Patterns of Forest Phylogenetic Community Structure across the United States and Their Possible Forest Health Implications

Kevin M. Potter and Frank H. Koch

The analysis of phylogenetic relationships among co-occurring tree species offers insights into the ecological organization of forest communities from an evolutionary perspective and, when employed regionally across thousands of plots, can assist in forest health assessment. Phylogenetic clustering of species, when species are more closely related than expected by chance, suggests a process of evolutionary niche conservatism. Because such communities share much evolutionary history and an affinity for similar environmental conditions, they may be particularly susceptible to threats such as insects and diseases and shifting climate conditions. Meanwhile, a pattern of phylogenetic evenness, in which the species are less closely related than by chance, may indicate competitive exclusion or interspecies facilitation. The ecological integrity of such communities may be less at risk because they may encompass a wider variety of evolutionary adaptations. Using a network of more than 100,000 forest inventory plots across the conterminous United States, we tested whether community phylogenetic structure was significantly clustered or even at multiple scales. Clustering predominated across most of the study area, indicating the widespread significance of evolutionary niche conservatism, except in areas of the west. Phylogenetic structure varied along environmental gradients, suggesting that clustering predominates in more favorable locations and evenness predominates in areas with harsher environments. These results have implications for broad-scale forest health monitoring.

Keywords: evolutionary ecology, forest community, forest ecology, forest health assessment, landscape ecology

Recent broad-scale forest health assessments have incorporated measurements of the evolutionary relationships among tree species within forested communities of the United States. Specifically, such relationships are associated with biomass accumulation in some circumstances (Potter and Woodall 2014), whereas regional changes in tree seedling evolutionary diversity may indicate the future effects of climate change on forested communities (Potter and Woodall 2012). Such analyses are part of an emerging synthesis of community ecology and evolutionary biology (Johnson and Stinchcombe 2007) that provides a framework for quantifying evolutionary and ecological patterns in ecological communities as well as for testing some of the central mechanisms important in structuring those communities (Swenson et al. 2007, Emerson and Gillespie 2008, Burns and Strauss 2011). These analyses of phylogenetic relationships among tree species also offer new insights into the ecological organization of forest communities from an evolutionary perspective and have the potential to assist in the assessment of the health of forest communities from the perspective of biodiversity and resilience to stress (Potter 2009).

Specifically, two general types of phylogenetic community analysis are useful for quantifying evolutionary relationships within forest tree communities across broad scales: the generation of phylogenetic diversity statistics and the quantification of phylogenetic community structure. Phylogenetic diversity statistics (Faith 1992, Webb et al. 2006) are meaningful measures of biodiversity because, unlike species richness and related metrics, they account for the evolutionary history within a set of co-occurring species. Such analyses of phylogenetic relationships among species should better represent the diversity of species traits within that community, because taxonomically distinct species are expected to contribute more to the diversity of functional traits existing within the community (Faith 1992, 2002). Recent
publications have assessed regional relationships between phylogenetic diversity statistics and species richness in North American forests (Paquette and Messier 2011, Potter and Woodall 2012, 2014).

The central methodology of community phylogenetics, meanwhile, is to define the phylogenetic structure of a local community (evenness or clustering) relative to random species assemblages drawn from a broader species pool (Swenson et al. 2006). Stated another way, community phylogenetic methods test whether the species present in a community are more or less clustered on the tree of life compared with a random selection of species from a wider geographic area that could potentially be present within the community. Whereas analyses of plant community structure generally have been conducted at relatively local scales (for example, Kembel and Hubbell 2006, Silvertown et al. 2006, Webb et al. 2006), the examination of phylogenetic structure across broad geographic regions should prove useful (Pennington et al. 2006). Recently, such studies have quantified forest tree phylogenetic community structure across broad regions in North America (Qian et al. 2013), South America (Giehl and Jarenkow 2012), and Australia (Kooyman et al. 2011). Within the context of forest ecology and forest health, plot-level forest phylogenetic community structure across large scales can be used to test how environmental variables are associated with patterns of relatedness among co-occurring tree species and can allow for assessments of which forested areas are at greater risk from threats such as pests, pathogens, nonnative invasive plants, and changing climate conditions.

Analyses of phylogenetic structure in species occurring within and across communities provide insight into the ecological and evolutionary processes that organize these communities (Webb 2000, Bryant et al. 2008). A clustered phylogenetic distribution of taxa, in which species are more closely related than expected (for example, Tofts and Silvertown 2000, Webb 2000, Verdu and Pausas 2007), suggests that habitat use may be a conserved trait among the species in the community (Webb et al. 2002). Such a pattern of environmental filtering is expected to occur when closely related species share similar tolerances to the abiotic environment, as a result of shared functional traits that arose in a common ancestor, and exhibit evolutionary niche conservatism (Cavender-Bares et al. 2004, 2006). This is the expectation that, all else being equal, related species will tend to occupy the same sorts of environments (Donoghue 2008). Because such forest communities share much evolutionary history and an affinity for similar environmental conditions, they may be particularly susceptible to threats such as generalist insects and diseases, nonnative plant invasion, and shifting climate conditions.

On the other hand, an even phylogenetic distribution of taxa, in which species are less closely related than expected by chance (for example, Cavender-Bares et al. 2004, Slingsby and Verboom 2006, Bryant et al. 2008), may suggest competition among related taxa for similar niches within the community (Webb et al. 2002). This competitive exclusion, a form of biotic interaction, should limit the coexistence of closely related species if they share limiting resources (Cavender-Bares et al. 2004, 2006). Phylogenetic evenness also may be caused by abiotic habitat filtering that acts on distantly related species that exhibit convergent niche evolution toward fitness for a particular set of environmental conditions (Kraft et al. 2007, Donoghue 2008). Ecological facilitation (Vandermeer 1989), in which positive interactions occur between phylogenetically distant species, could also result in this type of phylogenetic community structure (Valiente-Banuet and Verdu 2007, 2008). The ecological integrity of such communities could be less at risk from environmental threats such as insects, disease, nonnative plants, and climate change because the species contained within the communities may encompass a wider variety of evolutionary adaptations to respond to changing conditions.

Using a network of more than 100,000 standardized forest inventory plots, we tested whether the forests of the conterminous United States exhibit significant phylogenetic structure (clustering or evenness) nationally and regionally. These permanent plots are a consistent size, are inventoried in a consistent manner across the United States, and are sampled at a consistent spatial intensity across that area (Bechtold and Scott 2005). The result is a standardized forest tree inventory data set that samples from an extensive area at an intensity that allows for robust analysis at large scales and reduces the likelihood of an incorrectly rejected hypothesis of clustering or evenness in cases where the community is very species poor or constitutes a very large proportion of the regional species pool (Kraft et al. 2007). To our knowledge, no other such standardized plot-level forest community data set is available at a continental scale. The plots are small enough (0.067 ha) that species may interact on the plots, allowing for the possibility that patterns of phylogenetic evenness in community structure could be the result of interspecific competition (Vamosi et al. 2009).

To operationally define the spatial extent of the species pools to which forest tree communities are compared in our study (Swenson et al. 2006), we analyzed the forest inventory data within the framework of a multiple-scale nested system of ecoregions, in which the ecoregions at each descending scale are defined in an increasingly finer fashion based on geology, climate, soils, potential natural vegetation, and potential natural communities (Cleland et al. 1997). We conducted these analyses across three ecoregion scales for three taxonomic groupings and identified which environmental factors help explain patterns of phylogenetic relatedness across broad regions.

Methods
Forest Tree Community Data Assembly

The Forest Inventory and Analysis (FIA) Program of the US Department of Agriculture (USDA) Forest Service systematically collects tree inventory information on all forestland ownerships in the United States and is the primary source for information about the extent, condition, status, and trends of the nation’s forest resources (Smith 2002). FIA maintains a network of more than 100,000 permanent ground plots of forested land across the conterminous United States, with a sampling intensity of approximately 1 plot per 2,428.11 ha. Forested land is defined as being at least 0.405 ha in extent and at least 10% stocked by forest trees of any size, or as land formerly having such tree cover and not currently developed for a nonforest use. Each FIA plot consists of four 7.32-m fixed-radius plots, allowing for the possibility that patterns of phylogenetic evenness in community structure could be the result of interspecific competition (Vamosi et al. 2009).

To operationally define the spatial extent of the species pools to which forest tree communities are compared in our study (Swenson et al. 2006), we analyzed the forest inventory data within the framework of a multiple-scale nested system of ecoregions, in which the ecoregions at each descending scale are defined in an increasingly finer fashion based on geology, climate, soils, potential natural vegetation, and potential natural communities (Cleland et al. 1997). We conducted these analyses across three ecoregion scales for three taxonomic groupings and identified which environmental factors help explain patterns of phylogenetic relatedness across broad regions.

Methods
Forest Tree Community Data Assembly

The Forest Inventory and Analysis (FIA) Program of the US Department of Agriculture (USDA) Forest Service systematically collects tree inventory information on all forestland ownerships in the United States and is the primary source for information about the extent, condition, status, and trends of the nation’s forest resources (Smith 2002). FIA maintains a network of more than 100,000 permanent ground plots of forested land across the conterminous United States, with a sampling intensity of approximately 1 plot per 2,428.11 ha. Forested land is defined as being at least 0.405 ha in extent and at least 10% stocked by forest trees of any size, or as land formerly having such tree cover and not currently developed for a nonforest use. Each FIA plot consists of four 7.32-m fixed-radius subplots, in which all trees having ≥12.7 cm or greater dbh are measured. In addition, each subplot contains a microplot with a radius of 2.07 m in which saplings of 2.54–12.45 cm dbh are inventoried (Bechtold and Scott 2005).

Using SAS version 9.2 (SAS Institute, Inc., Cary, NC) and MATLAB version 7.01 (The Math Works, Natick, MA), we assembled FIA tree and sapling inventory data for each of 102,304 plots and containing approximately 2.5 million trees (Woudenberg et al. 2010). At the time of analysis, data were not yet available from forested lands in central and west Texas. Using ArcMap 9.2 (ESRI, 2003).
Inc., Redlands, CA), we assigned each plot to the appropriate ecoregion domain, division, province, and section under the USDA Forest Service hierarchical ecoregion system (Figure 1), using the most recent geographical information system (GIS) mapping of these nested scales (Cleland et al. 2007). Domains and divisions, the two broadest scales, are based on large ecological climate zones, whereas each division is divided into provinces based on vegetational macrofeatures. Provinces are further divided into sections, encompassing areas similar in their geology, climate, soils, potential natural vegetation, and potential natural communities (Cleland et al. 1997).

By law, the exact coordinates of FIA plots are slightly altered to protect the privacy of forest landowners, with most of the adjusted coordinates located within 0.8 km and all within 1.6 km of the actual plot coordinates. In addition, a small proportion of plot coordinates from privately owned locations may be “swapped” with those of another private plot within the same county with similar attributes, such as forest type, stand-size class, latitude, and longitude (Woudenberg et al. 2010). Obscuring the original plot coordinates should have little effect on the results of this study because plot-level results are summarized by ecoregions. Ecoregion sections, the smallest scale of the analyses, are generally thousands of square kilometers in extent and generally encompass hundreds to thousands of plots each. In addition, each sample in the national FIA system is independently selected (Bechtold and Patterson 2005), so autocorrelation among samples is not an issue regardless of the proximity of the most closely neighboring plots (generally about 5 km apart).

**Supertree Construction**

Determining community structure statistics required constructing a hypothesized phylogenetic supertree of 311 tree species included in the FIA data (Potter and Woodall 2012). In brief, we used the online phylogenetic database and toolkit Phylomatic (Webb and Donoghue 2005) to generate a backbone phylogenetic supertree topology based on the Angiosperm Phylogeny Group II classification of flowering plant families (Angiosperm Phylogeny Group 2003). We then used the BLADJ (Branch Length ADJustment) module in the Phylocom version 3.41 software package (Webb et al. 2008) to assign ages to nodes in this supertree based on the fossil and molecular estimates reported by Wikström et al. (2001), with undated nodes spaced evenly between dated nodes to minimize variance in branch lengths. The resulting “pseudochronogram” can be used to estimate phylogenetic distance, in units of time, between taxa (Webb et al. 2008). To improve the resolution of the phylogenetic supertree to the species level and to expand it to incorporate gymnosperms, we surveyed recent molecular systematic and paleobotanical studies of the families and genera that encompass North American tree species and of the higher-level gymnosperm groups (Potter and Woodall 2012). We added dated node constraints to the supertree topology where possible and then re-ran the BLADJ algorithm in Phylocom to again set the ages of undated nodes evenly between dated nodes. Such supertree phylogenies approximate the actual evolutionary relationships among species and can be improved with additional molecular systematic research and with enhanced understanding of the fossil record.

**Community Phylogenetic Analysis**

Community phylogenetic methods test whether coexisting species are more or less clustered on the phylogenetic supertree compared with a random selection of species from a wider geographic area that could potentially be present within the community (Figure 1). Boundaries for the USDA Forest Service hierarchical system of ecoregions for divisions, provinces, and sections. For the names of ecoregion divisions, provinces, and sections, see Table 1, Supplemental Table S1, and Supplemental Table S2, respectively. The forest cover area, depicted in gray, was derived from MODIS imagery by the USDA Remote Sensing Application Center.
We used Phylocom 3.41 (Webb et al. 2008) to calculate community phylogenetic statistics at the plot level with random null communities drawn from the species present in each of three ascending ecoregion scales (section, province, and division) to examine how the geographic scale of the null community species pool affects the signal of phylogenetic clustering or evenness in plot-level species co-occurrences. Plots containing only one species cannot be interpreted using these measures of phylogenetic structure, so we used plots containing two or more species (Swenson et al. 2007). Ecoregion units containing fewer than 25 plots were excluded from the analysis to ensure reasonable sampling. In addition to analyzing all tree species inventoried, we conducted separate analyses of angiosperm and gymnosperm species.

For each plot, we calculated mean phylogenetic distance (MPD), in millions of years, between each species on a plot and all other species on the plot, and mean nearest neighbor distance (MNND), the mean minimum phylogenetic distance between each species on the plot and the most closely related species also on the plot, again measured in millions of years (Webb et al. 2006). The MPD and MNND statistics were then used in two indices that characterize the phylogenetic community structure of a set of co-occurring species (Figure 2). The net relatedness index (NRI) is a standardized measure of total (whole tree) within-community relatedness; and nearest taxon index (NTI), which quantifies nearest-relative (branch-tip) relatedness in a community. For both, the phylogenetic relationships of trees from a forest plot (A) are depicted on a phylogenetic tree (B). For NRI, the MPD from each tree to every other tree is calculated (C); for NTI, the MNND from each tree to its closest evolutionary neighbor is calculated (C). (The dashed lines on the hypothetical phylogenetic trees in B depict the evolutionary distances included in each measure for a single tree species on the plot; after this evolutionary distance is determined for each species, a plot mean is calculated.) The MPD and MNND values for a plot are then compared with the MPD and MNND values from a set of phylogenetic trees created from species drawn randomly from those present in the region (D). The results are the NRI and NTI statistics (E), which measure whether coexisting species are more or less phylogenetically clustered than expected by chance. NRI values are negative when species are less phylogenetically related than expected by chance across the entire tree (phylogenetic evenness) and are positive when they are more related (phylogenetic clustering), whereas NTI values are positive (clustered) when species occur with other closely related species and are negative (even) when species do not occur together with closely related species.

$$\text{NRI} = \frac{\text{MPD} - \mu_{\text{MPD}(0)}}{\sigma_{\text{MPD}(0)}}$$

where $\mu_{\text{MPD}(0)}$ is the mean phylogenetic distance taken from 1,000 simulated null communities and $\sigma_{\text{MPD}(0)}$ is the SD of MPD in those null communities. The nearest taxon index (NTI), meanwhile, is a standardized measurement of nearest-relative relatedness in a community. It quantifies the terminal (branch tip) clustering of co-occurring species regardless of the clustering of the higher-level clades on the tree (Webb 2000, Webb et al. 2002). Similar to NRI, NTI is calculated as

$$\text{NTI} = \frac{\text{MNND} - \mu_{\text{MNND}(0)}}{\sigma_{\text{MNND}(0)}}$$

where $\mu_{\text{MNND}(0)}$ is the mean of the MNND measurements from 1,000 simulated null communities and $\sigma_{\text{MNND}(0)}$ is the null community MNND SD. NRI values are negative when species are less phylogenetically related than expected by chance across the entire tree (phylogenetic evenness) and are positive when they are more related (phylogenetic clustering), whereas NTI values will be positive (clustered) when species occur with other closely related species and will be negative (even) when species do not occur together with closely related species (Kembel and Hubbell 2006).

To generate the null communities, we used the independent swap algorithm (Gotelli and Entsminger 2003) in Phylocom (Webb et al. 2008), which randomizes patterns of co-occurrence of species.
occurring on plots within each ecoregion but does not introduce species from the reference phylogeny into the null communities. Each of the 1,000 null communities used for each plot was generated through 50,000 swaps. For each scale of our analyses (11 ecoregion divisions, 35 provinces, and 165 sections), we used the species present across a given ecoregion unit as the pool for the creation of null communities.

We used a single-sample Wilcoxon signed-rank test (Conover 1971) to test whether the community phylogenetic structure across each ecoregion was significantly clustered or even, that is, whether NRI and NTI values differed significantly from 0. To investigate smaller-scale patterns of clustering, we interpolated plot-level NRI and NTI values determined with province-level species pools, using inverse distance weighting interpolation in ArcMap 9.2, at a resolution of 10 km with a 100-km search radius. Each province was interpolated independently using a polyline ecoregion province coverage as a search barrier.

Using PROC REG in SAS version 9.2, we generated single-variable linear regression models to determine which environmental variables were the best predictors of mean plot-level NRI and NTI across ecoregion sections, the smallest scale of analysis, with the species pool defined at the same scale. The models were ranked using Akaike’s information criterion (Akaike 1974), and only models significant at $P < 0.05$ and with $R^2 > 0.1$ were included. The climate variables for each plot were derived from the parameter elevation regression on independent slopes (PRISM) climate mapping system (PRISM Climate Group 2008), which had a spatial resolution of 4 km² and encompassed mean 30-year annual precipitation and monthly maximum and minimum temperatures across the years 1971–2000. Seven soil variables were included: available water capacity, bulk density, depth to bedrock, pH, permeability, plasticity, and porosity, all from CONUS-Soil (Miller and White 1998), a multilayer soil characteristics data set, at a resolution of 1 km², based on the USDA State Soil Geographic Database (STATSGO). Mean elevation was derived from a digital elevation model with a resolution of approximately 800 m² (US Geological Survey 1996), and plot latitude and longitude were provided in the FIA database. Mean values for each variable were calculated across plots within each ecoregion section.

Results

Our analyses identified statistically significant scale-associated differences in phylogenetic clustering metrics across 102,304 plots when we used the species present within hierarchically nested ecoregion sections, provinces, and divisions as the species pools (Figure 3). Across all scales, NRI and NTI were significantly positive (phylogenetically clustered) for the two sets of analyses that encompassed all tree species and angiosperms only. For the analyses encompassing gymnosperm species only, NRI and NTI were negative (phylogenetically even) across scales, with the exception of the NRI province-level scale analysis, which was neither significantly clustered nor even. The results further suggest a trend toward higher mean positive values at larger scales for both the all-species and angiosperm analyses and toward lower negative values for gymnosperms in the NTI analysis (Figure 3). Wilcoxon matched-pair tests comparing section, province, and division NRI and NTI (not shown) demonstrated that all were significantly different from each other for the analyses of all species and of angiosperms. The mean NRI and NTI values increased in magnitude with increasing scale, indicating an increasing degree of clustering (Figure 3). For gymnosperms, the section and division NRI values were not significantly different, but the NTI values for all three scales were significantly different from each other and were increasingly negative (evenly dispersed) with increasing scale of analysis.

In addition, our analyses demonstrated regional differences in phylogenetic clustering and evenness at multiple scales, as well as differences between taxonomic groups. NRI, the treewide measure of structure, was significantly clustered for most ecoregion divisions, the largest scale of analysis, with the exception of three: Tropical/Subtropical Steppe (310) was evenly dispersed, whereas Tropical/Subtropical Desert (320) and Savannah (410) were neither clustered nor evenly dispersed (Table 1). The same general pattern existed for NTI, the branch-tip measure of structure, except that the Subtropical (230) and Temperate Desert (340) divisions were also
evenly dispersed. In several cases, division-scale NRI and NTI for angiosperm species were neither clustered nor evenly dispersed. Only the Marine (240) division on the Pacific Coast was significantly evenly dispersed. Division-level NRI and NTI for gymnosperms, meanwhile, were more likely to be significantly evenly dispersed.

The geographic pattern of the clustering indices was further apparent when NRI and NTI values were calculated with province-level species pools, the scale immediately smaller than divisions. When all tree species were considered, ecoregion provinces with a preponderance of phylogenetically even plots tended to be most common in the Interior West, particularly in the Southwest (e.g., 313, M313, and 321), the Great Basin (M341 and 342), and the Northern Rockies (M332) for NRI (treewidth structure), whereas phylogenetically even plots for NTI (branch-tip) were also common in the Southeast (M231 and 232) (Figure 4A and D; Supplemental Table S1).

### Table 1. Standardized measures of forest tree community phylogenetic structure, calculated as means across plots in ecoregion divisions and provinces, with null community species pools determined by the species present in each ecoregion.

<table>
<thead>
<tr>
<th>Ecoregion name (Code)</th>
<th>All species</th>
<th>Angiosperms</th>
<th>Gymnosperms</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Plots NRI</td>
<td>NTI</td>
<td>Plots NRI</td>
</tr>
<tr>
<td>Warm Continental (210)</td>
<td>24,688 0.321* 0.259*</td>
<td>22,883 0.178* 0.196*</td>
<td>17,623 −0.042* −0.041</td>
</tr>
<tr>
<td>Hot Continental (220)</td>
<td>24,710 0.070* 0.188*</td>
<td>24,467 0.195* 0.194*</td>
<td>9,480 0.149* 0.139*</td>
</tr>
<tr>
<td>Subtropical (230)</td>
<td>28,736 0.078* −0.059*</td>
<td>26,416 0.087 0.057*</td>
<td>22,091 −0.004* −0.01*</td>
</tr>
<tr>
<td>Marine (240)</td>
<td>2,586 0.166* 0.147*</td>
<td>1,016 −0.048* −0.042*</td>
<td>2,534 0.03* 0.03*</td>
</tr>
<tr>
<td>Prairie (250)</td>
<td>2,846 0.068* 0.088*</td>
<td>2,818 0.178* 0.121*</td>
<td>547 0.182* 0.176*</td>
</tr>
<tr>
<td>Mediterranean (260)</td>
<td>3,733 0.335* 0.121*</td>
<td>2,201 −0.048 −0.036</td>
<td>3,213 −0.087* −0.135*</td>
</tr>
<tr>
<td>Tropical/Subtropical Steppe (310)</td>
<td>3,175 −0.117* −0.146*</td>
<td>1,343 0.076* 0.062*</td>
<td>2,935 −0.258* −0.313*</td>
</tr>
<tr>
<td>Tropical/Subtropical Desert (320)</td>
<td>99 0.031 −0.024</td>
<td>59 −0.009 0.024</td>
<td>48 −0.147* −0.183*</td>
</tr>
<tr>
<td>Temperate Steppe (330)</td>
<td>1,030 0.602* 0.294*</td>
<td>515 0.067 0.031</td>
<td>705 −0.125* −0.21*</td>
</tr>
<tr>
<td>Temperate Desert (340)</td>
<td>10,645 0.052* −0.053*</td>
<td>3,210 0.001* 0.001*</td>
<td>10,045 −0.045* −0.148*</td>
</tr>
<tr>
<td>Savannah (410)</td>
<td>56 0.369 0.142</td>
<td>40 −0.014 −0.02</td>
<td>46 0.102 0.102</td>
</tr>
</tbody>
</table>

- NRI is a tree-wide measurement of clustering, whereas NTI quantifies clustering at the tips of the branches on the phylogenetic tree. Positive values indicate phylogenetic clustering, and negative values evenness.
- Values significantly different from 0, based on a Wilcoxon single-sample signed rank test.
across the plots of only four provinces for NRI and for NTI, all but one in the West (Figure 4B and E; Supplemental Table 1). For gymnosperms, on the other hand, phylogenetic evenness was the prevailing pattern across the country both for NRI and NTI (Figure 4C and F; Supplemental Table 1). Results at the section scale, the smallest scale of analysis, were largely consistent with those at the province scale (Supplemental Table S2).

The interpolation of plot-level NRI (Figure 5A) and NTI (Figure 5B) revealed areas that exhibited strong spatial gradients in plot-level measures of phylogenetic clustering, including those in the Warm Continental (210) ecoregion division, where forests demonstrate decreasing clustering from south to north, and in the Marine (240) and Mediterranean (260) divisions, which are increasingly clustered at higher elevations. These patterns suggest a relationship between clustering and environmental variables such as elevation and precipitation. This relationship was further underscored by the explanatory power of environmental parameters in accounting for variations in the mean section-level NRI and NTI values (Table 2). For both metrics, calculated across all species, soil porosity, available water capacity, rock depth, and longitude (higher for more easterly longitudes) were associated with increased clustering, and greater soil bulk density and higher elevation were associated with increased evenness. Soil permeability and acidity (lower pH) were associated with increased treewide (NRI) clustering, but not with branch-tip clustering (NTI). Longitude (again, higher for more easterly longitudes) was the only variable for which the linear regression model of angiosperm clustering exceeded an $R^2$ value of 0.1, but several environmental variables were strongly associated with gymnosperm clustering, including available water capacity, depth to bedrock, longitude, and precipitation. Higher elevation was associated with greater gymnosperm phylogenetic evenness and was the best predictor of gymnosperm NTI (branch-tip) evenness ($R^2 = 0.374$).

Figure 5. Interpolation of plot-level NRI (A) and NTI (B) across forested areas of the conterminous United States at a pixel resolution of 10 km², with the null community species pool determined by the species present in each province. Interpolations were conducted independently for each province. The forest cover area was derived from MODIS satellite imagery by the USDA Remote Sensing Applications Center. See Supplemental Table S1 for the names of the ecoregion provinces.
The environmental associations were generally stronger for the gymnosperm branch-tip metrics (NTI) than for the gymnosperm tree-wide metrics (NRI).

**Discussion**

Our results demonstrate that the relative degree of phylogenetic clustering or evenness in the phylogenetic community structure of North American forests is strongly influenced both by the spatial scale of the species pool from which the community is drawn and by the taxonomic delineation of the species within the communities (Figure 3). The increasing phylogenetic clustering at larger scales, for all tree species and for angiosperms, is consistent with a study of forested areas in Central America (Swenson et al. 2006) and with expectations that clustering might be more common in communities with large regional species pools (Vamosi et al. 2009) as a result of biogeographic processes, such as limited dispersal, and evolutionary processes, such as local radiation, that result in co-occurring regions (Pennington et al. 2006, Swenson et al. 2007). Niche conservatism and therefore phylogenetic clustering are expected to emerge as the dominant patterns with increasing spatial scale, as the decreasing strength of competition and other density-dependent mechanisms, which prevent close relatives from co-occurring, give way to the increasing influence of biogeographical history, including the tendency of species within a clade to be concentrated in the clade’s region of origin (Cavender-Bares et al. 2009).

The opposite pattern among gymnosperms (Figure 3), however, is perhaps not unexpected, given that studies encompassing differing taxonomic groups have detected dissimilar phylogenetic patterns within communities (Swenson et al. 2007). Cavender-Bares et al. (2006), for example, found that communities that are more broadly defined phylogenetically are more likely to exhibit phylogenetic clustering as a result of trait conservatism and environmental filtering, whereas narrowly defined communities are more likely to exhibit phylogenetic evenness as a result of trait convergence, trait evenness, or both. In the current study, the plot-level phylogenetic distance per species is higher for gymnosperms than for angiosperms (220.1 million years per species versus 135.9 million years), suggesting that gymnosperms, at the plot level, generally represent more widely divergent lineages and are therefore more likely to tend toward a pattern of phylogenetic evenness across multiple scales. Given that the Pinales underwent major radiation during the Triassic period (~245–208 million years ago [mya]) (Willis and McElwain 2002) and that all the extant conifer families were in existence by 184 mya (Cheng et al. 2000, Eckert and Hall 2006), there may have been adequate time for conifers to evolve convergent traits that adapted distantly related taxa to similar environments and therefore allowed them to coexist at smaller spatial scales. The shorter time since the explosive diversification and radiation of angiosperms during the late Cretaceous (~100–65 mya) (Willis and McElwain 2002) may have hindered convergent evolution between distant lineages, which remain more closely associated with conserved niches and thus generated the pattern of phylogenetic clustering at the plot scale.

**Regional Patterns of Phylogenetic Structure**

Evolutionary and geographic histories are components of ecological systems that can be revealed in part through phylogenetic analyses, providing insight into the development of diversity patterns and unique aspects of biological communities in different regions (Ricklefs 2006). Our analyses demonstrated that the plot-level phylogenetic structure was generally clustered across ecoregion divisions, the largest scale of analysis (Table 1), suggesting a tendency toward niche conservatism in most areas. Kraft et al. (2007) suggested that such phylogenetic clustering generally arises only from filtering processes acting on conserved traits, so clustering provides provisional evidence of niche conservatism. In the analysis of community composition at large scales, this pattern could also result from biogeographic processes of local radiation and limited dispersal that cause close relatives to co-occur at a regional scale (Pennington et al. 2006). Our study, however, analyzed plot-level tree community composition, so such local radiation and limited dispersal...
might result in an unstructured or evenly dispersed community structure at the plot level relative to that at the clustered regional pool. This may in fact be the case for plots within two ecoregion divisions, the Subtropical (230) in the Southeast and the Temperate Desert (340) in the West, where plot-level structure for all species was clustered for the treewide or basal index (NRI) and evenly dispersed according to the branch-tip metric (NTI). These regions may have been occupied in the distant past by a relatively phylogenetically constrained set of tree species, which radiated into many forms via allopatric speciation, diversifying the regional species pool but reducing the likelihood that closely related species with similar traits would coexist at the plot level (Johnson and Stinchcombe 2007).

Interpreting phylogenetic evenness at the plot level is generally more complicated than understanding the forces that result in clustering, because evenly dispersed communities can arise from two quite different processes: competitive exclusion and filtering combined with convergent traits (Kraft et al. 2007). In the former, closely related species exclude each other as a result of competition for limiting resources, whereas in the latter, distantly related species have become similarly adapted to a given set of environmental conditions as a result of convergent evolution. In our analysis encompassing all tree species on FIA plots, only one ecoregion division was evenly dispersed by both structure metrics, the Tropical/Subtropical Steppe division (310) in the Southwestern United States (Table 1). In the analysis limited to angiosperms, only the Marine division (240) of the Pacific Northwest was evenly dispersed for both metrics. This pattern, as well as those at the smaller ecoregion province and section scales, is consistent with the findings of Qian et al. (2013), who observed that angiosperm tree species were phylogenetically clustered in the large majority of 12,100-km² quadrats across North America, a finding they attributed to phylogenetic niche conservatism. Meanwhile, our analyses of gymnosperms found phylogenetic evenness in four western ecoregion divisions and in a large proportion of the ecoregion provinces and sections, particularly in the West. These results suggest an important linkage between broad geographical region and phylogenetic structure, with the tendency toward evenness being generally limited to the western half of the conterminous United States.

Environmental Associations with Phylogenetic Structure

The results of our univariate regression analyses across ecoregion sections (Table 2) suggest that phylogenetic clustering and therefore niche conservatism may be more common in lower-elevation areas with moister and less compact soils. This finding is consistent with the expectation that biotic interactions shift from competition at more favorable, lower-elevation sites, to interspecies facilitation at harsher, higher-elevation sites (Bryant et al. 2008, Valiente-Banuet and Verdu 2008). Our results are also consistent with the finding that phylogenetic evenness and phylogenetic diversity were more highly correlated with live aboveground tree biomass accumulation on poorer sites, suggesting that the coexistence of functionally different tree species is associated with higher forest productivity in less productive and more stressful environments (Potter and Woodall 2014).

In addition, phylogenetic evenness in more extreme environments, such as the deserts and mountains of the Interior West of the United States, could result from the convergent evolution of necessary complex adaptations occurring in only a few lineages, which are then able to colonize and diversify in a relatively competition-free environment (Prinzing et al. 2001). The fact that branch-tip clustering (NTI) has a stronger relationship than treewide clustering (NRI) with higher available water capacity and with lower elevation suggests that such dynamics may have historically occurred in the forests of the United States, particularly among gymnosperms. Among woody plants, in fact, drier and higher-elevation sites across the globe have been found to contain proportionally greater numbers of genera and families at a given level of species diversity than low-elevation mesic sites (Enquist et al. 2002).

Possible Implications for Forest Health

Phylogenetic community structure measures, such as those described here, may serve as useful community-level indicators of potential resilience and susceptibility to forest health stressors. For example, phylogenetic community structure is expected to be associated with resistance to invasion (Strauss et al. 2006, Diez et al. 2008, Carboni et al. 2013), ecological impacts of disease (Gilbert and Webb 2007, Liu et al. 2012, Schweizer et al. 2013), and responses to climatic changes (Willis et al. 2008, Cavender-Bares and Reich 2012, Senior et al. 2013). Communities of species with even phylogenetic distributions possess greater-than-expected evolutionary diversity and may therefore encompass a higher proportion of species unaffected by a given stressor or able to adapt to it and may possess a greater diversity of functional traits. Given that community ecology theory predicts that an invasive species will be unlikely to become established when a species already exists with similar traits (Funk et al. 2008), a community containing a wider variety of traits, such as one in which species are phylogenetically evenly distributed, may be less susceptible to invasion. In contrast, species that are more closely related tend to be more ecologically similar (Burns and Strauss 2011). Phylogenetically clustered communities, therefore, are likely to contain less evolutionary and ecological diversity and may therefore be more vulnerable to stressors.

The combination of phylogenetic community structure for both the treewide (NRI) and branch-tip (NTI) measures may also have important forest health implications (Potter 2009). For example, a community may consist of species evenly distributed across the phylogenetic tree (negative NRI) while clustered at the branch tips (positive NTI). Such a community may contain fewer species susceptible to a given threat but may be more at risk of losing important unduplicated ecological functions were a set of closely related species to be eliminated. This was, however, not the general phylogenetic community pattern across any ecoregion division (Table 1), province (Figure 4; Supplemental Table S1), or section (Supplemental Table S2). On the other hand, a community clustered across the phylogenetic tree (positive NRI) while evenly distributed at the branch tips (negative NTI) may be able to retain its overall ecological functionality when species are eliminated as a result of a threat, but might encompass a greater number of more closely related species at risk of elimination. This was the pattern of phylogenetic structure for the Subtropical (230) and Temperate Desert (340) ecoregion divisions and in the Outer Coastal Plain Mixed Forest (232) and Coniferous Forest-Alpine Meadow (M331) ecoregion provinces. Forest communities may be least at risk of losing ecological function and more species when both NRI and NTI are evenly distributed (e.g., the Tropical/Subtropical Steppe [310] ecoregion division and ecoregion provinces 313, M313, and M341). Finally, forest communities may be most at risk of losing both greater proportions of species and ecological function when both metrics are clustered (e.g., the Warm Continental [210], Hot Continental
CAVENDER-BARES, J., AND P.L. PATTERSON. 2005. CAVENDER-BARES, J., A. KEEN, AND B. MILES. 2006. Phylogenetic structure analyses demonstrate that phylogenetic clustering predominates in forests across most regions of the country, indicating the widespread significance of evolutionary niche conservatism, except in areas of the West where competitive exclusion, convergent evolution, or facilitation may be more common processes. The degree of clustering is associated with environmental gradients in a manner suggesting the niche conservatism predominates in more favorable locations, whereas competitive exclusion, convergent evolution, or interspecies facilitation predominate in areas with harsher environments. It therefore appears that forest communities in areas possessing more favorable environments for forest trees may be most susceptible to the ecological affects associated with threats such as climate change, nonnative plant invasion, and insect and disease infestation. More research is clearly needed to assess the degree to which the phylogenetic structure of ecological communities, including forests across broad scales in the United States, is associated with susceptibility and resistance to such threats (Cavender-Bares et al. 2009).

**Literature Cited**


PENNINGTON, R.T., J.E. RICHARDSON, AND M. LAVIN. 2006. Insights into


PRISM CLIMATE GROUP. 2008. 30-Arcsec (800 m) gridded 30-year normal climate data (1971–2000). Oregon State University, Corvallis, OR.


US GEOLOGICAL SURVEY. 1996. GTOPO30—Global topographic data. EROS Data Center Distributed Active Archive Center, Sioux Falls, SD.


