region within continent” sampling.
Figure 6.3 Comparison of an ordination (using multidimensional scaling, MDS, axes 1 and 2 shown for both (a) and (b)) of 28 plots in lowland granite forest at Gunung Palung based on (a) Euclidean presence/absence distance versus (b) mean phylogenetic distance between plots (nearest taxon method). Plot edaphic classes: r, ridge; g, gully; p, plateau. Note the greater separation of habitats in (b).

less related than expected by chance. Again, the scale of samples (patches) will determine both the processes acting and the assemblage outcome. For example, it is possible that competitive, antagonistic forces among similar taxa may act on very small, neighborhood scales, for example among seedlings, and could lead to local exclusion in sub-habitat sized patches ("1a" in Figure 6.1), but that at the integrated scale of a habitat patch (ridge-top, river-bank, etc.) abiotic niche filtering appears to draw similar taxa together (Webb et al. 2006). The processes that determine habitat-scale assemblages occur on a spatial scale at which sample locations (habitats) are linked via seed dispersal, and at which any taxon could possibly occur in any sample within a few reproductive cycles. The appropriate pool of taxa is thus the sum of all taxa in all habitats. In tropical rainforest, where seed dispersal by vertebrates is common, this community scale is ca. 1–10 km.

When phylogenetic structure of habitat-sized samples has been examined in complete forest communities, a variety of results has been found. Webb (2000) found the tree taxa in 0.16 ha plots in Borneo to be more closely related than expected by chance, across the whole angiosperm phylogeny, which can result only from overall conserved traits and phenotypic "attraction" (Webb et al. 2002). H. Steers (personal communication) found a positive correlation between taxa co-occurrence in small plots in Mexican dry forest and their phylogenetic relatedness. Cavender-Bares et al. (2006) found phylogenetic clustering in plots in subtropical Florida woodland when all woody taxa were included in the analysis, but she found samples to contain taxa evenly distributed on a phylogeny when only the oaks were considered (Cavender-Bares et al. 2004). Kembel and Hubbell (2006) found that taxa in the 50 ha plot at Barro Colorado Island on some habitats (plateaus and secondary forest) were phylogenetically clustered, while in swamps the taxa were evenly distributed. Only Cavender-Bares et al. (2004) have simultaneously assessed phylogenetic structure and the distribution of ecological characters.

If we consider the dominant process structuring local assemblages to be abiotic niche filtering (above), then we can interpret these mixed patterns as implying that most niche traits are conserved phylogenetically, though a few are significantly convergent. This mixed pattern may correspond to "deep" and old fixedness of β-niches (habitats) and more lability in α-niches (intra-habitat niches; Pickett and Bazzaz 1978, Ackerly et al. 2006, Silvertown et al. 2006; see also Streelman and Dunley 2003). Better phylogenetic
resolution and trait data are needed: our lack of resolution at the tips of community phylogenies may be hiding more character convergence than currently detected, or may be preventing the detection of the local exclusion of the most closely related taxa on single habitat patches, taxa that share a phylogenetically conserved ecological niche.

Within the limits of our phylogenetic resolution, we can observe the overall distribution of clades associated with different habitats using the node-loading method described above. This can be thought of as an analysis of the association of habitat with species, genera, and families simultaneously. As an example, we see that myristicaceae/, phyllanthaceae/, malvaceae/, and sapindaceae/ clades are significantly associated with ridge-top plots in lowland hill forest on granite at Gunung Palung (we use “/” as a mark to indicate the name of a rank-free clade), while species in fagaceae/ occur less often than expected on this habitat (Figure 6.2). An ordination of mean phylogenetic distances among species in different plots, as opposed to their Euclidean presence/absence distance, indicates that taxa in plateau and gully habitats differ at a deeper phylogenetic level than either do with taxa in the ridge habitats (Figure 6.3). At a larger spatial and elevational scale, but still within the general model of “habitat-within-community,” we note that myristicaceae/, fagales/, shorea/, and sapotaceae/ are over-represented in forest on granite, relative to an area incorporating five habitat types (data from Cannon and Leighton 2004; Figure 6.2).

Figures 6.2 and 6.3 offer a phylogenetic characterization of the observation that the family composition of forests on different habitats appears to be quite predictable (Ashton 1988, Gentry 1988). For example, one of the most striking patterns in taxonomic turnover is the repeated change in family dominance with increasing elevation and/or decreasing fertility, from Euphorbiaceae, Meliaceae through Lauraceae (and Fagaceae in the Old World) to Ericaceae (Gentry 1982, 1988, Lieberman et al. 1985, Ashton 1988). This predictable floristic turnover occurs because many of the species in each of these groups share ecological characters. This may be due to repeated homoplasy, but we suggest it is more likely a consequence of one or more clades within the group being symplesiomorphic for a particular niche. We must of course be very careful to recognize the sampling bias inherent in making evolutionary conclusions based on taxa in an ecological sample. If most of the members of a clade present at a particular site (e.g., the euphorbiaceae/ at Gunung Palung) occur on rich, lowland soils, we cannot make deductions about the ecology of ancestral euphorbs, because the family (even in its modern definition) contains thousands of species, most of which are not rainforest inhabitants. It is not that the approach of inferring ancestral states from contemporary characters is flawed, but that our sampling must be a random subset of the clade (Ackerly 2000). The bias in a sample depends on the global biogeographic extent of the clade. Fortunately, we are in a better position to assess ancestral ecologies using tropical forest species than we would be in, say, a temperate grassland, because the ancestors of many of the lineages present in the former, unlike the latter, probably originated in environments and biotic habitats not too dissimilar from present day conditions (e.g., Davis et al. 2005, but see Schrire et al. 2005 for a seasonally dry, rainforest margin origin for Fabaceae).

Despite these caveats, studies that attempt to incorporate ecological data for the full extent of a clade can begin to make deductions about its ecological nature (e.g., Davis et al. 2005), an approach which we feel will be increasingly powerful when global ecological databases (e.g., the Center for Tropical Forest Science [CTFS] network; Ashton et al. 2004) are joined with the various emerging Tree of Life projects.

Assembly of regional pools from the continental pool

Regional species pools (10–1000 km) reflect a combination of the remnants of intra-continental speciation patterns and similar species responses of distribution to recent climatic changes. This is the “phylogeographic scale,” at which population and species phylogenies and networks can be reconstructed using genetic markers (Cannon and
Manos 2003, Dick et al. 2003), and spatial history of range expansion and contraction and population movement can be inferred (Templeton et al. 1995). The analysis of multi-clade assemblages at these scales has taken two main approaches. The first approach is non-phylogenetic similarity analysis, which interprets variation among local areas as the result of permanent range restrictions in some species, and climatic responses in other species that move more freely on a landscape (e.g., van Balgooy 1987). Terborgh and Andresen (1998) used family abundances as units in an analysis of patterns of Amazonian tree distribution. They were unable to use species or genera because there was so little overlap in taxonomic composition using these ranks. They suggested that the strong differentiation they found among regions in the Amazon basin was a result of both historical biogeographic factors (the historical isolation of the Guianan shield) and the interaction between species autecology and contemporary conditions, where families tend to contain many species with similar autecology.

The second approach is explicitly phylogenetic, requiring phylogenies for all members of the assemblages included, and interprets conserved congruence among the “area cladograms” for different lineages to reflect the history of land splitting (vicariance), and convergence, or homoplasy, to reflect historical dispersal events (and/or extinction and sympatric speciation; e.g., Brooks Parsimony Analysis [BPA], Brooks and McLennan 2001; Turner et al. 2001, van Welzen et al. 2003). However, because of its assumption that vicariance is the primary driver of clade-area association, area-cladogram congruence is not suitable for spatial scales where the signal from migration (or dispersal) resulting from climate cycles may be stronger than residual vicariance events, a situation likely to happen within continents. “Event-based” biogeographic methods, however, can be parametrized to allow for frequent dispersal (“trees-within-trees”; Dispersal Vicariance Analysis [DIVA], Ronquist 1997; Page 2002). The related questions of how much mixing has occurred, how fast species can move on a landscape, and the balance between residual allopatric speciation signal and recent (<10,000 years) climate tracking are fundamental to understanding tropical forests at regional scales. For example, are diversity gradients and taxonomic turnover patterns in Borneo due to a post-glacial expansion of species out of a refugium in the northwest (Ashton 2003b) or to contemporary west–east rainfall gradients (see Slik et al. 2003)? While north temperate forests appear to have re-established generic composition and diversity fairly rapidly after each recent glacial retreat (100,000 years to the present), perhaps soon reaching an equilibrium, is it likely that the species composition of high-diversity tropical forests would similarly rebound?

The analysis of phylogenetic structure (as described above) at this continental/regional scale, with the pool being the continental flora and the sample being a regional flora, may help answer some of these questions. Intra-continental diversification would be observed on the pool phylogeny as numerous small phylogenetic clusters of taxa in a regional-scale sample (high NTI). However, deeper association of clades with different regional samples would indicate that samples differ in their edaphic and climatic factors and taxa within the associated clades share suitable ecological characters. Using inter-plot phylogenetic distances to ordinate samples can extract ecological signal from the data that would be missed in ordinary presence/absence or abundance ordination, because across continental scales there may be sufficient species turnover that few species are shared by sample units (e.g., Terborgh and Andresen 1998). Finally, if the taxa in regional-scale samples are evenly distributed on the pool phylogeny (i.e., sibling taxa seldom occur together on the regional scale; low NTI), then this would indicate either extensive regional competitive exclusion or the persistent signal of allopatric speciation in all clades. The latter is unlikely, given the repeated mixing at continental scales caused by climate cycles.

As an example, we examined species turnover between three 50 ha Asian plots in the CTFS network: Lambir (on Borneo; Lee et al. 1999), Pasoh (Peninsular Malaysia: Manokaran et al. 1992), and Huai Kha Khaeng (HKK; Western Thailand; Davies et al. unpublished data). We treat the plots as samples of the regional floras, within the continental context of Southeast Asia; they share
a number of taxa by historical mixing of the Southeast Asian flora, but there is also a strong seasonality gradient from north to south. We note first that the Euclidean distance (based on species presence/absence) between HKK and Pasoh is 0.83 times the distance between Pasoh and Lambir, and HKK–Lambir is 0.97 times Pasoh–Lambir, indicating relatively low species similarity between Pasoh and Lambir, and overlap of common species among all three plots. However, the mean nearest taxon distance (the phylogenetic distance, or age, between most closely related taxa; Webb et al. 2002) between HKK and Pasoh is 3.0 times the distance between Pasoh and Lambir, and HKK–Lambir is 4.4 times Pasoh–Lambir, indicating more phylogenetic and therefore ecological similarity between Pasoh and Lambir, which are both aseasonal rainforests. Lambir also showed a significant association with rainforest clades, such as annonaceae/, sapindaceae/, and anacardiaceae/ (Figure 6.2), and a significant under-representation of taxa more common in seasonal forests: phyllanthaceae/, myrtales/, and asteraceae/.

Global biogeography and assembly of continental-scale biota

Most of the species diversity within continents is generated by intra-continental speciation, and the phylogenetic structure of continental samples on a global pool will show clustering of taxa within separate clades, which does not necessarily reflect any ecological signal. The larger question is, are deeper angiosperm clades distributed more evenly around the globe than expected by chance? Clearly, some deep clades are restricted geographically (e.g., dipterocarpaceae/ in Southeast Asia). However, in an influential paper, Gentry (1988) noted that the family-level taxonomic structure of rainforest plots was very similar across the three tropical zones. This could occur by (1) a more similar than expected phylogenetic composition of continental pools or (2) continental-to-regional, regional-to-local, or local-to-habitat processes that selectively create habitat-scale plots with a globally similar deep phylogenetic structure. The tricky part of addressing this question is deciding what the appropriate null models should be. An appropriate null model for the global assortment of plant lineages must take into account the history of land movements and climatic zones (e.g., Morley 2000), the emerging understanding of the biogeographic tracks of major lineages (e.g., Davis et al. 2002), and the possibility of extensive intercontinental dispersal (Pennington and Dick 2004). The time is nearing when a grand review of the movement of higher plant lineages in space–time will be possible, an update of Raven and Axelrod’s (1974) landmark paper. However, we know of no processes that would cause an over-dispersed (or even) global distribution of major clades. The stems of clades that compose tropical rainforests are old enough that there has been time for members of most major clades to disperse globally, but we expect and observe significant variation among continents in the phylogenetic composition of each clade.

Any similarity in composition of forest plots around the world must therefore lie either in selective intra-continental diversification patterns, such that continental species pools have a more similar than expected clade composition, or in continental-to-local processes that may cause local plot composition to be globally more similar than expected. The former is a possibility, given the conservatism of reproductive characters and their potential influence on speciation rates in different clades. Alternatively, what is the possibility that the taxonomic components of a forest “fit together” — that, for example, a “rubieae/ and lauraceae/ and sapotaceae/” set is more stable in some way than a “rubieae/ and rubieae/ and rubieae/” set? This resembles the old question of whether there are hundreds of niches in a forest so that hundreds of species can coexist. Rather than hundreds of niches, there may be a smaller number of niches that do occur regularly on the scale of a forest sample plot (Valencia et al. 2004), which, combined with clade-wide ecological characters, stabilize the higher-level taxonomic and phylogenetic structure of a forest. For example, we might eventually attempt to deconstruct forest phylogenetic composition in terms of conserved functional characters (Wilson 1999), for example: (1) N-fixers (fabaceae/), (2) emergents (some clades of...
dipterocarpaceae/), (3) understory, bird dispersed (some clades of rubiaceae/), etc. Within these broad “niche–clade” associations there are many equivalent species, the abundance of which may be determined predominantly by biotic density dependence or even by chance (Carson et al. Chapter 13 this volume, Chave Chapter 2 this volume).

TOWARDS PREDICTING TROPICAL CLADE COMPOSITION

As climate models become more detailed and powerful and the rate of forest conversion (hopefully) stabilizes, predictive models of forest species composition could be used to estimate the vulnerability of various forest areas to species loss and invasion. However, at a first examination, predicting which species are likely to occur in an unvisited forest community appears impossible, because the composition of tropical forest, at the local scale we perceive as we walk through it, is influenced by the complex interplay of abiotic environment acting on ecological characters, biotic interactions, historical causes of species range, and chance (e.g., Connell 1980, Ricklefs and Schluter 1993, Hubbell 2001). With $l$ species in a local community from which the $h$ members of a plot on a particular habitat are drawn, we might appear to need to know every entry in an $l \times E$ matrix of $E$ ecological attributes, an $h \times h$ matrix of biotic interactions. To predict the $l$ community members from a regional species pool of size $r$, we would need to know the contents of a vector of dispersal probabilities of length $r$, and so on up to the global scale. Beyond being unrealistic because of the amount of ecological data necessary, this approach still would not allow the prediction of taxonomic structure in unvisited areas for which the species are unknown.

A phylogenetic approach to ecological prediction helps us because it reduces the dimensionality of the problem. For $n$ taxa, a fully resolved phylogeny requires approximately $(n^2/2)$ bits of information, but the work of systematists is making this level of resolution a possibility. Once we have the phylogeny, the distribution of character states for the terminal taxa (e.g., species) can be coded with as few as a single change somewhere in the tree, or as many as $(n/2)$ changes (for a highly homoplasious character), but never approaching the $n$ states required if the taxa were all independent. Similarly, biogeographic distributions can be reconstructed on the tree, and there are good reasons to believe that interspecific biotic interactions should also have a phylogenetic signal (Webb et al. 2006). Overall, the dimensionality of the problem is greatly reduced. The phylogenetic approach provides the capacity to make reasonable predictions about the characters of unknown taxa, propagating information outwards on the tree of life (D. Ackerly personal communication).

For example, one potential algorithm for predicting the clade composition, or phylogenetic structure, of a habitat in tropical forest (but not the precise species composition) is:

1. Identify the ancestral ecological condition of $c$ clades of tropical forest trees: $\mathbf{ε}_c$.
2. Class the clades as having evolved ecologically in a particular fashion; for example, silverswords of Hawaii, with high divergence and homoplasy, or Rhizophoraceae with high conservatism: $\mathbf{ϕ}_c$.
3. Modify each clade by the general pattern of ecological evolution to give an expected vector of ecological character for the extant members of the clade:
   \[
   \mathbf{Ec} = f(\mathbf{ϕ}_c, \mathbf{ε}_c).
   \]
4. Filter the species in all clades through a clade-specific biogeographic $\beta_c$ and general climatic filter ($\kappa$) to give the clade composition of a local species pool:
   \[
   \mathbf{L} = \sum f(\beta_c, \mathbf{Ec}|\kappa).
   \]
5. Identify the niche template of a habitat-scale site: $\mathbf{ν}$.
6. Combine the above to give a potential species composition at a habitat-scale site:
   \[
   \mathbf{H} = f(\mathbf{ν}, \sum \mathbf{Ec}, \mathbf{L}).
   \]
7. Add the modifying roles of chance ($\theta$), and of phylogenetically correlated biotic interactions ($\varsigma$).
in influencing abundance within an ecologically equivalent group (and therefore presence/absence in small plots), to give a final vector of species composition:

$$\vec{h} = f(\vec{H}, \theta, \varsigma).$$

The species in vector $\vec{h}$ should have the same phylogenetic structure as the observed community, although more detailed information on the evolution of ecological character in all clades would be required to predict the precise species composition.

**CONCLUSIONS**

The characters of taxa that determine the outcome of local ecological interactions can be examined in a phylogenetic context. The nature of diverse tropical forest communities suggests that these characters may have generally evolved slowly (over millions of years) and essentially randomly, probably leading to a phylogenetically conservative pattern of autecology for most species, although ecologists and systematists need to work together to better document the precise pattern of character evolution. Local ecological biotic interactions are probably diffuse and lead to the co-occurrence of many related taxa in particular habitats, but the exclusion of similar taxa at smaller scales. Preliminary analysis of the phylogenetic structure of communities at various spatial scales reveals patterns of relatedness within habitats, regions, and continents, and supports the general theory of the evolutionary ecology of tropical forest trees outlined here. Phylogenetic models of ecological characters and biotic interactions may soon permit us to predict taxonomic composition of tropical forest communities. Such models will become increasingly important as the effects of both local and global human activities become more profound.

**ACKNOWLEDGMENTS**

We thank Peter Ashton, David Ackerly, Michael Donoghue, Mark Leighton, Gary Paoli, David Peart, and Miles Silman for discussions that helped form this chapter. Stefan Schnitzer and two anonymous reviewers offered very helpful comments. Sarayudh Bunyavejchewin, Somboon Kiratiprayoon, Md. Nur Supardi Noor, Abdul Rahman Kassim, and Sylvester Tan kindly provided access to the Thai and Malaysian CTFS plot data. The first author was supported by an NSF grant (DEB-0212873).

**REFERENCES**


