Phylogenetic and functional diversity area relationships in two temperate forests

Xugao Wang, Nathan G. Swenson, Thorsten Wiegand, Amy Wolf, Robert Howe, Fei Lin, Ji Ye, Zuoqiang Yuan, Shuai Shi, Xuejiao Bai, Dingliang Xing and Zhanqing Hao

Phylogenetic diversity (PD, the diversity of lineages) and functional diversity (FD, the diversity of functional traits or groups in a biological community) reflect important yet poorly understood attributes of species assemblages. Until recently, few studies have examined the spatial variation of PD and FD in natural communities. Yet the relationships between PD and FD and area (termed PDAR and FDAR), like the analogous species–area relationship (SAR), have received less attention and may provide insights into the mechanisms that shape the composition and dynamics of ecological communities. In this study, we used four spatial point process models to evaluate the likely roles of the random placement of species, habitat filtering, dispersal limitation, and the combined effects of habitat filtering and dispersal limitation in producing observed PDARs and FDARs in two large, fully mapped temperate forest research plots in northeast China and in north-central USA. We found that the dispersal limitation hypothesis provided a good approximation of the accumulation of PD and FD with increasing area, as it did for the species area curves. PDAR and FDAR patterns were highly correlated with the SAR. We interpret this as evidence that species interactions, which are often influenced by phylogenetic and functional similarity, may be relatively unimportant in structuring temperate forest tree assemblages at this scale. However, the scale-dependent departures of the PDAR and FDAR that emerged for the dispersal limitation hypothesis agree with operation of competitive exclusion at small scales and habitat filtering at larger scales. Our analysis illustrates how emergent community patterns in fully mapped temperate forest plots can be influenced by multiple underlying processes at different spatial scales.

A fundamental goal in ecology is to understand the processes and mechanisms influencing spatial patterns of species diversity in natural communities. Although species diversity is an important and easily interpreted community attribute, phylogenetically- and functionally-based studies are increasingly employed by ecologists to explore community-level processes (Webb et al. 2002, Cavender-Bares et al. 2009, Swenson 2011). These studies have traditionally focused on phylogenetic and trait dispersion in local communities (Weiher et al. 1998, Webb 2000, Kraft et al. 2008, Swenson and Enquist 2009). Community ecologists have also been interested in phylogenetic diversity (PD) and functional diversity (FD), and their relationship because these variables may be better predictors of productivity and stability than species diversity (Tilman et al. 1997, Cadotte et al. 2009, Flynn et al. 2011, Helmus and Ives 2012). In addition to community ecologists, conservation biologists have long been interested in PD and FD as they represent an ecosystem’s conservation value and range of ecological services, which often are threatened by human activities (Faith 1992, Petchey and Gaston 2006, Cadotte et al. 2011).

Despite the rapidly increasing interest in PD and FD, the quantitative study of biodiversity still generally focuses on patterns of species richness or diversity (Swenson 2011). For example, hundreds of publications have explored the species–area relationship (SAR), which describes how the number of species changes with sampling area. The SAR is one of the most important statistical patterns in biodiversity research, and the increase in species richness with area is often considered one of the few ‘rules’ in ecology (He and Legendre 2002). The SAR links local scales to several hectares and tends to reflect ecological processes such as random placement, habitat filtering, dispersal limitation and species interactions (Plotkin et al. 2000, He and Legendre 2002, Rosindell and Cornell 2009, Shen et al. 2009, Wang et al. 2011). However, the relative importance of these ecological processes in explaining SARs remains largely controversial. For example, the assertion
that species are placed independently is a fundamental assumption of several theories of biodiversity, such as neutral theory (Hubbell 2001), that yield agreement with observations (McGill 2010) but contradict studies that outline the importance of species interactions (Chesson 2000). Additionally, the observed spatial aggregation patterns of species may be created by habitat filtering and dispersal limitation, the two major drivers of community assembly underlying niche theory and Hubbell’s neutral theory, respectively (Hubbell 2001, Tilman 2004).

One explanation for the current failure to discriminate among competing ecological processes is that species diversity treats all species as evolutionarily independent and ecologically equivalent and may therefore not provide enough information about the mechanisms of community patterns (Swenson 2011, Swenson et al. 2012). However, alternative axes of biodiversity such as phylogenetic and functional diversity may have the ability to convey a non-independence and complement species–area relationships in determining the degree to which processes such as random placement, habitat filtering and/or dispersal limitation contribute to the spatial distribution and assembly of species in plant communities (Webb 2000, Swenson et al. 2007, Kraft et al. 2008, Swenson and Enquist 2007, Helmus and Ives 2012). It is therefore somewhat surprising that the SAR concept has only rarely been extended to include phylogenetic diversity–area relationships (PDAR; Morlon et al. 2011, Helmus and Ives 2012) and virtually no published study has explored functional diversity–area relationships (FDAR) in detail.

Information regarding the phylogenetic and functional relationships of species may improve our understanding of the ecological processes underlying the spatial patterns of biodiversity in temperate forest tree communities. For example, distantly related or functionally dissimilar species may co-occur on local spatial scales where individuals directly interact (e.g. <30 m) due to competitive exclusion (i.e. more similar species are outcompeted when they meet each other), whereas more closely related or functionally similar species co-occur on the within-habitat scales (e.g. 30–200 m) encompassing the same habitat type if habitat filtering occurred (Weih and Keddy 1995, Cavender-Bares et al. 2006, Swenson et al. 2006). Thus, the influence of phylogenetic and functional traits on the spatial pattern of species in plant communities may differ at different spatial scales (Webb 2000, Swenson et al. 2007, Kraft et al. 2008, Swenson and Enquist 2009) and should therefore impact the shape of the PDAR or FDAR relative to the SAR in a predictably way: the PDARs and FDARs should initially show a steep increase at areas where individuals interact, then flatten down when covering for areas of the same habitat type, and showing again a steeper increase when the area covers different habitat types.

In this study, we develop theoretical expectation for the PDARs and FDARs built on data from two 25 ha fully mapped temperate forest plots from four competing models of community assembly: 1) the random placement hypothesis where species show no habitat specification and no mechanism of intraspecific aggregation (such as dispersal limitation); 2) the habitat filtering hypothesis where the distribution pattern of species in the community is only determined by their habitat specifications but not by mechanisms of intraspecific aggregation; 3) the dispersal limitation hypothesis where only intraspecific aggregation but no habitat specification is invoked; and 4) the combined habitat filtering and dispersal limitation hypothesis. Note that all four ecological process hypotheses assume that species have no direct interactions (McGill 2010), although shared habitat associations may create larger-scale spatial dependencies among species distribution patterns (Wiegand et al. 2007, 2012). This occurs if two species share the same habitat patches (positive larger-scale association) or if they have contrasting habitat preferences (negative larger-scale associations).

Our analyses utilize a recently developed technique in spatial point pattern statistics that allows fitting point process models corresponding to these four ecological process hypotheses to the spatial pattern of a given species (Waagepetersen 2007, Shen et al. 2009). A stochastic realization of the fitted point process is generated for each species, and the simulated patterns of all species are superimposed to yield one 'null community'. We generated for each hypothesis 100 null communities and averaged the PDARs and FDARs estimated from the 100 null communities to yield expectations for the different hypotheses (AIC; Supplementary material Appendix 1, Fig. A1). Finally, we evaluated the fit by using the Akaike’s information criterion and compared the observed and predicted PDAR and FDAR using simulation envelopes.

Based on expectations of how competitive exclusion, habitat filtering, and dispersal limitation may impact the co-occurrence of species at different spatial scales (Supplementary material Appendix 1, Table A1) we can qualitatively describe how the PDARs and FDARs generated by the competing hypotheses may differ (Fig. 1). Because the random placement hypothesis 1) does not consider intraspecific clustering it will overestimate species richness at the local scale (Wang et al. 2011), which will increase the PDAR and FDAR and because it does not consider habitat filtering it will also overestimate the PDAR and FDAR at the within-habitat scale (magenta curve in Fig. 1). The habitat filtering hypothesis 2) should yield correct predictions of PDAR at the inter-habitat scale because the driving mechanism at this scale (habitat filtering) is considered. However, because hypothesis 2) does not consider intraspecific aggregation it will overestimate species richness and therefore PDAR and FDAR at the local scale (green curve in Fig. 1). In contrast, the dispersal limitation hypothesis 3), which conserves the observed intraspecific aggregation of species should yield at the local scale lower PDAR values than the habitat filtering hypothesis 2) (because it yields at this scale null communities with lower species richness), but then overestimate the PDAR and FDAR at intermediate to large scales because the habitat filtering mechanism is missing (Fig. 1, red line). Finally, consideration of both habitat filtering and dispersal limitation should only show the departure expected to arise from biotic interactions on local scales via an underestimation at smaller areas.

Model selection together with an interpretation of the departures from the observed PDAR and FDAR allows us to infer on the relative importance of different mechanisms.
generating the observed PDAR and FDAR. For example, if model selection would favor the dispersal limitation hypothesis we can infer that other processes not represented by this hypothesis (i.e. species interactions and habitat filtering) may be relatively unimportant for generating the observed spatial diversity patterns (Plotkin et al. 2000, McGill 2010, Wiegand et al. 2012) as already observed for the SAR curves in these forests (Wang et al. 2011). In this case, phylogenetic and functional traits may not be the dominant processes influencing the spatial pattern of species in these forests.

**Methods**

**Study areas**

This study utilized and compared the spatial patterns of tree communities in two large temperate forest plots. The first plot is located in the Changbaishan (CBS) Nature Reserve, one of the largest biosphere reserves in China, located in north-eastern China (42°23′N, 128°05′E) near the border of China and North Korea. The 25-ha (500 × 500 m) forest plot has not experienced logging or other human disturbance for at least 300 yr. The mean elevation of the CBS plot is 801.5 m a.s.l. with a range of 791.8 to 809.5 m. The mean annual temperature is 3.6°C, and the mean annual rainfall is approximately 700 mm. During each census of the plot all individuals with a diameter at breast height ≥1 cm are mapped, measured and identified. Altogether 38 902 individuals belonging to 52 species, 32 genera and 18 families were recorded during the first census of the CBS plot in 2004 (Wang et al. 2010). The second study area, the 25.2-ha (300 × 840 m) Wabikon Forest Dynamics Plot, is located in the Chequamegon-Nicolet National Forest in north-eastern Wisconsin, USA (45°33′N, 88°48′E). The Wabikon forest is a secondary forest, which was partly logged in the early 1900s. The mean elevation of the plot is 498.1 m a.s.l. with a range of 488.3 to 514.2 m. The mean annual temperature is about 4.1°C, and the mean annual rainfall is approximately 819 mm (U.S. National Oceanic and Atmospheric Administration Satellite and Information Service). During the first plot census in 2008 a total of 48 849 individual trees (dbh ≥1 cm) belonging to 38 species, 26 genera and 16 families were mapped, measured and identified.

In order to examine the effect of habitat filtering on the PDARs and FDARs of the CBS and Wabikon plots, we followed Wang et al. (2011) and evaluated three topographical variables (i.e. elevation, slope and aspect), and for the CBS plot, measurements of 8 soil properties (i.e. pH, organic matter, total N, total P, total K, and available N, available P and available K). Soil variables were not measured for the Wabikon plot. Elevation was measured at the corner of each 20 × 20 m quadrat in the two plots. Elevation, slope and aspect values for each 5 × 5 m subquadrats were then interpolated using ordinary kriging. At the CBS plot, soils were sampled using a regular grid of points every 50 m. Two additional sample points at 2, 5, or 15 m were selected in a random compass direction from the grid point to capture variation in soil nutrients at finer scales (Wang et al. 2012). At each point 500 g of topsoil (0–10 cm depth) was collected, and the eight soil properties were then analyzed. Spatial predictions for 5 × 5 m subquadrats were calculated using using ordinary kriging. We also used the tree density in each quadrat (5 × 5 m) as a variable to capture the overall biotic conditions in the 5 × 5 m quadrats.

**Phylogenetic construction**

A phylogenetic tree was constructed for the two forest plots using the informatics tool Phylomatic (Webb and Donoghue 2005). Phylomatic utilizes the Angiosperm Phylogeny Group III (APG III 2010) phylogeny as a backbone. Known relationships from additional systematic work and taxonomic relationships are pasted onto this backbone to estimate the phylogenetic tree. Branch lengths were estimated for each tree using the BLADJ algorithm implemented in the software Phylocom (Webb et al. 2008) using estimated node dates from Wikstrom et al. (2001). The phylogenetic trees used did have soft polytomies, but they were few given the low number of species per genus in the study systems. Previous work by Swenson (2009) has shown that the PD metric utilized in this research is largely insensitive to terminal soft polytomies suggesting that the few soft polytomies in the phylogenies used had little-to-no effect on the results.

**Functional trait collection**

The following functional trait data were collected for tree species in each plot: maximum height, leaf area, specific leaf
area, wood density, leaf nitrogen, and leaf phosphorus. While these traits do not represent a full representation of all aspects of plant function, they provide robust indicators of where species fall along several major axes of plant functional strategy such as the adult light niche (maximum height), light capture (leaf area), leaf economics spectrum (specific leaf area, leaf nitrogen and leaf phosphorus), and trade off between structural investment and growth and mortality rates (wood density) (Swenson et al. 2012). The trait data were collected from greater than 10 individuals when possible, but sample sizes were lower for the rarest species. Trait collection protocols followed that of Cornelissen et al. (2003).

Here we are interested in quantifying the overall functional diversity instead of the diversity of individual traits. Because some of our traits are correlated, we calculated the principal components (PCs) from these traits after standardization (i.e. subtracting the mean value of the traits of all species and dividing by 1 standard deviation), and used the first five components to generate a functional distance matrix because they together explained more than 96% of the total variance in traits. This approach avoids the over- or under-representation of some axes of functional strategy in the distance matrix. We constructed a functional distance matrix by computing the Euclidean distances of the PCs between all species. This matrix and hierarchical clustering were then used to generate a dendrogram with branch lengths representing the functional similarity of species. A dendrogram was used to calculate the FD statistic (Petchey and Gaston 2006) analogously to the PD statistic (Faith 1992) used for calculating phylogenetic diversity.

Phylogenetic and functional area relationships

The PDARs and FDARs were constructed by randomly selecting quadrats with increasing sizes in the two temperate forest communities. The same method was used to compute the SARs (Wang et al. 2011). The phylogenetic diversity and functional diversity in each quadrat were calculated using the PD (Faith 1992) and FD (Petchey and Gaston 2006) metrics which are identical to one another, except that PD utilizes a phylogeny and FD a dendrogram derived from trait data. We used the R function pd() including the root in the package ‘picante’ to calculate both PD and FD using the phylogenetic tree and trait dendrogram. The PD and FD metrics yield the total phylogenetic (or functional) branch-length joining the basal node to the tips of all the species in the sample (Faith 1992). This metric is proportional to species richness for a dendrogram where species share no branch-length. The PD and FD metrics were also chosen because alternative metrics that utilize pairwise distances (Chave et al. 2007, Hardy and Senterre 2007, Cadotte et al. 2010), for example, are not monotonic with species richness making inferences from these curves less tractable.

Ecological processes

Point process models are mathematical models that can be used to characterize the stochastic properties of point patterns and to simulate point patterns based on these characteristics (Illian et al. 2008). By fitting point process models to the observed species distribution patterns we can quantify certain properties of the patterns (e.g. an observed spatial association to environmental variables or observed clustering) and use the fitted model to simulate patterns with these properties. We used four point process models that incorporated the stochastic mechanisms of our four hypotheses. This allowed us to generate null communities that are expected under each hypothesis. To this end, we first fitted a given point process model to the distribution pattern of each species, then generated for each species one stochastic realization of the fitted point process model, and finally superimposed the simulated patterns of all species to yield one null community (Supplementary material Appendix 1, Fig. A1). Because we superimposed the simulated patterns of all species independently none of our hypothesis considers the effects of species interactions. From each null community we generated a PDAR and FDAR and compared it to the observed ones.

The random placement hypothesis

We used the homogeneous Poisson process to represent the random placement hypothesis. The homogenous Poisson process model is characterized by two fundamental properties (Stoyan and Stoyan 1994): a) the points (individuals of species) are independently scattered which means that there is no interaction between the individuals (i.e. no clustering), and b) the intensity, \( \lambda \), of the pattern (i.e. the mean density of the species in a unit area) is constant which means that the number of individuals in study plots of area \( A \) follows a Poisson distribution with mean \( \lambda A \) (i.e. no habitat association). In this case the probability of finding an individual in an infinitesimally small disk of centre \( x \) and area \( dx \) yields \( \lambda dx \). Note that this point process model has no parameter to be fitted (except the number of individuals of a given species which is already determined by the data). This point process therefore generates for each species a completely random pattern (i.e. random placement) that does not conserve the observed habitat association, clustering or species interactions.

The habitat filtering hypothesis

We used the heterogeneous Poisson process to represent the habitat filtering hypothesis. This point process model holds the condition a) described above (i.e. no clustering), but the intensity of the model depends now on location \( x \) (i.e. the probability of finding an individual in an infinitesimally small disk of centre \( x \) and area \( dx \) is \( \lambda(x) dx \)). The intensity \( \lambda(x) \) may be influenced by topographic and edaphic factors. We used techniques presented in Waagepetersen and Guan (2009) based on log-linear regression models to estimate the intensity function \( \lambda(x) \) based on environmental covariates (topography, neighboring tree species, and soils) available for our two plots. This point process therefore generates for each species a pattern that conserves the observed habitat association, but not the observed clustering or species interactions. However, the heterogeneous Poisson process model is able to generate indirect (larger-scale) species associations if two species share the same habitat or have opposed habitat requirements.
**The dispersal limitation hypothesis**

Given that most species show aggregation at one or several scales (Condit et al. 2000, Wiegand et al. 2009), the independence property a) of the two Poisson process models is unrealistic for most data. We therefore used the homogeneous Thomas process (representing the dispersal limitation hypothesis) to approximate the observed aggregated patterns of individual species. The homogeneous Thomas process is the simplest point process model that creates clustered patterns. It consists of a number of randomly and independently distributed clusters. The position of the cluster centers follows a homogeneous Poisson process with intensity \( \varrho \) (i.e. \( \lambda A \) is the number of clusters in a study region of area \( A \)). The individuals belonging to the clusters are then distributed following two simple rules. First, the number of individuals that belong to a given cluster follows a Poisson distribution (with mean \( \mu = \lambda l \) where \( \lambda \) is the intensity of the pattern). Second, the actual locations of individuals around the cluster centers follow a bivariate Gaussian distribution with variance \( \varSigma^2 \) (Stoyan and Stoyan 1994). One advantage of the Thomas process is that an analytical formula exists for the most important summary statistic of point patterns, the pair correlation function \( g(r) \):

\[
g(r) = 1 + \frac{1}{\varrho} \frac{\exp(-r^2/4\varSigma^2)}{4\pi\varSigma^2} \quad (1)
\]

and for its cumulative version, the \( K \)-function which yields

\[
K(r) = 2\pi \int_{r=0}^r g(t) dt \quad (\text{Illian et al. 2008}).
\]

The pair correlation function describes the neighborhood density of trees at distance \( r \) of an arbitrary selected tree relative to the intensity \( \lambda \) of the pattern. If \( g(r) > 1 \) the pattern is aggregated and the second term of Eq. 1 (which is always larger than zero and describes the degree of aggregation of the pattern) can be fitted to a given data set.

The homogeneous Thomas process generates for each species a pattern that conserves the observed clustering, but not species interactions or habitat specialization. However, fitting a Thomas process to a heterogeneous pattern may formally produce a good fit if the environmental heterogeneity is patchy rather than varying across a spatial gradient. Thus, the homogeneous Thomas process may provide a good fit for species that show habitat specialization, but because no habitat association is considered this process cannot describe interspecific species associations that are mediated by shared or opposed habitat requirements. This will be noticed by summary statistics such as SAR or PDAR that operate at the community level.

**The habitat filtering and dispersal limitation hypothesis**

The heterogeneous Thomas process (representing the habitat filtering and dispersal limitation hypothesis) is slightly more complicated than the homogeneous Thomas process. The heterogeneous Thomas process results from ‘thinning’ a homogeneous Thomas process with an intensity function \( \lambda(x) \) (Waagepetersen 2007). The thinning operation deletes points, and the probability of a given point to be retained is given by \( \lambda(x)/\lambda^* \) where \( \lambda^* \) is the maximal value of \( \lambda(x) \) inside the plot. Biologically this means that dispersal limitation creates a pattern that can be described by a homogeneous Thomas process, but that habitat filtering removes trees from areas of low habitat suitability. If \( \lambda(x) \) is known, the parameters of the corresponding (pre-thinning) homogeneous Thomas process can be fitted to the observed pattern of a given species using the inhomogeneous \( K \)-function (Waagepetersen 2007). The heterogeneous Thomas process model is therefore a simple phenomenological description of aggregation, which however accounts explicitly for the effect of environmental heterogeneity and can therefore describe larger scale spatial dependency among species. The thinning perspective is also useful for simulation of this point process model: it is straightforward to simulate a homogeneous Thomas process model and then apply thinning to obtain a realization of the corresponding heterogeneous Thomas process. This point process therefore generates for each species a pattern that conserves the observed habitat association and clustering, but not species interactions.

**Fitting of point process models**

The parameters of the four point process models were estimated for each species using the methods presented in Waagepetersen and Guan (2009). For estimating the intensity function \( \lambda(x) \) we used the topographic variables and the tree density per quadrat for the Wabikon plot, and included the soil variables for the CBS plot as environmental variables. We computed the principal components (PCs) from the eight soil variables and used only the first two components as condensed variables because they explained 86.6% of total variance in soil variables. Variable condensing can reduce possible overfitting. The intensity function was then fitted for a given species pattern using maximum likelihood estimation to models of the form:

\[
\lambda(x) = \exp(\beta_0 + \beta_1v_1(x) + \ldots + \beta_nv_n(x))
\]

with coefficients \( \beta \) and the environmental variables \( v(x) \), where \( n = 6 \) for the CBS plot, and \( n = 4 \) for the Wabikon plot. To avoid potential overfitting due to small-scale aggregation (which is then fitted with the heterogeneous Thomas process model) we used a stepwise model reduction using Wald-tests (Waagepetersen and Guan 2009).

The parameters \( \varrho \) and \( \varSigma \) of the homogeneous Thomas model were fitted to the data using a minimum contrast method (Stoyan and Stoyan 1994) and the known analytical expression for the \( K \)-function. In case of the heterogeneous Thomas process model the fitting of the model parameters is done with the inhomogeneous \( K \)-function which incorporates the estimate of the intensity function \( \lambda(x) \).

**Evaluating the fit of the different hypotheses**

We calculated the FDARs and PDARs for the null communities generated from the four parameterized point process models in the same way we calculated them for the actual observed data obtained in the CBS and Wabikon plots. The predicted PDARs and FDARs for the four models were computed by averaging the PDARs and FDARs from 100 simulated communities and a 95% simulation envelope (CI) was constructed for each predicted PDAR and
FDAR. Two methods were used to evaluate these models (Supplementary material Appendix 1, Fig. A1). First, the observed PDARs and FDARs from the original data obtained in the CBS and Wabikon plots were compared with the PDARs predicted by the four models. A model is considered adequate if the observed PDARs and FDARs fall within the 95% simulation envelopes of the predicted PDARs and FDARs; otherwise, the model is rejected. Second, to identify the most parsimonious model we used Akaike's information criterion (AIC) to compare the four competing models based on the sum of residuals and the number of parameters in different models (Webster and McBratney 1989, Shen et al. 2009, Wang et al. 2011). All calculations were carried out in R ver. 2.12.2 (R Development Core Team), using the ‘spatstat’ package (Baddeley and Turner 2005).

**Results**

The two temperate plots (CBS and Wabikon) showed almost identical SAR curves at small neighborhoods up to 0.03 ha (Fig. 2). This area corresponds to circles with 10 m radius that contain at CBS and Wabikon on average 46 and 58 individuals, respectively. A small enlargement in area at this scale leads mostly to addition of a cluster of a new species (i.e. the ‘sampling phase’; Rosindell and Cornell 2009), which indicates that the small-scale spatial patterns of the two forests are similar. However, at areas larger than 0.03 ha, the species richness at the CBS plot increase quicker than that of the Wabikon plot despite a lower total number of individuals (Fig. 2A).

Phylogenetic diversity increased at both forests at small neighborhoods (< 0.04 ha) in the same way. At intermediate areas (i.e. 0.04–15 ha), phylogenetic diversity tended to be substantially higher in the CBS plot, and at the largest area both values coincided following a steep increase of the Wabikon PDAR (Fig. 2B). This indicates that species at distant parts of the Wabikon plot (which enter only at large quadrat sizes) are phylogenetically more diverse. Interestingly, the initial increase of phylogenetic diversity with species richness was almost identical at the two plots (up to 25 species) but then increased stronger at Wabikon (Supplementary material Appendix 2). The increase in functional diversity with area showed at both forests a very similar pattern (Fig. 2C). In addition, PDAR and FDAR, but not SAR, approximated a linear increase with log area over much of their range at both forests. Phylogenetic diversity and functional diversity showed a non-linear increase with species richness (Supplementary material Appendix 2).

In both temperate forest plots the random placement hypothesis received the least support among all four competing hypotheses in explaining the observed PDARs and FDARs (Table 1). It overestimated phylogenetic diversity at most scales (Fig. 3), although differences at the CBS

![Figure 2](image-url)  
Figure 2. Species, phylogenetic and functional diversity area relationships in CBS and Wabikon plots. (A) Species–area relationship. (B) Phylogenetic diversity–area relationship. (C) Functional diversity–area relationship.
further improved the prediction of PDAR (Table 1). However, the habitat filtering hypothesis still overestimated phylogenetic diversity at most areas (Fig. 3) as expected by our conceptual model (Fig. 1). This indicates that habitat filtering alone is not sufficient for explaining the PDARs in the studied temperate forests or that we missed important environmental variables.

Accounting for the effect of dispersal limitation yielded significantly improved predictions of PDAR in both forests compared with those made by the random placement and habitat filtering hypotheses (Fig. 3). This was also indicated by minimal AIC values among all four competing models for both forests (Table 1). Unexpectedly, this

plot were relatively small (Fig. 3A). Although the random placement hypothesis produced seemingly accurate values near the two ends of the PDARs, i.e. near 0 and 25 ha in the CBS plot or 0 and 25.2 ha in the Wabikon plot, this was an artifact because the total phylogenetic diversity and plot area were fixed regardless what models were applied. This is true for all other point processes in our analysis.

As expected, the habitat filtering hypothesis provided a better fit to the observed PDARs than random placement (Fig. 3; Table 1) because it considered the information from three topographic factors (elevation, slope and aspect) and one bioenvironmental factor (i.e. tree abundance in each quadrat). Considering soil factors in the CBS plot further improved the prediction of PDAR (Table 1). However, the habitat filtering hypothesis still overestimated phylogenetic diversity at most areas (Fig. 3) as expected by our conceptual model (Fig. 1). This indicates that habitat filtering alone is not sufficient for explaining the PDARs in the studied temperate forests or that we missed important environmental variables.

Accounting for the effect of dispersal limitation yielded significantly improved predictions of PDAR in both forests compared with those made by the random placement and habitat filtering hypotheses (Fig. 3). This was also indicated by minimal AIC values among all four competing models for both forests (Table 1). Unexpectedly, this

Table 1. Comparison of Akaike’s information criterion (AIC) among the four ecological process hypotheses: random placement hypothesis, habitat filtering hypothesis, dispersal limiting hypothesis, and the joint habitat filtering and dispersal limitation hypothesis.

<table>
<thead>
<tr>
<th></th>
<th>Random placement</th>
<th>Habitat</th>
<th>Dispersal</th>
<th>Habitat and dispersal</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>CBS-PD</strong></td>
<td>560.73</td>
<td>512.62, 494.42</td>
<td>373.06</td>
<td>417.39, 378.37</td>
</tr>
<tr>
<td><strong>Wabikon-PD</strong></td>
<td>646.18</td>
<td>577.62</td>
<td>492.87</td>
<td>525.73</td>
</tr>
<tr>
<td><strong>CBS-FD</strong></td>
<td>206.16</td>
<td>174.08, 155.8</td>
<td>65.46</td>
<td>83.69, 54.07</td>
</tr>
<tr>
<td><strong>Wabikon-FD</strong></td>
<td>265.83</td>
<td>198.02</td>
<td>119.84</td>
<td>125.18</td>
</tr>
</tbody>
</table>

Note: †without soil data, ‡with soil data.

![Figure 3](image-url)

Figure 3. The observed (dots) and predicted (colored lines) phylogenetic diversity area relationships for the data from the CBS and Wabikon plot. Panels (A) and (C) show species absolute phylogenetic diversity whereas panels (B) and (D) show the relative phylogenetic diversity (i.e. observed – predicted) for the different hypotheses. The vertical bars are the 95% simulation envelopes arising from simulation of the point pattern models representing different hypotheses. For clarity we do not show the vertical bars for two hypotheses based on topographic habitat only. The vertical dashed lines indicate quadrat lengths of approximately 30 and 200 m. Topographic habitat: habitat filtering hypothesis based only on topographic variables and stem density, full habitat for CBS: habitat filtering hypothesis based on stem density and topographic and soil variables, and habitat + dispersal: the corresponding habitat and dispersal limitation hypothesis.
Discussion

In this study we used the additional information contained in phylogenetic diversity area relationships (PDARs) and functional diversity area relationships (FDARs) to discriminate among four competing hypotheses of community assembly at two large temperate forest plots. Our main result suggests that aggregation of species due to mechanisms such as dispersal limitation provides a good approximation of the observed accumulation of phylogenetic and functional diversity with area, just as it did for the species area curves (Wang et al. 2011). This result also suggests that the assumption of independent species placement (Plotkin et al. 2000, McGill 2010, Wiegand et al. 2012) yields reasonable predictions of PDARs and FDARs given the inherent factors that constrain dispersal. Thus, these models that inherently ignore biotic interactions can still provide reasonable approximations of the diversity–area relationships examined.

However, there was some evidence that habitat associations (i.e. habitat filtering), although not the primary drivers of the patterns analyzed here, may improve the predictive
power of the models that assumed dispersal limitation. Wang et al. (2011) found that correctly simulating species turnover (i.e., distance decay curves) required combined consideration of habitat filtering and dispersal limitation in these forests, and we found that the combined habitat filtering and dispersal limitation model best explained the FDAR at CBS when soil variables were used. This result is also consistent with studies in tropical forests (Plotkin et al. 2000, Shen et al. 2009). The relative failure of the habitat filtering hypothesis by itself may also be attributed to omission of important environmental variables, leading to insufficient predictions of the species’ habitat associations. Nevertheless, the good match of the dispersal limitation hypothesis at smaller areas indicates that competitive species interactions (which may be positively correlated with phylogenetic and functional similarity of species) are not primarily responsible for the observed spatial patterns at the two forests. The alternative explanation of this pattern (Fig. 1) is that phylogenetically distant species converged to similar habitat requirements and increased the PDAR at the local scale and the within habitat scale. However, if this were the case we would expect different patterns in the departures between observed and expected curves from PDAR and FDAR which were clearly not found (cf. Fig. 3B and 4B, 3D and 4D). Instead, the departures in almost all cases followed the SAR (Fig. 2 and 3 in Wang et al. 2011). This finding suggests that PDAR and FDAR were basically driven by the SAR which means that phylogenetic or functional similarity may be relatively unimportant in structuring these forests.

Although the dispersal limitation hypothesis provided a good approximation of the PDAR and FDAR and was selected as the most parsimonious model among the four competing hypothesis analyzed here, it showed distinct departures from the observations. Interpreting these departures may point to additional processes of importance. In both forests, the dispersal limitation hypothesis tended to underestimate phylogenetic diversity at intermediate spatial scales (10–100 m) and overestimated phylogenetic diversity at larger scales at Wabikon (Fig. 3). This pattern agrees with our expectation based on the conceptual model (Fig. 1). Thus, the two forests may show a weak signal of competitive exclusion at the local scale and an additional effect of habitat filtering at the within habitat scale.

Unexpectedly, the combined habitat filtering and dispersal limitation hypothesis did not produce a better fit than the dispersal limitation only hypothesis at the Wabikon plot, where it overestimated PDAR and FDAR consistently (as did the habitat filtering only hypothesis). This indicates that some important habitat or historical factors were missing in the construction of the corresponding null communities and/or that the Thomas process did not fully account for the spatial aggregation pattern. Indeed, the Wabikon forest plot contains a patch of successional forest that was logged approximately 40 yr ago. Although species diversity at the Wabikon plot seems to be related to topographic factors (e.g., elevation), we suspect that disturbances may mask the effects of topographic factors and reduce the predictive ability of the heterogeneous Thomas process, which represents the combined effects of habitat filtering and aggregation.

The observed SAR and PDAR in the Wabikon forest plot accelerate strongly at larger spatial scales. This pattern is consistent with the presence of the successional forest and a distinct lowland patch that contains gymnosperm species (Pinus strobus, Thuja occidentalis, Tsuga canadensis) and several others (Alnus incana, Ilex verticillata, Populus balsamifera) that are rare or absent from other portions of the Wabikon plot. Both of these areas are located near the edge of the forest plot and are typically not fully incorporated into the area relationships until the largest spatial scales are evaluated. The species of gymnosperms increase both the overall species diversity of the forest plot as well as the overall PD, but the increase is not generally evident until the largest spatial scales have been incorporated. Interestingly, this pattern was not evident in our analysis of functional diversity. When plotting functional diversity over phylogenetic diversity we found that the curve for Wabikon was for all quadrat sizes consistently above that of CBS (Supplementary material Appendix 2).

The random placement hypothesis performed relatively well at the CBS plot and overestimated phylogenetic diversity only by a maximum of 13% and functional diversity by 11% at the 0.6 ha scale (2291 PD and 20 FD predicted; 2028 and 18 observed). In contrast, at Wabikon PD and FD were overestimated a maximum of 48 and 33% at the 1 and 0.3 ha scales, respectively (2278 PD and 24 FD predicted; 1541 and 18 observed). This result and the steeper increase of the PDAR for larger areas at Wabikon suggests that the CBS plot showed a greater mixing of distantly related species compared with the Wabikon plot. Indeed, some species at the Wabikon plot showed very patchy distribution patterns that resulted in hotspots of species diversity, phylogenetic diversity and functional diversity (Supplementary material Appendix 3). These trends are much weaker at the CBS forest (Supplementary material Appendix 4). Consequently, the consideration of habitat association (i.e. the habitat hypothesis) improved the PDAR and FDAR prediction at Wabikon considerably (Fig. 2–3), but only moderately at the CBS plot. In particular, local disturbance and pronounced variation in elevation contributed to strong spatial heterogeneity of tree distributions in the Wabikon plot. Given the difference in history between the CBS plot (old growth) and the Wabikon plot (second growth), one is tempted to suggest that mixing of phylogenetic and functional species attributes increases during ecological succession, a result that also has been documented by Hardt and Swank (1997).

Conclusions

In this study we used spatial patterns of phylogenetic and functional diversity to assist the species area relationship in discriminating among competing hypotheses of community assembly of trees in two temperate forests. We found that a model that accounts for the observed aggregation of species (e.g. caused by dispersal limitation) provided the most parsimonious explanation of the observed tree distributions. Spatial patterns of phylogenetic and functional diversity tended to mirror the species area relationship rather closely,
suggesting that phylogenetic and functional relationships are not the primary drivers of the patterns studied here. However, the dispersal limitation hypothesis tends to underestimate phylogenetic and functional diversity at intermediate spatial scales and overestimated it at larger scales. This pattern agrees with the expectation of competitive exclusion at small scales and habitat filtering at larger scales. Phylogenetic and functional relationships therefore may play a minor but detectable role in shaping the observed patterns in spatial species richness and phylogenetic and functional diversity. Our study illustrates how phylogenetic similarity and functional traits can inform studies that aim to identify the mechanisms underlying spatial patterns of species diversity in natural communities.

Acknowledgements – We especially thank the CBS and Wabikon temperate forest plot census and data management teams. This work was supported by the Knowledge Innovation Program of the Chinese Academy of Sciences (KSCX2-EW-Z-5, KZCX2-YW-QN402 and KZCX2-EW-401) and National Natural Science Foundation of China (3101120470 and 40971286). NGS was supported by the United States National Science Foundation (DEB – 1046113). TW was supported by the European Research Council (ERC) advanced grant 233066. We thank Jens-Christian Svenning for his constructive comments on the manuscript.

References


